A SYNTHETIC COHORT ANALYSIS OF OCCUPATIONAL ACHIEVEMENT AND MIGRATION

Angela Lane, Department of Sociology, Indiana University

Introduction

The American Occupational Structure [1], aside from an enormous contribution to the general study of stratification, replaces previous scanty research on migration and stratification with comprehensive, systematic analysis. The major concern, unresolved by earlier work, is whether the migration process selects largely higher status persons or rather confers advantages to migrants which are separate from whatever advantages or handicaps they initially possessed. These matters are discussed, as is the effect of place of origin on status in the place of destination. The analysis initially shows that greater mobility from father's status among migrants appears due to superior origins rather than mobility per se. Second, comparing migrants who moved between communities of the same size with non-migrants who remained in such communities, superior occupational status of migrants is largely, although not entirely, due to superior education and status of first job. The residual differences may be explained by other background factors; or by variations in the occupational structure which were not controlled by holding community size constant. However, it is also plausible to interpret this residual as the effect of migration on occupational status, above and beyond the advantages originally possessed by migrants, regardless of an invariant occupational structure, i.e., given migrants and non-migrants with the same backgrounds, if migrants moved to a community with a similar occupational structure, they would do better than those they left behind.

When migration between places of different size (with a likely change in the occupational structure) is also considered, (1) migrants do better than those they left behind, regardless of where they move--with one exception, men who move from a rural, non-farm community to a farm community; (2) migrants do better or worse than non-migrants at their place of destination, depending on whether the migrants had urban or rural origins. These patterns occur without controlling for background differences between migrants and non-migrants. When education and first job are controlled, the magnitude of the differences drops but is not eliminated, and the pattern described above remains unchanged. Thus, while most of the migrant/non-migrant status difference is due to selectivity, some of it can be interpreted as due to migration per se. If so, dim prospects can be left behind via migration, and opportunities at the destination can be turned to the advantage of most migrants regardless of the initial advantages or handicaps with which the migrants began.

While Blau and Duncan's findings permit interpretation of an independent effect of migration, their analysis did not specifically allow for such an effect since it assumed that the relationships between father's occupation, son's education, first job, and 1962 occupation were the same, but that migrants differed from non-migrants with regard to the means of these variables. Second, their analysis did not go beyond the effect of migration on status at an unspecified point in a career--for instance, to life-long effects.

However, the life-time pattern of occupational achievement needs to be known first. The literature on occupational achievement over time, when cross-sectionally defined regression and correlation coefficients are compared, indicates no observable trends [7]. Given the obvious problems inherent in cross-sectional comparisons, it is unfortunate that stratification studies of real cohorts are largely non-existent. A minor exception, the panel data available from the 1957 Princeton Fertility Study, is unfortunately limited in important respects as to make any conclusions questionable. (Cf. [7] for discussion of these points; also [3,4] for a discussion of limitations in the data.)

A less suitable alternative to following real cohorts is to use a synthetic cohort approach. The most familiar use of the synthetic cohort approach is that of the life table, where experience at a given age (e.g., 60) is the cumulative product of all prior ages (e.g., 20, 30, 40, etc.). The approach taken in the synthetic cohort study of occupational achievement does <u>not</u> combine all previous experiences, only father's status, son's education, and one or more prior occupations. Further, the synthetic cohort is not one hypothetical cohort followed over time (as in the life table) but rather a number of age groups assumed to be a single group observed at regular intervals over time. Duncan [2, p. 14] describes this process in terms of four age cohorts as follows:

> "Suppose we thought of the four sets of data as pertaining to a single cohort, studied at four successive points in time, at decade intervals. Then, all the data should fit into a single causal or processual sequence."

Last, this approach estimates the relationships between background variables and occupational achievement by means of regression analysis. While the synthetic cohort approach suffers from methodological limitations, it does permit provisional findings which can be tested.

Blau and Duncan [1, pp. 177-188], find that for age cohorts in the 1962 OCG study, (1) the direct effects of father's status and son's education are considerably weakened as time passes while (2) prior occupation becomes more influential for subsequent occupation with the passage of time.[1] Further, since the cumulative effect of earlier occupational status is incorporated into the model, the proportion of explained variance increases with time, and thus implies greater determination of status later in a career than earlier. Similar results were found in a synthetic cohort model based on Chicago data from the Six City Survey [2, pp. 13-14]. The major difference between the two sets of findings is that the importance of prior occupational status increases at a fairly steady rate in the OCG data, while the Chicago data for cohorts 35 to 44 and older show a sharp jump in the importance of occupation ten years earlier.

However, the adequacy of the Blau-Duncan causal chain model is questionable. (Featherman [3] indicates that the correlations could not be reproduced using this model). A major limitation of the Blau-Duncan and Duncan synthetic cohort models may lie in their simplicity. Insofar as they consider only one occupation prior to the occupation of interest, they neglect the fuller aspects of work history. Occupational shifts are especially common during the early phases of a career. Interests or skills developed during youth may be acted upon later in life if initial job preferences ultimately prove less desirable than initially hoped or if preferable alternatives appear. Thus, a simple causal chain model, such as that used by Duncan may be too rigid a representation of occupational achievement processes since it fails to allow for either flux in careers indicated by high rates of job shifts between occupational and industrial groupings [9,10,5], or for correlations of about .22, .72, and .90 between occupations held five years apart for men 25-34, 35-44, and 45-64, respectively [8].

The models developed in this paper are distinguished from previous work in that: (1) the effects of two prior occupations rather than only one on occupational attainment are investigated; (2) life-time career patterns are represented rather than occupational status at any single age; (3) the models are constructed for migrants and non-migrants separately, in order to determine if the relationships represented by the models differ according to migrant status.

Data

The data for this study are taken from the Six City Survey of Labor Mobility, conducted in 1951 [9]. Complete work histories for the period 1940 to 1951 were obtained from all members in the sampled households who were employed at least one month in 1951. Altogether, over thirteen thousand sample cases of men and women were collected in New Haven, Philadelphia, Chicago, St. Paul, San Francisco, and Los Angeles. The subsample considered here comprises 6820 non-veteran males of non-farm origins between 25 and 64 years of age as of 1951. Data were collected on age as of 1951, respondent's reported occupation as of January 1940, December 1944, and December 1949, educational attainment, and length of residence in city of current residence. Status of father and son was measured with prestige scores assigned to each occupational title represented in our sample.[2] Migrant status was assigned to persons who had moved into one of the six cities since 1949, and non-migrant to those not moving since 1940. All analysis was done within age groups (since geographical mobility is highly dependent on age) and for groups of both migrants and non-migrants.

The migrants and the non-migrants were subdivided into two overlapping sets of age cohorts. The first set included all persons in each of the following age intervals as of January 1951: 25 to 34, 35 to 44, 45 to 54, and 55 to 64. The second set included those within each of the following age intervals, as of January: 30 to 39, 40 to 49, 50 to 59 (the cohort aged 60 to 69 was dropped due to the small case base). These seven age breaks match the points in the life cycle of the synthetic cohort model to be introduced below.

The cohorts overlap (i.e., some of those 25 to 34 are also present in the 30 to 39 cohort), and this overlapping membership is true for all age cohorts. The rationale for overlapping the age-cohorts lies in the larger case base which is thereby available for each age cohort. As a result, some of the correlations are not independent of each other; however, these correlations are averaged together to arrive at an estimate of a single correlation. Otherwise, all correlations are independent of each other.

Table 1 presents the age of each cohort at the initial date of the survey (January 1951) and at each of the earlier dates for which occupation was reported. These ages are approximate insofar as the time points are not separated by the same length of time. Because this difference in interval length is unlikely to confound results to any significant extent, it has been ignored, and the length of time periods between observations presumed equal. For notational convenience, it is assumed that age in 1949 is the same as age in 1951, and that ages in 1944 and in 1940 are, respectively, five years and ten years younger. This alteration has no analytic effect or meaning.

The Synthetic Cohort Model

In a simulated cohort model, people born at different times are presumed to have the same experiences at similar ages. That such an assumption is reasonable with regard to occupational achievement is suggested by the absence of time trends in cross-sectional comparisons of stratification measures published thus far (see literature review in [7]). However, including migration may not satisfy the synthetic cohort assumption as easily. Migration is selective by age. Presumably, those who migrate at age fifty are different from those who migrate at age thirty, and the life-time patterns of younger migrants may not recapitulate those of older migrants. If so, then, to an unknown extent an assumption of the synthetic cohort model will be violated. Even given this possible violation, however, something may still be learned from a necessarily hypothetical model which permits interpretations verifiable later with real cohorts. The purposes of this paper, therefore, are primarily illustrative and methodological rather than substantive with respect to the effects of migration.

Following Duncan [2], the sets of data specific to the seven age cohorts have been treated as if they were a single cohort studied at seven successive points in time. The synthetic cohort begins with men aged 25 to 34 and follows them throughout their work careers, observing their occupational status at five-year intervals. If X_i represents occupational status at some given age, then X_1 is the initial point of observation of occupational status for men 25 to 34, X_2 the occupational status of this group five years later, when they are 30 to 39. Likewise,

X_3 = occupational status at 35 to 44,

X_4 = occupational status at 40 to 49,

X_5 = occupational status at 45 to 54,

X_6 = occupational status at 50 to 59,

X_7 = occupational status at 55 to 64.

Figure 1 (consider only the schematic diagram, ignoring the values for the moment) represents this model where X_i is occupation at five-year intervals of age and E = educational attainment.[3] Occupational attainment at any age (X_i) is treated as a function of both education (E) and occupational attainment at two younger ages, five years and ten years younger (X_{i-1}, X_{i-2}, respectively).

This formal model can be represented as:

$$X_i = b^*X_{i-1} + c^*X_{i-2} + d^*E + e' \qquad (1)$$

Thus, for example, $X_5 = b^*X_4 + c^*X_3 + d^*E + e'$. Equation (1) is estimated for each X_i by means of least squares regression analysis,[4] where all variables are standardized (with mean zero and unit variance), b^*, c^*, d^* are standardized regression coefficients, and e' is an error term.

All of the age cohorts in Figure 1 are shown in Table 1. Since occupational status observed at each age in Figure 1 is a function of occupational status observed five years and ten years younger (as well as of education--which is ignored for the moment, to simplify discussion), the correlations used to estimate equation (1) are those between occupations held at earlier dates by a given real cohort, when the cohort was five and ten years younger. For example, examining Table 1 for those 35 to 44 in 1951, a correlation between occupations held in 1949 and 1944 is also a correlation between occupations held at ages 35 to 44 and 30 to 39. Equivalent procedures were used to extract correlations between education and occupation.

One final comment is necessary. In some instances, using the method just described, more than one correlation coefficient for pairs of occupations observed at different ages may be obtained, e.g. the correlation between occupations held by those 25 to 34 and 30 to 39 can be found in two different places--referring back to Table 1: (1) for those 30 to 39 in 1951, observed at 1949 and 1944, and (2) for those 35 to 44 in 1951 observed at 1940 and 1944. For cases in which two correlations were available, the assumption was made that the two observations were sampling variations around some true correlation; to remove these variations, the correlations were averaged.

Findings

The estimates from equation (1) are presented as path diagrams, in Figures 1 and 2, for migrants and non-migrants, respectively. The regression coefficients b^*, c^*, d^* are the values near the straight connected lines leading to any X. Thus, in Figure 1, $X_5 = .448X_4 + .274X_3 + .161E$.

One clear finding is that occupational level at any age is most strongly affected by occupation held five years earlier, a lesser net effect is contributed by the occupation held ten years earlier, and the least net effect is contributed by education. The sizes of the residuals are large, a suggestion that factors other than those considered here affect occupational status.

Perhaps more important, a comparison of the values in Figure 1 with those on Figure 2 suggests that migrants differ from non-migrants in two respects. First compare the effect of education on occupational attainment at any age (excluding the observation on occupational attainment at ages 55 to 64, X_7, where the regression coefficient is negative; the assumption in this case is that the effect of education is not really negative, but that it ceases to be important). For three out of four comparisons of the regression coefficients for education and occupational status from 35 to 44 (X_3) to 50 to 59 (X_6), the net effect of educational attainment is higher for migrants than non-migrants. The only exception to this finding is for occupational level attained at ages 40 to 49 (X_4). Further, the correlation between education and occupational attainment at ages 25 to 34 (X_1) and 30 to 39 (X_2) is higher for migrants than for non-migrants. Comparison of these regression coefficients not only indicates migrants differ from non-migrants at any age, but how they differ over time. The size of the coefficients for non-migrants roughly declines, moving from X_3 to X_7, suggesting that education becomes less important for occupational attainment over their lifetimes. By contrast, the coefficients for migrants show a slight curvilinearity, suggesting that among older men who migrate, education retains importance for occupational success.

What of the effects of prior occupations? Of the ten regression coefficients in Figures 1 and 2 connecting current occupation with previous occupations, seven are higher for non-migrants than for migrants. The correlation between occupational attainment at ages 25 to 34 (X_1) and at ages 30 to 39 (X_2) is higher for non-migrants than migrants. Thus, for non-migrants current job is more dependent on jobs held in the past than for migrants. Non-migrants appear to put down roots not only where they live but in their careers.

Interestingly enough, as they age, non-migrants' current occupation seems to rely increasingly more on a recent occupation than on an earlier occupation. This finding becomes apparent from a comparison of the steady increase in the influence of occupation five years earlier, while the importance of occupation ten years earlier first rises and then declines with increasing age of the cohort.

By contrast, migrants as they age appear to experience no increasing dependency of current occupation on past occupation. If anything, the dependency is lessened for occupation held five years earlier: the net coefficients first increase and then decrease with age. The effect of occupation held ten years earlier is fairly stable (.23 - .27), with one exception (.47 between X_5 and X_7). These patterns over time add support to the assertion of basic differences between migrants and non-migrants at any single point in time: the careers of migrants appear to be less rooted than those of non-migrants in earlier occupations.

This difference is also suggested by the patterns the residuals take. At every point in the life cycle, the residuals are larger for migrants than non-migrants, an indication that the force of outside factors not included in the model is more substantial for migrants than for non-migrants. In addition, the residuals suggest that the life-cycle pattern differs for migrants and non-migrants. For migrants, the size of the residuals forms a fairly stable pattern, although it undergoes some variation in moving from age groups 35 to 44 (X_2) to 55 to 64 (X_7). For non-migrants the residuals decline steadily with increasing age, suggesting that exogeneous factors become less important with increasing age and that, if anything, one's past career becomes more important to one's future career. In these data migrants at older ages seem able to escape their occupational histories almost as well as younger migrants; the same cannot be said for non-migrants.

In sum, the findings suggest that migration affects the process of occupational achievement with respect to: (1) the magnitude of the relationships among education, earlier jobs, and current occupation at any single age, and (2) the patterns the achievement process takes over the cohorts' career histories.

With regard to the first finding, for both migrants and non-migrants, at any age in a career history, occupations held in the past continue to exert an influence on current occupation, but migration appears to weaken this relationship. On the other hand, the impact of education is generally greater for migrants than non-migrants at any age.

The synthetic cohort result as such--the pattern these relationships assume as the cohort moves through its life cycle--appears to depend on migrant status. Education steadily becomes less important over the life-time of non-migrants, but for migrants, the influence of education first declines and then increases at later ages. The effect of occupation held five years earlier becomes steadily more important for non-migrants as they age. Migration weakens this influence.

The results presented above suggest the following interpretations. Migrants seem better able to escape the confines of their previous occupations than non-migrants. The model also suggests the means by which they escape--education. The inevitability of career histories can be altered. Once having pursued one line of work, overcoming the limitation on options which accompanies the choice of any particular line of work seems to be aided by migration to a new labor market where one can make a new career choice, different from past job experience. To do so, however, the migrant must rely more heavily on his educational attainment.

Such a reliance is likely to be true from the employer's perspective also. The applicant with a work history specific to a particular community can readily be checked out by a prospective new employer. Hence, the applicant's educational attainment might be disregarded, given the availability of other, possibly more relevant, information from prior employers. If an applicant is new to the labor market, however, the retrieval of information from a past employer may be more difficult, and an employer is more likely to rely on a standard means of evaluation such as education.

<u>Summary</u>

The analysis reported above has provided hypotheses on the relationship of migration and occupational achievement which can be checked when data on real cohorts become available in the future. While the synthetic cohort approach may have some drawbacks with respect to migration, the provisional findings are of intrinsic interest in their own right and worth repetition. Migration appears to weaken the dependency of current occupation on past occupations and to enhance the usefulness of educational attainment for current occupational status at any age. Over a life-time, the importance of education wanes while that of earlier occupation increases for non-migrants' current occupation. Migration allows escape from occupational experiences which cumulate over the life cycle and become more important with age, but this escape depends on assets provided by education.

Footnotes

[1] Featherman's [3, p. 123] real cohort data from the Princeton Survey, despite its limitations, show "a similarity between the synthetic OCG model and the model for the Princeton subset."

[2] A study conducted in 1964 by the National Opinion Research Center on the prestige accorded to occupations has produced a score for each detailed Census occupational title. These Census titles were used to classify occupation in the Six City Study. The NORC study was conducted under the direction of Robert W. Hodge; the major report of the study will be published in a forthcoming volume. Hodge, Siegel, and Rossi [6] have shown that occupational prestige scores and ratings are stable for at least fifty years. Thus, 1964-specific scores can be assigned to occupations held in 1949 or 1940 without concern for the instability of scores over time.

[3] The conventions of such a schematic diagram are that a value associated with: (1) a curved line is a correlation coefficient; (2) a straight connected line is a regression coefficient, in standardized form; and (3) the straight unconnected line leading away from the diagram is a residual. Cf. Duncan [2] for further clarification.

References

[1] Blau, P.M., and Duncan, O.D., The American Occupational Structure, New York: John Wiley, 1967.

[2] Duncan, O.D., "Path Analysis: Sociological Examples," American Journal of Sociology, 72(July, 1966), 1-16.

[3] Featherman, D.L., The Socioeconomic Achievement of White Married Males in the United States: 1957-1967, unpublished doctoral dissertation, The University of Michigan, Ann Arbor, 1969.

[4] ____________, "A Research Note: A Social Structural Model for the Socioeconomic Career," American Journal of Sociology, 77(September, 1971), 293-304.

[5] Gallaway, L.E., Interindustry Labor Mobility in the United States, 1957 to 1960, Washington, D.C.: Government Printing Office, 1967.

[6] Hodge, R.W., Siegel, P.M., and Rossi, P.H., "Occupational Prestige in the United States: 1925-1963," American Journal of Sociology, 70(November, 1964), 286-302.

[7] Lane, A., "The Process of Occupational Achievement Over Time: A Decomposition of Change," unpublished manuscript, 1972.

[8] Lane, A., unpublished research findings from analysis of Six City Survey data.

[9] Palmer, G., Labor Mobility in Six Cities: A Report on the Survey of Patterns and Factors in Labor Mobility, 1940-1950, New York: Social Science Research Council, 1954.

[10] Parnes, H., Research on Labor Mobility: An Appraisal of Research Findings on the United States, New York: Social Science Research Council, 1954.

Acknowledgements

The author is indebted to Robert W. Hodge for guidance during the research, and to George J. Stolnitz and Robert M. Hauser for their review of earlier drafts of the manuscript and helpful suggestions. Editorial assistance was provided by Carolyn Mullins. This study was conducted while the author was a National Institute of Mental Health Fellow, 1968-70. Part of the analysis was supported by a grant from the National Science Foundation to the National Opinion Research Center, University of Chicago (#GS-85 "Occupations and Social Stratification").

TABLE 1. -- Age of observed cohorts adjusted for notational convenience

Actual Age January 1951	Assumed Age December 1949	Assumed Age December 1944	Assumed Age December 1940
First Set of Age Cohorts			
25-34	25-34	20-29	15-24
35-44	35-44	30-39	25-34
45-54	45-54	40-49	35-44
55-64	55-64	50-59	45-54
Second Set of Age Cohorts			
30-39	30-39	25-34	20-29
40-49	40-49	35-44	30-39
50-59	50-59	45-54	40-49
60-69	60-69	55-64	50-59

E = Educational status
X_1= Occupational status at 25-34
X_2= Occupational status at 30-39
X_3= Occupational status at 35-44
X_4= Occupational status at 40-49
X_5= Occupational status at 45-54
X_6= Occupational status at 50-59
X_7= Occupational status at 55-64

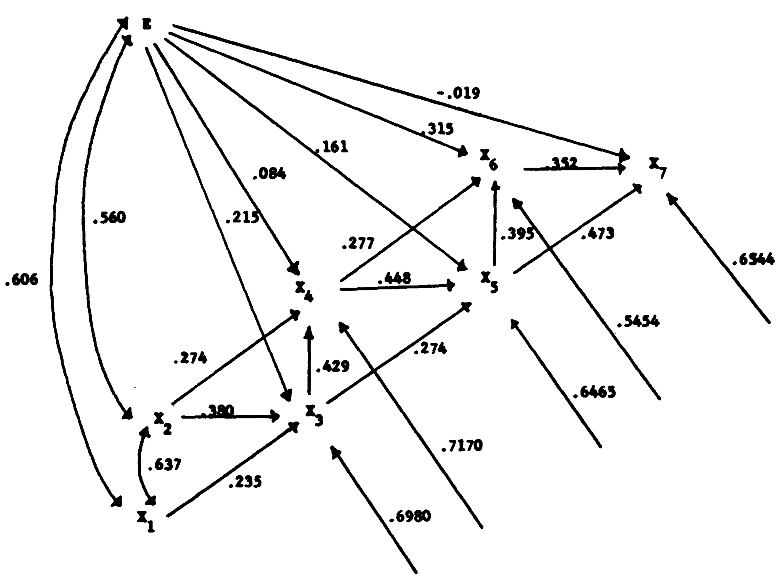


Fig. 1--Path analysis of intragenerational occupational achievement for migrants

E = Educational status
X_1= Occupational status at 25-34
X_2= Occupational status at 30-39
X_3= Occupational status at 35-44
X_4= Occupational status at 40-49
X_5= Occupational status at 45-54
X_6= Occupational status at 50-59
X_7= Occupational status at 55-64

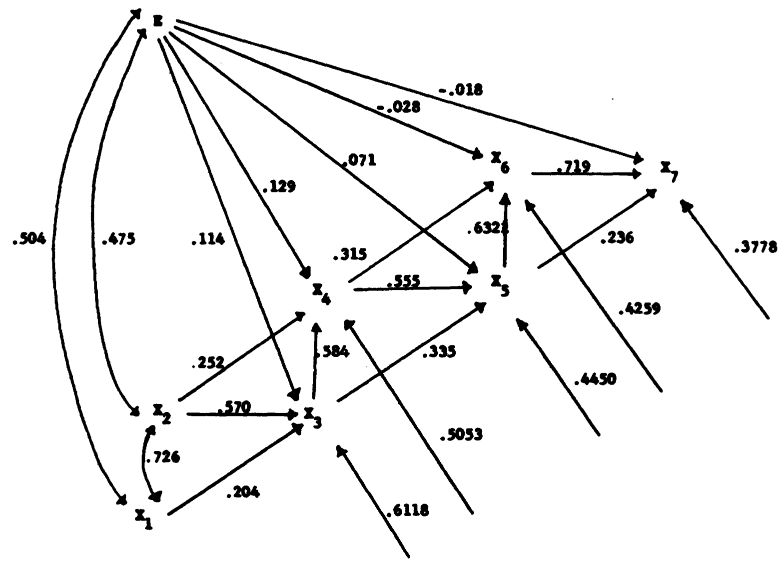


Fig. 2--Path analysis of intragenerational occupational achievement for non-migrants

TRENDS IN OCCUPATIONAL MOBILITY, 1962-1970

Robert M. Hauser and David L. Featherman, University of Wisconsin-Madison

In the past decade there has likely been as much concern among social analysts about trends toward "rigidification" in American society as in earlier periods. The first definitive measurements of trends in social mobility were taken in 1962 in the Current Population Survey (CPS) supplement, "Occupational Changes in a Generation" (OCG) [1] in a survey of 20,700 males aged 20-64. Despite earlier pessimism about rates and volume of mobility, the OCG data indicated substantial upward occupational mobility between generations, especially for younger cohorts [1] [3]. Furthermore, age-constant intercohort comparisons uncovered no appreciable changes in the regime connecting the occupations of fathers and sons [5].

Definitive estimates of trends since 1962 await a replication of the OCG survey in 1973 [6], but indirect evidence for changes in the pattern and volume of occupational mobility is presented here through an adaptation of a procedure used by Duncan [3] in analyzing trends between 1940 and 1962. Let $P = (p_{ij})$ be the transition matrix of an intergenerational occupational mobility table. Its elements represent the probability of a son's movement from the i^{th} category of father's occupation to a current occupation in the j^{th} category, and $\sum_j p_{ij} = 1.0$. Let $A = (a_i)$ be the origin vector of the mobility table, a row vector which gives the proportion of men who originate in the i^{th} occupation class, $\sum_i a_i = 1.0$, and let $C = (c_j)$ be the vector which gives the proportionate distribution of men over destination categories, $\sum_j c_j = 1.0$. Thus, we have the identity, $C = AP$. Likewise, we may also write $C = BQ$, where C is defined as before, while B is the vector of occupations of men in their first full-time jobs, and Q represents the matrix of transition probabilities from first to current jobs.

Functional notation identifies the vectors and matrices of men in a given cohort observed in a particular year. Thus, $C(r,s)$ is the occupation distribution of men in the r^{th} cohort in the s^{th} year, and so on. For a selected cohort and year, then, the transition from fathers' to current occupation distributions takes the form $C(r,s) = A(r,s)\ P(r,s)$. From the OCG survey we have estimates of C, A, P, B, and Q for cohorts within ages 20-64 in 1962. In order to make inferences about changes over time in P and Q we make the following assumptions: that within the prime working ages cohorts of U.S. males are closed with respect to mortality, net migration and movement into and out of the experienced civilian labor force and that the quality of data on current occupation, father's occupation, and first job does not vary with age or time. These assumptions have two pertinent consequences. First, for men born in year r, $A(r,s+t) = A(r,s)$ and $B(r,s+t) = B(r,s)$, where t may be greater or less than zero. This says that we may use the 1962 survey to estimate the origin vectors (fathers' occupations or first jobs) observed in any year for cohorts covered in the 1962 survey. Second, the assumptions imply that it is legitimate to compare observed destination distributions across years. Thus, we can make the age-constant intercohort comparison, $C(r,s)$ with $C(r+t,s+t)$, or the intracohort comparison $C(r,s)$ with $C(r,s+t)$. Obviously, our assumptions are not perfectly met, either as to population coverage or response quality, and our inferences are subject to substantial risks of measurement error.

Granting our assumptions, it becomes possible to make inferences about intercohort change in a mobility matrix. Consider the null hypothesis $P(r,1962) = P(r+t,1962+t)$, where we have observed only $P(r,1962)$. This says that the mobility matrix for men aged (1962-r) is unchanged t years later (or earlier). Under the null hypothesis we may write

$$\begin{aligned} C(r+t,1962+t) &= A(r+t,1962+t)\ P(r+t,1962+t) \\ &= A(r+t,1962+t)\ P(r,1962), \end{aligned}$$

which we can estimate by

$$_P\hat{C}(r+t,1962+t) = A(r+t,1962)\ P(r,1962),$$

since $A(r+t,1962+t) = A(r+t,1962)$ by assumption. We denote our estimate of the expected distribution here by $_P\hat{C}(r,s)$ in order to differentiate it from $_Q\hat{C}(r,s)$, the estimate based on the first job vector and the transition from first to current occupation. For example, we can estimate the 1972 occupation distribution (at age 35-44) of men born in 1927-36 (aged 25-34 in 1962) by applying the 1962 intergeneration transition matrix of men born in 1917-26 (aged 35-44 in 1962) to the origin vector of the younger cohort. The same logic applies to hypotheses about intercohort change in the intragenerational mobility matrix.

Comparisons among expected and observed distribution for recent years permit us to make limited inferences about change in mobility matrices in the past decade. While identity of destination vectors does not imply identity of transition matrices, differences between destination vectors clearly imply rejection of the null hypothesis (subject to the possibility that internal changes in the matrix are due solely to changes in the marginals and not at all to changes in interactions between rows and columns of the matrix).

It is possible to partition the net intercohort differences in occupation distributions for men of the same age into components attributable to intercohort changes in occupational origins, in the transition from father's occupation to first job, and in the transition from first job to current occupation. The necessary identity is

$$\begin{aligned} C(r+t,s+t) - C(r,s) &= [C(r+t,s+t) - {_Q\hat{C}}(r+t,s+t)] \\ &+ [{_Q\hat{C}}(r+t,s+t) - {_P\hat{C}}(r+t,s+t)] \\ &+ [{_P\hat{C}}(r+t,s+t) - C(r,s)]. \end{aligned}$$

The two terms in the first bracket on the right differ only because of intercohort differences in the transition matrix from first job to current occupation. This is,

$$C(r+t,s+t) = B(r+t,s+t)\ Q(r+t,s+t),$$

while

$${}_{Q}\hat{C}(r+t,s+t) = B(r+t,s)\ Q(r,s).$$

Thus, since $B(r+t,s) = B(r+t,s+t)$ by assumption, the difference between $C(r+t,s+t)$ and ${}_{Q}\hat{C}(r+t,s+t)$ is the effect of intercohort change in the transition from first job to current occupation on the net intercohort difference. To interpret the difference in the second bracket, denote the transition matrix from father's occupation to first job as $M(r,s)$. Then

$$P(r,s) = M(r,s)\ Q(r,s),$$

so

$${}_{P}\hat{C}(r+t,s+t) = A(r+t,s)\ M(r,s)\ Q(r,s).$$

Also,

$${}_{Q}\hat{C}(r+t,s+t) = A(r+t,s)\ M(r+t,s+t)\ Q(r,s)$$

since

$$B(r+t,s) = A(r+t,s)\ M(r+t,s+t)$$

by assumption. Thus, ${}_{P}\hat{C}(r+t,s+t)$ and ${}_{Q}\hat{C}(r+t,s+t)$ differ only because of intercohort change in the transition from father's occupation to first job, and their difference represents the effect of that change on the net intercohort difference.

Finally, the difference between the terms in the third bracket is the effect on the net intercohort difference of the intercohort shift in the distribution of sons by their fathers' occupations. Thus, $C(r,s) = A(r,s)\ P(r,s)$, while ${}_{P}\hat{C}(r+t,s+t) = A(r+t,s)\ P(r,s)$, which differs from the first expression only by virtue of changes between cohorts in the vector of occupational origins.

Because the CPS began using 1970 Census occupational coding materials in January 1971 [2], observed occupation distributions after that date are not strictly comparable with our expected distributions. Furthermore, had we been limited to tabulations by standard 10-year age-breaks, our efforts would have been stymied by the incompatibility of 1962 and 1972 occupation classifications. Since we have access to unit record tapes of the OCG survey, we have proceeded to make trend comparisons over a shorter period by varying the age-breaks in our origin vectors. Specifically, we have applied the transition matrices for those aged 35-44, 45-54, and 55-64 in 1962 to the origin vectors of those aged 27-36, 37-46, and 47-56 in March 1962 in order to generate expected distributions for men aged 35-44, 45-54, and 55-64 in March 1970. We obtained observed distributions in 1970 from the March 1970 Current

Table 1.-- Percentage distribution by occupation and net change, 1962-1970, by age: U.S. men in the experienced civilian labor force, March 1962 and March 1970

Occupation	35-44			45-54			55-64		
	1962	1970	Change	1962	1970	Change	1962	1970	Change
Professional, technical, and kindred workers									
Self-employed	1.91	1.85	-0.06	1.51	1.59	0.08	1.71	1.55	-0.16
Salaried	10.89	14.45	3.56	7.66	10.38	2.72	7.37	8.77	1.40
Manager, officials and proprietors, exc. farm									
Salaried	9.59	13.50	3.91	8.36	13.56	5.20	9.60	11.70	2.10
Self-employed	7.62	4.15	-3.47	9.94	5.42	-4.52	10.05	5.51	-4.54
Sales workers	5.14	4.93	-0.21	5.00	4.87	-0.13	3.99	5.63	1.64
Clerical and kindred workers	6.47	6.06	-0.41	6.66	6.78	0.12	5.92	6.47	0.55
Craftsmen, foremen and kindred workers	21.16	22.77	1.61	22.56	23.45	0.89	19.51	22.53	3.02
Operatives and kindred workers	19.10	18.93	-0.17	17.68	18.84	1.16	16.10	16.82	0.72
Service workers, including private household	4.86	4.69	-0.17	6.28	5.16	-1.12	7.91	7.57	-0.34
Laborers, except farm and mine	6.96	5.25	-1.71	6.53	5.24	-1.29	6.51	5.80	-0.71
Farmers and farm managers	4.92	2.46	-2.46	6.41	3.85	-2.56	9.22	6.05	-3.17
Farm laborers and foremen	1.39	0.96	-0.43	1.41	0.87	-0.54	2.11	1.60	-0.51
Total	100.00	100.00		100.00	100.00		100.00	100.00	
Number (1,000)	11,085	10,513		9,594	10,423		6,563	7,151	

Source: March 1962 OCG survey and March 1970 Current Population Survey (person tapes).

Population Survey person tape. In passing we should note that with freedom to vary age-breaks in both the OCG and CPS tabulations it is possible to make annual trend measurements at any desired ages.

NET INTERCOHORT SHIFTS, 1962-1970

The occupation distributions of men aged 35-44, 45-65, and 55-64 in 1962 and 1970 are compared in Table 1. The net intercohort shifts from 1962 to 1970 may be summarized as a fairly smooth continuation of the trends of earlier decades [4]. There were substantial intercohort shifts toward employment as salaried professionals and managers and smaller shifts toward employment as craftsmen, foremen and kindred workers. The former were largest at the two younger ages and the latter at the oldest age. Within the professional category there was no net shift toward self-employment; all of the net change was attributable to increases in salaried professionals. The growth among salaried managers was almost perfectly offset at each age by a substantial decline in the proportion of proprietors. (Our conclusions about shifts within the managers, officials, and proprietors category are unaffected by a 1967 procedural change in the CPS which improved the quality of self-employment reports. We estimate this change of procedure could account for a maximum shift of one percent of the male total from self-employed to salaried status.) A similar, but weaker pattern can be ascertained in net inter- and intragenerational shifts from 1952 to 1962 at younger ages in Duncan's 1965 paper on mobility trends (Table 4, p. 497). Only the decline in the proportion of farmers rivals that among self-employed managers, but the decline in the proportion of nonfarm laborers is also fairly large. The remaining categories show small downward shifts in their share of the occupation distribution.

While the March 1970 CPS estimated there were 7,151 thousand men aged 55-64 in the experienced civilian labor force, the number of men 45-54 in March 1962 estimated from the OCG survey was 9,104 thousand. The net loss of nearly 22 percent of the cohort, presumably due primarily to retirement and mortality, presents a serious threat to our assumption of closure. Thus, our findings for men aged 55-64 should be interpreted with great caution. In the sequel to this paper we expect to remedy this defect by changing our referent population and adding a category of "no occupation" to the destination distribution. In the two younger cohorts there is no prima facie evidence of severe violation of our closure assumption; the 1962 and 1970 estimated population totals differ by only 2.8 and 5.0 percent, respectively, for those aged 25-34 and 35-44 in 1962.

The components of intercohort change in the occupation distribution between 1962 and 1970 are shown in Table 2. The most striking feature of the table is the fact that virtually all of the net intercohort shifts in the occupation distribution are attributable to changes in the matrix of transitions from first jobs to current occupations. In no occupation group at any age is the effect of change in occupational origins or in the transition from origin to first job as large as one percentage point. Moreover, there are relatively few instances in which all three components are consistent as to sign.

With but one exception intercohort shifts in occupational origins at each age increase the chances that a man will become a professional, salaried manager, salesman or clerical worker, and they decrease the chances of his becoming a laborer or farmer. Shifting occupational origins have virtually no impact on the likelihood that a man will become a proprietor or a service worker. Since the occupation categories are listed in an order which approximates the socioeconomic ranking of major occupation groups from top to bottom, it is fair to conclude that the overall effect of intercohort shifts in occupational origins is to produce a slight upgrading of the occupation structure.

The transition from occupation origins to first jobs takes place over an interval in the life cycle which is roughly invariant with respect to calendar time. Thus, comparisons across ages of intercohort shifts due to changes in that transition matrix represent intertemporal change. At ages 35-44 changes in the origin-first job transition matrix place more men in professional and salaried managerial jobs and fewer as salesmen, clerical workers, craftsmen or operatives, while there are virtually no effects on the proportions of proprietors, service workers, laborers, or farmers. At ages 45-54 changes in the same transition matrix place more men as salaried professionals, proprietors, and craftsmen, and fewer are placed as salaried managers, salesmen, clerical workers, operatives, and farmers, while the remaining groups are virtually unaffected. At age 55-64 shifts in the origin-first job transition matrix lead to the placement of more men as proprietors, craftsmen and operatives and fewer as salaried professionals, salesmen, clerical workers, and farmers. In light of these observations and the modest size of the observed shifts we conclude that there are no consistent trends in the influence on the occupational structure of change in the transition matrices from occupational origins to first jobs.

Following the pattern of earlier decades [3], net intercohort shifts in the occupation distribution are largely attributable to changes in the transition matrix from first full-time jobs to current occupations. The components due to shifts in this transition matrix are similar across the age groups, and, of course, they are much like the net intercohort shifts described above. There are substantial positive shifts toward employment as salaried professionals and managers and as craftsmen, and there is a smaller positive shift into the operative category. There is a large shift away from proprietorship, and there are small, but consistent shifts out of the four lowest categories: service workers, farm and nonfarm laborers, and farmers. Finally, shifts involving self-employed professionals, salesmen, and clerical workers are generally small and form no consistent pattern across the age groups.

Overall, the components of intercohort change in the occupation distribution due to changes in the first job-current occupation transition matrix can be said to have increased opportunities for upward mobility. The seeming exception to this generalization, net movement out of the category of self-employed managers, may not be as much a contradiction as it appears. Proprietors are typically small businessmen, not the heads of large firms or corporations, and they have less

Table 2.-- Components of intercohort change in occupation distributions due to social origins and transitions from father's occupation to first occupation and from first occupation to current occupation: U.S. men in the experienced civilian labor force, March 1962 and March 1970

Occupation	35-44			45-54			55-64		
	Origins	Father's Occ to First Job	First Job to Current Job	Origins	Father's Occ to First Job	First Job to Current Job	Origins	Father's Occ to First Job	First Job to Current Job
Professional, technical, and kindred workers									
Self-employed	0.13	0.36	-0.55	0.02	-0.05	0.11	-0.02	-0.05	-0.09
Salaried	0.75	0.56	2.25	0.29	0.46	1.97	0.16	-0.32	1.56
Manager, officials and proprietors, exc. farm									
Salaried	0.27	0.17	3.47	0.18	-0.16	5.18	0.19	-0.09	2.00
Self-employed	0.01	0.01	-3.49	0.00	0.41	-4.93	0.10	0.28	-4.92
Sales workers	0.21	-0.21	-0.21	0.08	-0.13	-0.08	0.17	-0.15	1.62
Clerical and kindred workers	0.21	-0.30	-0.32	0.12	-0.23	0.23	0.10	-0.23	0.68
Craftsmen, foremen and kindred workers	-0.03	-0.35	1.99	-0.01	0.39	0.50	0.21	0.29	2.52
Operatives and kindred workers	-0.30	-0.28	0.41	-0.04	-0.14	1.33	0.20	0.19	0.33
Service workers, including private household	0.00	0.03	-0.20	0.07	-0.05	-1.14	0.05	0.11	-0.50
Laborers, except farm and mine	-0.28	0.07	-1.50	-0.06	-0.11	-1.12	-0.11	0.11	-0.71
Farmers and farm managers	-0.82	0.03	-1.67	-0.58	-0.27	-1.71	-0.87	-0.18	-2.12
Farm laborers and foremen	-0.17	-0.08	-0.18	-0.08	-0.11	-0.35	-0.18	0.04	-0.37

Source: March 1962 OCG survey and March 1970 Current Population Survey (person tapes).

Table 3.-- Indexes of dissimilarity representing components of intercohort change in occupation distributions at selected ages: U.S. men in the experienced civilian labor force, March 1962 and March 1970

Component of intercohort change	Age 35-44	45-54	55-64
Occupational origin	1.59	0.76	1.18
Transition from father's occupation to first job	1.22	1.26	1.02
Transition from first job to current occupation	8.12	9.32	8.71
Total intercohort change 1962-1970	9.08	10.16	9.43

Source: Tables 1 and 2

education and lower incomes than do salaried managers. The overall pattern of shifts due to change in the intracohort mobility matrices might be described as an upgrading of the occupational structure within both the manual and nonmanual sectors, accompanied by a smaller shift from manual to nonmanual occupations.

By 1970 the groups at the bottom of the occupation hierarchy from which there was net out movement during 1962-1970 contained 13.4, 15.1, and 21.0 percent of the experienced civilian labor force at ages 35-44, 45-54, and 55-64, respectively, compared to 18.1, 20.6, and 25.8 percent in 1962. By 1970 farm occupations included only 3.4 percent, 4.7 percent and 7.6 percent of the labor force at those ages. Thus, the possibilities for continued upward mobility are limited unless there appear new patterns of movement out of occupations in the middle of the hierarchy.

The differences between occupation distributions we have compared to form components of intercohort change are summarized using indexes of dissimilarity in Table 3. The index of dissimilarity is equal to the sum of positive percentage point differences between two distributions. It represents the percentage of cases in one distribution which would have to be shifted to a different category in order to make it identical to a second distribution. The relative sizes of the indexes on the first three lines in each row confirm our earlier observation that changes in occupational opportunities between cohorts are due primarily to changes in the transition matrix from first jobs to current occupations. The indexes for that transition are nearly as large as the indexes for the total intercohort comparisons, shown on the fourth line of Table 3. Further, the indexes of total intercohort change are almost as large as the sums of indexes over the component changes, which implies there is relatively little unnecessary net movement in the shifts between cohorts.

In Table 4 we present out estimates of components of intercohort occupational shifts during 1962-1970 due to changes in intergenerational and intragenerational mobility matrices along side Duncan's [3] estimates for men aged 35-44 and 45-54 in earlier periods. Note that the intergenerational effects shown here include the effects of changes in both the occupational origin-first job and first job-current occupation transition matrices. Unfortunately, we are unable to separate self-employed from salaried professionals prior to 1952.

The indexes of dissimilarity, shown at the base of each column, suggest that net changes in the mobility matrices had a larger effect on the occupational distribution during 1942-1952 than in 1952-1962 or 1962-1970. Because the professional and managerial categories are collapsed we have obviously under-estimated the decline in net occupational redistribution from 1942-1952 to the present, but the decline, if real, is surely not monotonic; shifts in the occupation distribution due to changing mobility patterns are clearly larger during 1962-1970 than in 1952-1962 both at ages 35-44 and 45-54.

At age 35-44 changing mobility matrices produced more movement into professional employment during 1952-1962 than in either 1942-1952 or 1962-1970. At age 45-54 there was no clear pattern of change between 1952-1962 and 1962-1970. There has been a clear shift away from the category of managers, officials and proprietors in the past three decades. At age 35-44 there was a net shift of 3.4 percent due to changes in intragenerational mobility during 1942-1952, but no net shift during 1962-1970. The apparent explanation is a continuing net movement into the ranks of salaried managers, compensated by net movement away from proprietorship, where both sorts of changes occurred more rapidly during 1962-1970 than in the preceding decade. There have been essentially no net movements into or out of sales or clerical occupations during the period covered by Table 4.

Table 4.-- Differences, in percentage points, between occupation distributions for men of specified ages produced by 1962 intergeneration and intrageneration mobility matrices and by matrices for earlier and later years

Occupation	Intergeneration mobility 35-44 1952 -1942	Intergeneration mobility 35-44 1962 -1952	Intergeneration mobility 35-44 1970 -1962	Intergeneration mobility 45-54 1962 -1952	Intergeneration mobility 45-54 1970 -1962	Intrageneration mobility 35-44 1952 -1942	Intrageneration mobility 35-44 1962 -1952	Intrageneration mobility 35-44 1970 -1962	Intrageneration mobility 45-54 1962 -1952	Intrageneration mobility 45-54 1970 -1962
Professional, technical, and kindred workers	0.9					1.7				
Self-employed		0.3	-0.2	-0.1	0.1		0.3	-0.6	0.0	0.1
Salaried		3.5	2.8	1.5	2.4		2.8	2.2	1.9	2.0
Manager, officials and proprietors, ex. farm	3.1					3.4				
Salaried		2.4	3.6	0.7	5.0		2.4	3.5	0.9	5.2
Self-employed		0.0	-3.5	1.0	-4.5		-0.4	-3.5	0.5	-4.9
Sales workers	-1.5	0.1	-0.4	0.3	-0.2	-1.3	0.3	-0.2	0.3	-0.1
Clerical and kindred workers	0.2	0.4	-0.6	0.6	0.0	0.5	0.7	-0.3	0.9	0.2
Craftsmen, foremen and kindred workers	3.7	-0.8	1.6	0.1	0.9	3.3	-1.0	2.0	0.2	0.5
Operatives and kindred workers	3.3	-2.0	0.1	0.0	1.2	2.7	-1.8	0.4	-0.4	1.3
Service workers, including private household	-1.1	-0.3	-0.2	-0.3	-1.2	-1.4	-0.3	-0.2	-0.2	-1.1
Laborers, except farm and mine	-3.1	-0.5	-1.4	-1.0	-1.2	-3.3	-0.4	-1.5	-1.1	-1.1
Farmers and farm managers	-3.5	-2.6	-1.6	-2.3	-2.0	-3.5	-2.3	-1.7	-2.5	-1.7
Farm laborers and foremen	-2.0	-0.5	-0.3	-0.5	-0.5	-2.1	-0.3	-0.2	-0.5	-0.4
Index of dissimilarity	(11.2)	(6.7)	(8.2)	(4.2)	(9.6)	(11.6)	(6.5)	(8.1)	(4.7)	(9.3)

Source: March 1962 OCG survey and March 1970 Current Population Survey (person tapes) and O. D. Duncan, "The Trend of Occupational Mobility in the United States," American Sociological Review 30 (August, 1965): Table 4, p. 497.

At age 35-44 there was substantial net movement into the ranks of craftsmen and operatives in 1942-1952, and there were small net shifts away from and into those categories in 1952-1962 and 1962-1970 respectively. At age 45-54 there were essentially no shifts in the craft and operative categories due to changing mobility regimes between 1952 and 1962.

There is a consistent pattern of net movement out of the four lowest manual occupation categories. The net shift away from the two farm categories appears to have declined continuously (along with the relative numbers in those categories) over the three decades. Shifts away from services and nonfarm labor were smaller in 1952-1962 than in the preceding decade, but the net movement may have increased again from 1962 to 1970.

In our view the major research issue posed by our findings is the structure of change in intragenerational mobility matrices. Our further explorations in this area will begin with an effort to adjust that matrix to take account of intercohort shifts in educational attainment. In addition we shall apply procedures like those used here to examine changing mobility patterns among black men.

ACKNOWLEDGMENTS

This research was supported by the National Science Foundation (Grant No. GI-31604X) and by the College of Agriculture and Life Sciences, University of Wisconsin, Madison. The authors wish to thank Otis Dudley Duncan for suggesting that we, rather than he write this paper and John M. Bregger for supplying unpublished tabulations of the Bureau of Labor Statistics, U.S. Department of Labor. Computations were carried out using facilities of the Center for Demography and Ecology by Peter J. Dickinson with the assistance of Hernando Gomez-Buendia and James R. Kluegel.

REFERENCES

[1] Blau, Peter M., and Otis Dudley Duncan, The American Occupational Structure. New York: John Wiley and Sons, Inc. 1967.

[2] Bregger, John E., "Revisions in Occupational Classifications for 1971" Employment and Earnings (February):5-8. U.S. Department of Labor, Bureau of Labor Statistics. 1971.

[3] Duncan, Otis Dudley, "The Trend of Occupational Mobility in the United States," American Sociological Review 30 (August):491-498. 1965.

[4] ________________, "Occupation Trends and Patterns of Net Mobility in the United States," Demography 3:1-18. 1966.

[5] ________________, "Patterns of Occupational Mobility among Negro Men," Demography 5: 11-22. 1968.

[6] Featherman, David L. and Robert M. Hauser, "Design for a Replicated Study of Social Mobility in the United States," Working Paper 72-6, Center for Demography and Ecology, University of Wisconsin, Madison. 1972.

LEAST-SQUARES ESTIMATION OF EFFECTS ON INTER- AND INTRAGENERATIONAL OCCUPATIONAL TRANSITION PROBABILITIES[1]

Richard A. Wanner and E. M. Beck, University of Wisconsin

Introduction

A great deal of attention has been devoted recently to the development of models for the estimation of the relative influence of a set of explanatory conditions on binary dependent variables [8,9,10,11,12,13,15,18]. The present paper presents the application of one of the suggested techniques to two social processes which lend themselves to conceptualization within this framework: intergenerational and intragenerational occupational mobility.

Individuals can be characterized as having certain probabilities of making a transition from a prescribed set of "origin" positions into some set of "destination" positions. We will utilize a model which provides for the introduction of independent conditions to explain variations in the probabilities of movement out of these origin positions.

Formulation and Estimation

The simplest approach to the problem is to employ an ordinary least-squares (OLS) model

$$Y_i = \beta_0 + \sum_1^j \beta_j X_{ij} + \varepsilon_i \qquad (1)$$

in which Y_i assumes a value of unity if the i^{th} person was a mover from a particular origin position to a specified destination position during the time period of interest [cf. 12,16]. Since the expected value of Y_i is the population transition probability between the origin and destination

$$E(Y_i) = Pr(Y_i = 1) = P_i$$

positions, equation (1) can be expressed as the linear probability function (LPF):

$$E(Y_i | \sum_1^j X_{ij}) = Pr(P_i | \sum_1^j X_{ij}) = \beta_0 + \sum_1^j \beta_j X_{ij} \qquad (2)$$

The unstandardized regression coefficients of the LPF are interpreted straightforwardly as the relative contribution of each of the independent variables to the transition probability [16].

As pointed out by Theil [18] and Neter and Maynes [13], among others, this model has two major defects. The first difficulty involves the distribution of the error variances. It can be shown that the error variances have a binomial distribution [7,8],

$$V(\varepsilon_i) = P_i(1 - P_i) = P_i Q_i$$

thus the homoscedasticity assumption of OLS can not be met. This problem is readily solved by employing a weighted least-squares (WLS) solution for the parameter estimates [7,8,15,18]. The WLS estimators are minimum variance linear unbiased estimators of the LPF.

The second problem with the LPF is that the expected value of the dependent variable, the predicted transition probability, can exceed unity or be less than zero for any given observation. This undesirable outcome is the result of the left-hand side of equation (2) being constrained to vary between 0 and 1 while the right-hand side is under no such limitation.

Two recommendations have been advanced for coping with this problem. Speare [15] and Huang [7] suggest that the simplest solution might be to truncate the predicted transition probability at 0 and 1. However, since the first derivative of the conditional expectation is no longer continuous, an exact solution for the parameter estimates is not possible and an iterative estimation procedure must be employed [15]. While this approach maintains the integrity of the interpretation of the unstandardized coefficients, it is costly in computational terms.

Theil [18] and Goodman [9,10], among others, recommend a second strategy: transform the dependent variable such that the 0,1 constraint no longer presents difficulties. Specifically, Theil suggests using the "logit" (logarithmic unit) transformation:

$$L_i = \log_e \frac{P_i}{Q_i}$$

which has the effect of producing a dependent variable, L_i, which approaches $+\infty$ as the probability of moving, P_i, approaches unity and $-\infty$ as P_i approaches zero. The statistical model then becomes:

$$E(L_i | \sum_1^j X_{ij}) = \beta_0 + \sum_1^j \beta_j X_{ij} \qquad (3)$$

Estimates of the parameters then measure the relative effects of the explanatory variables on the logit itself. After solving for the WLS estimates of the parameters, the expected transition probabilities can be readily retreived:

$$\hat{P}_i = (1 + e^{-\hat{L}_i})^{-1}$$

By comparing directly estimates derived from the LPF, the truncated LPF, and the logit models, Speare [15] concludes that the latter two are superior since they tend to assign larger effects to the independent variables thus allowing for the incorporation of a greater number of explanatory conditions in a given model. While the truncated LPF does maintain one advantage over the logit specification, the clear interpretation of the unstandardized coefficients, it is our position that the logit model is to be preferred on theoretical grounds as well as for its greater computational facility.

The truncated LPF implies that it is possible that the predicted transition probability for any given observation *may* be 0 or 1; it is difficult for us to believe that the probability of mobility would ever reach these extremes. It is more likely that there is always a small probability of movement out of an origin position, although this probability may be quite small. On the other hand, it is equally difficult for us to conceive of a situation in which it is certain that an individual will move regardless of the circumstances. The logit specification implies that the probability for stayers never attains zero, and likewise, the probability for movers never reaches unity. We find these implications more satisfying than those implied by the truncated LPF. The logit specification will be used in this paper.

City Effects on Intergenerational Mobility

The first application of the model is to an exploration of the extent to which the social context of urban centers influences intergenerational occupational mobility. Students of social mobility have increasingly turned to the city as the locus of change in the structure of stratification since it is the urban milieu which tends to generate the key processes having a direct effect on mobility rates (see, for example, Schnore [14]). In brief, we argue, as has Duncan [5] , that American cities are organized in a system in which functional specialization is overlaid with a set of dominance relationships based upon the relative concentration of financial and commercial activities in the city.[2] Regardless of functional specialization, as indexed by industrial composition, it is hypothesized that a more dominant place will incorporate a larger proportion of highly ranked coordinative and administrative positions. In addition, it is expected that such city effects will be more pronounced for those individuals who have remained in their place of origin, since only then can the differential opportunity structure be expected to operate at the individual's point of entry into the labor market. The introduction of a control for migration status is thus essential to establish the existence of such an effect.

The data to be utilized in the following analysis are those collected by the Bureau of the Census for Blau and Duncan's [2] major survey of the extent and sources of social mobility in the United States, "Occupational Changes in a Generation" (OCG). As an adjunct to the Bureau's monthly "Current Population Survey" (CPS), the OCG questionnaire, or a follow-up interview, was administered to a national sample of 20,700 males between 20 and 64 years of age.[3] Since the present analysis will utilize SMSA's as units of analysis, it is crucial that the sample be representative of intra-SMSA populations. Our best assurance of this is that in the 1962 CPS all SMSA's were coterminous with primary sampling units. Only the fifteen largest SMSA's for which OCG data were collected were coded separately, and it is to these that the analysis must be restricted.

The dependent variable is the observed probability of mobility from any origin status (father's occupation) upward into a destination status (son's occupation) where coordinative or administrative functions are performed; coordinative or administrative positions are operationalized here as professional, managerial, and clerical positions. This probability is calculated from the relative frequency of movers in categories determined by the independent variables:

$$X_1 = \begin{cases} 1 & \text{If the person resides in the same city he did at age 16} \\ 0 & \text{Otherwise} \end{cases}$$

TABLE 1

UPWARD MOBILITY INTO COORDINATIVE POSITIONS DETERMINED BY POSITION OF CITY OF RESIDENCE ON URBAN SYSTEM DIMENSIONS AND MIGRATION

Position on Industry Dimension	Migrants		Non-migrants	
	Position on Metropolitan Functions Dimension			
	High	Low	High	Low
	Probability of Upward Mobility			
Manufacturing	0.301 (923)*	0.320 (263)	0.287 (487)	0.234 (286)
Service	0.323 (646)	0.378 (130)	0.302 (463)	0.279 (53)
	Estimated Probabilities			
Manufacturing	0.305	0.297	0.267	0.260
Service	0.336	0.327	0.296	0.288
	Discrepancies**			
Manufacturing	-0.004	0.023	0.020	-0.026
Service	-0.013	0.046	0.006	-0.009

*Number of cases in cell. For purposes of inference, these N's have been adjusted to approximate simple random sampling.

**$\chi^2 = 4.614$ (NS)

$$X_2 = \begin{cases} 1 & \text{If the person resides in a city high on the metropolitan functions hierarchy} \\ 0 & \text{Otherwise} \end{cases}$$

$$X_3 = \begin{cases} 1 & \text{If the person resides in a manufacturing-oriented city} \\ 0 & \text{Otherwise} \end{cases}$$

The first panel of Table 1 indicates the observed values for the conditional probabilities of upward mobility into coordinative positions, along with the corresponding adjusted sample sizes.

Defining p_i as the observed probability of occupational mobility, in the i^{th} subset, the observed logits were obtained:

$$L^*_i = \log \frac{P_i}{1 - P_i} \qquad (5)$$

The following linear logit specification was employed:

$$\hat{L}^*_i = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3 \qquad (6)$$

Table 2 presents the WLS estimates of these coefficients and their asymptotic standard errors. As expected, migrants are significantly more likely to be upwardly mobile into coordinative positions, regardless of city of residence.[4] Our initial hypothesis, that residence

TABLE 2

WLS ESTIMATES OF THE MAIN EFFECTS OF EXPLANATORY CONDITIONS ON UPWARD MOBILITY INTO COORDINATIVE POSITIONS

Effect	Estimated Coefficient	Standard Error
Constant	-0.882	0.063
Due to non-migration	-0.185	0.077
Due to residence in a city low on metropolitan hierarchy	-0.041	0.088
Due to residence in a service-oriented city	0.142	0.078

in a city ranked high on the metropolitan functions hierarchy will enhance the probability of upward mobility into coordinative positions, is unconfirmed, since this estimate is exceeded in size by its standard error. It is interesting that a small, but nonsignificant interaction between migrant status and rank of city on the metropolitan hierarchy is evident in the discrepancies between predicted and observed probabilities presented in the third panel of Table 1.[5] Here, the relationship is in the hypothesized direction for non-migrants, but not for migrants.

An unanticipated finding is the marginally significant effect of residence in service-oriented cities evident in Table 2. It is our presumption that such an effect is explainable by variations in the occupational composition of manufacturing vs. service-oriented cities: places with a relatively high concentration of service functions simply contain a higher proportion of coordinative positions to be filled. To test this hypothesis we apply a procedure developed by Deming [3] and applied to mobility tables by Duncan [4] in which the entire table for each city is proportionally adjusted to a new set of marginals: the distributions of sons by their own occupations and by their fathers' occupations. In the present case these adjustments are to the marginals for all SMSA's with population over one million. The question this procedure enables us to answer is whether or not the observed effect of residence in a service-oriented city is produced by an opportunity structure possessing a greater demand for individuals to fill coordinative positions, i.e. will the proportional adjustment of the mobility matrices for individual SMSA's allow us to account for this effect, or are there some additional unmeasured variables which operate to produce nonproportional differences across the fifteen matrices.

As the first panel of Table 3 attests, the proportional adjustment of the mobility matrices rather drastically reduces variability among the observed probabilities. In Table 4 we can see that the previously observed relationship between the probability of upward mobility and residence in service-oriented cities disappears, indeed is reversed to a non-significant degree. Apparently the observed effect of residence in service-oriented cities is wholly explicable by differentials in the opportunity structures of such cities, as opposed to those of manufacturing cities. Again, the model contains no significant interactions.

Industrial Development and Intragenerational Occupational Movers and Stayers

The second example comes from a comparative investigation of the short-term effects of the introduction of a large-scale highly automated steel plant in an agrarian region of the Midwest in 1966.[6] The basic research problem is to assess the relative influence of several explanatory variables on the probability of a given worker changing position in the occupational structure during the period of industrial development, 1966 to 1971. In short, we would like to specify the conditions which affect the transition rates among occupations. Our major interest here is not upon transition rates between specific origin and destination occupations, but upon differentiating occupational movers from stayers. Data for the analysis came from retrospective work histories collected in 1971 from a sample of 778 employed heads of households in two areas of Illinois. The first area comprised townships surrounding the plant location and is designated as the "experimental" region; the second area was a "control" region that has not undergone industrial development.

TABLE 3

OBSERVED, PREDICTED, AND DISCREPANCIES IN PROBABILITY OF UPWARD MOBILITY BY POSITION OF CITY OF RESIDENCE ON URBAN SYSTEM DIMENSIONS AND MIGRATION, OCCUPATIONAL DISTRIBUTIONS ADJUSTED

Position on Industry Dimension	Migrants		Non-migrants	
	Position on Metropolitan Functions Dimension			
	High	Low	High	Low
	Probability of Upward Mobility			
Manufacturing	0.305 (923)*	0.305 (263)	0.293 (487)	0.299 (486)
Service	0.299 (646)	0.274 (130)	0.281 (463)	0.375 (53)
	Estimated Probabilities			
Manufacturing	0.303	0.307	0.295	0.300
Service	0.296	0.301	0.288	0.293
	Discrepancies**			
Manufacturing	0.002	-0.002	-0.002	-0.001
Service	0.003	-0.027	-0.007	0.082

*Adjusted sample N's.

** $X^2 = 2.344$ (NS)

TABLE 4

WLS ESTIMATES OF MAIN EFFECTS OF EXPLANATORY CONDITIONS ON UPWARD MOBILITY INTO COORDINATIVE POSITIONS, OCCUPATIONAL DISTRIBUTIONS ADJUSTED

Effect	Estimated Coefficient	Standard Error
Constant	-0.834	0.063
Due to non-migration	-0.038	0.077
Due to residence in a city low on metropolitan hierarchy	0.024	0.087
Due to residence in a service-oriented city	-0.033	0.079

Three explanatory variables were used in trying to account for the overall mobility rates: (1) the region in which the worker resided; (2) the length of time the worker has spent active in the labor force; (3) and whether or not the worker was employed in "blue collar" types of employment.[7] The following definitions were used:

$X_1 =$
- 1 If the worker resided in the "experimental" region during the period 1966 to 1971
- 0 Otherwise

$X_2 =$
- 1 If the worker had been active in the labor force for more than 10 years as of 1966
- 0 Otherwise

$X_3 =$
- 1 If the worker was employed in a "blue collar" occupation at the beginning of industrial development
- 0 Otherwise.

After partitioning the data by these three conditions, the proportions of occupational movers were computed for each subclassification. These proportions are presented in the first panel of Table 5.

After transforming these observed probabilities by equation (5), the linear logit model, (6), was estimated using WLS. The weighted least-squares estimates of the coefficients, and their asymptotic standard errors, are presented in Table 6. It was found that the effects of being in the experimental region and having been in the labor force for more than ten years tend to decrease the probability of occupational mobility whereas having a blue collar occupation at the beginning of the industrial development period tended to increase the probability of mobility. The largest influence of the transition probabilities was the length of time spent in the labor force while regional residence tended to have the smallest impact.

The third panel of Table 5 presents the discrepancies between the observed and predicted proportions of occupational movers. It should be noted from Table 5 that the linear logit model predicts fairly well for all the subclassifications except for white collar workers in the control region who have been in the labor force for less than ten years. For this group the model underestimates the observed transition

TABLE 5

OBSERVED, PREDICTED, AND DISCREPANCIES IN PROBABILITY OF OCCUPATIONAL MOBILITY BY REGION, OCCUPATIONAL CATEGORY, AND TIME IN LABOR FORCE

	Experimental Region		Control Region	
	Blue Collar	White Collar	Blue Collar	White Collar
	Probability of Mobility			
In Labor Force Less Than Ten Years	0.346 (78)*	0.250 (36)	0.333 (30)	0.471 (17)
In Labor Force More Than Ten Years	0.247 (291)	0.158 (152)	0.250 (120)	0.241 (54)
	Estimated Probabilities			
In Labor Force Less Than Ten Years	0.348	0.283	0.398	0.328
In Labor Force More Than Ten Years	0.234	0.184	0.274	0.218
	Discrepancies**			
In Labor Force Less Than Ten Years	-0.002	-0.033	-0.065	0.143
In Labor Force More Than Ten Years	0.014	-0.026	-0.024	0.023

*Number of cases in cell.

** $X^2 = 2.672$ (NS)

TABLE 6

WLS ESTIMATES OF MAIN EFFECTS OF EXPLANATORY CONDITIONS ON INTRAGENERATIONAL MOBILITY

Effect	Estimated Coefficient	Standard Error
Intercept	-0.719	0.238
Due to residence in the experimental region	-0.212	0.181
Due to being in the labor force more than ten years	-0.562	0.192
Due to having a "blue collar" occupation	0.305	0.183

probability rather severely.

Conclusions

In summary, this paper has focused on some of the problems associated with estimating the relative influence of a set of explanatory conditions on binary dependent variables and presented two applications of a linear logit model to the prediction of the probability of movement from some set of origin positions to a set of destination positions. These applications have demonstrated the utility of such a model in introducing independent variables to account for specific patterns of movement in both inter- and intragenerational occupational mobility matrices.

FOOTNOTES

[1]The first co-author was supported in this work by grants from the National Science Foundation (GI-31604X), David L. Featherman and Robert M. Hauser co-principal investigators, and the research committee of the Graduate School, the University of Wisconsin, Madison. Support for the second co-author was given by the College of Agriculture and Life Sciences at the University of Wisconsin, Madison, and PHS grant MH-19689 from the National Institute of Mental Health, Gene F. Summers and John P. Clark co-principal investigators.

[2]Applying suggestions by Galli [6] of a multiple variable approach to community structure and Winsborough [20] of a Q-factor analytic technique, we sought to test the expectation of Duncan [5] that a system of large urban places can be described by the two dimensions discussed above and rank the 55 largest SMSA's on them. See Wanner [19] for a more complete treatment of this procedure.

[3]For a fuller discussion of the survey design and characteristics of the data set, see Blau and Duncan [2] .

[4]As Blau and Duncan [2] indicate, the greater achievement among migrants is explainable primarily by their generally superior backgrounds, especially son's education and father's SES.

[5]Theil [18] suggests what is essentially a Chi-square goodness-of-fit test between the observed and predicted probabilities for testing the validity of the model.

[6]For further details concerning this investigation see Beck [1] and Summers [17].

[7]"Blue collar" occupations included farmers, craftsmen, operatives, service workers, and laborers.

REFERENCES

[1] Beck, E. M., A Study of Rural Industrial Development and Occupational Mobility, unpublished Ph.D. dissertation, The University of Tennessee, Knoxville, Tennessee, 1972.

[2] Blau, Petèr M. and Otis Dudley Duncan, The American Occupational Structure, New York: John Wiley and Sons, 1967.

[3] Deming, W. Edwards, Statistical Adjustment of Data, New York: John Wiley and Sons, 1943.

[4] Duncan, Otis Dudley, "Methodological Issues in the Analysis of Social Mobility," in Neil J. Smelser and Seymour Martin Lipset (eds.), Social Structure and Mobility in Economic Development, Chicago: Aldine, 1966, pp. 51-97.

[5] Duncan, Otis Dudley, W. R. Scott, S. Lieberson, B. Duncan, and H. H. Winsborough, Metropolis and Region, Baltimore: Johns Hopkins Press, 1960.

[6] Galli, Omer R., "Occupational Composition and the Metropolitan Hierarchy: The Inter- and Intra-Metropolitan Division of Labor," American Journal of Sociology, 69 (1963): 260-269.

[7] Huang, David S., Regression and Econometric Methods, New York: John Wiley and Sons, 1970.

[8] Goldberger, Arthur S., Econometric Theory, New York: John Wiley and Sons, 1964.

[9] Goodman, Leo A., "A Modified Multiple Regression Approach to the Analysis of Dichotomous Variables," American Sociological Review, 37 (1972): 28-46.

[10] Goodman, Leo A., "A General Model for the Analysis of Surveys," American Journal of Sociology, 77 (1972): 1035-1086.

[11] Grizzle, J. E., C. F. Starmer, and G. G. Koch, "Analysis of Categorical Data by Linear Models," Biometrics, 25 (1969): 489-504.

[12] Lazarsfeld, Paul L., "Regression Analysis with Dichotomous Attributes," Social Science Research, 1 (1972): 25-34.

[13] Neter, John and E. Scott Maynes, "On the Appropriateness of the Correlation Coefficient with a 0,1 Dependent Variable," Journal of the American Statistical Association, 65 (1970): 501-509.

[14] Schnore, Leo F., "Social Mobility in Demographic Perspective," American Sociological Review, 26 (1961): 407-423.

[15] Speare, Alden, "Alternative Models of Individual Mobility," Proceedings of the Social Statistics Section of the American Statistical Association, Washington, D.C., 1971: 364-368.

[16] Spilerman, Seymour, "The Analysis of Mobility Processes by the Introduction of Independent Variables into a Markov Chain," American Sociological Review, 37 (1972): 277-294.

[17] Summers, Gene F., R. L. Hough, J. T. Scott, and C. L. Folse, Before Industrialization: A Social System Base Study of a Rural Area, Bulletin No. 736. Urbana, Illinois: Illinois Agriculture Experiment Station, University of Illinois, 1969.

[18] Theil, Henri, "On the Estimation of Relationships Involving Qualitative Variables," American Journal of Sociology, 76 (1970): 103-154.

[19] Wanner, Richard A., Occupational Mobility and the Urban Hierarchy, unpublished Ph.D. dissertation, The University of Wisconsin, Madison, Wisconsin, 1972.

[20] Winsborough, H. H., "Review of Jeffrey K. Hadden and Edgar F. Borgatta, American Cities: Their Social Characteristics," American Sociological Review, 31 (1966): 433.

NEW APPROACHES TO CODING AND ANALYZING MORTALITY DATA

Robert A. Israel
Marvin C. Templeton
Marshall C. Evans
National Center for Health Statistics

During the past seventy-two years of collecting national mortality data from the vital registration system in the United States, the coding of the demographic and medical data has been a necessary clerical chore. In addition, the analytic possibilities involving cause of death have been limited by the principle of ascribing one and only one cause to each death, as proscribed by international agreement, thus resulting in a loss of the other medical information which may appear on the death certificate. This paper presents a discussion of computer systems that have been developed, and some that are in the process of being developed by the National Center for Health Statistics, that classify and code demographic and medical data items from death certificates by computer processes and as a by-product yield additional cause of death information which may be utilized for new and more extensive analyses.

One of the most complicated kinds of statistical data classification has been the assignment and coding of the underlying cause of death. This concept involves the determination of "the single disease or injury which initiated the train of morbid events leading directly to death or the circumstances of the accident or violence which produced the fatal injury." 1/

The World Health Organization has provided the statistical classification structure and the attendant rules for determining the underlying cause of death through the decennially revised International Classification of Diseases. 1/ While this classification system is comprehensive and invaluable from many points of view, it leaves for the user a number of serious problems. Some of these difficulties are: (1) the long period of training required to develop a competent nosologist schooled in the classification and coding of mortality medical data and the related problem of staff retention in an area of work that requires exactness and controlled productivity; (2) the difficulty in achieving desired levels of data comparability and accuracy brought about by complicated rules and classification structures, inexactness in reporting of some medical conditions, and human limitations in the broad range of knowledge required by a nosologist; and (3) the selection of a single underlying cause of death resulting in the loss of the remaining information reported on the medical certification portion of the death certificate. Much of the lost information may be relevant to current health problems.

To meet the expanding needs for greater utilization of mortality medical data and to take steps which would overcome some of the problems encountered in the traditional coding procedures, the National Center for Health Statistics has developed an automated computer system which provides statistical data on all medical information reported on death certificates. The system is designed to generate the underlying cause of death for the first time through computer methodology by applying the international coding rules to each medical relationship given by the certifier of the death in much the same manner employed by nosologists in manual operations. The system carries the name ACME which is the acronym for "Automated Classification of Medical Entities."

The system requires the entry onto computer tape of all diseases, conditions, accidents, and injuries given in the medical certifier's statement. The information is recorded onto tape utilizing codes based on the Eighth Revision International Classification of Diseases Adapted for Use in the United States (ICDA). Special instructions are used in applying the ICDA codes to all diseases, conditions, and injuries. Once the medical certification is placed on magnetic tape in coded form and in the same sequence as they appeared on the death certificate, the medical relationship of the entries are matched against a series of decision tables which relate all ICDA codes to one another in terms of each international rule. Defining the specific content of each international rule decision table presented one of the most difficult systems development problems in that it required a deliberate decision on the relationship of each detailed code to every other ICDA code relative to the purpose of each international rule. Traditionally, such decisions have been made by individual coders who have had limited guidelines on the causal relationship of diseases and on priorities of some conditions over others.

The computer program is written in PL1 Computer Programming Language. It provides for storage of all decision tables in core storage, minimizing reference time requirements, which, in turn, increases output speed. All functions are performed and the selection of underlying causes is accomplished at a rate of approximately seven records per second.

The system was developed, tested, and implemented on the NCHS IBM-360 Model 50 computer which had 256K core storage until the system's capacity was recently expanded. Storage requirements dictate that the system, in its present form, cannot be accommodated on a computer of less capacity.

Decision table content is updated by introducing individual changes when made by the staff. The tables are routinely maintained on disc storage and a special program provides for insertion or deletion of content coupled with a print-out of the updated tables with special notations of the changes for visual verification.

Approximately 5.5% of the data records are rejected for manual processing and are printed out in full detail with appropriate messages defining the reason, a means of handling the exceptional cases not conforming to the provisions of the system. In addition, whenever a disagreement is encountered in the quality control process, and there is manual coding of a sample of records serving as a quality control, then a full description of the detailed steps the computer system applied is printed out including each decision table reference and the order in which the decisions took place. This leaves no question as to how the computer system derived its final answer. This is one means by which adjustments to the decision tables are identified, defined, and incorporated into the system. Furthermore, if a particular decision is changed and is considered sufficiently significant to warrant correction on records previously processed, it can be accomplished by simply resubmitting the data file to the system based on corrected decision tables.

Manual application of the International Coding Rules is subject to varying degrees of inconsistency. The computer system provides an avenue of absolute consistency and a means of isolating troublesome certifications for further study. An important by-product is the full documentation for the first time of assumptions and decisions going into cause-of-death classification. Heretofore, documentation has been given in terms of guidelines with minimum reference as to how they are to be exercised in specific situations. Automating this element of the classification process gives users detailed insight into the data and permits more intelligent analyses of the end product.

The same data that serve as input to the ACME system for the assignment of the underlying cause of death, i.e., the ICDA codes for each diagnostic term appearing on the medical certification of death, can also serve as the data inputs to multiple cause of death analyses which would take cognizance of all conditions reported at death on the certificate. This approach is not in itself new but at the national level prior to 1968 the coding of more than just a single underlying cause of death has been undertaken only five times since 1900 - in 1917, 1925, 1936, 1940, and 1955. The underlying (or principal) cause of death and one associated cause were coded in 1917, 1925, 1936 and 1940; in 1955 all reported information was coded. A single table showing the cross tabulation of underlying and contributory causes was published without comment for the data years 1917, 1925 and 1940 in the annual vital statistics publications of the United States for the years 1918, 1925, and 1940 respectively. 2-4/ A paper presented to the American Public Health Association in 1923 presented a brief analysis of the 1917 data and strongly recommended additional work on multiple causes of death. 5/

Continued interest in multiple-cause tabulations was stimulated by the Fourth International Conference for the Revision of the International List of Causes of Death. International comparisons of the procedures for selection of the primary cause of death indicated that comparability of death rates could not be achieved on an international basis until there was more knowledge of the contributory causes of death. An extensive study of multiple causes of death was then undertaken for 1936. Two condensed reports arising from these data were published in 1939 and 1940 but they did not contain the full set of tables. 6,7/ Associated causes of death were again coded for 1940 and a table was included in the regular annual vital statistics volume for that year. 4/

During the early 1950's a number of activities both at the national and at the State level, were undertaken in connection with multiple-cause-of-death studies. By 1956, the National Office of Vital Statistics (now the Division of Vital Statistics of the National Center for Health Statistics) had developed a manual of instructions for the coding of multiple causes of death and work proceeded slowly in the coding of a sample of 1955 death records. The results of this activity were published in various journals and in a Supplement to the 1955 edition of Vital Statistics of the United States. 8-11/ This latter publication, however, did not appear until 1965. In addition, at the international level, two meetings - one in 1967 in London and one in 1969 in Geneva - brought together a number of interested nations for the purpose of exploring uses of multiple cause analysis and methodology but with heavy emphasis on minimum standards for international comparability.

For all of the activity relating to multiple cause analysis dating back to the early part of the century and continuing sporadically to the present, it might be assumed that the major problems have been solved, but this is not the case. There still remain a number of key issues.

First of all is the question of the appropriateness of the medical diagnoses appearing on the death certificate. Are they accurate? This question, of course, is applicable to not only multiple cause of death analysis but to the more conventional underlying cause of death as well. With all the attention focused on mortality statistics over the years, it is necessary to continually be concerned with this question. Studies in the past have indicated some problems in accuracy, but nowhere near enough evidence has been found to invalidate the usefulness of mortality data. However, not enough is known about this problem and more effort must be directed at studies of accuracy of medical certification on death certificates. Closely related to the problem of accuracy is the question of completeness of reporting of conditions. If attention is now to be focused on all of the conditions listed on the death certificate, what conditions should be listed? Over the years, educational efforts have been made to elicit from the certifying physician enough information to satisfactorily indicate the single primary or underlying cause to which the death is to be assigned. Any conditions listed on the certificate other than the underlying cause hopefully were useful in arriving at that assignment, but no other use was made of additional medical data. However, the multiple cause approach changes all of that. Now there is interest in each condition on the death record. How many conditions do physicians list? How many should they list? For underlying cause purposes a certificate which simply indicates "pulmonary tuberculosis" is statistically equivalent to one which indicates that the immediate cause of death was "congestive heart failure" which was due to "bronchiectasis" which in turn was due to "pulmonary tuberculosis". For multiple cause purposes there is an obvious difference in the amount of information available and yet the two cases may have been medically very similar. How uniformly do certifying physicians include conditions which contribute to death but are not related to the immediate cause? What about conditions present at death but unrelated to the immediate or underlying cause? Obviously a new educational effort will have to be made if any degree of consistency of reporting is desired. On the other hand, we are not ready to abandon the underlying cause concept. The multiple cause data are considered a useful supplement but not a replacement. Therefore any change in instructions will have to be carefully phrased so as not to upset the one approach for the sake of the other. Related to these problems is the format of the internationally recommended medical certification. Should this format be revised to better elicit the kinds of information desired? It is hoped that experimentation in this area leading to recommendations to the World Health Organization will be undertaken in various parts of the country and around the world.

The second major area of concern lies with the structure and content of the International Classification of Diseases (ICD). How should this important statistical tool be modified to accommodate multiple condition analysis? Presently, there are rubrics in the ICD which are already combination categories intended to provide for the joint reporting of certain combinations of diseases such as category 404, Hypertensive heart and renal disease. For multiple cause purposes, does this title represent one condition, two conditions, or three or more conditions? A closely related problem is the coding and counting of conditions such as hypertensive cardiovascular arteriosclerosis (one term with one code number in ICD) as opposed to the reporting of cardiovascular disease due to arteriosclerosis due to hypertension reported on three different lines of the death certificate. Should these two situations be handled differently or the same for statistical purposes? What kinds of coding rules should be developed to code multiple conditions on death records? There are numerous problems of this type which arise because of the structure and intended purpose of the ICD. What recommendations should be made to alleviate some of these problems? The third area of concern lies with the kinds of output or statistics that result from this approach. What kinds of tabulations and analyses are envisioned? What uses can be made of such data? Here, as with the other areas of concern, there is not complete agreement. However, some very basic tabulations have been generally agreed upon among the interested countries. These basic tabulations will for example, show for a selected list of causes of death the relationship between the underlying cause and the associated causes reported with the underlying cause, the number of deaths where each selected condition is mentioned whether or not the condition is the underlying cause, and the total number of times the selected conditions are mentioned on death certificates. It should be noted that the number of deaths involving a given condition and the number of times that the given condition is mentioned on death certificates are not necessarily the same. Other basic tabulations might present data on combinations of conditions that appear more or less often than might be expected if the conditions are independent of each other.

There have been a number of questions or problem areas raised and not many, if any, answers have been supplied. These questions are by no means exhaustive; it is hoped that they are sufficiently suggestive of the kinds of problems being faced and solutions that are being sought. This does not mean that the National Center for Health Statistics does not have any approaches to the technique of multiple cause analysis. On the other hand, ideas and suggestions from users or potential users of these data are encouraged. It is important to ask once again the question raised by Dorn and Moriyama: "Why do we want statistics on causes of death?" 8/

These authors stated that the uses to which statistical data are put are basically determined by the available data. They further stated that the types of information one might reasonably expect cause-of-death statistics to provide in order to maximize their usefulness are as follows:

a. accurate reflection of the conditions contributing to the fatal outcome in the opinion of the medical certifier.

b. the relative importance of the various diseases, injuries and acts of violence as causes of death and

c. reliable representation of the time trend of the frequency with which the various conditions are reported as bringing about death.

It is believed that multiple cause of death data will maximize the use of available diagnostic information. It will be possible to make a count of all reported conditions as well as an unduplicated count of deaths. There will be a basis for analyzing mortality trends for various diseases which is not now possible because of the loss of information resulting from the selection of a single condition as the underlying cause of death. Multiple cause tabulations will provide a great deal more data than are now available. They will avoid some of the long standing objections to the underlying cause concept. However, if multiple cause data are to realize their full potential and if the National Center for Health Statistics program of routine publication of such data is to be of value, then significant interaction between the producers and the consumers of the data must come about. For the immediate future, decisions are being made in order to produce data that can be reviewed and evaluated. In the long run, improvements in the basic input as well as in the output and analysis will hopefully add significantly to the vital and health statistics field.

Thus, the ACME system presents several benefits in the coding and analysis of mortality data. It provides for the automated selection of the underlying cause of death in a uniform manner, eliminating to a very large extent the inter-coder variations inherent in a complex manual system. It provides, through the same basic input codes, the ability to tabulate and analyze more than one cause per death. In addition, the system gives baseline information for evaluating the International Classification and the rules for the selection of the underlying cause, an invaluable set of data useful for the periodic revision of the ICD.

These benefits are derived with savings in training and overall manpower requirements. Previous attempts to code multiple causes as well as underlying causes have proved to be very costly. Coding of multiple causes requires manpower resources equivalent to that used in coding underlying causes alone. Manpower for detailed coding required by the ACME system is about one-fourth greater than that required to manually select the underlying cause. The contrast is attributable to simplification of the manual coding process. Training requirements for nosologists have been reduced from 12 to 18 months to from 3 to 6 months.

It should be noted that the ACME system has been introduced by our colleagues in the Vital Statistics Section of Statistics Canada and we jointly believe that it is a significant achievement in keeping vital statistics data handling and processing apace with advances in methodology and technology.

There are other challenging research projects in the data preparation and processing field under long range development in the National Center for Health Statistics.

One project underway is the development of a computer system named CONTEXT which will accept full text or standard abbreviations of the medical terms encountered on death certificates and will convert these terms automatically into their corresponding ICD code numbers. Such code numbers then could form the input to the ACME system, thereby eliminating the manual assignment of ICD codes for each condition as is now the practice.

Another such project has as its goal the automation and standardization of non-medical or demographic data items on vital records. The objectives of this project are:

1. To automate to the highest practical level the entire spectrum of data collection, handling, and processing of the nation's vital statistics data.
2. To develop standardized terminology, definitions, and data classification detail for vital statistics applications at various levels of government.
3. To provide a single data handling system which will meet the needs of Federal, State, and local vital statistics programs.

For conversational conveniences, we have dubbed this project ASSIST, the acronym for "Automated Standardized System for Indexing and Statistical Tabulations."

The state of the art in data processing has developed to a degree that technologically there should be no serious barriers in meeting the objectives of ASSIST. Data needs are clear, which leaves to human resolve and commitment the matter of developing standard terminology, definitions and classification detail.

The rising costs of manpower and equipment and the growing gap between vital statistics methods and available technology bring considerable pressure for upgrading our procedures through efforts such as ACME, CONTEXT and ASSIST. These new approaches to coding and analyzing mortality data will provide opportunities for the production of more uniform, standardized statistics on a more timely basis, and will eliminate much of existing duplication of effort between the various levels of government, while at the same time provide for sufficient flexibility in the system to meet the needs of federal, State, and local health statistics programs.

REFERENCES

1. Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death. Volume 1, World Health Organization. Geneva, Switzerland, 1967.
2. U.S. Bureau of the Census, Mortality Statistics, 1918, U.S. Government Printing Office, Washington, D.C., 1920
3. U.S. Bureau of the Census, Mortality Statistics, 1925, Part I, U.S. Government Printing Office, Washington, D.C., 1927.
4. Vital Statistics of the United States, 1940, Part I, U.S. Government Printing Office, Washington, D.C., 1943.
5. Dublin, Louis I., and Van Buren, George H., "Contributory Causes of Death--Their Importance and Suggestions for Their Classification," American Journal of Public Health, Vol. 14, No. 2, February 1924, pp. 100-105.
6. U.S. Bureau of the Census, "Deaths from Alcoholicm, United States: 1936," Vital Statistics - Special Reports, Vol. 7, No. 59, 1936.
7. Janssen, Theodore A., "Importance of Tabulating Multiple Causes of Death," American Journal of Public Health, Vol. 30, No. 8, August 1940, pp. 871-879.
8. Dorn, Harold F., and Moriyama, Iwao M., "Uses and Significance of Multiple Cause Tabulations for Mortality Statistics," American Journal of Public Health, Vol. 54, No. 3, March 1964, pp. 400-406.
9. Moriyama, Iwao M., "Chronic Respiratory Disease Mortality in the United States,"

Public Health Reports, Vol. 78, No. 9, September 1963, pp. 743-748.

10. National Cancer Institute, "New Numerators for Old Denominators--Multiple Causes of Death," by D. E. Krueger, National Cancer Institute Monograph, No. 19, Washington, D.C., 1966, pp. 431-443.

11. Vital Statistics of the United States, 1955, Supplement, U.S. Government Printing Office, Washington, D.C., 1965

INTERPRETATION OF VITAL STATISTICS

Anders S. Lunde, National Center for Health Statistics
Nathan Keyfitz, Harvard University

The field of vital statistics is as open to investigation, analysis and experimentation as it was three hundred years ago when John Graunt (1620-1674) introduced the observations on the London Bills of Mortality. The research possibilities are sometimes overlooked by statisticians and demographers because the registration systems now established in developed countries seem routinized and sterile. Also, in developing countries where vital statistics registration systems are yet a century or more away the focus has been on survey techniques and surveillance systems with their own special problems and methods of analysis.

This paper considers some conceptual schemes which may help in the understanding of population structure and change, theoretical constructions which serve to explain population dynamics, and methods for assessing vital statistics and census data. The purpose of the paper is to show that, through the reconstruction of the past using relatively simple models, vital statistics can be continuously reinterpreted for the improvement of data and for more refined analysis.

Population Projection

The interpretation of vital statistics can be aided by a hypothetical projection into the future, a familiar technique for making sense of birth and death data. Such projection is distinguished from forecasting the future which is too difficult a task at the present state of knowledge. Demography does well to confine itself to accounting for past demographic phenomena, a difficult enough enterprise. Another technique, with which we are concerned here, is to break down the phenomenon under investigation into simple elements and then see how well the phenomenon can be reconstructed by putting these elements together into a suitable model.

If statistical data speak for themselves, they speak in a very low voice. Most of what they say is inaudible without some theoretical and practical amplification. As an example of how vital statistics are made to talk think of the birth data for the United States and for Mexico. In recent years Mexican births have been running about 2,500,000 per year, while those of the United States have been about 3,500,000 per year. The direct comparison of these numbers seems to tell us that the United States is higher than Mexico by about 1,000,000 births or 40 per cent. Nearly anything that one could say on the basis of this fact alone would be wrong.

Interaction starts when we note that the population of the United States at the present time is 210,000,000 and that of Mexico about 55,000,000; that in terms of these the birth rates are 17 per thousand and 45 per thousand respectively; that overall death rates are similar, Mexico's higher age-specific rates being offset by its much younger population. By a succession of such derived facts we gradually approximate to the true meaning of the birth totals of the two countries. We go on to age specific rates of birth and death; make the death rates into a life table; use the life table along with the birth rates to calculate intrinsic rates; find that the intrinsic rate of natural increase, the growth of the population that would result in due course if the age-specific rates of birth and death were to persist, is 7 per thousand for the United States and 35 per thousand for Mexico. Then we can assert that continuance of the Mexican rates would ultimately result in a doubling of the population each 20 years, of the United States rates only once each 100 years. At the end of a century of such rates the United States would contain 400,000,000 people, while Mexico would have multiplied by 32, and so contain 1,600,000,000 persons.

Now the figures begin to talk. For the above very rough calculation tells us that the Mexican rates of increase cannot continue for 100 years, while the United States rates can continue. They tell us that anyone who opposes birth control as a matter of principle for Mexico may claim to be taking a high moral position, but that he is in fact proposing a course that can only lead to an increase in the death rate.

The technique used in this argument amplifies a difference by projecting into the future. When we use raw births to ascertain age-specific birth rates in the two countries, and then go on to calculate intrinsic rates of about 0.007 for the United States and about 0.035 for Mexico, we find Mexico to be higher but the difference looks trifling; in any one year it could hardly have any effect. An amplification is provided by the hypothetical projection into the future: the future that enters the calculation is not in real time, but is rather in the special kind of time on which a conditional prediction is based.

The censuses help explain vital statistics. Births of the United States would be much greater, without any change in the chance of childbearing for individual women, if the population was as young as that of Mexico. With such a young population our birth rate would be 20 per thousand instead of 17 per thousand--merely because of the different mix of ages. On the other hand if Mexico was of the age distribution of the United States it would show only about 42 births per thousand rather than the actual 46. A principle of inertia applies in demographic affairs -- having a high birth rate in the past is in itself an influence making for a high birth rate in the future. A high past birth rate combined with a low death rate similarly reduces deaths, and so makes for an especially rapid population increase, simply by virtue of the way that high births and low deaths generate a young age distribution.

We have seen that a hypothetical future is one of the demographer's instruments for understanding the present. But he also needs the past for instance in explaining how the age distribution as shown in the 1970 census came to be what it is. Here the past vital statistics serve as the means of explanation of the present cross-sectional distribution. We proceed to look into this aspect.

Explaining Population Characteristics

To understand a phenomenon, we must break it down into its simplest elements, then put these elements together again in such a way as to reconstruct the phenomenon. This at least was the method of Descartes, and it seems the way to make intelligible the population characteristics presented as census and other data. Such characteristics as age, sex, marital status, birthplace, occupation, and industry, are suited to Cartesian treatment, but we will stay with characteristics that can be explained by vital statistics.

Past rates of birth and death are especially convenient for explaining age distributions. By partitioning and reassembling the observations from their constituent elements, the stable model can be utilized. By use of this or some other suitable model we can say that the ages of any country would be if previous age-specific birth and death rates had been of certain amounts.

As an example we explain some differences between Venezuela in 1965 and Sweden in 1803-7 so comparing a present developing country with an underdeveloped country of the early 19th century. We observe that in early 19th century Sweden about 30 per cent of the population was under 15 years of age, while in Venezuela now about 45 per cent is under 15 years of age. The difference is great and of serious consequence for the economy and society of the two countries. How much of it is due to difference in birth rates and how much to difference in death rates?

An answer is given by putting together an artificial population based on the stable model. (What model to use is a question that has to be considered afresh according to the circumstances of each occasion.) On this model Venezuela shows 47.7 per cent under 15 and Sweden 31.3 per cent under 15. If we make up an artificial population with the birth rates of contemporary Venezuela and the death rates of early 19th century Sweden we find a proportion under 15 of 43.6. Thus holding birth rates fixed we find that the effect of the difference of death rates--especially early 19th century Sweden's higher infant mortality--is to lower the under 15 by 4 percentage points out of the 16 percentage points of total difference. Presumably the remaining 12 percentage points are accounted for by the difference in birth rates. We can check this by doing the calculation in the other direction (with Venezuelan death rates and Swedish birth rates) and find the same result. Apparently the two sequences in our experiment leads to an unambiguous result: interaction between birth and death rates is small. Of the total difference between the 47.7 per cent under 15 for Venezuela and the 31.3 per cent for Sweden, three quarters are due to the difference in births and one quarter to the difference in deaths.

It is more surprising to learn that about the same is also true for the differences between the two populations in proportion of old people, in dependency ratios, and in rates of natural increase. The calculation can be improved by allowing for changing rates of birth and death. But where such complications give essentially the same result we can gladly dispense with them.

Demographers concede that our ability to predict the future is limited to those cases where no sharp turning appears. Non-demographers agree heartily; the difference between them comes in whether population science ought to be able to foretell the future. Forecasting of complex concrete situations is too difficult a task to be the test of a science. A less severe test, though one difficult enough, is to account for the past. Where the stable population or other mathematical model is too simple to be useful we ought to be able to set up a random birth process on the computer that would produce a set of vital statistics and censuses indistinguishable by the methods of statistical inference from the vital statistics we observe. This more modest ideal of accounting for the past by a mathematical or simulation mechanism can serve as a target for demographic theory until we are bold and strong enough to hope for precise forecasts.

Along these lines we have the development of several population simulation models using Monte Carlo techniques. One of these, POPSIM, is a dynamic demographic model designed for computer simulation of the principal demographic processes occurring in human populations.1/ It is classed as a microsimulation model because it generates a vital event history, including the dates of birth, marriage, divorce, widowhood, remarriage, and death of each individual in the computer population. While it is a two-sex model, it may be used for simulating cohort as well as period data. It is a stochastic model in the sense that random sampling from probability distributions is used to determine which events occur to an individual and when they occur; and it is a dynamic continuous time model permitting the probabilities to change with time. The original version of the computer program was not designed to distinguish between social segments of the population; the program has now been modified to permit distinctions by hypothetical segments in terms of race, residence, and family income.

One point to be made here is that while the technique may be developed, it should be more widely tested with published results so that more statisticians and demographers may become concerned with the technique and its applicability. The problems inherent in the appraisal of the technique only foreshadow the many technical-analytical problems which computer based simulation is only now beginning to compel us to resolve.

Some simplistic estimates of population changes have been used, for example, to determine the effect of the so-called Zero Population Growth; they may be scientifically inappropriate because they take into account only a few variables. On the other hand, using hundreds of variables, one study has already provided implications of considerable social and economic consequence; such findings are too often given as final interpretations whereas the method is essentially theoretical and should be considered only as basically useful for continued testing and experimentation. Probability studies infer experimentation.

Simulation models can be used not only for projection from today on but also, in the light of the observations of this paper, for testing past experience with both census and vital statistics data. One can incorporate the details of the demographic history of the past into the actual development of various population futures to see how these were derived and what the alternate consequences might have been.

<u>Verification</u>

But prior even to drawing conclusions about the future, and tracing population characteristics into the past, is checking whether the vital statistics are correct. There is no substitute for the checking that goes on in vital statistics offices, but that checking needs some supplementation, in particular from the attempt to interpret and use the vital statistics that are published.

As for error, for example, there are a number of questions regarding the possible errors in the death certificate. More persons seem to have a hand in the production of a death certificate than in the production of most statistical reporting forms. From the point of view of what constitutes good reporting or survey technique too many uncontrolled persons are involved in the process. In some hospitals the form goes through several hands in different departments; the physician fill out the medical certification, and/or signs it; the funeral director fills in what is called the demographic information from various sources. Apart from the question of error on medical certification in terms of selectivity and subjective judgement as made by the physician is the fact that from a purely scientific point of view the process is unreliable. Yet the process itself and questions concerning the derivation of selected items have not been subject to the intensive investigation they deserve.

A set of death statistics may look satisfactory, but when one makes a life table from them it may turn out that the expectation of life at birth is 110 years. At the very least then one would want to re-examine the data, both those referring to the deaths and to the exposed population, and see whether anything could have gone wrong. Such a test is very gross, and would reveal only the most inexcusable errors.

A finer test is available if one is looking at mortality for a number of countries. Grouping the countries according to the factors that have a known influence on mortality--income per head, for example--one could see how the rates and expectations for any one group of countries, say a group that is homogeneous in respect of income, are distributed. If one of the groups is far outside of the distribution then its vital statistics need looking into. This test can be especially effective for particular causes of death; in a recent unpublished work Samuel Preston has shown that a given cause shows similar age profiles at very different levels of fertility.2/

In other instances we can check a birth rate by asking whether the persistence of that rate in the past would have accounted for the present age distribution as shown by a census. If the age distribution is very steep, then the births must have been correspondingly high. When we see that Colombia (1964) reported about 46 per cent of its population under age 15, and registered births were only 38 per thousand population, we suspect omission in the vital statistics. Fitting a stable model to the ages 5 to 29 gives over 47 births per thousand population. But the births could have been falling in the preceding years, so we fit (to ages 5 to 70) a modified stable distribution that allows for falling births. This gives over 44 births per thousand. The ratio of children 5-9 to women 20 to 50, less dependent on the stable assumption, shows about 45 births per thousand. The virtual agreement of all these ways of looking at the age distribution reinforces our suspicion that many births, possibly as many as one quarter, were unregistered.

Demographers do not come to their data without preconceptions. They carry around in their heads a set of anticipations of what birth rate ought to appear for developed and for developing countries. They know what age distributions ought to look like for different kinds of populations. A question worth raising is whether these anticipations can be more formally used as <u>a priori</u> probabilities in a Bayesian scheme. On a Bayesian way of doing demography one would sketch out his conclusions in advance, and indeed draw up a complete picture of his subject, necessarily vague in that each of the parameters would be expressed as a fairly wide range with a probability attached. Then the new data would be brought in, and would serve to narrow the ranges within which the diagram was sketched out. It would not always merely narrow an <u>a priori</u> range; it could move the range to some other position altogether, and even increase it.

The Truth Attained by Assembly and Selection

The assembly of material from diverse sources can help us to discern the defects of any one source. Births and deaths are nearly always underregistered in newly established vital statistics systems, and often in old ones as well. All censuses undercount certain categories, for example men about age 20. Yet defective vital statistics can be used along with defective censuses to produce a complete account.

The method is based on the fact that as individuals are born and pass through their lives they are noted at various statistical checkpoints. A child may be missed by the birth registration network, and again by the decennial census taken when he is age 3, but then caught in the next census at age 13, missed in the census of age 23, and so on. The statistical system has enough redundancy that being caught some of the time suffices, and Coale and Zelnick have ingeniously taken advantage of this for the United States.3/ We may have to match the sources name by name in order to catch everyone, and this is often done in historical demography.

But if there are certain ages in which most individuals are caught in the census, then we can use purely demographic or aggregate methods, without matching. The census level at ages 5 to 14, for example, may be perfectly complete, and adding an estimate of preceding deaths to it enables us to infer the births. From the improved birth total so obtained we can go on to the distribution of births, say by age of mother or parity, for which the vital statistics may be adequate. This is an example where the vital statistics are assessed and improved by the census.

In other instances a pair of censuses can check one another and provide at least a hint of what is happening to births and deaths. Suppose a sharp fall in a 1967 census population enumerated as 10-14 from the enumerated as 5-9, a notch in the age distribution occurs, such as actually shown for Indonesia. It could be due to a real rise in successive cohorts, if for instance to a drastic fall in infant mortality occurred about 1950, or else it could be due to a simple understatement by enumerators of the age of children, exaggerating the 5-9 total and reducing the 10-14. In the second case the notch would be wholly imaginery, and if the same census method was used in 1971 the notch would reappear at the same ages. If, on the other hand, the notch was real it would have moved along to ages 15-19 and 20-24 by 1971. Real people get 10 years older in the course of 10 years; imagining people do not age. We now have the 1970 Census of Indonesia, and its preliminary count shows that some of both (a large youth cohort and age misstatement) are present.

Fertility

In the same way the vital statistics can serve to check the census as in the comparison of cumulative fertility rates and vital statistics period rates. Census data on children ever born can be compared with vital statistics accumulations and the differences analyzed. There are special problems here, deriving from underenumeration rather than under-registration and from survivorship. Yet, while not be an easy matter to interpret, significant differences in magnitude and in estimated rates should stimulate new questions as to causality and validity. Also, data by age of mother, marital status and race can be compared. Hypothetical projections of what the cumulative data will be, as based on vital statistics estimations, or even what actual experience shows, can be developed and used as baselines for interpretation.

Rearrangement of Data

The interpretation of vital statistics information often requires its drastic rearrangement. The best known example is the translation of period into cohort material. We may wish to follow groups of individuals rather than take cross-sections of events across time, and consider the births to women aged 18 in 1930, aged 19 in 1931, 20 in 1932, ... , assembling these together and speaking of them as the cohort born in 1912. Their reproductive activity can be compared with the corresponding activity of those born in 1913, 1914, ... The total children born to cohorts may vary less than do the total children born from one calendar year to the next. An unusually high death rate when a cohort is young may eliminate less hardy members, so that its subsequent mortality is lower than that of another cohort not subject to as rigorous selection.

The difference between cohorts and periods is unimportant as long as age-specific birth and death rates are constant through time. But once fluctuations occur in births then cohorts and periods are no longer equivalent. Suppose a dip in the United States economy followed by immediate recovery. If parents typically control their births with a view of having a certain number of children, there may be a dip in births but it will be followed by a subsequent offsetting rise. The delay in the formation of some families has had only a trifling effect on the growth of the population. But if the economic dip is prolonged, so that some parents pass the prime childbearing ages, then they are likely not to make up all the lost births, and the economic recession will have caused a net loss of population. The full meaning of the vital statistics is not revealed until we know how long an economic dip causes what loss in population. Full knowledge would show how the birth rate mimics the ups and downs of the economic cycle.

Combinations of Elements

We have been discussing studies in which new approaches develop as inherent in the data. There are other needed investigations in which insight emerges from new combinations of data. One of these is the manipulation of medical certification of death to provide combinations of diseases or multiple causes of death, and all other related factors reported on the death certificates, in simple tabular form. In countries of populations in the millions where the total number of deaths are in the hundreds of thousands or millions, and thousands of variables (causes) are involved, the problem heretofore has remained unsolveable. By the use of the computer both as a combining and as a selective tool, the problem is now solved. After some years of work the Automated Classification of Medical Entities (ACME) is a reality; the great potential of this system now awaits the work of scientific analysts.4/ While the system automatically selects the underlying cause of death for national totals and international comparisons, any combination of causes can be determined by prescribed programming. This may throw new light on what mortality itself is from the statistical standpoint. To carry the method one step further; it should be possible to use ACME to develop a model for more refined study of the past effects of selected morbidity and combinations of causes of death on population growth. Whenever something like this comes along we must ask ourselves what other combinations of elements can be investigated.

Extension of Life Table Techniques

The model population distribution which we know as the life table has a number of uses other than that of indicating life expectancy on the basis of established fertility and mortality rates. It has been used to assess the effect on the population of the elimination of selected diseases, for example, and has indicated the number of persons, especially older ones, who would have survived had these diseases not taken their toll. It is possible to consider the model as a form for testing various estimations of fertility and mortality by age, race, and sex. In other vital statistics areas the technique has been successfully used as a theoretical device as, for example, where marriages are considered as births and divorces as deaths. These have apparently been more successful than the epidemiological model of using pregnancy as a disease and contraceptives as a cure but both approaches indicate that techniques have no limits as regards conceptual development.

Sampling from Vital Records

Because of the cost factor and demands on limited staff, it may become necessary to consider how to simplify the record processing task by reducing the one hundred percent sample of births or deaths to a direct 50 percent sample, or to a selective item sampling of 5, 10, or 25 percent. Sometimes decisions to affect these procedural changes are not evaluated until the data year is completed and the results are in. At other times decisions to do selective sampling are made on the basis of statistical inference, without testing. With the vital statistics data available for past years it should be possible to experiment extensively to determine the effect of differential sampling on final frequencies and rates.

For an example, the decision in the United States to go to the 50 percent sample of births made impossible the production of detailed tabulations on multiple births. Other means had to be developed to produce estimates for twins, triplets, and so on, but these data have not been considered adequate by some programs. What else was eliminated, not so visible as twinning, has not been fully explored. On the other hand, to see if legitimacy rates could be produced for those registration areas which did not have the item on the birth certificates, a retrospective test, the Inferential Method, was used.5/ In this method, data derived from complete reporting for a past year was compared to that from incomplete information from the same certificates which inferred illegitimacy. The rates were found to be comparable. Unfortunately, the method was also found to be too costly for processing on a regular basis.

The use of vital records of any past period for laboratory purposes generally is probably limited because of matters respecting confidentiality and accessibility but local and national offices may wish to consider the scientific potential inherent in these investigations. The use of vital records for any past year as a sampling frame for retrospective studies in fertility, mortality and nuptiality, has already proven fruitful.

Conclusion

Before anyone can hope to forecast the future he must learn to interpret the past. The interpretation can consist in breaking down the phenomenon under investigation into simple elements, and then seeing how well past instances of the phenomenon can be reconstructed by putting these elements together. One example is the explanation of an age distribution in terms of the birth and death rates of the preceeding 90 years. Another is referring the present increasingly higher death rates of males to the different incidence of causes for the two sexes. A third is showing how cohorts defer and advance their childbearing in accordance with economic conditions, thus allowing a relatively constant number of children per family to give rise to widely fluctuating period rates. Such reconstruction facilitates causal analysis, and tells what would occur if one of the underlying elements were different. It can be used to say what increase in the proportion of old people would result from eliminating cancer or heart disease; what the ultimate population would be if couples immediately dropped their childbearing to bare replacement; what effect emigration has on the subsequent birth curve. The reconstruction of the past through the use of models, some relatively simple and others complex, can be carried further and will help in the interpretation of vital statistics.

There are other areas for experimentation as in the case of comparing vital statistics and census data, the rearrangement of data such as in cohort fertility, in new combinations of data elements, and an extension of established techniques, such as life table applications, and so on.

There is more to come from vital statistics analysis. In the past 25 years the new techniques, measures, and methodologies developed by Whelpton, Greville, Campbell, Lee, Linder, and a number of others, have shown that this area of investigation is by no means dead. This paper is a reminder that the field of thought opened by John Graunt three centuries ago is still wide open for stimulating statistical investigation and analysis.

1. Daniel G. Horwitz, "Presentation by Research Triangle Institute Team of Report on Methods of Building a Health Manpower Simulation Model," Washington, D.C.: NIH, Bureau of Health Manpower Education, Proceedings and Report of Conference on a Health Manpower Simulation Model. December 1970. Vol. II, pp. 7-23.

2. Samuel Preston, unpublished paper.

3. Ansley J. Coale and Melvin Zelnick, The Enumeration of the White Population of the United States and Its Fertility. Princeton, N.J.: Princeton University Press, 1964.

4. Israel, R. A., Templeton, M. C., and Evans, M. C., "New Approaches to Coding and Analyzing Mortality Data". Unpublished paper, 1972.

5. Clague, A. J., and Ventura, S. J., "Trends in Illegitimacy United States 1940-1965". Washington, D. C.: DHEW, National Center for Health Statistics, Vital and Health Statistics, Series 21, No. 15, February 1968, pp. 72-81.

"EXPERIENCE WITH ABORTION REPORTING IN NEW YORK CITY FOLLOWING LIBERALIZATION OF THE LAW."

Frieda G. Nelson, Jean Pakter, Donna O'Hare, N.Y.C. Health Dept.

Introduction

On April 10, 1970 the New York State Legislature passed a law which vastly liberalized the performance of abortions in New York State. Heretofore, the law permitted abortion only if, in the physician's judgement, the life of the expectant mother was in jeopardy. The new law made the abortional act legal if the woman consented and if the abortion was performed by a licensed physician within 24 weeks of the commencement of pregnancy. The law was silent on the question of residency requirement. Its effective date was July 1, 1970.

The short interval between the passage of the new law and its effective date was a time of conjecture and surmise.

Opponents of abortion reform, as might be expected, prophesied serious difficulties. But others, many favorably disposed towards a liberalized abortion policy, began to ask serious questions. What would the demand be? Estimates varied widely from 50,000 to as high as one-quarter of a million. What about the influx of non-residents and would they overwhelm the medical facilities of our city and deprive residents of services? Some predicted flatly that New York City would become the "abortion capital of the nation"! Would minority groups be served? Would they even come forward to request service or would abortion be repugnant to them and viewed as a form of genocide? Would readily available abortion replace contraception as the main mode of family planning? Would women come for abortion early in pregnancy? Could the medical facilities of New York City "tool up" quickly enough to meet the demand? Would legal abortion be done safely - with a minimum of risks? And what about the beneficial effects of safe abortion? Proponents of legal abortion had urged it on grounds that it would reduce maternal morbidity and mortality. Would this really occur, or would criminal, unclean abortions persist unchecked with deaths from the clean, legal procedures merely superimposed on the mortalities of the criminal ones. Would there be any impact on the birth rate and on infant mortality.

The following is a report of the experience in N.Y.C. under the new abortion law for the first year and one half based on statistical data collected.

Method of Collecting Statistical Data

Data were collected by means of certificates designed for reporting all terminations of pregnancy, induced or spontaneous. The filing of these certificates is mandated by the New York City Health Code. These provide data regarding number of abortions, weeks of gestation, reason for the termination, age and parity of the female, and whether a previous termination had occurred since July 1, 1970. The mode of termination and the facility where termination took place, are recorded; also the legal residence of the patient.

However, the certificates which are supposed to be filed within 48 hours after termination were not designed to yield information on complications. A separate weekly reporting system was therefore instituted under which a special report form was to be forwarded every week to the Department of Health by each facility performing abortions. This form enabled the department to obtain data on the number and type of complications.

In addition, the municipal hospitals provided at first a daily and then a weekly telephone report on abortions performed.

Each method of reporting served as a check on the other and helped us estimate the degree of underreporting known to exist.

The certificates of termination of pregnancy provided our primary source of information. Most of the analyses which follow are based on them.

In the 18 months after abortions were made legal in New York State, certificates for 250,632 abortions were filed with the New York City Health Department. Of these 86,381 were residents, 164,251 were non-residents. Taking into account other sources of supplementary information; e.g.; telephone reports from municipal hospitals and weekly abortion reports from participating hospitals, it is estimated (probably conservatively) that 277,230 abortions were done the first year and one half and that about 65% were to non-residents.

In the entire State of New York, 325,881 abortions were performed in the 18 months. Thus, 76.9% of this number were done in New York City and the remainder, 23.1%, outside New York City.

Ten Leading Areas For Abortions To Non-Residents

Women came to New York City for abortions from every one of the 50 states as well as New York State residents living outside New York City.

Table 1 lists the 10 most frequent areas from which the non-residents came. New Jersey geographically adjacent to New York heads the list and some in the mid-west such as Illinois, Michigan and Ohio are included. Canada, our foreign neighboring country is 10th on the list.

~~Facilities~~ For Abortions In New York City: Types Of Provider

For the residents of New York City the chief source of abortions were the municipal and the voluntary hospitals which together provided 73% of abortions for residents of the city. For the non-residents the chief source of abortions was the free-standing clinics - some affiliated with hospitals. Non-residents also were largely served by proprietary hospitals some of which were the New York City termini of active abortion referral service for out-of-towners.

It was the stated policy of the municipal hospitals to accept only city residents. However, more than 1000 non-residents were treated at these hospitals. (Table 2).

Abortions Performed By Weeks Of Gestation: Residents And Non-Residents

At the outset of the program, there was much concern that the demand for abortions would overtax the available facilities and staff and would create delays and unduly long waiting lists. Another concern was that many women might come in the later stages of pregnancy requesting abortion either through ignorance or fear. It was reassuring then to find that the concerns were groundless since our statistics show that over 3/4 of all women came in the first 12 weeks or less gestation. The non-residents tended to have a higher percent (78%) as compared with the residents (75.6%). (Table 3).

Method Of Termination Of Pregnancy By Facility

As noted in Table 4A of the total number of abortions performed, the suction method far exceeded all other methods and accounted for over 60% of the total. The Dilatation & Curettage method accounted for almost 25%. The saline instillation method accounted for 14% and the method least in use was the hysterotomy, less than 1%.

The various types of facilities differed in the proportionate use of methods. In great measure, this was a reflection of the distribution pattern of stages of gestation. For example, if the municipal hospitals had a relatively high proportion of saline instillations (26%), it was an indication of the relatively greater proportion of women coming in after the first trimester of pregnancy.

Conversely as might be expected, the free standing clinics reported over 92% abortions by suction method since all patients in these clinics are supposed to be 12 weeks or less gestation.

The trends in mode of abortion by facility from the onset of the program through the ensuing months showed the following changes:

In municipal hospitals the percentage use of suction kept increasing while percentage of D & C's was dropping.

The suction method went up from 25% to 63% whereas the D & C's declined from 35% to 15%.

The saline instillation method dropped from 37% of the total to 21%.

A similar pattern was noted for the ward service of voluntary hospitals where suction method accounted for 57% rather than 43%, the percent at the beginning of the program. Here too D & C's declined from 29% to 23%. The saline method declined from 25% to 18%. Likewise in the private services of voluntary hospitals the use of suction rose to 53% from 43% while the D & C method declined from 38% to 27%.

The saline method appeared to increase proportionately from 16% to 19%. Hysterotomy as a mode declined.

The trend in the proprietary hospitals has been quite revealing. In this group, the suction method has had a somewhat lower percent, 35% compared to the other facilities. The D & C method declined from 50% to 36% but the most striking change is the increased proportion of saline instillations which went up from 12% to 28% of all abortions in the proprietary hospitals.

In the free standing clinics the pattern of methods utilized has changed least. Suction still accounts for over 90% of all methods. However, a slight increase is observed in the proportion of D & C's performed (up to 8% from 1.5%).

Tables 4B shows the number and percent of method of termination by facility for residents. The residents receiving saline terminations went chiefly to municipal hospitals and second to voluntary hospitals but non-residents went chiefly to proprietary hospitals for saline terminations, exceeding twice the number of New York City residents' saline abortions in municipal hospitals. In fact, many more saline abortions were done for non-residents (22,033) than for the residents (13,473). (Tables 4B and 4C).

Age Of Women Undergoing Abortions

The greatest proportion of women undergoing abortions fell in the age group 20-29 years accounting for approximately 54% of the total. (Table 5). For residents this age group comprised 58.5% and for non-residents 51%. The teenage group accounted for 26% of the total. The residents teenage group was about 17% of the total and the non-residents about 31%. Thus, it would appear that the non-residents tend to be a proportionately younger group than the residents. In fact, 11.5% of the non-residents were 17 years or less in age.

Women who were 35 years or more accounted for 9% of the total. Among the residents the group 35 years and over comprised over 10% of the total whereas among non-residents this proportion was less, namely about 8%.

Race And Ethnic Distribution Of Women Aborted

As noted in Table 6, the white women accounted for approximately 74% of the total. This high proportion stemmed chiefly from the non-residents of New York City who were predominantly white (88.5%). The non-white comprised 22% of the total and the Puerto Rican 4%.

The ethnic and racial distribution of residents was quite different from that of non-residents. In New York City, 45% of the total abortions for residents were white; non-white comprised about 44% and the Puerto Rican about 11%. Whereas for non-residents the white was 88%, non-white was approximately 11% and the Puerto Rican less than 1%.

These figures reveal clearly that legalized abortion is being sought by women of all ethnic groups in the city including the minority groups such as the black and Puerto Rican. These groups were the ones who had been largely the victims of crude attempts at abortions by unskilled non-medical individuals, or self-induced by dangerous and desperate measures. In fact, deaths among these women comprised the largest component of our pregnancy associated deaths year after year.

Short of death, thousands of these women entered the hospitals of the city every year in sepsis and shock following these attempted abortions.

It is noteworthy that women of all ethnic groups are becoming enlightened and are availing themselves of the services if they desire to terminate a pregnancy.

In 1971, about 30% of all births in New York City were to non-white whereas almost 44% of abortions among residents were to non-white. The ratio of abortions to live births in the past 18 months has been higher for the non-white than any other ethnic group.

415.5/1000 live births for white

737.2/1000 live births for non-white

331.4/1000 live births for Puerto Rican

The relatively small proportion of non-white and Puerto Rican among the non-residents is in part a reflection of the fact that a non-resident must have funds for travel as well as medical care. Thus, the poor non-resident is handicapped in obtaining an abortion in New York City. However, the resident who is poor can obtain an abortion through Medicaid which covers the cost or at a minimal cost if she cannot afford private care.

Parity Of Women Undergoing Abortions In New York City

A difference has been noted between the residents and non-residents of the city undergoing abortions regarding histories of previous pregnancies or parity.

Among residents almost 59% had had previous births or pregnancies. In fact, over 12% had a pregnancy of fifth or more. Whereas among non-residents for the majority this was a first pregnancy (63.7%). (Table 7). In contrast, among residents only 41% were experiencing a first pregnancy.

The trend in the 18 months since the program commenced is a noticable decline in the proportion of women of first pregnancy order and an increase in proportion of multiparae. This is true for both the residents and non-residents.

The relative differences between resident and non-resident parity distribution remain. Specifically, the proportion of primiparae for residents dropped from 46% to 39%; for non-residents it dropped from 68% to 62%.

The proportion of multiparae increased for each order of parity for the residents of New York City as well as the non-residents with the passage of time.

Complications Reported (By Weekly Bulletins) Following Abortions

The overall reported rate of complications (Table 8) was 7.3/1000 abortions. This figure is realistically an underestimate of the true incidence of complications particularly those of a minor nature which may not be recorded or reported. Underreporting may also be due to the fact that cases done outside hospitals, on an ambulatory basis, especially among non-resident women are lost of follow-up and accordingly complications may not be reported.

The data strikingly show the relative seriousness of later terminations versus early terminations as measured by the difference in complication rates. For terminations 12 weeks or less the rate was 3.8 whereas for those over 12 weeks the rate was 23.7/1000 abortions! This is more than a sixfold difference in rates of complications.

For the early terminations, perforated uterus was reported as the major complication with a rate of 1.3/1000. Infection and hemorrhage were next in order of frequency.

For the late terminations, the most frequent complication was retained tissue (11.0/1000) followed by infection and hemorrhage (4.5 and 3.5/1000).

Failure was reported more often for late termination than for early (1.8/1000 compared with less than 0/05/1000).

The nature of the complications is closely associated with the method employed as will be described.

The rates of complications were recorded over the 18 months for the various quarters of the year to observe whether with the passage of time, experience and skills acquired might have led to a diminution in the frequency of complications. This was indeed the finding particularly for the early terminations.

The incidence of perforation of the uterus dropped from 1.5 initially to 1.0 in the last quarter of 1971.

Infection rate declined from 1.9 to 1.1 and hemorrhage from 1.5 to 0.8. On the other hand, the complication of retained tissue rose from 2.3 to 3.0/1000.

The overall rate of complications did decline from 8.5 to 6.7/1000 and with the exception of complication of retained tissue all specific complications as reported declined.

In comparing the first year's complications (July 1970 - June 1971) with the second year to date (July 1971 - April 1972) a decline was noted for complications of terminations 12 weeks and under with a decline in the rate from 4.6 to 3.1 but for those terminations beyond 12 weeks the decline was relatively slight from 26.8 to 24.8/1000.

Complication Rates By Type And Method Of Termination

In discussing the varying rates of complications by length of gestation, allusion was made to the association of complications by method of termination employed. Tables 9A and 9B show clearly that saline instillation method carried the highest risk, closely seconded by hysterotomy (28.3 and 27.8 respectively). The most frequent specific complication recorded for the saline instillation method was retained tissue (14.7/1000) and second by infection (5.4/1000) while the suction method had the lowest rate of complications with an overall rate of 3.7/1000. The D & C method carried a somewhat higher risk overall, namely 5.1/1000 abortions with slightly higher frequencies for perforation of the uterus (1.9/1000), hemorrhage and infection (1.0 and 0.9/1000 respectively) than in cases where suction is employed. It was observed that D & C was utilized in more advanced gestations where it was felt that suction was not advisable. It has been established that risks increased with suction and or D & C as gestations go beyond the 10th week and considerably increased between the 12th and 15th weeks of gestation.

Hysterotomy is relatively not as safe a procedure as judged by the incidence of serious complications. Infection rate is highest for this method (11.0/1000 as compared with 5.4 for saline, 0.9 for D & C and 0.8 for suction).

In analyzing the trends in complication rates by method, it was observed that the rates declined for all except the saline method which in the last quarter was recorded as 31.3/1000 abortions.

The various categories of hospitals showed declines in rates of complications except the proprietary hospitals. This group showed increases reflecting most likely the increased proportion of late terminations or saline instillations being performed in proprietary hospitals.

The free standing clinics also showed some increase though slight. This may indeed indicate need for intensified surveillance of free standing clinics since these are not permitted to undertake terminations beyond the 12th week of gestation.

Reasons For Terminations Of Pregnancy

Before the advent of the liberalized law the reason most commonly cited was "psychiatric." After the law took effect the "psychiatric" indication became the least common. The most frequently cited indication has become "social or sociologic."

Medical reasons were cited in only 1% of the cases of the residents and even less, 0.4%, for non-residents.

It was quite apparent that the strictures of the old law led to subterfuges for indications for so-called therapeutic abortions. In great measure these were available only to private patients and not to poor ward patients. The liberalized law has put aside the veil of intellectual rationalizations and permitted a fair and honest appraisal of indications for termination of pregnancy.

Pregnancy Associated Deaths

One of the most compelling reasons for the public health officials and most of the medical community to favor liberalized abortions was the mortality following the so-called illegal "criminal" type of abortion.

Data tabulated from 1960 thru 1971 (Table 10) for deaths associated with pregnancy show a significant decline after 1970, both for the total and the component associated with abortions. The 12 abortion associated deaths include 7 legally induced, 3 "illegal" or self induced and 2 spontaneous abortions. The total abortion associated death rate achieved a record low of 0.9/1000 live births.

It should be pointed out that from 1966 on a decline in maternal mortality (Figure 1) was noted. This may well have been associated with the availability of family planning services which had become officially endorsed and readily available in the late 60's. Curbing unwanted pregnancies would appear to have a salutary effect on mothers and babies in terms of mortality.

Effect Or Association Of Liberalized Abortions And Birth And Birth Rates

It had been predicted by many population experts that births would increase in the 70's because of the expansion of the child bearing age group resulting from record high births in the late 1940's.

Accordingly, the decline noted in Table 11 is significant. The drop from 149,192 to 131,920, a 12% decline exceeded the national decline of 4.3%. (Figure 2).

Thus far in 1972, the decline has accelerated even further.

It is our belief that the liberalized abortion program as well as family planning services have had a major impact. Other factors, such as economic recession, draft of young men for military services may be deterrents for child bearing. Nevertheless, the acceleration of the decline sustains our conviction that abortions as well as availability of family planning measures are playing the role in the decrease.

Out-Of-Wedlock Births

Another effect of a liberalized abortion program under scrutiny was the impact it might have on out-of-wedlock pregnancies.

Our data showed that for the first time since records were maintained that the steady increase of out of wedlock births year by year has actually been reversed and a decline noted by 11.8% which was even greater than the decline 11.5% noted for the in-wedlock births. (Table 11).

Since out of wedlock births have been associated with higher rates of premature births and infant mortality, the decline is reassuring and should bring about a lower incidence of prematurity and infant mortality. Thus far this has proved to be the case (Table 12). All ethnic groups shared in this decline.

A record low in neonatal mortality was achieved in 1971 with a rate of 14.9/1000 live births.

Summary And Conclusions

The experience during the first 18 months in N.Y.C. following the passage of a liberalized abortion law effective July 1, 1970 for the State of New York has been reviewed.

It is estimated that 277,000 abortions were performed in this period of 18 months and that almost 2/3 of them were for women who did not reside in the City. They came not only from adjacent states but from great distances including other countries such as Canada.

The residents of the city turned to the municipal and voluntary (non-profit operated) hospitals for abortions. As time elapsed about 10% resorted to free-standing clinics for early terminations. On the other hand, the non-residents resorted primarily to proprietary hospitals and free standing clinics, especially the latter which have proliferated in number and have been serving the largest number of women in recent months.

The majority of women requesting abortions had terminations early, i.e. 12 weeks or less gestation, over 3/4 of the total. A greater proportion of residents with the passage of time came in earlier and most recent data indicated that 79% of them obtained abortions at 12 weeks or less gestation whereas in the beginning it was about 70%.

The non-residents tended to come earlier for abortions than the residents at first but not the proportion is about the same for both. If anything, a slight increase in the proportion of late terminations among non-residents has been noted recently.

Of all modes of termination employed, the suction method was most often used. Dilation and curettage was second. For later terminations (16 weeks and over) the saline method was used far more frequently than any other method.

The proportion of saline terminations in a given facility reflected the proportion of later terminations performed at the facility. Thus, municipal hospitals had a relatively higher proportion of saline instillations for termination. Now as proprietary hospitals have been serving a higher proportion of later terminations, the percent and number of saline terminations have increased considerably in this category of hospitals.

The greatest proportion of women receiving abortions was in the 20-29 year age group (54%). This was true for residents and non-residents. However, teen-agers were in higher proportion among non-residents than residents 31% compared to 17% for the latter. For both groups, the proportion of teen-age women has been on the increase.

Women of all ethnic or racial groups sought abortions. Among residents 45% were white, 44% were non-white and 11% Puerto Rican.

For the non-residents, 88% were white, 11% were non-white and less than 1% Puerto Rican. These data clearly show that as the program has been organized, women of all ethnic groups seek abortion and that among residents there is no financial problem since the poor can obtain services. Thus the minority groups, the black and Puerto Rican, who are more likely to be poor, can now obtain abortions under safe legal circumstances. For the non-resident however, who have no financial resources, coming to N.Y.C. for an abortion, may pose problems.

The parity of women undergoing abortions as shown in our data appeared to be increasing as time went on. The proportion of primiparae for residents dropped from 46% to 39% - for non-residents from 68% to 62%. Residents appeared more likely to be terminating pregnancies because of desire to limit size of family or space childbirth. Whereas among non-residents more often it was a first pregnancy which was deemed unsuitable for them to continue.

The rates of complications although likely under-reported in our data, nevertheless revealed significant features. Late terminations carried a much higher complication risk, 23.7/1000 abortions compared to 3.8/1000 for early terminations--over 6 times greater frequency.

Perforation of the uterus was reported as the most commonly occurring major complication of early terminations, with a rate of 1.3/1000. Infection and hemorrhage followed next in order of frequency. Suction appeared to carry less risk than D & C if employed in cases of 12 weeks or less gestation.

The late termination methods were chiefly saline and hysterotomy. Both these methods carried relatively high risks for the saline instillation method, retained tissue was the most frequently reported complication.

In the course of the 18 months, the rates of reported complications declined, more notably for early terminations and comparatively little for the late terminations.

All deaths associated with pregnancy, both abortion associated or non-abortion associated declined in 1971 to a record low of 2.9/10,000 live births. The rate for abortion-associated deaths in that year was 0.9/10,000 live births. This includes abortions which were legal, illegal and spontaneous.

Births have declined in New York City in 1971 by 11.6%, whereas for the U.S.A. the decline was 4.3%. The accelerated rate of decline of births in New York City is attributable to the liberalized abortion program. Out-of-wedlock births declined for the first time in 1971 after years and years of constant increases. In fact, the decrease for out-of-wedlock births was even greater than for the in wedlock births.

Prematurity rates (births under 2501 grams) declined in 1971. (9.2%). The neonatal mortality (under 28 days of life) achieved a record low (14.9/1000 live births).

Conclusions

Based on 18 months experience, the liberalized abortion program has had a favorable impact on "maternal" mortality, including abortion-associated deaths. Numbers of women admitted to hospitals following illegal abortions, septic "incomplete" abortions have declined markedly.

The drop in births and birth rates with declines in premature and out-of-wedlock births has been noteworthy.

A record low in neonatal mortality has been established.

These data are believed to have been significantly influenced by the liberalized abortion program as well as the extensive family planning services available since 1965 to women in New York City.

The medical community of the city deserve commendation for the high standards established and maintained thus far and for their desire and effort to eliminate commercialism in a legalized abortion program.

Referrals for non-residents as well as residents are available through reputable non-profit sources - Planned Parenthood of New York City in cooperation with the Health Services Administration of the City of New York, the Clergymen's Counselling Services, as well as the New York County Medical Society Services.

Continued efforts must be made to urge women to seek safe legal abortions early if it is desired to terminate the pregnancy and to utilize family planning services to avoid the need for an abortion and especially to avoid the repeated use of abortion as a means of curbing unwanted pregnancies.

TABLE 1

TEN LEADING AREAS FOR ABORTIONS TO NON-RESIDENTS
July 1, 1970 - December 31, 1971

Station Area	Number	Percent of Total Non-Resident Abortions
New Jersey	22,831	13.9
Illinois	14,506	8.8
Michigan	14,288	8.7
Ohio	13,668	8.3
Pennsylvania	13,379	8.2
Florida	9,702	5.9
Massachusetts	9,103	5.5
New York State (excl. N.Y.C.)	8,005	4.9
Connecticut	7,045	4.3
Canada & other countries	5,525	3.4
Total	118,052	71.9

TABLE 2

NUMBER AND PERCENT OF INDUCED ABORTIONS IN NEW YORK CITY
TO RESIDENTS AND NON-RESIDENTS BY TYPE OF PROVIDER
July 1, 1970 - December 31, 1971

Provider	NUMBERS			PERCENTS		
	Total	Residents	Non-Residents	Total	Residents	Non-Residents
Municipal	31,945	30,589	1,356	12.7	35.4	0.8
Voluntary - Service	14,273	12,247	2,026	5.7	14.2	1.2
Voluntary - Private	35,552	20,207	15,345	14.2	23.4	9.4
Proprietary	80,360	12,546	67,814	32.1	14.5	41.3
Free Standing Clinics	88,502	10,792	77,710	35.3	12.5	47.3
Total	250,632	86,381	164,251	100.0	100.0	100.0

TABLE 3

NUMBER AND PERCENT OF INDUCED ABORTIONS IN NEW YORK CITY
TO RESIDENTS AND NON-RESIDENTS BY WEEKS OF GESTATION
July 1, 1970 - December 31, 1971

Weeks of Gestation	NUMBERS			PERCENTS		
	Total	Residents	Non-Residents	Total	Residents	Non-Residents
12 weeks or less	193,871	65,264	128,607	77.4	75.6	78.3
13 - 15	19,280	6,403	12,877	7.7	7.4	7.8
16 - 20	26,880	9,793	17,087	10.7	11.3	10.4
21 - 23	5,279	2,285	2,994	2.1	2.7	1.8
24 weeks and over	2,957	1,229	1,728	1.2	1.4	1.1
Not Stated	2,365	1,407	958	0.9	1.6	0.6
Total	250,632	86,381	164,251	100.0	100.0	100.0

TABLE 4A

TOTAL NUMBER AND PERCENT OF INDUCED ABORTIONS IN NEW YORK CITY
BY METHOD OF TERMINATION AND PROVIDER
July 1, 1970 - December 31, 1971

Provider	NUMBERS METHOD OF TERMINATION					PERCENTS METHOD OF TERMINATION				
	Total	D & C	Suction	Saline	Hysterotomy	Total	D & C	Suction	Saline	Hysterotomy
Municipal	31,945	7,430	15,639	8,361	515	100.0	23.2	49.0	26.2	1.6
Voluntary-Service	14,273	3,588	7,366	3,070	249	100.0	25.1	51.6	21.5	1.8
Voluntary Private	35,552	11,361	17,067	6,564	560	100.0	32.0	48.0	18.4	1.6
Proprietary	80,360	32,897	29,704	17,327	432	100.0	40.9	37.0	21.6	0.5
Free Standing Clinics	88,502	6,721	81,588	184	9	100.0	7.6	92.2	0.2	0.0*
Total	250,632	61,997	151,364	35,506	1,765	100.0	24.7	60.4	14.2	0.7

*Less than 0.05

TABLE 4B

NUMBER AND PERCENT OF INDUCED ABORTIONS TO RESIDENTS OF NEW YORK CITY
BY METHOD OF TERMINATION AND PROVIDER
July 1, 1970 - December 31, 1971

	NUMBERS					PERCENTS				
	METHOD OF TERMINATION					METHOD OF TERMINATION				
Provider	Total	D & C	Suction	Saline	Hysterotomy	Total	D & C	Suction	Saline	Hysterotomy
Municipal	30,589	7,114	15,096	7,882	497	100.0	23.2	49.4	25.8	1.6
Voluntary-Service	12,247	3,279	6,265	2,472	231	100.0	26.8	51.1	20.2	1.9
Voluntary-Private	20,207	7,361	10,431	2,086	329	100.0	36.4	51.7	10.3	1.6
Proprietary	12,546	4,869	6,552	1,013	112	100.0	38.8	52.2	8.1	0.9
Free Standing Clinics	10,792	1,635	9,135	20	2	100.0	15.2	84.6	0.2	0.0*
Total	86,381	24,258	47,479	13,473	1,171	100.0	28.1	55.0	15.6	1.3

*Less than 0.05

TABLE 4C

NUMBER AND PERCENT OF INDUCED ABORTIONS TO NON-RESIDENTS OF NEW YORK CITY
BY METHOD OF TERMINATION AND PROVIDER
July 1, 1970 - December 31, 1971

	NUMBERS					PERCENTS				
	METHOD OF TERMINATION					METHOD OF TERMINATION				
Provider	Total	D & C	Suction	Saline	Hysterotomy	Total	D & C	Suction	Saline	Hysterotomy
Municipal	1,356	316	543	479	18	100.0	23.3	40.1	35.3	1.3
Voluntary-Service	2,026	309	1,101	598	18	100.0	15.3	54.3	29.5	0.9
Voluntary-Private	15,345	4,000	6,636	4,478	231	100.0	26.1	43.2	29.2	1.5
Proprietary	67,814	28,028	23,152	16,314	320	100.0	41.3	34.1	24.1	0.5
Free Standing Clinics	77,710	5,086	72,453	164	7	100.0	6.6	93.2	0.2	0.0*
Total	164,251	37,739	103,885	22,033	594	100.0	23.0	63.2	13.4	0.4

*Less than 0.05

TABLE 5

NUMBER AND PERCENT OF INDUCED ABORTIONS IN NEW YORK CITY
TO RESIDENTS AND NON-RESIDENTS BY AGE OF WOMEN
July 1, 1970 - December 31, 1971

	NUMBERS			PERCENTS		
Age of Women	Total	Residents	Non-Residents	Total	Residents	Non-Residents
17 or less	24,164	5,395	18,769	9.6	6.2	11.5
18 - 19	41,549	9,161	32,388	16.6	10.6	19.7
20 - 29	134,633	50,551	84,082	53.7	58.5	51.2
30 - 34	26,778	12,122	14,656	10.7	14.1	8.9
35 & over	22,351	8,715	13,636	8.9	10.1	8.3
Not Stated	1,157	437	720	0.5	0.5	0.4
Total	250,632	86,381	164,251	100.0	100.0	100.0

TABLE 6

NUMBER AND PERCENT OF INDUCED ABORTIONS IN NEW YORK CITY
TO RESIDENTS AND NON-RESIDENTS BY RACE AND ETHNIC GROUP
July 1, 1970 - December 31, 1971

	NUMBERS			PERCENTS		
Race and Ethnic Group	Total	Residents	Non-Residents	Total	Residents	Non-Residents
White	184,585	39,152	145,433	73.6	45.3	88.5
Nonwhite	55,822	37,942	17,880	22.3	43.9	10.9
Puerto Rican	10,225	9,287	938	4.1	10.8	0.6
Total	250,632	86,381	164,251	100.0	100.0	100.0

TABLE 7

NUMBER AND PERCENT OF INDUCED ABORTIONS IN NEW YORK CITY
TO RESIDENTS AND NON-RESIDENTS BY ORDER OF PREGNANCY
July 1, 1970 - December 31, 1971

Pregnancy Order	NUMBERS Total	Residents	Non-Residents	PERCENTS Total	Residents	Non-Residents
First	140,365	35,746	104,619	56.0	41.4	63.7
Second	36,039	16,040	19,999	14.4	18.6	12.2
Third	29,502	13,603	15,899	11.8	15.7	9.7
Fourth	20,010	8,946	11,064	8.0	10.3	6.7
Fifth	11,648	5,331	6,317	4.6	6.2	3.8
Sixth or more	10,870	5,435	5,435	4.3	6.3	3.3
Not Stated	2,198	1,280	918	0.9	1.5	0.6
Total	250,632	86,381	164,251	100.0	100.0	100.0

TABLE 8

COMPLICATIONS FOLLOWING ABORTION, BY TYPE AND PERIOD OF
GESTATION, NUMBER AND RATE PER 1,000 ABORTIONS-NEW YORK CITY
July 1, 1970 - December 31, 1971

TYPE OF COMPLICATION	Total Number	Total Rate	12 Weeks and Under Number	12 Weeks and Under Rate	13 Weeks and Over Number	13 Weeks and Over Rate
Hemorrhage	252	1.1	112	0.6	140	3.5
Infection	329	1.5	153	0.8	176	4.5
Perforated Uterus	298	1.3	246	1.3	52	1.3
Anesthesia	18	0.1	13	0.1	5	0.1
Shock	10	*	2	*	8	0.2
Retained Tissue	520	2.3	87	0.1	433	11.0
Failure	78	0.3	7	*	71	1.8
Lacerated Cervix	49	0.2	41	0.2	8	0.2
Other	89	0.4	49	0.3	40	1.0
Unspecified	14	0.1	9	*	5	0.1
Total Complications	1,657	7.3	719	3.8	938	23.7
Total Abortions	226,722		187,099		39,623	

Source: Weekly Abortion Reports

*Less than 0.05

TABLE 9

COMPLICATION RATES PER 1000 ABORTIONS BY TYPE AND
METHOD OF TERMINATION, NEW YORK CITY
July 1, 1970 - December 31, 1971

METHODS OF TERMINATION

Type of Complication	Total	Dilation and Curettage	Suction	Saline	Hysterotomy	Other Including Hysterectomy
Hemorrhage	1.1	1.0	0.5	4.0	3.7	0.9
Infection	1.5	0.9	0.8	5.4	11.0	1.5
Perforated Uterus	1.3	1.9	1.3	0.2	5.1	2.1
Anesthesia	0.1	0.1	0.1	0.1	0.7	-
Shock	*	*	*	0.2	0.7	-
Retained Tissue	2.3	0.5	0.5	14.7	1.5	0.6
Failure	0.3	*	*	2.4	-	-
Lacerated Cervix	0.2	0.3	0.2	0.1	0.7	0.3
Other	0.4	0.3	0.3	1.0	4.4	0.6
Unspecified	0.1	0.1	*	0.2	-	0.6
Total	7.3	5.1	3.7	28.3	27.8	6.6

Source: Monthly Abortion Reports

TABLE 10

PREGNANCY ASSOCIATED DEATHS, TOTAL AND DUE TO ABORTIONS:
NEW YORK CITY, 1960-1971

Year	Number of Deaths: Total	Number of Deaths: due to Abortions	Ratios per 10,000 Live Births: Total	Ratios per 10,000 Live Births: due to Abortions
1960	115	46	6.9	2.8
1961	130	55	7.7	3.3
1962	121	53	7.3	3.2
1963	116	42	6.9	2.5
1964	74	34	4.5	2.1
1965	104	41	6.6	2.6
1966	80	31	5.2	2.0
1967	76	20	5.2	1.4
1968	66	21	4.7	1.5
1969	77	24	5.3	1.6
1970	68	22	4.6	1.5
1971	38	12	2.9	0.9

TABLE 11

TOTAL LIVE BIRTHS, OUT-OF-WEDLOCK AND IN-WEDLOCK BIRTHS
NEW YORK CITY, 1960-1971

Year	Total Live Births	Out-of-Wedlock Births	In-Wedlock Births
1960	166,300	13,901	152,399
1961	168,383	15,723	152,660
1962	165,244	16,412	148,832
1963	167,848	18,436	149,412
1964	165,695	20,223	145,472
1965	158,815	20,980	137,835
1966	153,334	22,714	130,620
1967	145,802	24,336	121,466
1968	141,920	26,262	115,658
1969	146,221	29,325	116,896
1970	149,192	31,903	117,289
1971	131,920	28,126	103,794

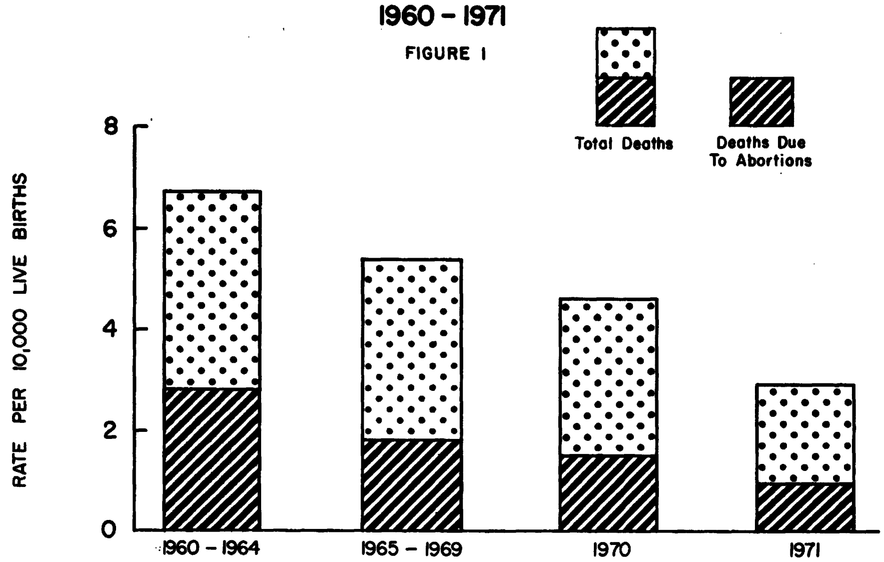
TABLE 12

LIVE BIRTHS BY WEIGHT GROUP - NUMBERS AND PERCENTS
NEW YORK CITY 1960 - 1971

Year	Numbers: Total	Numbers: Under 2501 Grams	Numbers: 2501 and Over	Numbers: Not Stated	Percents: Total	Percents: Under 2501 Grams	Percents: 2501 and Over	Percents: Not Stated	Neonatal Mortality Rates*
1960	166,300	15,544	149,945	811	100.0	9.3	90.2	0.5	19.2
1961	168,383	16,085	151,840	458	100.0	9.5	90.2	0.3	19.4
1962	165,244	15,979	148,956	309	100.0	9.7	90.1	0.2	20.0
1963	167,848	16,692	150,800	356	100.0	9.9	89.9	0.2	19.0
1964	165,695	16,747	148,778	170	100.0	10.1	89.8	0.1	20.0
1965	158,815	16,108	142,459	248	100.0	10.1	89.7	0.2	19.2
1966	153,334	15,995	136,953	386	100.0	10.4	89.3	0.3	18.6
1967	145,802	14,763	130,714	325	100.0	10.1	89.7	0.2	18.3
1968	141,920	14,153	127,373	394	100.0	10.0	89.7	0.3	17.1
1969	146,221	14,531	131,485	205	100.0	9.9	89.9	0.2	18.1
1970	149,192	14,449	134,515	228	100.0	9.7	90.2	0.2	16.2
1971	131,920	12,078	119,672	170	100.0	9.2	90.7	0.1	14.9

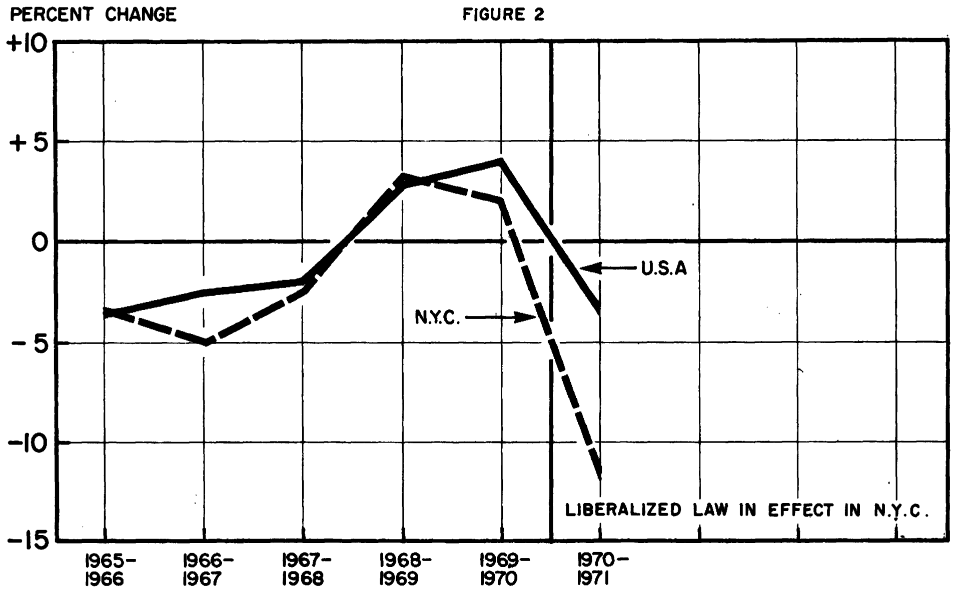
PREGNANCY ASSOCIATED DEATHS
TOTAL AND DUE TO ABORTIONS
RATE PER 10,000 LIVE BIRTHS
NEW YORK CITY
1960 - 1971

FIGURE 1



ANNUAL PERCENT CHANGE IN THE NUMBER OF LIVE BIRTHS
UNITED STATES AND NEW YORK CITY, 1965 THRU 1971

FIGURE 2



DISCUSSION

Carl L. Erhardt, Consultant, New York City

Israel, Templeton and Evans inform us of some real breakthroughs in processing vital statistics that can speed up delivery of data and may also even introduce cost savings. Not alone that, but new data can be at the same time put out. In connection with the multiple cause tabulations, earlier New York City experiments with this mode indicated, for example, that among those aged 65 or more, 8.7 percent of the deaths were ascribed to vascular lesions of the central nervous system as underlying cause but at least 17.1 percent had suffered such an accident. But the plaguing question in my mind is whether the present format of the death certificate is appropriate to elicit the information really needed for determination of multiple causes. Mr. Israel has properly mentioned the need for experimentation along these lines.

To shed some light on the question we did a small follow-back study many years ago on a sample of 73 deaths from cardiovascular diseases. In 53.4 percent of the cases, no additional information turned up, but the percentage of deaths with cardiovascularrenal disease as the sole cause dropped from 65.8 percent to 28.8 percent. The proportions of decedents among whom specified other conditions were found was as follows:

	On certificate	On Follow-back
Hypertension present	26.0%	58.9%
Arteriosclerosis present	79.5	89.0
Diabetes present	9.6	15.1

Obviously, the frequency of arteriosclerosis as a condition among these decedents is understated on the certificates, but such conditions as hypertension and diabetes are even more seriously underestimated and will be, unless special arrangements are made to obtain the information. To investigate possible etiological associations, deficiencies of this kind can introduce telling bias. In view of the rather poor success of educational efforts to improve cause-of-death reporting, it seems that such efforts have to be bolstered by a format that tends to extract the required information. There are, as Mr. Israel points out, many more problems to be resolved for full utilization of multiple causes. Other projects of NCHS, such as CONTEXT and ASSIST give promise of aiding processing at each level of government -- local, state and national -- and hence hastening the demise of the present system of doing everything two or three times, with only a single product, albeit differently limited geographically, as an end result. These efforts of NCHS deserve plaudits. Their early implementation will be welcome.

The report presented by Mrs. Nelson and her colleagues on the experience of the first year and a half with the new abortion law presents a fascinating array of data. They emphasize, among other items, the service to nonresidents provided by the proprietary hospitals and free-standing clinics, showing that between 40 and 50 percent of nonresidents are aborted in such facilities. (Table 2) It may be helpful to look at the data the other way; that is, the proportions of total patients who are nonresidents in each type of institutional setting:

	Percent nonresident
Municipal	4.2
Voluntary - Service	14.2
Private	43.2
Proprietary	84.4
Free Standing Clinics	87.8

Emphasis has been laid on the extra hazards of late abortion and Table 3 indicates that late abortions among nonresidents are relatively fewer than among residents. One must recall that, as reported here, the free standing clinics, where 47.3 percent of the nonresident abortions occur, are restricted to intervention in early pregnancy. Since practically all these cases are, by definition, early, the small differences between residents and nonresidents in the proportions of late abortions suggests that in the hospitalized cases, the situation may be quite different. It is worth noting in this connection that the Abortion Surveillance Report of the National Center for Disease Control for the second quarter of 1971 indicates that of the eight areas included in the report only two areas have lower proportions than New York City of saline instillations. The other five areas have higher proportions for this procedure by rather considerable margins. The ratios of abortions to live births among residents is cited as 415.5 per 1,000 among whites, 737.2 among nonwhites and 331.4 among Puerto Ricans. It is of some interest that the last published ratios on therapeutic abortions before the change in the law were 4.2 for whites, 1.8 for nonwhites and 0.9 for Puerto Ricans. The factor of change is a multiple of a hundred.

The point is made in the paper that "psychiatric" indications became the least common after liberalization whereas it was most common previously as indication for therapeutic intervention. These observations are based on percentages and obviously with "social or sociologic factors" the large component among indications now, the percentage for "psychiatric" reasons must decline. One wonders whether the <u>ratio</u> for "psychiatric" indicators also declined. It is noted that medical reasons are cited for about one percent of the residents; this would mean about 864 instances in the 1½ years of experience, or about 576 a year, a number that is greater than the number of therapeutic abortions before the passage of the new law. The findings here seem compatible with findings reported for other areas where the law has been changed.

For this audience it may be worth calling attention to the paper by Evard (<u>Am. J. Obstet. Gynecol.</u> 113:415, June 1, 1972) who discusses the effects of legal abortions on maternal, fetal, neonatal and perinatal rates. He points out that both numerators and denominators may be affected and suggests that statisticians should give attention to these effects. Evard figures that each legal abortion, for example, subtracts one from the number of live births that would otherwise have occurred and may also have effects on numerator data, such as infant deaths that did not occur because of lack of the preceding live birth. His argumentation is clearly stated but, if as seems to be generally held, legal abortions are to a large extent replacing illegal abortions, then the live births would not have taken place anyway. Hence, there is not a one-to-one relationship.

Live births in New York City declined by 17,272 in 1971 from the 1970 total. In the last half of 1970, there were 19,349 resident abortions and in the first half of 1971, there were 31,570 additional. Under Evard's assumption, these would have resulted in live births in 1971. This constitutes a difference between the actual number of births and an "expected" number of 33,647. If one assumes 20% of the aborted pregnancies would have been lost spontaneously anyway, then there is still a residual of 23,463.

Is the substantial residual a measure of what would otherwise have been illegal abortions? Or how many illegal abortions continue to occur; for one reason or another, that add to this total? If legal abortions have driven illegal abortions out of the marketplace, then their number has been far less than has sometimes been estimated, if the figures here cited actually represent fulfillment of "need". The decline in births is certainly real. Does this decline represent women who otherwise would not have had abortions, despited disenchantment with childbearing (i.e., a measure of unfilled need, so to speak?) To assume a one-to-one relationship between abortions and live births would provide a roughly estimated total number of anticipated live births in 1971 (again allowing for a 20% depletion because of spontaneous loss) of 172,655, or more than the all-time high of 1947. Such a total would imply a sudden rise in fertility, which seems unlikely in this magnitude in one year.

In any case, it is obvious, despite the crudity of this formulation, that legal abortions and subsequent live births do not have a one-to-one relationship for a variety of reasons. Otherwise, one would have to assume live births would have increased in 1971 to the neighborhood of 172,655 rather than decreasing by more than 17,000. It's an interesting game that I hope someone may be inclined to play.

It is a joy to hear someone say, as do Keyfitz and Lunde, that the field of vital statistics is still open to investigation. I suspect two factors (and perhaps others) are involved in the apparent "routine and sterile" appellation. First, support for the program has inappropriately diminished and, second, many analyses are probably made for internal use of agencies and they never see the light of day outside.

Their exposition of methods of using and evaluating vital statistics and modification of procedures for differing purposes is a useful reminder that we need not, and should not, be satisfied with the same old things in the same old way. I agree with the authors that by accounting for the past we may shed light on methods of predicting the future. But assumptions must be made in which human behavior is involved, and human behavior is not precisely predictable; witness the rise in birth rate in the forties after dire prediction of population declines and the current decline in fertility after prediction of inevitable increases because of the changing age structure of the population. Probably the best we can hope for is to select operational assumptions of various directions and magnitudes. With the aid of the computer and suitable models, we can set forth ranges of prediction that may be even more helpful than single predictions by indicating the uncertainties implicit. A range of prediction may forestall the arguments that often arise over a single prediction because no one reads the fine print!

DISCUSSION

Yves Péron, University of Montreal

Harry Page's paper allows us, I think, to have a good understanding of the feeling of frustration which affects, in our time, a great number of Canadian demographers. Let us have a close look at two particular points:

On the one hand, Vital Statistics only give obviously too summary informations on the demographic evolution of Canada. The progress of demographic analysis, during the past thirty years has thought us that, in order to study the occurrence of demographic events, the best approach consists in taking into account their necessary succession in an individual's history. In other words, if the occurrence of an event B is to be studied, a demographer shall first try to find out what is the event A which is the most immediately and necessarily anterior to B:

- for divorces, widowhood, first order legitimate births, the anterior event is marriage;
- for second order legitimate births, the anterior event is the first order legitimate birth... etc.

Then the demographer will try to calculate:

a) the average number of occurrences of B per person having lived through event A;

b) the distribution of time-intervals between the occurrence of events A and B.

Now, for want of pertinent questions on Vital Statistics forms, such an approach is most often not possible in Canada. For instance, we don't know how many brides of a given marriage cohort have a first child; and of course we don't know either the time-interval since marriage.

On the other hand, we know that the administration of the Canadian system of family allowances might supply, at least for fertility, the very informations we are lacking, if only these informations were treated by an appropriate statistical process.

Thank God, Mr. Harry Page just announced the near end of our misfortune. It seems that he has taken good note of certain provinces irreducible opposition to any drastic revision of vital statistics forms, and he offers us today with our sheet-anchor: the automatic linkage of records, which will at last supply the informations that we have waited for so long.

I had the opportunity to follow the work of Hubert Charbonneau and Jacques Légaré who, at the Department of Demography of the Université de Montréal, are precisely doing the automatic linkage of vital statistics records of New France, during the 17th and 19th centuries. And we know that such an operation requires, besides talent, a great deal of time, and of technical and financial means. No doubt the researchers of Statistics Canada have the talent. But will they also have, easily and rapidly, all the desirable technical and financial means? More precisely — and this is a question for Mr. Page — in how many years from now will all necessary informations be easily and regularly available?

****

The automatic linkage process raises delicate problems of identification and Newcombe's experience shows it clearly. In an article published in 1969 in "Population", New combe stated that these identification problems would be reduced if the date of birth of spouses was recorded on the marriage record and if the date of birth of parents was registered on the birth certificates. Then it seems that a fundamental revision of vital statistics records could only contribute to improve the efficiency of automatic record-linkage. Consequently I wonder if, whatever hope one places in record linkage, one should not start the battle again for the revision of vital statistics certificates.

The opposition to adding new questions, and particularly questions relating to demographic events, cannot be maintained, it seems to me, in all good faith. Why would it be politically and administratively more delicate to ask the date of birth instead of age, as was done at the 1971 Census? Why would mothers refuse to give the date of their last confinement, when they give, at the census, the date of birth of their children? And is it more puzzling to ask the date of marriage of the mother, than to ask, as it is now done, if the father and the mother are married to each other? If only the recording of these three dates was obtained from the provincial governments, a great improvement in Canadian Vital Statistics would be realized. And it should be noted that this improvement would have a very low cost.

****

Whatever may be our personal training, we will all agree, I think, that the innovations contemplated by Statistics Canada are to be welcomed. But I would like, at the end of these comments, to suggest one supplementary innovation which seems necessary.

There is general agreement that in Canada, the recording of demographic events is practically exhaustive. But completeness of recording does not necessarily mean completeness of tabulation. Because of the very long delays between the occurrence of an event and its local recording, and also between local and central recording, a certain number of records are not received on time and cannot be submitted to statistical processing. That is a regrettable situation for two reasons:

a) the magnitude of the under-utilization of data. For the Province of Quebec only, 3 to 4000 births, out of an annual total of 100,000, are lost in this way for the analysis.

b) the time-selection of this under-utilization. The Late records predominantly relate to events occurring during the last months of the year, so that any analysis of seasonal variations from published data is bound to be biased.

If all that is right — and Mr. Page can tell us about it — wouldn't it be possible to envisage two types of vital statistics tabulation:

a) a current tabulation of records which are received on time;

b) a revised tabulation, one or two years later, which would integrate late records?

In their paper, Nathan Keyfitz and Anders Lunde have sketched with brio a wide panorama of the potential analyses offered by vital statistics, and they have underlined, in the conclusion, that there is a long way to the full exploration of these possibilitites. Since they have said almost everything, I shall limit myself to two complementary remarks.

****

My first comment is related to the authors' insistance on placing vital statistics analysis under the heading of past reconstruction. Without denying that this is an important aspect, I would like to mention that the analysis of vital statistics also permits, in certain circumstances, to get some insight of the future.

It may happen, indeed, that the simple calculation of demographie indices leads without ambiguity to the conclusion that a certain increase (or decline) of nuptiality or fertility, during a recent period, cannot be interpreted as a lasting but an obviously transitory phenomenon. Such indices as parity-progression ratios and cumulative age-specific current marriage rates(1) have been of great help in allowing demographers to point out catching-up periods or periods during which the occurrence of events is postponed. Such successes in the diagnosis of a situation have certainly played an important role in the reputation of seriousness that demography has acquired.

(1) Called "mariages réduits" in French.

Most certainly also, such successes require particularly refined vital statistics data. One particular refinement which is required is the possibility to calculate time-intervals between consecutive events. For this, we need the date of occurrence of the most important previous events. In other words, it is required that the events observed during a given year can be associated easily to their respective type of cohorts: generation, marriage cohorts, parity-X women... etc.

****

My second comment is motivated by the lack of reference, in the paper, to the possible use of the vital statistics system, as a support for a follow-up type of observation.

To my knowledge, the most extended example of that kind of utilization, is the French survey of mortality by occupation. The 1954 Census has been used to draw a sample stratified by occupation. Since then, year after year, death certificates of the sampled individuals are searched by the vital statistics personel and periodically, the survivors are contacted to control the validity of the survey.

A similar process might be used for surveys of a different nature, fertility surveys for instance. Whenever a personal identification number has been selected, it is easy, through the vital statistics system, to collect informations related to the individuals of the sample in a longitudinal way. One of the advantages of such a procedure — it is worth while mentionning it — is to eliminate selectivity effects and to avoid the problems related to the interpretation of ex-post variables.

There is no doubt that the utilization of vital statistics in this way will be of great help to future demographic research.

****

To conclude, if it is true that vital statistics still have more to teach us on the past, as Keyfitz and Lunde have very well shown, they can as well become a more efficient tool for demographic future prospection, if they are properly managed.

An Examination of Optimization in the Missouri Master Sample

Bernard Lazerwitz, University of Missouri

1. Optimization

The specific sample examined was selected from the Missouri master sample design at an overall sampling fraction of 1 in 1250. Within that portion of the sample selected from segments and chunks, there was an average of 5.7 sample hu's per chunk, 4.1 sample hu's per segment, 8.3 sample hu's per secondary selection; and 20.6 sample hu's per county. Within that portion of the sample selected from city directories (and block supplements),[1] there was 2.7 sample hu's per city directory cluster. Within sample hu's, one adult respondent was selected by means of an adult selection table technique.[2]

Kish (2) gives the following two equations to use in determining optimum occupied sample hu size per primary sampling unit.

1. $\text{deff} - \left[1 + \text{roh}\ (b-1)\right]$

where:

a) deff = cluster sample design efficiency. For the entire sample, deff. is 1.56. For that portion of the sample selected through chunks and segments, deff. is 1.96; for that portion of the sample selected from city directories (and block supplements), deff is 1.44.

b) roh is the intraclass correlation coefficient.

c) b represents the average number of occupied sample housing units per primary sampling unit.

2. $\text{optimum } b = \left[\dfrac{C_a\ (1-\text{roh})}{(c)\ \ (\text{roh})}\right]^{\frac{1}{2}}$

where:

C_a = the average cost per primary sampling unit of training, planning, travel time, listing, mileage, and miscellaneous expenses.

C = the average cost per occupied sample housing unit of actual interviewing, other field time, and editing.

Applying these equations, in turn, to the entire sample; sample hu's from the non-metropolitan areas (primarily from chunks and segments); and sample hu's from the metropolitan areas (St. Louis and Kansas City) --almost exclusively from city directories and block supplements--gives the information of Table 1.

The design effect is somewhat larger in the non-metro areas than in the metro areas. This is reasonable since the metro area sample hu's were primarily selected in small city directory clusters while the non-metro area sample hu's were in the larger clusters of a county-town -chunk design. For the same design reasons, the non-metro area primary sampling units have larger b's than the metro areas. The roh factor is almost twice as large in the metro areas as in the non-metro areas. The C_a costs per primary sampling unit are considerably larger in the non-metro areas which make up 47% of the sample. Again, this is to be expected because of the greater travel distances in these small towns and rural areas. The average costs per occupied sample hu of interviewing, other field time, and editing does not vary too much between the two parts of the sample. Note that the optimum b figures are consistently larger than the actual b's for both parts of the sample. The metro area sample design is a one stage selection of city directory clusters (apart from the small block supplement). Hence the optimum b for metro areas not only refers to the desired clustering per primary sampling unit, but also gives the optimum level for the actual final stage selection clusters.

The data of Table 1 indicate that it should be possible to introduce additional field work savings by increasing sample cluster sizes. This can readily be done by selecting larger clusters of city directory sample lines from the master sample city directory clusters. It can be done in the chunk-segment sample portion of the master sample by selecting clusters of segments for any particular survey instead of individual setments. For the next statewide survey, we shall double previous city directory and segment clustering. This would raise directory selections from clusters of five lines to clusters of ten lines. We shall double the within-chunk rate and then select segments in clusters of two.

1. OPTIMIZATION FACTORS FOR PROJECT 030, Missouri Master Sample, 1971

Sample Category	deff	b	roh	C_a	C	optimum b
Entire Sample	1.56	6.4	0.100	$32.47	$2.92	10.0
Non-Metro Areas	1.96	14.0	.074	$70.54	$3.03	17.1
Metro Areas	1.44	4.2	1.1375	$21.05	$3.53	6.1

2. GENERALIZED SAMPLING ERROR OF PERCENTAGES[a] - PROJECT 030, 1971.

(in percentages)

Reported Percentages	Number of Interviews								
	100	200	300	400	500	600	700	800	900
50	10.0-12.5	7.1-8.9	5.8-7.2	5.0-6.2	4.5-5.6	4.1-5.1	3.8-4.7	3.5-4.4	3.3-4.1
30 or 70	9.2-11.5	6.5-8.1	5.3-6.6	4.6-5.7	4.1-5.1	3.7-4.6	3.5-4.4	3.2-4.0	3.1-3.9
20 or 80	8.0-10.0	5.7-7.1	4.6-5.7	4.0-5.0	3.6-4.5	3.3-4.1	3.0-3.7	2.8-3.5	2.7-3.4
10 or 90	6.0-7.5	4.2-5.2	3.5-4.4	3.0-3.7	2.7-3.4	2.4-3.0	2.3-2.9	2.1-2.6	2.0-2.5

[a]The figures in this table represent two standard errors. Hence, for most items the chances are 95 in 100 that the value being estimated lies within a range equal to the reported percentages, plus or minus the sampling error.

In order to enable survey users to employ correct statistical inference procedures with these multi-stage sample survey data, we have developed generalized sampling error tables for individual percentages and for the difference between two percentages for varying numbers of interviewers. Here, I shall present just Table 2 for individual percentages. In Table 2 the low level estimates found in the cells give the 95 per cent confidence limits based upon the usual simple random sample formula. The high level estimates take into consideration the additional amount of variance derived from the use of a clustered sample. The procedures and statistical formulas used to obtain these sampling errors can be found in Kish (2) or Lazerwitz (3). The necessary computer program has been obtained from the Sampling Section of the Survey Research Center of the University of Michigan.

To illustrate the use of the table, let us find the sampling error for that 29% of the women of the survey who feel that "professors who advocate controversial ideas have no place in a state supported university." Since the total number of female interviews is 502, we enter the column of Table 2 headed "500" and the row headed "30 or 70". This tells us that chances are 95 out of 100 that this 29 per cent is subject to a sampling error of plus or minus 5.1 per cent (using the high level estimate).

Frequently, the difference between two percentages of the data of the statewide survey exceeds their proper high level estimate of sampling error. Hence two such percentages can be considered significantly different at a 95 per cent confidence level. Occasionally, some of the survey data are based upon percentages whose differences do not exceed their low level estimates. In all such cases, the percentages cannot be considered significantly different. When the difference between two percentages falls between their low and high level estimates of sampling error, the question of significance is considered unresolved. In such situations, it would be best to compute the specific sampling error of the involved difference rather than try to work with generalized tables.

III. Yield and Coverage Expectations

How well did this new sample design turn out with regard to actual sample hu coverage? On the whole, there is a good match between an expected yield of 1328 sample hu's and an actual yield of 1357 sample hu's. Here the excess of 29 sample housing units are primarily a result of the block supplement sample yield in St. Louis City. The very nature of the block supplement sample exposes one to the risk of encountering large clusters of new construction or of unlisted housing units in older structures missed by city directories. It would take extensive field work to avoid such situations which can be better handled by allowing more sample size variation and the technique of a "surprise stratum" (which was utilized for the St. Louis supplement sample).

Footnotes

[1]The block supplement yield on this survey was just 65 hu's, many of which were vacant.

[2]See Kish (1) for these selection tables.

References

(1) Kish, Leslie, "A Procedure for Objective Respondent Selection Within the Household," Journal of the American Statistical Association, 44 (September 1949), 380-87.

(2) ______________, Survey Sampling, New York: John Wiley, 1965, 206-17, 268-70, 282-99.

(3) Lazerwitz, Bernard, "Sampling Theory and Procedures," in Methodology in Social Research, (edited by H. Blalock), New York: McGraw-Hill, 1968, 298-313.

DESIGN EFFECTS IN THE NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS SURVEYS

J. R. Chromy, R. P. Moore, and Anne Clemmer
Research Triangle Institute

I. INTRODUCTION

The National Assessment of Educational Progress [1] is a long-range study of the knowledge, skills, understandings, and attitudes of certain specified subpopulations of young Americans. National Assessment's two major goals are (1) to make available the first census-like data on performance of these specified subpopulations on exercises designed to measure specific educational objectives within selected subject matter areas, and (2) to measure changes in performance on these exercises over time.

To achieve these objectives, a system of annual surveys of four national age groups has been developed. In each of the first three annual surveys, special exercises have been developed in two or three subject matter areas. The Year 01 survey, begun in 1969, involved Science, Citizenship, and Writing; the Year 02 survey involved Reading and Literature. The first opportunity to measure change over a period of time will occur in Year 04 with the reassessment of Science. Some design effects [2] for Science and Writing exercises from Year 01 are presented in this paper. In addition, some preliminary results from selected Year 02 exercises are presented. The design effects in this paper relate to national age group estimates or smaller subpopulation estimates of the proportion of the subpopulation answering specific exercises correctly.

The emphasis in National Assessment reporting on the performance of subpopulations on individual exercises is unique in educational measurement. Because of this feature, it was not necessary to require each respondent to participate in all exercises in any annual survey, but it was necessary only to require that a probability sample of the target population be obtained for each exercise. As a result, the exercises for each year of National Assessment were assembled into several packages and each respondent participated by completing only one of the packages on a probability basis. For respondents assessed through the school sampling frame, packages were administered on a group basis and on an individual interview basis. Individual administration methods were used with respondents contacted through the household sampling frame.

Because of the multiple package (questionnaire) approach to National Assessment and other special requirements for adequately representing subpopulations defined by region, by community characteristics, and by socioeconomic status characteristics, the sample design for National Assessment was extremely complex when considered in its entirety and can only be accurately described by going into considerable detail. A brief description of the Year 01 sample design and estimation procedures, however, is considered necessary before the Year 01 design effects may be discussed.

II. THE YEAR 01 SAMPLE DESIGN

The four target age populations were 9-year-olds, 13-year-olds, 17-year-olds, and young adults of ages 26-35 residing in the United States. The 9-, 13-, and 17-year-olds enrolled in elementary and secondary schools were sampled through a school sampling frame. Out-of-school 17-year-olds and young adults of ages 26-35 were sampled from a household sampling frame; out-of-school 17-year-old data are not included in this paper. In general, the sample design for either frame could be described as a multistage stratified design.

A sample of 208 primary sampling units was selected. Primary sampling units were defined conceptually within larger listing units consisting of cities, counties, portions of counties, or groups of counties. In Year 01, the same sample draw of primary sampling units was used for both the school and household frame. The subsequent stages of the sample were selected independently within these primary sampling units. The primary sample was stratified by region, by community-size characteristics, and by income characteristics. In most cases, two primary sampling units were selected per stratum; in a few smaller strata, only one primary sampling unit was selected.

The sample design for a particular National Assessment package can be described more easily than the entire design for all packages. For the school sample, the design for a group-administered package involved one group of twelve students randomly selected from one school in each primary sampling unit. The design for an individually administered package involved nine students per primary sampling unit (PSU) selected from (typically) 2 to 5 schools.

In the household sample, the second-stage sampling units were area clusters of housing units. Ten second-stage sampling units were selected in each PSU. The package sample size per second-stage sampling unit under this design was a random variable, and averaged less than one. Table 1 gives a quick view of the sample design for both sampling frames.

An attempt was made to sample from certain subpopulations at a higher rate than from others in order to provide adequate subpopulation sample sizes for reporting purposes. The disproportionate sampling methods were applied at both the first stage and subsequent stages of the sample selection whenever the necessary identifying data were obtainable.

As previously stated, the estimates reported from the National Assessment data are the

Table 1

Year 01 Sample Design Summary

Sampling frame	Number of PSU's	Within-PSU Sample
School Frame		
Group-administered packages	208	12 students selected from one school
Individually administered packages	208	9 students selected from 2 or more schools
Household Frame		
Individually administered packages only	208	10 second-stage area clusters. Respondent sample size per package is a random variable.

estimated proportions of subpopulations who would respond correctly to specific exercises. These estimates (P-values) were computed as combined ratio estimates. The denominator of the ratio estimate was an estimate of the size of the subpopulation of interest, and the numerator was an estimate of the number of persons in the subpopulations who would perform in an acceptable manner on the given exercise.

The Horvitz-Thompson estimator was used for estimating both the numerator and denominator values in the ratio estimate; each sample response was given a weight or expansion factor equal to the inverse of the respondent's probability of selection for the package containing the particular exercise. These weights were further adjusted, where necessary, to account for nonresponse.

Variances of these estimates of proportions were estimated by using a first-stage jackknife estimator and ignoring the finite correction factor. Since the finite correction factors were small, this estimator would be expected, on the average, to give a small overestimate. Some collapsing of strata was also necessary to include the variance contributions from strata where only one primary sampling unit had been selected.

III. YEAR 01 DESIGN EFFECTS

National design effects were estimated for 149 science and writing P-values. The median design effect estimate for the 149 exercises examined was 2.38, with the majority of the 149 design effects falling between 1.50 and 3.00. As Table 2 shows, 82 percent of the design effects were 3.00 or less, 87 percent were 3.50 or less, and 94 percent were 4.00 or less. Table 3 shows median design effects and ranges in design effects for various subgroups of national design effects classified by age group, administration mode, and subject matter area.

The design effects for group-administered exercises were higher than those for individually administered exercises due to more clustering of the sample respondents. As stated, each group package was administered once in each PSU to a group of 12 students selected from a single school. For each individual package, 9 respondents per PSU were selected from several schools.

The design effects for 13-year-olds were smaller than those for 9-year-olds, while the 17-year-old exercise design effects were smaller than those for either 9- or 13-year-olds. This should be expected since the older the students are, the more likely they are to be enrolled in larger schools serving larger, more heterogeneous populations. It may be that high schools are more heterogeneous in terms of students than are junior high schools, and both junior high and senior high schools are more heterogeneous than the elementary schools.

The design effects for adults 26 to 35 years of age were about equal to those for 9-year-olds, possibly reflecting a similar intracluster correlation for the household sample due to small, compact clusters and variable housing patterns within PSU's.

There was no apparent difference in design effects for science exercises as compared with writing exercises. The comparison is difficult because of the small number of group-administered writing exercises and the fact that no individually administered science exercises were examined in this study.

Table 2

Distribution of National Design Effects

Design Effect	Number	Percent
< 1.00	1	1%
1.00 - 1.50	16	11%
1.51 - 2.00	29	19%
2.01 - 2.50	43	30%
2.51 - 3.00	32	21%
3.01 - 3.50	8	5%
3.51 - 4.00	10	7%
4.01 - 4.50	5	3%
4.51 - 5.00	3	2%
> 5.00	2	1%
Total	149	100%

Table 3

Median Design Effects for National P-Value Estimates

Age	Administration Mode	Subject Area	Number of Exercises	Median Design Effects	Range of Design Effects	Mean Number of Respondents
9	Group	Science	30	2.68	1.92-4.94	2,442
13	Group	Science	27	2.26	1.31-6.01	2,415
17	Group	Science	10	1.81	.90-2.51	2,122
17	Individual	Science	1	1.13	1.13	579
26 to 35	Individual	Science	16	2.57	1.38-4.08	878
9	Group	Writing	24	2.81	1.51-3.80	2,426
13	Group	Writing	5	4.36	1.93-10.88	2,416
9	Individual	Writing	13	2.21	1.45-2.68	1,817
13	Individual	Writing	23	1.89	1.24-2.88	1,863

Tables 4, 5, and 6 show median design effects for subpopulations defined by regional strata, and for sex and size of community subpopulations defined by poststratification. The median design effects for subpopulation estimates are of about the same magnitude or slightly smaller than the median design effects for national estimates.

The median design effects for 9- and 13-year-old writing exercises tended to be highest for the Southeast region (Table 4).

No consistent differences were noted in the median design effects for males and females (see Table 5).

Table 6 shows median design effects for size of community (SOC) subpopulations defined by poststratification. As with national design effects, the median design effects for SOC subpopulations are higher for group-administered exercises than for individually administered exercises. There is possibly a tendency for the metropolitan and urban area median design effects to be smaller than those for more sparsely populated medium city and rural (small place) subpopulations.

The design effects shown in this paper reflect the combined effects of clustering of the sample, unequal weighting of sample respondents, stratification, and other sample design and estimation factors. The effect of unequal weighting of sample respondents has been estimated to be from 1.3 to 1.6, depending upon the exercise.

IV. SOME PRELIMINARY YEAR 02 RESULTS

Some major revisions in the National Assessment sample design occurred in Year 02. The first principal change involved doubling the within-primary sampling unit sample and halving the number of primary sampling units. The planned number of administrations per individual package was increased to 20 per primary sampling unit.

The second change involved the use of controlled selection of the primary sample to permit stratification by State as well as by the previously discussed set of stratification variables. Due to this change, some approximate variance estimation procedures were required and only partial results have been obtained.

An approximate variance estimation formula based on the Yate-Grundy variance estimator was used to study design effects <u>in one region</u>. To isolate the design effect associated with clustering from the design effects associated with unequal weighting, stratification, and other design factors, the same variance estimation formula was also applied to systematic half-samples of the ordered data file for selected items. The variance estimates from the two half-samples were averaged. Design effects were then computed for both the whole sample and the systematic half-samples.

If a positive intraclass correlation existed, the design effect would be expected to be larger using the whole sample data. With a negative intraclass correlation, the opposite result would hold. Table 7 gives some results for ten selected exercises. A measure of homogeneity is also given for the group-administered exercises. This measure, based on the ratio of the change in design effect to the change in cluster size, should behave similarly to the intraclass correlation coefficient under a self-weighting cluster design. If this condition continues to hold under more complex sample designs, then the method described provides a convenient means of describing the behavior of the variance as cluster size is varied.

Table 4

Median Design Effects for Regional Subpopulation P-Value Estimates

(9- and 13-Year-Olds Only)

Age	Administration Mode	Subject Area	Number of Exercises	Median Design Effects by Region			
				Northeast	Southeast	Central	West
9	Group	Writing	24	1.89	2.93	2.32	2.65
13	Group	Writing	5	3.05	3.65	3.50	2.65
9	Individual	Writing	13	2.34	1.30	1.85	2.17
13	Individual	Writing	23	1.64	2.11	1.61	1.35

Table 5

Median Design Effects for Sex Subpopulation P-Value Estimates

Age	Administration Mode	Subject Area	Number of Exercises	Median Design Effects by Sex	
				Males	Females
13	Group	Science	20	2.57	2.25
26 to 35	Individual	Science	15	2.08	2.20
9	Group	Writing	24	2.74	2.54
13	Group	Writing	5	2.95	4.38
9	Individual	Writing	13	2.27	2.03
13	Individual	Writing	23	1.80	1.84

Table 6

Median Design Effects for Size of Community (SOC) Subpopulation P-Value Estimates

Age	Administration Mode	Subject Area	Number of Exercises	Median Design Effects for SOC Categories			
				Big City	Urban Fringe	Medium City	Small Place
9	Group	Science	30	2.26	2.01	2.56	3.38
13	Group	Science	27	2.43	2.20	2.14	1.90
26 to 35	Individual	Science	16	1.91	2.25	1.47	1.86
9	Group	Writing	24	2.04	2.18	2.41	2.86
13	Group	Writing	5	3.79	2.95	3.82	3.69
9	Individual	Writing	13	1.75	1.97	2.66	1.91
13	Individual	Writing	23	1.22	1.37	1.75	2.38

Table 7

Year 02 In-School Design Effects For 10 Selected Exercises

Administration Mode	Age	Exercise	Sample Size	Estimated Design Effects: Half-Sample	Estimated Design Effects: Whole Sample	An Approximate Measure of Homogeneity
				(D_1)	(D_2)	(Δ)*
Group	9	1	618	2.90	5.81	.485
		2	618	.82	.61	-.035
	13	1	655	2.57	3.78	.202
		2	655	1.89	2.57	.113
	17	1	603	2.98	4.52	.256
		2	603	.15	.69	.090
Individual	9	1	549	1.07	.74	---
		2	549	1.42	2.36	---
	17	1	541	1.79	1.34	---
		2	535	.30	2.90	---

*$\Delta = (D_2 - D_1)/(12 - 6)$ for group-administered exercises only.

The only immediate conclusion one can draw is that the estimates of standard errors and design effects vary considerably from exercise to exercise. Further study is needed to determine if the wide variation in design effects is a real phenomenon or whether such results are obtained empirically only because the variance estimates themselves have a high variability.

REFERENCES

[1] A project of the Education Commission of the States.

[2] Leslie Kish, Survey Sampling. New York: John Wiley and Sons, 1965.

SEPP (SAMPLING ERROR PROGRAM PACKAGE)

Martin R. Frankel, Baruch College, CUNY
and University of Michigan

INTRODUCTION

The SEPP (Sampling Error Program Package)[1] consists of three computer programs which produce sampling errors (BRRP, PSALMS and ABSERD). BRRP (Balanced Repeated Replication Package) computes sampling errors using the method of balanced repeated replications. At present it can compute these sampling errors for ratio means, simple differences of ratio means, simple correlation coefficients, multiple regression coefficients (standardized and unstandardized), (p-1) order partial correlation coefficients, and multiple correlation coefficients.

PSALMS (Paired Selection Algorithm for Multiple Subclasses) computes sampling errors using a method based on the Taylor approximation (also known as the δ or delta method, or linearization or propogation of error, and sometimes the Keyfitz method). At present it can compute sampling errors for simple ratios, and linear combinations of ratios (e.g., differences of ratios). Unlike BRPP it can easily deal with many ratios from different subclasses of the sample in a single run. In addition, unlike BRRP it is designed to deal automatically with poststratification class weights, although this option has not yet been checked out and cannot be used at present. ABSERD (A Basic Sampling Error Program for Ratios and Differences) is a simpler and more modest program appended as a useful addition to the two major programs.

ABSERD

ABSERD (A Basic Sampling Error Program for Ratios and Differences) was written for users who do not have access to large computers and need only a limited number of sampling errors for ratio means and their differences. ABSERD is written entirely in basic FORTRAN, and can be compiled and run on machines with a minimum core size of 16K. Unlike PSALMS and BRRP, the original data cases do not have to be sorted by strata and psu within strata before entry into the program. On a single program set-up and run sampling errors may be computed for a pair of ratio means and their difference. It is possible to compute sampling errors over different subclasses of the total sample in this single run. Because of its simplicity, this program has no provisions for handling missing data (blanks) or missing data codes and values. The user must filter out cards or records that contain these missing data indications to assure that blanks are not interpreted as zeros and missing data codes are not interpreted as data values.

BRRP

BRRP (Balanced Repeated Replication Package) was designed primarily to compute sampling errors for regression coefficients and correlation coefficients (simple, partial and multiple) associated with a single multiple regression equation. As a byproduct, the program also computes sampling errors for means and differences of means for all variables that enter the regression equation.

Sampling errors are computed by the method of balanced repeated replication (also known as pseudo-replication). [1,2] The program allows the user the option of specifing the pattern of replication to be used; fully balanced, partially balanced, or unbalanced. The program can compute sampling errors using either half-samples alone, or half as well as complement-half samples.

Communication between the user and the program is accomplished by the use of standard ISR variable description. This is described in more detail in the section on PSALMS that follows.

BRRP is modularized to simplify the computation of sampling errors for different subsets of independent variables in a reduced regression equation. The program requires an IBM 360 running under OS (Operating System).

PSALMS

The overwhelming majority of estimates produced from survey samples are ratio means (including proportions) and differences between means. PSALMS was designed to efficiently and economically compute sampling errors for these most often used estimators. Unlike other sampling error programs, PSALMS can compute variances for many different means over different

---

[1] *A more complete description of the programs is given in: Kish L.; Frankel, M.R.; and Van Eck, N. SEPP: Sampling Error Program Package. Ann Arbor: Institute for Social Research, 1972. Information about obtaining the programs is available from Neal Van Eck, Institute for Social Research, University of Michigan, Ann Arbor, Michigan, 48106. Support for the writing of SEPP came from the National Center for Health Statistics (PH-43-67-767), the National Science Foundation (GS-3191X), and from the Institute for Social Research.*

subclasses of the sample in a single run. Specification, by the user, of the ratio means and differences, as well as the sample subclasses for which sampling errors are desired is done in three steps.

First, each variable which is to be used in the formation of a mean or which enters into the definition of a sample subclass is assigned a unique identification number. This identification number is associated with a particular data field by the use of an ISR T-card. For each such variable, the user may assign values which he wishes to be interpreted as "missing data". After this assignment of identification numbers these original data variables are referred to by their number preceeded by the letter 'X'.

All means, or more correctly ratios, for which sampling errors are desired are formed as the ratio of two sample sums, a numerator and a denominator. Sometimes these numerators and denominators can be formed directly by summation of original data or X-type variables. But often, particularly if we are interested in a mean for a subclass of the total sample, summation across X-type variables will not give us the needed numerators or denominators. These are created by an internal recode language L1. Using information from the original X-type variables, the user may create 0-1 (B-type) variables or more general (Y-type) variables. For example, the L1 statement B1 IF X32=1-4; creates a denominator for a mean based on the subclass of the sample having values 1 through 4 for original variable 32. The statement Y1=X56 IF B1; forms a numerator containing the value of original variable 56, for this sample subclass.

The last step of the man-machine communication with PSALMS is the specification of the ratio means and differences for which sampling errors are desired. Here the user puts together numerators and denominators, based on either original data variables (X-type) or user created variables (B- and Y-type). For example, the statement R1=Y1/B1, tells the program to compute sampling errors for the mean of variable 56, over the subclass of the sample consisting of all cases with codes 1 through 4 on variable 32.

Like BRRP, PSALMS requires an IBM 360 running under OS.

REFERENCES

1. Kish, L. and Frankel, M., "Balanced Repeated Replications for Standard Errors." Journal of the American Statistical Association, 65 (1970), 1071-1094.

2. McCarthy, P.J., "Pseudo-Replication: Half Samples," Review of the International Statistical Institute, 37 (1969), 234-264.

NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS: AN EDUCATIONAL INDICATOR

Robert C. Larson, W. Todd Rogers, Donald T. Searls,
National Assessment of Educational Progress

In the past, the value of the educational system has most often been equated with the number of people educated by that system. However, recent major emphasis has been placed on the outputs of the educational process. Data is now required that indicates the extent to which skills, knowledge, and attitudes are acquired by young people. Both the Organization of Economic Cooperation and Development [1] and The Office of Management and Budget [2] recently identified the National Assessment of Educational Progress as an educational indicator that provides this desired information. Although a relatively young project, these acknowledgements indicate the potential of National Assessment as an educational indicator of output.

A project of the Education Commission of the States, National Assessment has the purpose of providing dependable information describing what young Americans (9-, 13-, 17-year-olds and young adults 26-35) know and can do. More specifically, the assessment is designed 1) to obtain, at regular periodic intervals, census like data on the knowledge, skills, understandings, and attitudes possessed by various subpopulations in the United States and 2) to measure the growth or decline in educational attainments that takes place over time in learning areas of educational concern. Ten subject areas are assessed: Art, Career and Occupational Development, Citizenship, Literature, Mathematics, Music, Reading, Science, Social Studies and Writing. The assessment timetable through 1980 is summarized in Exhibit 1. Within each subject area, National Assessment determines the proportions of the subpopulations of interest who can acceptably answer a question or successfully perform a task judged important by scholars and lay people and accepted by educators as something that should be taught in our schools. Successive assessments in a subject area will determine changes in these proportions which reflect the magnitude of change in the educational attainment of the various subpopulations for which data has been collected. Thus data will be available that will be of value not only for allocating public and private funds to the educational enterprise, but also to all levels of the educational decision making community and general public.

1. ASSESSMENT TIMETABLE

1969-70	SCIENCE, WRITING, CITIZENSHIP
1970-71	READING, LITERATURE
1971-72	MUSIC, SOCIAL STUDIES
1972-73	MATHEMATICS, SCIENCE*
1973-74	CAREER AND OCCUPATIONAL DEVELOPMENT, WRITING*
1974-75	ART, CITIZENSHIP*
1975-76	READING*, LITERATURE*
1976-77	MUSIC*, SOCIAL STUDIES*
1977-78	MATHEMATICS*, SCIENCE**
1978-79	CAREER AND OCCUPATIONAL DEVELOPMENT*, WRITING**
1979-80	ART*, CITIZENSHIP**

*SECOND ASSESSMENT
**THIRD ASSESSMENT

The intent of this paper is to provide a brief overview of the structure and methodology[1] of National Assessment and to present some of our output measures collected to date and how these might be used as indicators of educational attainment and change.

NATIONAL ASSESSMENT MODEL--METHODOLOGY

Subject matter areas. As mentioned above, ten different subject areas are included in the assessment. The first stage in our procedure is to develop objectives and subobjectives within each subject area. These objectives, developed and reviewed by panels of scholars, educators, and concerned lay people, must be considered important to subject matter specialists, reflect acceptable teaching goals, and be of value and important for young people in a modern society to learn.

Once an appropriate and acceptable set of objectives has been written, several exercises are developed to represent each objective. These exercises are the instruments with which NAEP gathers data to be used to indicate the extent to which the objectives are being met by each of the various subpopulations of people for which results are desired. Unlike more common educational instruments, the exercises used by National Assessment are not designed to discriminate between individuals or groups of individuals. Like objectives, each exercise is developed and thoroughly reviewed by specialists in exercise construction within the area to be assessed and other scholars, educators and concerned lay people.

From the pool of developed exercises

[1]For a more thorough discussion of National Assessment procedures, see references [3].

representing the area to be assessed, a set of exercises is selected and used in the actual assessment so that the number of minutes of administration time at each age level takes between 150 and 200 minutes. No student enrolled in school is allowed to spend more than 45 minutes in the assessment. Therefore, the selected exercises are divided into about 12 different sets of packages. Typically, nine of the packages include those exercises that can be group-administered to twelve respondents at once. The remaining packages are administered to one respondent at a time. These individually administered packages include performance exercises (such as a science experiment or playing a musical instrument) and exercises which require a "mini-interview" of the respondent. (The number of different packages varies from year to year and from one age to another.) In each package there are both exercises which are unique at each age and exercises which overlap across two or more adjacent age levels.

Sampling and administration. National Assessment policy is to provide data at a national level and avoid making comparisons among states, schools and school districts, teachers, or individuals. This policy is reflected in our sampling. A national probability sample of approximately 2500-2600 individuals per group administered package and 2100-2200 individuals per individually administered package are assessed each year. The samples, selected using a multi-stage design, are stratified by region, size of community, and socio-economic status. A school frame is used for the 9-, 13- and about 87 percent of the 17-year-old samples. Young adults are obtained through a household frame; 17-year-olds not enrolled in school are assessed through the adult household frame and a specially designed frame of dropouts and early graduates. The total sample size required for the three younger ages is approximately 29,000. Since young adults are permitted to take up to four packages each, approximately 4800 adults are assessed each year. In the last two assessments, approximately 95% of the selected schools and 82% of the selected eligible adults have taken part.

The packages are administered by specially trained exercise administrators. Group administered exercises are presented on tape to standardize procedures and to facilitate understanding for young people with poor reading ability. Individually administered packages are presented in an interview mode according to well documented instructions for administration. Thirteen year olds are assessed during October and November, 9s during January and February, and 17s in-school during March and April. The supplementary frame for 17s out-of-school is completed during May-July; the adult assessment runs from October through May.

Analyses. The completed packages are carefully edited for errors. Some of the exercises are in multiple choice format in which respondents record their answers by darkening ovals; others require scoring by specially trained people with subject matter background in the areas assessed. This data, weighted for disproportionate sampling, forms the base for all analyses. For each exercise the estimated proportions of acceptable and unacceptable responses within each population of interest is computed.

Approximately fifty percent of the assessed exercises are retained for future assessments and are therefore not released. These exercises provide the basis for measuring change across time and are selected to provide reasonable coverage across all objectives, difficulty levels and populations of interest.

Since no respondent receives all the packages, the concept of a "total score" is inappropriate. Instead, results are reported by individual exercise, as well as by a summary value (e.g., the median) for small subsets of exercises which measure the same objective, or have some other common attribute.

OUTPUT MEASURES AND POTENTIAL INDICATORS

Output measures. Present reporting categories include age, and within age, geographic region, sex, size and type of community, educational level of parents, and color.[2] The age levels correspond to

2. REPORTING CATEGORIES PRESENTLY USED BY NATIONAL ASSESSMENT

CLASSIFICATION	NUMBER OF SUBGROUPS	SUBGROUPS
Age level	4	9, 13, 17, 26-35 years
Sex	2	Male, Female
Geographic region	4	Northeast, Southeast, Central, West
Level of Parental Education		No high school, Some high school, Graduated high school, Post high school
Size and Type of Community	7	Extreme inner city, Extreme affluent suburb Extreme rural, Inner city fringe, Urban fringe, Medium-size city, Small city
Color	2	Black, White, other

the end of primary, intermediate, secondary, and formal post-secondary education. The categories Region and Sex have traditionally shown large differences. School districts are thought to vary most

[2]See the list of references for released reports by subject area.

by Size-and-type of Community, and Color and Parents' Education are thought to do well in differentiating socio-economic levels and home and family environments.

Each output measure for an exercise is an educational indicator which possesses meaning for the factors reported. One should be careful not to overemphasize the effect of the identified factor used in labeling the reporting category. Each label for each group represents not only the factor indicated by its name, but also a variety of other factors which may influence the performance of that group.

Several examples are presented which represent the various output measures we have been reporting over the last two years. The examples are presented in a mixed fashion across subject areas and objectives in order to give a general view of our output measures as well as illustrate various aspects of our data.

3. OUTPUT MEASURES FOR HAWK EXERCISE (SCIENCE). PERCENT OF 17 AND YOUNG ADULT RESPONDENTS.

OBJECTIVE: II. Possess the abilities and skills needed to engage in the processes of science

SUBOBJECTIVE: F. Ability to check the logical consistency of hypotheses with relevant laws, facts, observations, or experiments

IN A PARTICULAR MEADOW THERE ARE MANY RABBITS THAT EAT THE GRASS. THERE ARE ALSO MANY HAWKS THAT EAT THE RABBITS. LAST YEAR A DISEASE BROKE OUT AMONG THE RABBITS AND A GREAT NUMBER OF THEM DIED. WHICH OF THE FOLLOWING PROBABLY THEN OCCURRED?

17	ADULT	RESPONSES
4%	2%	⊂⊃ THE GRASS DIED AND THE HAWK POPULATION DECREASED.
1%	1%	⊂⊃ THE GRASS DIED AND THE HAWK POPULATION INCREASED.
68%	52%	⬬ *THE GRASS GREW TALLER AND THE HAWK POPULATION DECREASED.
4%	4%	⊂⊃ THE GRASS GREW TALLER AND THE HAWK POPULATION INCREASED.
20%	30%	⊂⊃ NEITHER THE GRASS NOR THE HAWKS WERE AFFECTED BY THE DEATH OF THE RABBITS.
2%	10%	⊂⊃ I DON'T KNOW.
1%	1%	⊂⊃ NO RESPONSE.

4. OUTPUT MEASURES FOR GOVERNMENT POWER EXERCISE (CITIZENSHIP). PERCENT OF SUCCESS FOR 9, 13, 17, AND YOUNG ADULT FOR SEVERAL REPORTING CATEGORIES

OBJECTIVE: IV. Know the main structure and functions of our governments.

SUBOBJECTIVE: B. Recognize the main functions and relations of governmental bodies; recognize that governmental powers are limited by the people through the constitution.

SCORING: Examples of acceptable reasons to Response C: people could stop him; elected officials could stop him; checks and balances system of government; laws stop him; country would be a dictatorship; not the democratic way.

A. DOES THE PRESIDENT HAVE THE RIGHT TO DO ANYTHING AFFECTING THE UNITED STATES THAT HE WANTS TO DO?

⊂⊃ YES
⊂⊃ NO
⊂⊃ I DON'T KNOW

B. (IF YES) WHY?

C. (IF NO) WHY NOT?

	Stated that the President does NOT have the right to do anything affecting the United States that he wants (No to A).				Stated that the President does NOT have the right and gave an acceptable reason (on C).			
	9	13	17	AD	9	13	17	AD
National	49.2	72.3	78.4	89.0	18.4	51.7	70.8	80.0
Region								
Northeast	50.2	72.2	79.6	89.6	22.6	57.0	75.2	80.5
Southeast	47.3	70.3	76.9	85.9	16.7	43.9	66.8	74.1
Central	49.2	72.7	79.4	87.6	15.1	51.1	69.5	79.6
West	49.7	73.7	77.2	92.8	19.7	53.8	71.3	84.9
Sex								
Male	49.9	78.9	82.7	91.8	18.4	57.9	76.1	81.7
Female	48.5	65.9	74.6	86.5	18.4	45.7	66.1	78.5
Parents Educ								
No H.S.	45.7	57.2	75.6	84.9	18.7	31.4	70.1	72.3
Some H.S.	32.2	66.5	68.5	92.7	10.8	43.3	58.9	83.5
Grad H.S.	46.5	66.4	76.1	90.5	16.1	47.8	69.5	82.7
Post H.S.	58.7	79.3	84.5	92.0	25.2	57.9	77.2	86.4

The percentages reported for the two exercises above are measures of the level of performance of an entire age group (nation) or reporting category. By themselves these percentages are not useful indicators unless reference points can be defined upon which meaningful judgments about the performance can be based.

Several meaningful reference points can be constructed from the data itself. The national level of performance of an entire age group, for example, provides individual reference points for each exercise. The difference between subgroup performance and the national level of performance is useful for describing how uniformly learning is spread among the subgroups defined on one exercise. (See Exhibit 5.)

For exercises retained for later assessments, a similar reference point is defined for the second cycle. A comparison of the differences between subgroup performance and the national performance from one cycle to the next indicates how the relative distribution of learning across subgroups of interest is changing over time on a single task. Also, the difference between proportions of success from assessment to assessment determines if more or fewer people in an age group or subgroup of interest are able to respond acceptably.

For exercises administered to more than one age group during an assessment year, similar kinds of comparisons between ages can be made in the form of indicators which answer such questions as: At what age have most of the people gained the understandings measured by the exercise? or Is the distribution of learning across subgroups for a given exercise similar at different ages?

All the output measures considered so far have been for single exercises. Often it is more convenient to summarize the output measures for all exercises of one type or classification (such as the set of exercises reflecting one objective, typed as physical science or biological science, classified by goal in citizenship or by theme in reading and literature). The median of the differences between subgroup performance and national performance for a single classification of exercises is the basic summary statistic we have reported. In a sense, this median value is "typical" of the performance of the respondents in one subgroup relative to all respondents of the same age for exercises of one type included in the assessment. Thus, generalizations about the "tendencies" of a subgroup with respect to certain types of exercises can be formed.

5. OUTPUT MEASURES FOR POEM EXERCISE (READING). DIFFERENCE IN PERCENT CORRECT BETWEEN RESPONDENTS IN A REPORTING CATEGORY AND ALL RESPONDENTS.

Read the poem and answer the questions which follow it.

THE WAYFARER

The wayfarer,
Perceiving the pathway to truth,
Was struck with astonishment.
It was thickly grown with weeds.
"Ha," he said,
"I see that no one has passed here
In a long time."
Later he saw that each weed
Was a singular knife.
"Well," he mumbled at last,
"Doubtless there are other roads."

A. What do the following lines from the poem mean?

"I see that no one has passed here
In a long time."

- The way of truth is popular.
- People are fascinated by the truth.
- Truth comes and goes like the wind.
- The truth is difficult to recognize.
- Few people are searching for the truth.

- I don't know.

THEME: VII. Receding and Drawing Inferences.
A. Drawing Inferences from Information Given

OBJECTIVES: IB3a: Comprehend what is read. Interpret nonliteral and figurative language. Recognize the connotations as well as the denotations of the words read.

IIA2. Analyze what is read and be able to trace sequences. Follow the development of an author's idea.

	13	17	Adult
National	61.5	69.2	66.9
Region			
Northeast	1.2	-0.8	4.9
Southeast	-2.7	-4.1	-13.8
Central	0.6	1.8	-1.8
West	0.9	2.2	6.0
Sex			
Male	-4.9	-3.6	-0.5
Female	5.2	3.8	0.5
Parents' Educ			
No H.S.	-15.0	-12.7	-9.3
Some H.S.	-6.0	-10.7	-2.3
Grad H.S.	0.0	1.5	1.0
Post H.S.	8.3	7.5	16.1

Potential Indicators. The output measures and indicators discussed in the previous section are ones which National Assessment has used in their reports and are derived from the data. Different reference points, independent of the data, can be defined, each reflecting concerns and interests of special groups, such as teachers, curriculum people, policy makers or lay groups. For example, the estimation of expected levels of performance establishes a meaningful and relevant base for different interest groups to judge the performance on an exercise. The task of setting expectations has been undertaken by a group of prominent science educators representing the National Science Teachers Association. They are interested in making National Assessment data on Science exercises more meaningful to science educators. Thus, part of their effort has been to make subjective (expert) judgments of an expected level of performance as a basis for evaluating the assessed performance. National Assessment will continue to study procedures for establishing expectation levels of different groups of people. Of particular concern is the "public's" expectation of how our nation should perform and to provide information which makes comparison of performance with public expectation convenient.

Several states have recently conducted state assessments in which National Assessment exercises were used. Thus, they are able to compare directly their performance with the performance of similarly defined reporting categories in the nation--e.g., it is important to find out if those in the rural areas of a state perform as well as all rural areas in the nation.

National Assessment is currently investigating interactions between subgroups as an additional output measure and potential indicator. It is informative to know, for example, if Blacks in the Southeast do as well as those in the Southeast in general, and if Blacks in the Northeast do worse than those in the Northeast in general. These pieces of information give different implications for allocation of resources to the schools in those regions. The first, if true, would imply that we allocate resources to all schools in the Southeast, regardless of color composition; the second implies that we allocate resources to schools in the Northeast which are predominantly Black.

Aggregate indicators are useful to relieve the problem of too much information which obscures important results. However, there are certain sizeable problems inherent in the development of aggregate indicators from National Assessment data which make such indicators hard to visualize. These problems include differential weighting of exercises, determining units of measure, and the location of meaningful reference points and dimensions. As a possible solution, we have been investigating the use of a singular value decomposition of a non-square array of exercises by reporting categories (cells contain output measures). This could lead to various meaningful linear combinations of subgroups or exercises falling along single interpretable orthogonal dimensions. These linear combinations of exercises could be formed at the objective, theme, goal, subject, between subject levels or even between subjects assessed in different years. Each combination providing different, useful summary measures on "sets" of exercises. Similarly, linear combinations of reporting categories weight exercises along orthogonal "cultural", "psychological", "socio-economic", and "behavioral" dimensions.

National Assessment has a relatively short research history. Much of our time has involved "doing the business of assessing." We have just begun to generate ideas for educational indicators. The task is not small, and the challenge is there. We invite your participation.

REFERENCES

[1] A Goal-Oriented Framework of Educational Statistics. Office of Economic Co-operation and Development, 1971.

[2] Social Indicators: Education, Office of Management and Budget, Working Paper, March, 1972.

[3] Report 02-GIY Reading and Literature--General Information Yearbook (Preliminary report). Available from National Assessment, 1860 Lincoln Street, Denver, Colorado 80203.

[4] Report 1 Science National Results; Report 2 Citizenship National Results; Report 3 Writing National Results; Report 4 Science Results (by region, sex, size of community); Report 5 Writing Results (by region, sex, size of community); Citizenship Results (by region, sex, size of community); Report 8 Writing Mechanics. Available from the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402.
Report 7 Science Results (raw and balanced by region, sex, size and type of community, color, parental education); Report 02-R-00 Reading Summary; Report 02-R-20 Reading (Selected Exercises). Available from National Assessment, 1860 Lincoln Street, Denver, Colorado 80203.
Literature, Music, Social Studies and Writing Reports will be released during the next year.

MONITORING AND INTERPRETING TURNING POINTS IN EDUCATIONAL INDICATORS

Abbott L. Ferriss, Emory University

In 1965, when the 125th annual meeting of the American Statistical Association was held in Philadelphia, Wilbert E. Moore and Eleanor Bernert Sheldon presented a paper entitled, "Monitoring Social Change: A Conceptual and Programmatic Statement" (8). In it they called for contributions to understanding "large-scale structural change in American society" and the collection and analysis of new and better data. Their focus on social change, "large-scale structural transformations", was to monitor indicators of "structural alternations" so as to use "such information for entry into the system, to alter the magnitudes, speed, or even direction of change in terms of explicit, normative criteria" (p. 144). The scheme they presented at that meeting eventually grew into the volume, Indicators of Social Change (11). Since then a kind of "social indicators movement" has arisen. The program of the current meeting of the A.S.A. schedules at least five sessions explicitly on social indicators, and the N.S.F. program of research support for social indicators exceeds two million dollars (FY71).

I wish to point out the critical role played by turning points in the process of monitoring the social systems, with particular reference to education, and to illustrate some of the problems of interpreting current turning points and comment on the importance of the monitorship function.

The Monitorship Function

In the literature on social indicators (1) (10), it appears that insufficient attention is devoted to the monitorship function called for by Moore and Sheldon. There is concern with finding appropriate statistical indicators, with tracing the interrelationship of an indicator through a social system, with developing models that will help to predict the future, but insufficient attention to identifying critical turning points in available statistical indices and tracing out the implications of such changes. Perhaps, in the field of education, what passes for monitorship comes with the announcement each year of the results of a new periodic survey, showing that the new data reveals a percentage increase or decrease over the last year, with reference only to the evidence of the survey being reported. Monitorship should be more than this.

What is needed to monitor a social system? Statistical time series of characteristics of the system are the first requirement. For maximum clarity in interpretation, these time series should be the elements in a theoretical model that is capable of expressing cause-effect or, as a minimum, associated sequences of sufficient reliability to predict a future state of the system. The evidence of past trends in the major statistical elements of the system should be at hand. Predictions of trends in the major statistical elements, also, are quite useful, in that they serve to foretell possible turning points.

The monitorship function, itself, consists of placing the new observation of the current year in its context with the past. Has a turning point occurred? Does the new observation deviate markedly from past trends? If so, does the change have significant consequence for the future, particularly for other normatively significant elements in the system?

In presenting such implications the monitor must draw information, not merely from the "new" datum of the current survey, but he must incorporate knowledge from the entire range of studies that have attempted to understand the phenomena of the system. Parenthetically, it is for the latter reason that those who conduct and report periodic surveys are not always the best interpreters and monitors of the periodic statistics that they generate. Reports of new data from periodic surveys seldom incorporate data from other studies.

If the monitor judges that a turning point indicates changes in other significant elements of the system, then the consequences of the change must be specified. The monitorship function has been completed when these consequences are identified and the information placed in the hands of decision makers whose decisions bear upon the phenomena under study.

Educational Indicators

The field of education can boast that it displays a long list of statistical time series (6). Not all, however, are theoretically nor even normatively significant. For example, the vast arrays of financial statistics cannot be interlocked with changes in the amount of learning that has taken place. Models also have been constructed of various features of the educational system, and some of these models, even, enable prediction, such as predictions of enrollment, of teachers required, and of degrees granted. Unfortunately, not all of the normatively relevant observations in education have been periodically and systematically made: witness the lack of time series on the quality of elementary and secondary school learning, on a measure of education other than the time devoted to it, on indicators of the dis-functional results of the secondary school's emphasis upon achievement among those who lack an achievement orientation, the lack of a measure of quality by field in post-baccalaureate education, and others too numerous to mention. Even with these deficiencies, however, the educational system displays a greater wealth of time series than other institutional areas, except, of course, the economic area.

The potential consumers of the results of monitoring the educational system include all those who make decisions relative to educational policies or programs. Educational staffs, boards of control, the legislators, the departments of education, and the associations of teachers, etc., in short, all who hold a stake in the educational process should be informed of turning points in educational indicators, and alerted to the implications of the change.

Educational indicators are statistical time series that reflect the state of some feature of the educational system. While educational indicators respond to major events that affect the social system, such as wars, major depressions or prolonged periods of prosperity, they are not strictly cyclical. Many series display a long-term growth that is characteristic of our expanding population and economy. Major structural changes of the society have direct effect upon educational indicators. One of these is the shift from farming to industrial and commercial occupations, and the continuing transition to service occupations; these structural transformations have influenced the content of the curriculum and the specialization of the educational output. National policies concerning education have unquestioned bearing upon the educational system: the decision to support or to limit the support of graduate education, the Supreme Court decision ordering school desegregation, the support of pre-school educational centers --these and many other national policies have influenced trends in various educational indicators.

I would like to illustrate several turning points in educational indicators and suggest interpretations and implications in keeping with the monitorship function, outlined above.

Example 1: Continuation Ratio, 10th to 11th Grade

The ratio of school enrollment in successive grades provides an indicator of school retention (6, pp. 29-44). Over the nearly sixty-year span shown in Figure 1, the continuation ratio of 10th graders presents almost a sinewave. The secular trend, however, has been strongly upward despite the numerous turning points. The two major World Wars and the Korean Conflict have been distinctly negative influences. From inspection it appears that some economic depressions adversely affect continuation in school while others do not. Partial correlation coefficients, however, clarify these relationships. With the long-term trend held constant, the continuation ratio is positively associated with unemployment (.45), negatively associated with the per capita gross national product (-.36), and, of course, negatively associated with the expansion of military forces (-.26). That dropouts are more likely to occur during prosperous periods has also been observed by Beverly Duncan (5, p. 128). The effect of military expansion might be interpreted as entrance into the civilian labor force as well as joining the Armed Forces. In any event, these two factors and the linear trend explain 88 percent of the variance in the 10th to 11th grade continuation ratio. While other influences undoubtedly stimulate retention or affect dropouts, a more elaborate model would be required along with additional time series to adequately explain the turning points in school continuation and enable their prediction.

Example 2: Ratio, Baccalaureates to High School Graduates Four Years Earlier

The ratio of baccalaureates to high school graduates four years earlier (6, pp. 107-112) presents several turning points of interest in the study of major events, societal trends and policies affecting the educational system. The ratio itself is an imprecise measure, since its assumption is not uniformly met across time: the assumption that high school graduates uninformly proceed immediately to college and graduate four years later with a bachelor's degree. This defect in the statistic is illustrated by the artifically rapid increase after World War II, reaching its peak in 1950, as a mass of veterans, delayed in their education, graduated. The turning point about 1954 signifies a return of the system to more normal student input-output relationship.

Before 1895 baccalaureate degrees were

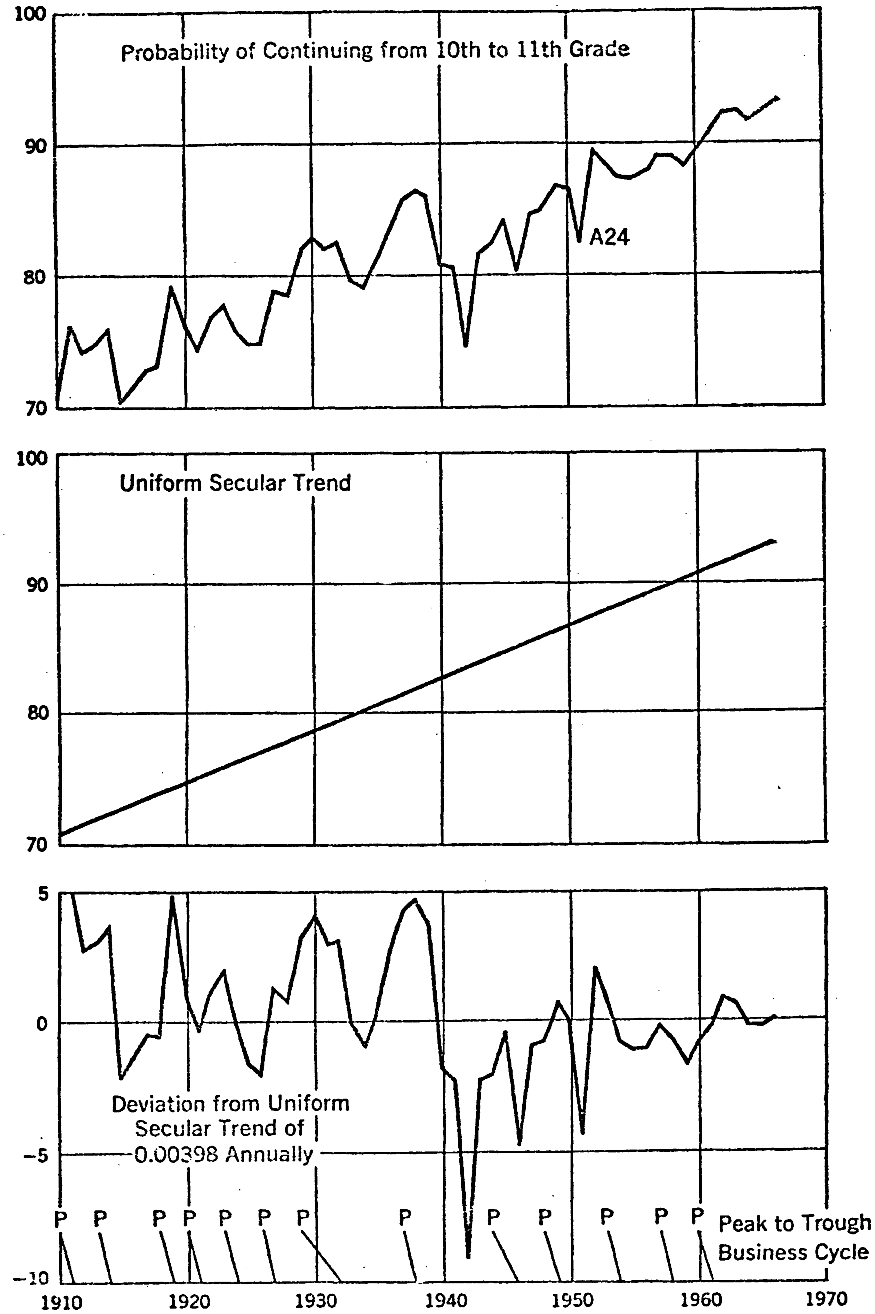


FIGURE 1. Tenth to Eleventh Grade Continuation Ratio, 1910-1966

Source: (6. p. 30)

granted to approximately half of the high school graduates four years earlier. The ratio dropped after 1895 because of the increasing rate of attendance in secondary schools. The need for vocationally trained employees that came as a result of industrialization following the Civil War, finally, in the 1890's, reached secondary education. The number of high school graduates, expressed as a percent of the population 17 years of age, began to increase. However, they did not continue to college and bachelor's degrees four years later did not then begin to increase. The secondary school was "transformed", as Martin Trow has reminded us (9), into a vocational system, not for the elite few as it had been, but for the masses. The turning point in this ratio, then, in the mid-1890's set in motion a trend that did not reach its trough -- and hence its upturn -- until after World War I. Industrial and commercial expansion during and after World War I created a demand for college trained engineers, accountants, business executives, and the like. Educational policies responded and a mild upturn in the ratio was the result. As the Depression of the 1930's came on, per capita baccalaureate production at first faltered and then, as the prospects of an international conflict became apparent about 1937, began to increase and peaked in 1940. With World War II college and universities produced far fewer graduates. Under these conditions, the ratio declined and bottomed-out about 1947.

The ratio appears to have reached a modest level of 28 per hundred high school graduates four years earlier. Projections suggest that the ratio will remain at this level in the immediate future: no turning point is anticipated.

Aside from the adjustment of the educational system to World War II, the educational policies that have brought about these major turning points in this educational indicator were the consequence of industrial, economic, and social changes of the last half of the last century, changes that created a demand for vocational education. The vast increase in the proportion of clerical and kindred workers experienced by the occupational structure verifies the change. American secondary education has now returned to the need for a college preparatory system of education and, as Trow says, continues to contend with a number of policy issues and problems that hinge upon this change.

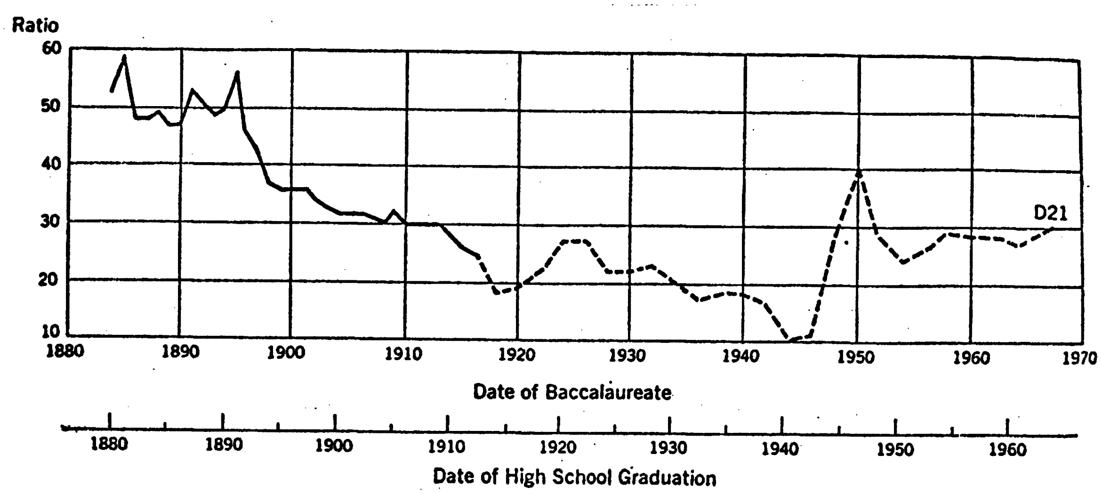


FIGURE 2. Ratio, Baccalaureates to 100 High School Graduates Four Years Earlier, 1884-1967

Source: (6, p. 110)

Example 3: School Enrollment of 18-19 Year Old Males

School enrollment is the subject of a supplement to the October Current Population Survey (2). The percent of 18 and 19 year olds enrolled in school by sex is shown in Figure 3. The decline in enrollment of this age group of males bears significantly upon future college enrollment and the future educational attainment of the labor force. In two years, 1968 to 1970, the percent of white 18-19 year old males enrolled in school declined 5. 5 percentage points while non-white males, 18-19, declined 9. 6 percentage points.

To investigate the influencing factors, the analyst would require information in time series on the number of 18 and 19 year olds drafted into military service, by month, the marriage rate of 18 and 19 year olds, the number of 18 and 19 year olds by sex in the labor force, unemployed, seeking work, and so forth, the number of 18 and 19 year olds by sex receiving unemployment benefits, and other such indicators.

Only an approximate assessment of these influences is possible. Between October 1969 and October 1971, the number of 18 and 19 year old males enrolled in school declined by approximately 53, 000. In 1968, 60. 4 percent were in school and an estimated 15. 6 percent were in the Armed Forces. The next year the percent in the Armed Forces dropped to 12 percent and the next year to below 11 percent. Deferments from Selective Service to attend school have dropped about 1 million since 1968 (4). If Selective Service policies prior to the adoption in December 1970 of the lottery system encouraged school attendance (youth enrolling to avoid military service), then the lottery system and the reduced size of recent draft calls have tended to discourage school attendance.

Meanwhile, the labor force participation

18-19 Year Old School Enrollment

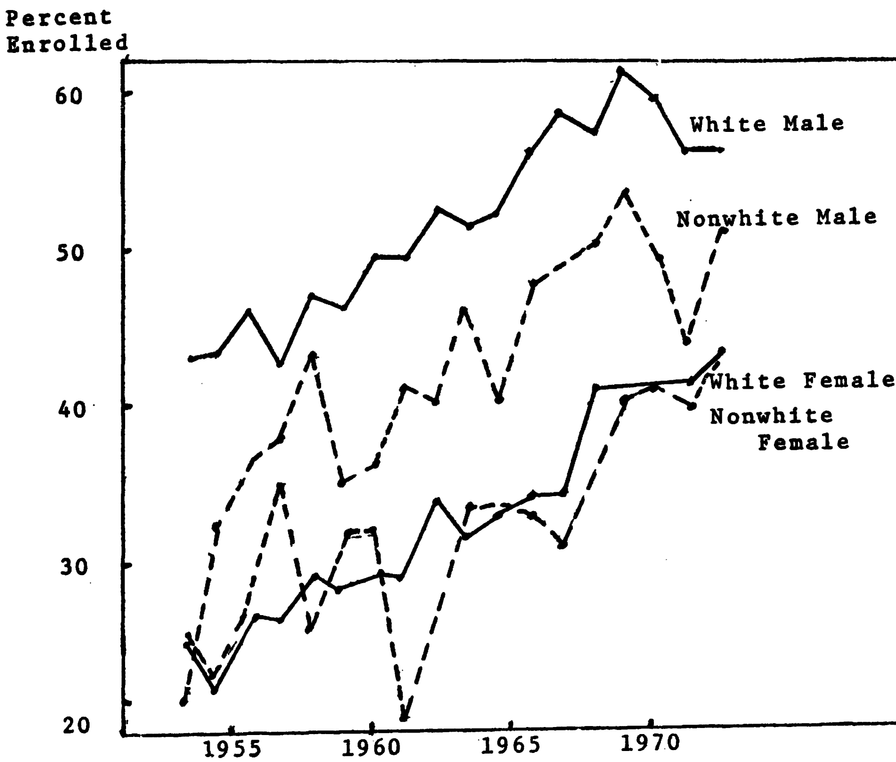


FIGURE 3. Percent of the Population 18 and 19 Years of Age Enrolled in School, 1954 - 1971 (October), by Sex and Color

Source: (7, p.23) (2)

rate (of 18 and 19 year old males) has remained fairly constant at 70 percent. The labor force participation rates of the enrolled 18-19 year old males and the rates of the non-enrolled males both dropped from 1969 to 1970. During 1971 the unemployment rate averaged 15 percent.

In sum, the percent enrolled in school is declining. The percent in the Armed Forces is declining. The labor force participation rate is constant, and the unemployment rate is high.

Other influences may be discouraging school attendance. There may be a negative reaction among youth to the achievement orientation within education. There also may be a reaction against the "Establishment" that is manifested in rejecting school attendance. There may be other interpretations.

Identifying the underlying cause is perhaps the most important function of monitoring social indicators, but estimating the future consequences of the turning of an indicator, also, is important. In this case it is quite obvious that a decline in educational attainment of the cohort at a future age will take place. The increments in enrollment of the 18-19 year olds are followed two years later by increments in the educational attainment of 20-21 year olds, as Table 1 illustrates. This educational deficiency is likely to follow this cohort through its life cycle.

TABLE 1. Incremental Change in School Enrollment and Educational Attainment For Cohorts

Year of Enrollment Increment	Increment in Percent Enrollment of 18-19 Year Old Males	Increment in Percent with at least 1 Year of College of 20-21 Year Olds (Two Years Later)
1966 to 1967	-1.5	-1.3
1967 to 1968	+1.4	+1.8
1968 to 1969	-1.0	-2.1
1969 to 1970	-5.0	?

Source: (2) and (3)

Concluding Remarks

The turning points in three educational indicators that were examined have illustrated the process of interpretation of educational indicators. The multivariate analysis of the 10th to 11th grade continuation ratio produced a definitive interpretation but it would be improved through the addition of predictors. The socio-historical mode of interpretation of the ratio of baccalaureates to high school graduates four years earlier rests upon the validity of a number of logical and factual assumptions. The analysis undoubtedly could be augmented by additional time series that would support the interpretation. Finally, the drop in the percent enrolled of 18-19 year old males can only partially be explained by the change in Selective Service policies. The search for more substantial explanations has not thus far been rewarding.

This paper has not attempted to elaborate upon the models implied in the examples presented. Neither has it attempted to review problems of forecasting turning points in education. Both topics, however, are important steps in developing adequate interpretations. Finally, monitoring educational indicators is for the purpose of informing decision makers, and nothing has been said as to how decision makers in education are to be identified and the signals implied by turning points relayed to them. There is, however, a tested method for accomplishing this task.

Continuous monitorship of educational indicators, conducted in the manner described above, would facilitate the development of policy and programs to adjust to disfunctional educational trends. Interpreting turning points through identifying the directing forces that undergird the change would lead directly to knowledge of the cause-effect sequences that affect the educational system. Many educational indicators are available. Studies of interaction among significant elements within the educational system have been and continue to be made. The time now has come when systematic monitorship of the educational system may produce real contributions to policy and program development.

REFERENCES

(1) Agocs, Carol, "Social Indicators: Selected Readings," The Annals of the American Academy of Political and Social Sciences, March, 1970, 127-132.

(2) Census, U.S. Bureau of, School Enrollment, Series P-20, Washington: Government Printing Office, for dates indicated.

(3) Census, U.S. Bureau of, Educational Attainment, Series P-20, Washington: Government Printing Office, for dates indicated.

(4) Director, Selective Service System, Semiannual Report of the Director of Selective Service, Washington, D.C.: U.S. Government Printing Office, 1971.

(5) Duncan, Beverly, "Dropouts and the Unemployed," The Journal of Political Economy, LXXIII (April 1965): 121-134.

(6) Ferriss, Abbott L., Indicators of Trends in American Education, New York: The Russell Sage Foundation, 1969.

(7) ———, Indicators of Trends in the Status of American Women, New York: The Russell Sage Foundation, 1971.

(8) Moore, Wilbert E., and Eleanor Bernert Sheldon, "Monitoring Social Change: A Conceptual and Programmatic Statement," in Proceedings of the Social Statistics Section, The American Statistical Association, Washington: American Statistical Association, 1965, pp. 144-149.

(9) Sheldon, Eleanor Bernert, and Wilbert E. Moore, Indicators of Social Change: Concepts and Measurements, New York: The Russell Sage Foundation. 1968.

(10) Trow, Martin, "The Second Transformation of American Secondary Education," International Journal of Comparative Sociology, II, 1961, 144-166.

(11) Wilcox, Leslie D., Ralph M. Brooks, George M. Beal, and Gerald E. Klonglan "Social Indicators: Recent Trends and Selected Bibliography," Sociological Inquiry, 42 (1) (1972): 37-50

WHY HEALTH INDICATORS

Karl D. Yordy, Institute of Medicine

The perspective of this paper is the use of health indicators in the context of policy formulation and decision-making, with particular reference to the U. S. Government experience. I leave to others the important technical issues involved in the generation and use of health indicators. This policy perspective will not emphasize those valid motivations for health indicators that are related to long-range research objectives or man's innate curiosity. If the pursuit of social research were the main motivation for developing health indicators, I suspect that the course and pace of their development might be quite different. This paper also considers policy formulation in its entirety, including the political process. This carries beyond the domain of many policy analysts who consider their task completed when rational alternatives have been presented to the decision-makers. I believe that the scope of this perspective is important since no discussion of why health indicators are desirable is complete without including the political dimension of decision-making. If health indicators are to have a real role in affecting decisions, they must be a factor in the final formulation or re-formulation of policy, not just in the presentation of the issues.

Let us start then with the basic question of why the policy maker or manager should be interested in supporting the development and use of health indicators. From the perspective of the policy maker, we should accept Kenneth Land's definition of a social indicator which emphasizes "that the criterion for classifying a social statistic as a social indicator is its informative value which derives from its empirically verified nexus in a conceptualization of a social process." 1/ The policy maker's conceptualization is based on his defined responsibilities. He will be interested in data to the extent that he can perceive that data as helping him to carry out these responsibilities. He therefore looks for data that will be useful in predicting the future course of events, in assessing the impact of policy alternatives, in justifying a decision to others who must approve, and in evaluating the results of those decisions over time. All of these uses press him toward a concern for measurement of output and establishing a cause and effect between inputs and that measurable output.

This ideal statement of policy maker's concerns, however, is usually modified substantially amidst the pressures for immediate decisions based on very imperfect information. The recognition by both government officials and social scientists that decisions are too little informed by relevant data has led in recent years to a rising interest in the development of social indicators. Often the economic policy arena is cited as an example to be emulated. One specific example of this interest was the effort of the Department of Health, Education, and Welfare several years ago to develop a social report. That effort had very limited success, but the potential usefulness was described as follows:

"A social report with a set of social indicators could not only satisfy our curiosity about how well we are doing, but it could also improve public policy making in at least two ways. First, it could give social problems more visibility and thus make possible more informed judgments about national priorities. Second, by providing insight into how different measures of national well-being are changing, it might ultimately make possible a better evaluation of what public programs are accomplishing." 2/

The field of health is a particular case of this general interest, and the pressures to better inform the health policy process are steadily growing. I believe that most policy makers now view health as one of the desirable end results of the society, not just as an intermediate contributor to other social ends such as economic prosperity. This view of health, however, is usually a general feeling rather than a quantifiable definition. The realization of how ill-defined the objective is becomes one of the strong motivations for the policy maker's interest in health indicators.

There is a recent urgency in the policy maker's desire for quantifiable indicators of health, direct or indirect. This heightened concern results from a change towards interventionist public policies in the health field. How health services are organized, financed, and distributed, the impact of environmental and social factors on health, and our understanding of disease processes and their cures have all become specific public policy concerns in recent years. The fact that deficits exist in the health data now recognized as necessary for sound decision making, should come as no surprise when we consider how recent is this intensity of public interest in health policies. Prior to the last ten to twenty years, the narrow scope of public responsibilities generated relatively little demand for policy-relevant health data. Without that need to know, there was insufficient motivation to incur the costs associated with requiring relevant information and to overcome those political pressures which often make a virtue out of not knowing. Some concern for public policies affecting health existed for many years, especially in the control of communicable disease, the support of medical research, and some limited concerns about safety and environmental hazards. However, dramatic changes in public policy took place during the middle 1960's. Some illustrative highlights are: the community mental health centers program; Medicare and Medicaid; the establishment of neighborhood health centers focused on the poor; organized comprehensive health services for mothers and children; regional medical programs to combat heart disease, cancer, and stroke; comprehensive health planning; and the establishment of the National Center for Health Services Research and Development. This was a period of identifying deficits in the health of the nation and of high expectations that those deficits could be overcome by explicit public action.

It is instructive to look at the use of health data during that period of rapid policy formulation. Data was used extensively in the justification of the new health service programs that I have mentioned, including much use of data on morbidity and mortality from the National Health Survey and the Vital Statistics System. That such data existed at all is due to the foresight of those health data gatherers who anticipated the desires of policy makers and the changes in the political and social climate which permitted legislative action. To cite just a few examples of the use of data in identifying problems:

-- the higher infant mortality rates in areas of concentration of poor people;

-- higher morbidity rates among the poor for many diseases;

-- lower utilization of services by lower income groups;

-- the rising toll of deaths from heart disease, cancer, and stroke;

-- the burden of health care costs for the elderly.

The use of extensive data in justifying these actions breeds a future demand for more data and for a refinement of that data. Once interventions have been undertaken the pressures to document results, to indicate quantitative changes in identified social problems, and to create at least the perception of success or failure is inevitable. The data has now become the currency of the policy and political process, and the need to allocate scarce resources and justify policy decisions focuses attention on the data and forges them into indicators.

A precursor of this demand for more data was passage of Comprehensive Health Planning legislation in 1966 and broadened authority for health services research in 1967. These actions were concrete expressions of faith that data and analysis could help bring solutions to health problems, but an important effect of implementing these programs was to reveal the inadequacies of current data sources and to increase further the policy maker's appetite for relevant information about the nation's health.

If this is a reasonable historical description of the rise in interest in health indicators, what is the nature of that heightened interest likely to be? First, there will be a demand for specific indicators that are relevant to the social values singled out in the policy process. This demand may not be for a balanced view or even place much emphasis on the technical soundness of the data. The demand may even be quite unreasonable in its expectation that those designing data-gathering instruments anticipate the interest of the policy makers. The strength of the policy maker's interests may also lead to intervention in the actual nature of the indicators. For example, the strong public concern over the nutritional status of the American population led to an allocation of always scarce statistical resources to the inclusion of additional nutritional data in the health examination survey.

The strengthened interest of the policy maker also presses toward a definition of quantifiable output data. The competition for public resource inputs leads to continued pressures to define output measures and to relate the input to the output. Since I believe that health is increasingly defined as a social end in itself, resource allocations should reward those programs which, over time, are more successful in defining measurable outputs and the relationship of input to output. That conclusion may appear naïvely optimistic, but my observation of the policy and the political process indicates that a good output indicator, such as a reduction in infant mortality or the restoration of normal functions, will eventually overcome rhetorical justification of the desirability of an input, such as more hospital beds or more health manpower, especially as ultimate resource limitations are perceived more sharply.

Another effect of the policy maker's interest in health indicators will be a strong pressure for simplicity and the ability to explain the meaning of the indicator to a person not trained in statistics or mathematics. Since I have already indicated the willingness of policy makers to intervene in the nature of data gathering processes, the virtue of simple explanation deserves attention among those who are developing health indicators.

The policy maker will also exert strong pressure for timeliness and fine-grained geographic detail in the health indicator. The policy maker is likely to be willing to trade off precision for timeliness, and the local variations in this diverse country make the ability to provide geographic comparisons a highly desired data characteristic. The strength of both of these concerns grows with the specificity of the policy intervention to overcome deficits in health services and the need to relate effect to that cause.

The policy maker's interest in the use of health indicators also leads to a growing tension between knowing and doing. After the success of using indicator data to identify problems and to justify new actions, there is a growing frustration not only with the seeming lack of results from many of these interventions but also with the difficulties in developing more refined indicators of output. Those difficulties were acknowledged by those involved in the very limited effort within HEW to develop a social report. 3/ The recent strong advocacy of well-designed social experiments as the basis for policy formation by such knowledgeable persons as Alice Rivlin would seem to be a reflection of this frustration and an expression of doubt that social indicators of sufficient precision can be developed to be the sole guides of public action. 4/ Such deliberate social experimentation is unlikely to bring a halt to wide-scale social action decisions until the results of the experiments are in. For example, the recent OEO contract to the Rand Corporation for a carefully designed experiment of consumer response to different configurations of health insurance will not delay the legislative action on some form of national health insurance. However, such social experiments may play an important role in defining and refining relevant measures of output by establishing clearer understanding of cause and effect. Both social experiments focused on a limited population and improved health indicators giving information about changes in the entire population should contribute to the development of what Donald T. Campbell has called, "the experimenting society." 5/ The policy maker is not afforded the luxury of a choice between knowing and doing. He must act, usually on the basis of very imperfect information, but the more he acts and is judged for those actions, the more he will come to demand information both prior to decisions and as a feedback on the results. Enthusiasm and firm belief can carry him a long way but the increasing level of frustration about the results of public actions and the considerable disillusionment with earlier attempts to make decision-making more rational are often attributed to the lack of relevant and timely information.

All of these pressures from the policy makers can be viewed by those responsible for generating health indicators as the price of success. As data is used in the policy process, decisions about the generation of that data increasingly become policy decisions rather than purely statistical decisions. Some of these policy decisions will involve compromise, including pressures to leap ahead of sound statistical methodology. As statistical indicators become linked to specific policy and managerial decisions, we will have to learn to cope with open policy and political debate over the desirability of changes in statistical series, a debate which may often pay little attention to methodological arguments and concentrate instead on the policy impact of changes in the indicator numbers.

In spite of this view that the policy maker's interest in health indicators is not an unmixed blessing, I would strongly advocate that these prices of policy relevance be paid and that the pursuit of the development of health indicators be vigorously pursued. The policy process is too often ill-informed, and the crucial role of the political process in adjudicating value conflict is often confused and diluted by the lack of sound descriptive information. These are the vital roles that health indicators must play in the policy process:

First, health indicators are useful in identifying problems, both in terms of health status and in the use of resources. Existing health data has been used extensively by policy makers for this purpose.

Secondly, health indicators can be used to monitor changes over time. This information is often useful to the policy maker whether or not he can attribute cause to the change. For example, if objective measures of nutritional status improve, especially if there is a lessening in the differences among population groups, that information is useful to the policy maker per se.

A third function for health indicators is to provide a context for the evaluation of operational statistics. There is a clear trend toward more explicit planning, regulation, and management of health services and other health related activities. Most of these activities will generate operating statistics upon which important decisions are based. But operating statistics are too narrow an information base for many types of decisions. The decision maker often needs to compare operating statistics concerning particular program beneficiaries with patterns for the population as a whole. For example, the operator of the health care program providing comprehensive services should be interested in whether the patterns of disease incidence or mortality in the population being served by his program vary from those same characteristics in the whole population for that geographic area.

Another important function of health indicators is to put pressure on policy makers to refine the objectives of their programs in terms which can be related to expected outcomes. It is easy to underestimate the difficulty of this task, both technically and in terms of organizational behavior. But the increasing use of health indicators by policy makers, often in adversary positions during the political process, creates strong pressures to justify the performance of programs in terms of outcomes. Many decisions will continue to be made in terms of inputs, but we should stress the need to specify the assumed models of cause and effect which transform inputs into outputs. Without the availability of appropriate health indicators, we have little basis to press for specification in these models of transformation. Land has referred to the specification of these models as "the major unsolved problems in social indicators." 6/

A further role of health indicators is to contribute to the evaluation of the effectiveness of particular programs or policies. Program evaluation is often cited as a reason for health indicators. However, it should be noted that unless the specification of objectives, which I have just discussed, has taken place, the link of health indicators to program evaluation will be much more difficult. Indeed, it is unlikely that general health indicators alone can be the basis for determinant program evaluation. But health indicators should provide invaluable assistance in designing the program evaluation studies by providing clues to cause and effect relationships.

I believe that much can be achieved by continued improvements in the current types of health data. However, it is clear that the development of an index of health that can be used as a common output measure will be highly desirable for some kinds of decision-making. An index will make explicit the value weighting of different types of indicators and will further focus policy debate on results in terms of health as an ultimate objective of the society. The development of a health index is also a methodological necessity in applying certain rational techniques of resource allocation. The index clarifies the imperative for difficult choices amidst the strong advocacy for various health programs. The work of Fanshel, Bush and others shows excellent progress in this direction. 7/ The development of a health index now needs to be subjected to the heat of the policy process, especially since such an index will only be accepted by the policy makers if they understand the values contained within it.

Finally, at this stage of major public debate on future health policies, a fundamental usefulness of health indicators is to focus that debate on changes in health status, or at least on such social goals as equity of access, rather than just inputs and costs. To the extent that we can define health status indicators and the relationship of our actions to those indicators, we will not only shape the nature of the public debate but we may have profound impact on the way policy decisions are made and on the actual conduct of medical care and other health related activities. There is a curious and disturbing dichotomy in much of the current discussion of health policy. Many of the important policy decisions focus on cost, utilization, distribution of resources, and concerns about efficiency. These concerns are prominent in the current debate over national health insurance and health maintenance organizations. But much of the actual performance in the health care system is focused on concerns about quality and effectiveness. In fact, health care providers are strongly criticized for a lack of cost consciousness or conern with the efficient use of resources. There is clearly a gap of understanding and action -- a gap often filled with ideological debate.

I believe that this gap should be filled by wider use of health indicators in policy formulation, planning, regulation, and management of health activities. Much of the efficiency debate is empty without an agreed upon measure of output unless one adopts a totally nihilistic attitude about the effectiveness of health services. On the other hand, health providers often show too little concern for the use of resources as they pursue the goal of effectiveness, and indeed there is often little concern for true measures of effectiveness in terms of changes in health status. A. L. Cochrane has dealt with this lack of relationship between efficiency and effectiveness in health care in his recent book and he concludes with an assessment that there is too little use of scientific evidence in the operation of the British health service, causing much wastage of resources and deficits in effectiveness. 8/ That conclusion is certainly transferable to the U. S. health care system. While Cochrane places his hopes on increased use of randomized control trials, I believe that better use of health indicators in the development of a health index could move us strongly in the same direction.

Unless we emphasize output indicators, we will continue to have attempts to limit cost per se, with very unequal results. We need more precision and rigor in our decisions, based on evidence, and an increase in policy debates on the nature of that evidence. Only in this direction lies progress.

1/ Land, Kenneth C., "On the definition of social indicators," The American Sociologist, Vol. 6, November, 1971, p. 323.

2/ U.S. Department of Health, Education, and Welfare, Toward a Social Report, Washington, 1969, p. xii.

3/ Ibid., p. xiv.

4/ Rivlin, Alice M., Systematic Thinking for Social Action, The Brookings Institution, Washington, 1971.

5/ Campbell, Donald T., "Critical problems in the evaluation of social programs," a paper prepared for the annual meeting of the Division of Behavioral Sciences, May 19, 1972, processed, p. 39.

6/ Land, op. cit.

7/ Fanshel, Sol and J. W. Bush, "A health-status index and its application to health-services outcomes," Operations Research, Vol. 18, Nov.-Dec. 1970, pp. 1021-1066.

8/ Cochrane, A. L., Effectiveness and Efficiency, The Nuffield Provincial Hospitals Trust, London, 1972.

SOCIAL INDICATORS FOR HEALTH BASED ON FUNCTION STATUS AND PROGNOSIS

J. W. Bush, M.D., M.P.H.
Milton M. Chen, Ph.D.
Donald L. Patrick, Ph.D.

1.0 INTRODUCTION

Widespread interest in social indicators has focussed attention and effort on various proposals for a general indicator of health status [Sanders, 1964; Chiang, 1965; Division of Indian Health, 1966; Sullivan, 1971; Gitter and Mostofsky, 1972]. Difficulties in developing such an indicator have led to several series of criteria for an ideal index, but many of the criteria overlap and conflict depending on the use conceived [Sullivan, 1966; Moriyama, 1968; WHO, 1971; Bush, et. al., 1972]; Goldsmith, 1972.

In our efforts to develop rational health planning models, we also recognized the need for a health index, but we have not previously offered a precise formulation of our model as a community-wide health indicator [Fanshel and Bush, 1970; Chen and Bush, 1971]. The purpose of this paper is to propose two indices, derived from a central concept of health, that we consider necessary to describe different aspects of community-wide health status.

Before presenting our two indices, we would like to review briefly some of the criteria and models that have been proposed. The strengths and weaknesses of the proposals can then be seen in sharper perspective.

1. The indicator should be of direct normative interest. Although this criterion has been criticized from different points of view [Sheldon, 1970; Land, 1972], we believe that the indicator should have a welfare orientation. This requires an implicit or explicit value component so that if the indicator improves, all other things being equal, society can be considered better off. Where only one number (indicator) is involved (mortality rate, bed-disability days, disease incidence rates), the value problem is rarely critical. But a complex social construct like health must certainly be represented by a weighted combination of many indicators, so that the weighting or social value problem becomes crucial. Most previous efforts to construct health indicators have tried to circumvent the value problem, and we believe that type of effort has led to many unsatisfying results.

2. The weighted index should be useful for priority setting, planning, and evaluation. Since the major point in collecting social statistics is to assist in policy-making, the index should be adaptable for social optimization. We should construct a health indicator in such a way that not only the direction but also the magnitude of the change in the indicator value is significant. With other relevant information on the component of the indicator responsible for the change, or estimates of the probable effect of some policy, the health indicator can then be used for priority setting and program planning.

We agree with Land [1972] that we should strive for indicators that fit meaningfully into larger social system models [Orcutt, 1960, 1970; Isard, 1969]. Within such models, even data that is not directly normative takes on meaning because it has a specified relation to other elements of the model. In health, this would be particularly helpful, since our abundant activity and utilization data can be given little normative interpretation without a larger model of health outcomes.

Adaptability for social optimization raises acutely the problem of constructing a social welfare function [Arrow, 1963; DeMeyer and Plott, 1971]. But refusing to construct a normative index does not make the problem disappear. It may be possible in the health field to mitigate some of the serious and thorny theoretical problems by the very process of index construction [Whitmore, 1972].

3. The indicator should permit study of the probable effects of different social policies. Admittedly the structural links between policies and indicators would be less developed in health than in economics, and the data to establish causal relations may be difficult to obtain, but the difficulty should not be overemphasized. If the links could be postulated theoretically, then they could be studied empirically--an almost inviolable methodological sequence in the cumulative sciences. Sheldon and Freeman's conclusion [1970:99] that it is impossible to use social indicators for setting priorities and developing a social balance sheet is probably overdrawn.

This does not mean that the health indicator should reflect policy directly or even be sensitive to a single change in policy. But the relation between the policy and the indicator should be specifiable, and the change (or lack of it) should be amenable to investigation.

Although we raise adaptability for social optimization as a desirable criterion for a health status index, here we will not discuss planning primarily, but will focus on methods for constructing objective health indices for time series and cross-sectional comparisons of total populations.

4. The health indicator should be useful for evaluative research. Although comparison of the health status of a given group over time is of great interest, comparison among different groups or subgroups at a given point in time gives important information for the public decision-making

process. Only rarely can a community-wide indicator give any indication of the impact of a policy change, but the follow-up of "treated" population subgroups in conjunction with control communities provides some evidence of causal relations [Campbell, 1969]. Such use of a social indicator for evaluative research is already in effect using our Index [Lawrence County Project, 1972].

5. The health indicator should be sufficiently sensitive to detect most of the significant changes in health status. The mortality rate and life expectancy are no longer adequate as indicators of health in western societies. A classification of multiple states of health must be developed that is refined enough to realistically reflect the array of conditions that afflict human populations, but it must be simple enough that data can be collected reliably without complex medical evaluation.

Furthermore, the set of weights reflecting the relative well-being of each of the states should not depend solely on criteria of economic productivity. Although the capacity to transform the Index values (or the underlying data) to economic criteria is desirable, earning capacity in itself is almost a dichotomous (work/no work) measure, and if systematically applied, would discriminate against many low or non-wage earning groups. It may be better to avoid the criterion at the outset than to try to retrench later from the socially unacceptable policy implications.

6. A community-wide health indicator should consist of "clearly defined component parts and each part should make an independent contribution to variations in the phenomenon being measured" [Moriyama, 1968:593]. A simple relation between the aggregated indicator and its parts will indicate what caused the change in the aggregate by locating the change in one of the component indicators. Component indicators may consist of the usual demographic subgroups, political jurisdictions (state, regional, local), or populations defined by other criteria.

7. An indicator of health should be derived from observable data and be easily reproduced. Ideally, the index should require no expensive new venture in data collection, but this case too should not be overstated. A social indicator for health that is intuitively acceptable and that satisfies the criteria outlined above would justify a significant new data collection effort. It may be possible to implement our suggestions without a drastic overhaul of current systems.

There may be other important criteria; this list is far from complete. Since few indicators will satisfy all the criteria, they must be judged as compromises among the different desirable properties. We will now compare several proposed indices with the above criteria.

Although Sanders [1964] discussed "effective life years", he did not make a concrete proposal for measuring community health levels. Operationally, his definitions lead to an economically productive man years index which is difficult to accept because of its insensitivity and the value judgements involved. Although computable from the life table, he did not propose a link to policy formulation.

Chiang [1965] proposed an index computed by the following formula:

$$H_x = 1 - \bar{N}_x \bar{T}_x - 1/2m_x \text{ where}$$

H_x = mean duration or average fraction of the year the individual is "healthy" in age group x

$\bar{N}_x$ = the average number of illnesses per person in [age group] x

$\bar{T}_x$ = the average duration of an illness for x

m_x = the age-specific death rate for the year.

The index can be computed from available data, but it is insensitive, since the single state of illness defined covers the entire range of illness conditions. Furthermore, death is weighted as equivalent to illness, a socially unacceptable value judgement. Finally, no method is defined for relating the index to policy decisions.

Although the Indian Health Service [1966] and the Pan American Health Organization [Ahumada, 1965] developed formulae for computing program priorities, their indices include terms related to reference populations or expert judgements of vulnerability, and are not well adapted to serve as direct indicators of health status in comparative and time series analysis.

The most intriguing recent proposal is Sullivan's "single index of morbidity and mortality", which is based on the concepts of "expectation of life free of disability" and "expectation of disability" [1971]. To compute the expectation of life free of disability, the conventional ${}_nL_x$ in the life table is weighted by a disability factor $I_x = 1 - \frac{W_x}{365}$ to obtain ${}_nL_x^*$ where W_x = number of days of disability per person per year in the interval beginning at age x. W_x is derived from health interview and institutional surveys. Values of T_x^* are computed in the same manner as the conventional life table. The resulting expectation of life free of disability, $\mathring{e}_x^*$, gives an appropriate index of health. The expectation of life free of disability for civilian white males for the U.S. in the mid-1960's was 62.5 years, and the expectation of disability was 5.3 years.

The model as proposed has several advantages: (1) it corresponds closely to a comprehensive quantification of the social construct of health; (2) it is related directly to hard, available data; (3) it bypasses many difficult value questions; (4) it could be augmented without major changes, and (5) it could be related indirectly to policy choices. We shall discuss the shortcomings of this index as well as its special meaning after presenting our own model.

2.0 HEALTH AND FUNCTION STATUS

2.1 Operational Definition of Health

Inherent in the social construct called "health" are two dimensions: (1) function level, an individual's level of function at a point in time, and (2) prognosis, his expected transitions to other levels, more or less favorable, at future times

(Fig. 1). For measurement purposes, these two dimensions require separate specification.

Function status is the primary value dimension of health. Optimum function is defined as conformity to society's standards of physical and mental well-being, including performance of the activities usual for a person's age and social role. As defined here, disturbances in function are not only social, but also include pain, and other physical and mental symptoms that are considered deviations from norms of well-being, even when there is no interference with social role performance. Such deviations from societal standards of well-being are value judgements. Standards of well-being can be defined and deviations can be classified into a series of function levels ranging from complete well-being through various levels of dysfunction to death. Social values can be measured for these function levels to produce a scale of well-being with a unit from 0 for death to 1 for complete well-being.

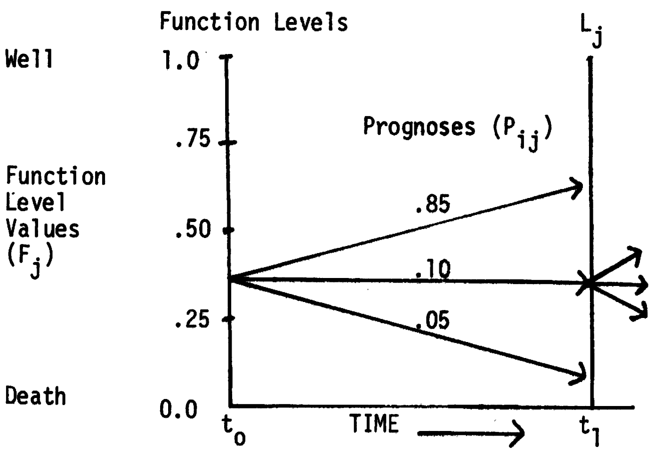


Figure 1. Function levels (L_j) represent states on the continuum from well-being to death. Function level values (F_j) are the social preferences representing the relative value of the levels between 0.0 and 1.0. Prognoses (P_{ij}) are the transitional probabilities for movement among the function levels over time.

Health status, on the other hand, is a composite of an individual's level of well-being at a point in time and his expected transition to other levels, more or less favorable, at future times. This view sharly distinguishes between the desirability of the immediate level of function and the probability of being in other levels as they change over time. These two dimensions, labelled function level and prognosis, have traditionally been confused in discussions of health and illness; both are necessary to describe the health status of an individual or population.

Treating the two variables as analytically distinct allows them to be quantified separately and to vary independently for different populations. Health status can then be described as a joint function of the two variables. Precisely stated, health status is the product (expected value) of the social preferences assigned to levels of function and the probabilities of transition among the levels over the remaining life of an individual

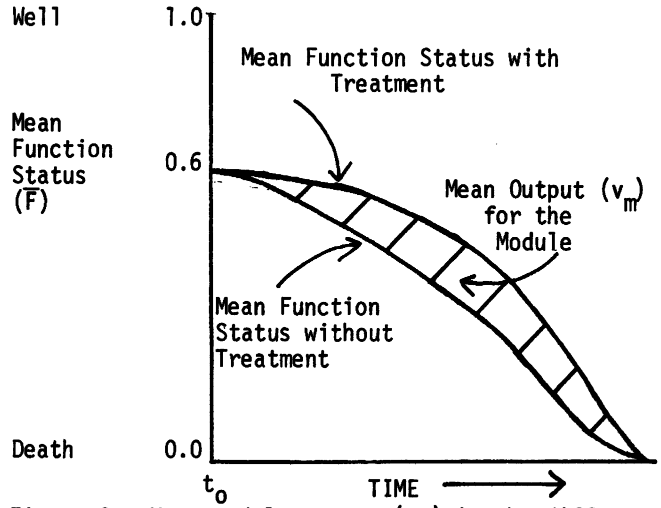


Figure 2. Mean module output (v_m) is the difference (ΔQ) between the quality adjusted life expectancy of the cohort with and without treatment.

or group. Conceptually, health status may be expressed as follows:

$$H = \frac{\Sigma F_j Y_j}{\Sigma Y_j} \quad \text{where}$$

H = health status of an individual or group

F_j = level of well-being (value, preference) that society assigns to function level j

j = index for function levels = 0, 1, 2, ..., 30

$Y_j = \sum_{t=1}^{k} \pi_{j,t}$ = total expected duration in function level j over all time periods

$\pi_{j,t}$ = proportion of time spent in level j between time periods t-1 and t, derived from the product (expected value) of the distributions and transition matrices for each time period

t = index of time periods = 0, 1, ..., k where k is the last time interval.

Given methods for incorporating various attributes of function, for measuring their values, and for incorporating the prognoses, we believe this formulation realistically incorporates the critical features of what society means by health [Patrick, et. al., 1972a]. From this conceptual definition, we shall derive formulae for quantifying the function and health status of a community that can be made operational with current data collection and value measurement techniques.

2.2 Operational Definition of Function Levels and Value Measurement

Using items from the Health Interview Survey of the National Center for Health Statistics, the Survey of the Disabled of the Social Security Administration, and several rehabilitation scales and ongoing community health surveys, we constructed three scales with mutually exclusive and collectively exhaustive steps to describe function status (Table 1). These different scales largely cover the spectrum of objective disturbances that diseases and disabilities can cause in role performance. Changes in these factors can occur not only because of physical disabilities, but also because of symptoms, sensory disturbances, mental retardation, and mental illness. These disturbances were summarized in an independent set of 42 symptom/problem complexes.

Combining different steps of the norms for social activity, mobility, and physical activity, and omitting rare or impossible combinations, 31 levels of function were created that can be used to describe the function status of an individual or a population. Members of a population may fall in a particular level for a variety of reasons, all having in common the defining features of that level. A matrix bounded by the function levels, 5 age groups, and symptom/problem complexes generates a universe of function status descriptions, as follows:

TABLE 1. SCALES AND DEFINITIONS FOR THE CLASSIFICATION OF FUNCTION LEVELS*

SOCIAL ACTIVITY

A	Performed major and other activities
B	Performed major activity but limited in other activities
C	Performed major activity with limitations
D	Did not perform major activity but performed self-care activities
E	Required assistance with self-care activities

MOBILITY SCALE

A	Travelled freely
B	Travelled with difficulty
C	In house
D	In hospital
E	In special unit

PHYSICAL ACTIVITY SCALE

A	Walked freely
B	Walked with limitations
C	Moved independently in wheelchair
D	In bed or chair

*Definitions and sources of scale items available from the authors.

Age 40-64.

Walked with limitations.

In hospital.

Did not perform major activity but performed self-care activities.

Had burn over large areas of face, body, or extremities.

Using a probability sample of such case descriptions to represent the function status universe, we have undertaken a series of studies to find optimum methods for measuring the social values or preferences associated with the function levels. Examples were derived from the rating of the 400 case descriptions by 62 graduate students and nurses using the method of equal-appearing intervals.

We have also compared the validity and reliability of category rating, magnitude estimation, and equivalence across different orders of method presentation, across individual and group testing situations, and across students and health leaders [Patrick, et. al., 1972b]. The results of the study indicate the feasibility of measuring levels of well-being that constitute a unidimensional, equal-interval preference continuum. The invariance of the values assigned to the function status conditions across the different judge groups provides evidence for the validity and reliability of the measurement methods. Along with this study, progress in the development of social value measurement techniques generally [Coombs, 1964; Stevens, 1966; Sellin and Wolfgang, 1964; Bock and Jones, 1968; Anderson, 1972] indicates that the previously immeasurable value dimension of health may be integrated into an empirically verifiable social indicator.

2.3 Function Status Index

Given a valid set of social values and the distribution of the population among the set of function levels from appropriate surveys, the distribution can be weighted by the values to summarize the function status of the population at a point in time in the Function Status Index.

Algebraically, the Function Status Index ($\overline{F}$) is expressed as:

$$\overline{F} = \frac{\sum_j N_j F_j}{N} \quad \text{where } 0 \leq F \leq 1 \text{ and}$$

N = total number of persons in a population

N_j = number of persons in function level j

F_j = weight or social preference for function level j

j = index for the function level = 0, 1, 2, ..., 30.

Table 2 illustrates a simple calculation of the function status index for a population at a given point in time.

If every member of the population were at the level of optimum function, then the function status index for the population would be 1.000.

Data from health surveys can be used to determine the function level distribution of a population. Then the FSI, an index that integrates prevalence data on disability levels of different social values, would have several advantages: (1) it gives a concise summary of the current level of physical and mental well-being in a community; (2) it is sensitive since multiple function levels are defined; (3) the data is observable and can be easily reproduced without spefial medical knowledge or memory; (4) each respondent is classified into one and only one level between death and well-being; (5) longitudinal studies of the same population can be used to construct time series; (6) different populations can be compared cross-sectionally (7) the Index can be disaggregated into component parts by subdividing the population to study different distributions, causes of dysfunction, or possibilities for intervention; (8) it can be related to respondent reportable causes of dysfunction (arthritis, shortness of breath, etc.) and to medical diagnosis by auxilary surveys; (9) mechanisms already exist in the National Health Survey for new data collection; and (10) treatment and control groups can be monitored over time with the FSI in evaluative research studies to determine if changes in function status are attributable to the program.

TABLE 2. ILLUSTRATIVE DISTRIBUTION OF PERSONS AMONG DIFFERENT FUNCTION LEVELS AND COMPUTATION OF THE FUNCTION STATUS INDEX

Function Level (L_j)	Number of Persons (N_j)	Function Level Values (F_j)	(F_jN_j)	$\bar{F} = \frac{\Sigma F_jN_j}{N}$
30	95,000	1.00	95,000	
27	3,000	.69	2,070	$\bar{F} = \frac{98,070}{100,000}$
17	1,000	.59	590	$\bar{F} = .9807$
10	700	.44	308	
2	300	.34	102	
Total	100,000(=N)		98,070(=ΣF_jN_j)	

While the FSI provides a more clear-cut outcome indicator for evaluative research and makes it possible to integrate data on different function levels, it does little to resolve the difficult problems of research design that remain essential for establishing causal relations. Clearly the level of the FSI as a social indicator is affected by many intervening variables, and these must be disentangled in the usual ways, primarily by decomposing the indicator to isolate the contributing factors to the change, by establishing global correlations with other indicators (such as income, crowding, or housing quality) over long periods of time with many different populations, and by performing specific studies to choose among competing hypotheses.

For all its advantages, the FSI is an incomplete indicator, since it does not include the prognoses inherent in our concept of health and may be misleading about the overall health status of the community. For example, an increase in the mortality from function levels with low values would cause the FSI of the remaining living population to rise; on the other hand, an increased probability of survival in the same low levels would cause the FSI to fall.

Even among the living population, existing changes in transition probabilities may not be reflected in the population distribution among the function levels and the FSI for many time periods, perhaps several decades. This delay in the response of the FSI, despite changes that have already occured, may obscure both the magnitude and the direction of the change in health status. Although Sullivan's index avoids the paradoxical influence of mortality (and we could so supplement the FSI), the lag problem, common to practically all social indicators, is inherent in his use of currently observed distribution vectors. For a realistic view of the health of the population, we require a more sensitive, comprehensive, and dynamic indicator, an indicator that detects the changes currently in process.

3.0 VALUE ADJUSTED LIFE EXPECTANCY

3.1 Prognosis and Function Level Expectancy

In addition to function status, the concept of health incorporates prognosis, or the expected movement from one level to other levels over subsequent time intervals. This movement can be described as a probabilistic process, where the transition probabilities are the prognoses. The mortality rate is the probability of moving to death from any higher level.

We have used the medical term "prognosis" since it connotes the health-related meaning of the transitions among the function levels and disease states. From the definition of health status above, it is not the momentary level of function, but the outlook for the future that primarily determines what medical specialists and the public mean by "health" status. Diseases are "serious" or "not serious" depending on the associated probability of severe impairment or death, sometimes without much regard for the immediate comfort of the patient.

Like the mortality rate, the transitions among the other function levels should be determined empirically by population monitoring. If the transitions occuring within the memory span obtained on a single interview are not adequate for the computations, then interviews of the same sample will be required on at least two occasions to obtain reliable data. A few questions added to panel studies such as those of the Current Population Survey, would be more than adequate to produce the data without substantial new effort. From such data, transition rates among all the function levels can be computed and integrated with the mortality rate for each age and demographic subgroup.

These prognoses determine the time to be spent in each function level, or the <u>function level expectancy</u>. Analogous to life expectancy, tables can be constructed for populations with given demographic characteristics (<u>modules</u>) to determine the life time expected duration in each of the function levels. Function level expectancy is nothing more than the distribution of the life expectancy among the various function levels. Also, like the current life table, or more properly, the mortality experience table, the function level expectancy is not a projection, but a convenient summary statistic of actual data that treats the current transitions as if they were persistent over the life of a synthetic cohort.

This strategy permits us to define a series of function level expectancies that, like the life table, are independent of the age structure of the population described and facilitates population comparisons. In the model of function level expectancy outlined here, population subgroups can be defined by any relevant demographic or disease characteristics for which data are available, including mental health.

We believe that the controversy over positive versus negative measures of health rests partly on the confusion between prognosis and level of function. For many years, some groups have argued that positive mental health is not simply the absence of disease or symptoms, but the possession of certain attributes associated with high function level expectancies, and that only a subset of the total population enjoys such "positive" health.

We believe that subgroups can be reliably defined that possess not only physical attributes, but also personality characteristics, that are associated with longer expectancies in high levels of function. Such population groups would indeed have a higher value-adjusted life expectancy as computed below. The terms "positive" and "negative" in which the controversy has been couched cannot be given mathematical specificity and should be superseded by the more precise and flexible terminology of prognoses, transition probabilities, and function level expectancies, where the function levels are defined to include the presence or absence of anxiety, depression, and other emotional disturbances.

Many concepts and methods developed in the life table are relevant in the construction of a function level expectancy table. Because of the multiple levels involved, however, the process is more complicated than the usual multiple decrement table, and resort to a refined stochastic model may be necessary. Essentially, the life table constitutes a synthesis of many Markov processes with a separate matrix applied to each age group.

We may define as many levels or states as we find useful and can reliably distinguish. As outlined previously, we now identify 30 levels, but these should be consolidated into a shorter list on the basis of further value studies. Since Markovian states will be defined by additional criteria in further computations, we shall regularly refer to this basic list as function "levels" rather than as states.

If we standardize the age intervals as, say, one year, and define each Markovian state by both age group and function level characteristics, we can construct a grand matrix that will encompass all the age specific transition matrices. As an illustration, we may specify four function levels: A = well, B = non-bed disability, C = bed disability, D = death. A person in age group 1 in disability level A will be denoted as A_1, etc. Thus formulated, transfers would occur only from one age group to the next age group. The transition matrix would appear as follows:

$$
\begin{array}{c}
 \\ A_1 \\ B_1 \\ C_1 \\ D_1 \\ A_2 \\ B_2 \\ C_2 \\ D_2 \\ \cdot \\ \cdot \\ \cdot
\end{array}
\begin{array}{c}
\begin{array}{ccccccccccc} A_1 & B_1 & C_1 & D_1 & A_2 & B_2 & C_2 & D_2 & \cdot & \cdot & \cdot \end{array} \\
\left[\begin{array}{ccccccccccc}
0 & 0 & 0 & 0 & X & X & X & X & \cdot & \cdot & \cdot \\
0 & 0 & 0 & 0 & X & X & X & X & \cdot & \cdot & \cdot \\
0 & 0 & 0 & 0 & X & X & X & X & \cdot & \cdot & \cdot \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \cdot & \cdot & \cdot \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdot & \cdot & \cdot \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdot & \cdot & \cdot \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdot & \cdot & \cdot \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot
\end{array}\right]
\end{array}
$$

X represents the transition probability with $0 \leq X \leq 1$. Persons remain identified in the death state by age permitting us to integrate morbidity with mortality. The process under this formulation becomes a Markov chain.

A realistic and sensitive index of health status requires that we determine the P_{ij}, i.e., the transition rates among all the Markov states jointly defined by function level and age group. We must ascertain where each respondent has transferred from, or follow him to determine where he transfers to, in order to complete each cell of the transition matrix. This matrix gives us knowledge of the process that is currently operating.

Moreover, the equilibrium distribution vector Y^*, that is expected from continued operation of the P_{ij}, gives a unique and superior indication of health status over single distribution vectors observed at arbitrary points in time. Even a drastic change in P_{ij} might not have its full effect for many time intervals, and we cannot know whether the observed distribution vector at a given point (as in Sullivan's model) represents the equilibrium distribution or not.

In fact, the observed vector might be quite different from Y^*, especially if the process is in its early phase of development. An indicator based on current distribution vectors is insensitive to changes that are already emerging under the operation of P_{ij}. The equilibrium distribution vector of the matrix P_{ij} is analogous to the stable population of the life table, since it points to the steady state of the process if the transition probabilities persist. But the matrix specified earlier does not have the properties required to produce the equilibrium distribution that can be interpreted as the expected duration of stay in each of the Markovian states.

To converge to an equilibrium vector when raised to successively higher powers, a stationary transition matrix must be irreducible and aperiodic [Hillier and Lieberman, 1968; Parzen, 1962]. A matrix is irreducible if and only if all states communicate with each other, i.e., any state can be reached from any other state. An irreducible chain may not converge to a unique equilibrium distribution unless at least one of the states in the system is aperiodic. An arbitrary state is said to have period s (s>1) if $P^t_{ij} = 0$ whenever t is not divisible by s and s is the smallest integer having this property. If a state can be entered only at time intervals 0, 2, 4, ..., then that state has a period of 2. On the other hand, if a state can be entered for two successive time intervals, then that state has a period of 1 and is aperiodic. If any state i in a set is aperiodic, then all states in the set will be aperiodic and the matrix derived from the set will be aperiodic [Bush, et. al., 1971].

To assure the irreducibility and the aperiodicity of the chain, we can artificially create a reservoir state (S_0) that is analogous to the assumption of a constantly renewed population in life table construction. If we assume a dummy entry rate from the reservoir state to the initial state A_0, and if, at the end of the last age interval, every member is transferred back to the reservoir state, then we can treat the problem of entry and exit under a closed system. In this way, all the states will communicate with a periodicity of 1. Since the reservoir state is a dummy device, it can be deleted from the equilibrium vector. The transition matrix would now appear as follows:

$$
\begin{array}{c}
\\ S_0 \\ A_0 \\ A_1 \\ B_1 \\ C_1 \\ D_1 \\ A_2 \\ B_2 \\ C_2 \\ D_2 \\ \cdot \\ \cdot \\ \cdot \\ A_{n-1} \\ B_{n-1} \\ C_{n-1} \\ D_{n-1} \\ D_n
\end{array}
\begin{array}{c}
\begin{array}{cccccccccccccccc}
S_0 & A_0 & A_1 & B_1 & C_1 & D_1 & A_2 & B_2 & C_2 & D_2 & \cdots & A_{n-1} & B_{n-1} & C_{n-1} & D_{n-1} & D_n
\end{array} \\
\left[\begin{array}{cccccccccccccccc}
X & X & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & X & X & X & X & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & X & X & X & X & \cdots & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & X & X & X & X & \cdots & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & X & X & X & X & \cdots & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0
\end{array}\right]
\end{array}
$$

where

S_0 = reservoir state

A_0 = entry state of the system

n-1 = last age interval

As an illustration, we constructed a grand transition matrix from the mortality rate and hypothetical transition rates among the other levels. We computed the equilibrium distribution vector by raising the power of the matrix, deleted the dummy states S_0, A_0, and D_n, and aggregated the proportions for the function levels A, B, C, and D over the age intervals to yield an equilibrium vector of $\pi^* = [\pi_A^*, \pi_B^*, \pi_C^*, \pi_D^*] = [.72413 \quad .04985 \quad .02134 \quad .20468]$. Since we used a total of 9 ten-year age intervals, the total of age intervals I = 90 years. The equilibrium function level expectancies are computed as follows:

$$Y^* = I\pi^* = 90\ [.72413 \quad .04985 \quad .02134 \quad .20468]$$
$$= [65.17 \quad 4.49 \quad 1.92 \quad 18.42]$$

Subtracting the 18.42 years duration in the death states, we obtain the life expectancy:

$$Y_A^* + Y_B^* + Y_C^* = 65.17 + 4.49 + 1.92$$
$$= 71.58 \text{ (years)}.$$

As indicated earlier, this equilibrium distribution vector Y^* is a unique summary of the process represented by the current transition matrices and provides a statistically reproducible estimate of the function level expectancies required for computing the Value-Adjusted Life Expectancy (Q^*).

An index based on Y^* will detect changes in health status that are ignored in the life table. Given the same mortality, different rates of transfer among levels A, B, and C represent significant differences in health status. In the life table, all such changes are disregarded. In fact, even small changes in rates for a single age group will change the equilibrium distribution vector.

Our notion of function level expectancies is similar to Sullivan's expected disability and disability free years. The disability and disability free states correspond roughly to our function levels. But Sullivan's computation of the disability expectations are not derived from the transition matrices, P_{ij}, carried to equilibrium; they are cumulated across all the age groups from currently observed distribution vectors, π_j, from household and institutional surveys.

Thus, Sullivan's disability and disability free life expectancies may be different from the equilibrium distribution of Y_j^*. For instance, from our hypothetical data, if the distributions were observed at the 8th period, the resulting estimate of the function level expectancy under Sullivan's model would be [70.654 4.023 1.650]. This distribution vector is different from the long-run equilibrium distribution, and gives no indication of the

TABLE 3. FUNCTION LEVEL EXPECTANCIES CONSTRUCTED FROM NATIONAL DATA [SULLIVAN, 1971]

Function Level L_j	Function Level Expectancies Y_j
Disability free	64.9
Non-Bed Disability	3.3
Bed Disability	2.0
Live Expectancy	70.0 (years).

direction or magnitude of the changes already in process. The advantage of Sullivan's index is that it can be computed using currently available data, and it intuitively provides an estimate of the Y_j in our framework, as illustrated in Table 4.

3.2 Value-Adjusted Life Expectancy and the Health Status Index

By applying the function weights to the function level expectancies, we obtain the value-adjusted life expectancy. This can be accomplished with a standardized set of weights by the operation

$$Q^* = \Sigma F_j Y_j^* \quad \text{where}$$

Q^* = the value-adjusted life expectancy.

Table 5 illustrates the computation by applying a set of values to the previously computed distribution vector.

TABLE 5. ILLUSTRATIVE COMPUTATION OF VALUE-ADJUSTED LIFE EXPECTANCY

Function Level Expectancies Y_j^*	Function Level Values F_j	$Y_j^* F_j$
65.17	1.00	65.17
4.49	.59	2.65
1.92	.34	.65
Value-Adjusted Life Expectancy (Q^*)		68.47

The value-adjusted life expectancy (Q^*) may be regarded as the equivalent of expected dysfunction free years of life or the expected quality adjusted years of life. This value-adjusted life expectancy can be used as a health indicator for comparing: (1) population groups of all ages cross-sectionally, (2) the health status of the same population over time. For example, Q^* for the U.S. resident in 1970 might be Q^*_{70} = 68.000 years, whereas the same figure for 1960 might have been Q_{60} = 67.527 years; the difference would be a precise composite expression for the total change in the population's health status. Even better than the Function Status Index, Q^* could be monitored over time as a social indicator that incorporates the dynamic aspects of health status, and could be correlated with other social indicators for social systems models and analyses.

For most purposes, Q^* is the best indicator of the health status of total populations. Since it is constructed from the same types of data and values as the Function Status Index (F), it has all the previously outlined uses and advantages. But Q^* remedies the major difficiency of the FSI and Sullivan's expectation of disability free life; it reflects immediately changes in prognoses, and yet is observable without medical diagnosis. It also corrects the insensitivity of current life expectancy measures, since it integrates multiple levels of function with mortality and is not subject to paradoxical change as other morbidity indicators have been. Furthermore, Q^* may be computed for age-specific comparisons.

Q is transformed to the 0-1 scale by the ratio to the sum of the remaining time intervals during which transitions might occur. If this standard life (S) is defined as 100 years, then at birth (Age 0) health status is the decimal transform of Q, as follows:

$$H = \frac{\Sigma F_j Y_j}{\Sigma Y_j} = \frac{Q}{S-A} = \frac{68.47}{100-0} = .6847$$

As age increases, however, H becomes increasingly sensitive to incremental changes, and it is not clear that these changes correspond to any clear interpretation of health status. Using the age-specific life expectancy in the denominator does not completely resolve the problem. Further study may permit us to devise a more meaningful ratio for comparisons between age groups.

The value-adjusted life expectancy is not very helpful in evaluation research, since cohorts under various treatments can rarely be followed over their life expectancy. Where the Markovian states are also defined by disease forms, however, the total disease history may be synthesized from currently observed transitions in a further expansion of the stochastic process outlined above [Bush, et. al., 1971].

A major advantage of Q^* is that it combines morbidity with mortality in a single number that is independent of both age and medical diagnoses. It can be determined using currently available survey and value measurement techniques, and replicated from year to year and from one population to another. Although more complex models may be possible, the discrete Markov model captures the important transitions of the life table for all the function levels, permits us to incorporate a standardized set of social values, and summarizes them in a unique scalar value for time series and cross-sectional comparisons, the major function of a social indicator.

3.3 Health Planning and Program Analyses

An ideal social indicator should be useful not only for monitoring but also for social decision-making. One of the most powerful uses of Q^* comes in planning and program analysis, where we project the probable impact of our policies on health status. It is the potential or expected difference in the value of Q^*, with and without the program, that drives the health system, and makes society willing to allocate resources to health services.

The output (value or benefit) of a health program can be defined as the increment of value-adjusted years of life (ΔQ) added to a target population by the program's intervention (Fig. 2). Different subgroups of the target population are affected differently by interventions. To accurately describe the different effects, we must disaggregate the target population into modules, that is, into subgroups that are homogenous with respect to prognoses and expected function status. The average value of treating the members of a particular module, m, is given by the difference in their expected value-adjusted years of life with and without the treatment ($v_m = \Delta Q$).

The total output of an entire program is the sum of the outputs of the individual modules, a linear function of the numbers of persons serviced, a function that can be maximized across multiple disease and population subgroups, as follows:

$$V = v_1 n_1 + v_2 n_2 \ldots v_m n_m \ldots v_z n_z \quad \text{where}$$

V = total program output

$v_m = \Delta Q$ = mean value of treating a member of module m

v_m = number of persons serviced in module m

z = index of the final module.

With the appropriate estimates of v_m, this model directly relates activity data about the numbers of patients treated to a meaningful output estimate, and provides a number with the required mathematical properties for use in cost/effectiveness and mathematical programming models. It is also amenable to relating resource inputs for defined services to a measure of output through production functions where the treatment of n_m is a function of the numbers of physicians, nurses, other technical personnel, drugs, laboratory equipment, space, and resources consumed.

The fact that the same basic concepts can be used both as a social indicator and as a planning model means that we can quantitatively analyze the contribution of one health program to the social indicator (Q^*) through its target population. Although we can never assert definitively that an observed change in Q^* for a population is entirely due to a particular program, we can investigate quantitatively the probable effects of the program to see if they could account for the change. Similarly, we can examine other possible causes of the change, such as demographic shifts. An examination of the precise number and composition of the program target populations, their expected health status without a program, and the time lag before the treatments take effect, make immediately evident why the effects of most health programs are not detectable in most social indicators for health.

From a theoretical standpoint, it should be possible to construct a macromodel of the health system using the concepts of health output described here. Overlaps in programs would be handled by creating new cells for the intersections of two or more target groups. Certainly it would be informative to examine the effects of eradicating various categories of diseases both on the expected value-adjusted years of life and on resource utilization. An analytical framework that connects a comprehensive social indicator for health status with a production function for health services will make it possible to examine such questions in detail.

CONCLUSION

We have proposed a Function Status Index ($\overline{F}$) that could be constructed using current data collection mechanisms and feasible value studies that would significantly augment our knowledge of the level of well-being of the population and facilitate cross-sectional and time series comparisons.

The Value-Adjusted Life Expectancy (Q^*) would give a reasonable approximation of an ideal health status index, but would require collecting new kinds of data on the function level transitions on a community or nation wide basis. Such data would be transformed into a single comprehensive index of physical and mental health status that incorporates both its value and prognostic dimensions. Such an index would be constructed from empirically determined components that would be replicable over time and among different population groups. In addition, the model provides criteria for evaluative research and an analytical framework for estimating the output and contribution of health programs to the health status of the population. We believe that methodological research should be intensified to make this measure a practical social indicator for health.

REFERENCES

Ahumada J, et al: Health Planning: Problems of Concept and Method. Publication No. 111, Pan American Health Organization, Wash., D.C., 1965.

Anderson NA: Information Integration Theory: A Brief Survey. Technical Report No. 24, Center for Human Information Processing, University of California at San Diego, La Jolla, Calif., 1972.

Arrow KJ: Social Choice and Individual Values. Yale University Press, 1963.

Bock RD, Jones LV: The Measurement and Prediction of Judgment and Choice. San Francisco, Holden-Day, 1967.

Bush JW, Chen MM, Zaremba, J: Markovian analysis of disease history and the problem of equilibrium. AJPH 61:2362-2375, 1971.

Bush JW, Fanshel S, Chen MM: Analysis of a Tuberculin testing program using a health status index. Socio-Econ Plan Sci 6:49-68, 1972.

Campbell DT: Reforms as experiments. Amer Psych 24:409-429, 1969.

Chen MM, Bush JW: A mathematical programming approach for selecting an optimum health program case mix. Paper presented at the 40th National ORSA Meeting, Anaheim, California, 1971.

Chiang CL: An Index of Health: Mathematical Models. Vital and Health Statistics, Series 2, No. 5, National Center for Health Statistics, Washington, D.C., 1965.

Coombs CH: A Theory of Data. New York, John Wiley, 1964.

DeMeyer F, Plott CR: A welfare function using "relative intensity" of preference. Quart J Econ 85:179-186, 1971.

Division of Indian Health: Principles of program packaging in the Division of Indian Health. Public Health Service, U.S. Department of Health, Education, and Welfare, Silver Springs, Md., 1966.

Fanshel S, Bush JW: A health-status index and its application to health-services outcomes. Oper Res 18:1021-1066, 1970.

Gitter AG, Mostofsky DI: Toward a social indicator of health. Soc Sci Med 6:205-209, 1972.

Goldsmith SG: The status of health status indicators. Health Serv Rep 87:212-220, 1972.

Hillier FS, Lieberman GJ: Introduction to Operations Research. San Francisco, Holden-Day, 1968.

Isard W: General Theory: Social, Political, Economic, and Regional. Cambridge, The MIT Press, 1969.

Lawrence County Health Care Project: Lawrence/Lamar health interview survey questionnaire. Instrument available from the project, Moulton, Alabama, 1972.

Michael J, Spatafore W: A basic information system for health planning. Pub Health Rep 83:21-28, 1968.

Miller JE: Criteria for evaluating the application of health status indices in a management context. Paper presented at the ORSA National Meeting, New Orleans, 1972.

Moriyama IM: Problems in the measurement of health status, in Sheldon EB and Moore W, eds. Indicators of Social Change: Concepts and Measurement. New York, Russell Sage, 1968, pp. 573-600.

Orcutt GH, et al: Macroanalytic simulation: a demographic model. Paper presented at the Annual Meeting of the American Statistical Association, Detroit, 1970.

Orcutt GH, et al: Microanalysis of Socio-Economic Systems: A Simulation Study. New York, Harper & Row, 1961.

Parzen E: Stochastic Processes. San Francisco, Holden-Day, 1962.

Patrick DL, Bush JW, Chen MM: Toward an operational definition of health. J Health Soc Behav, forthcoming 1972a.

Patrick DL, Bush JW, Chen MM: Measuring levels of well-being for a health status index, 1972b. Paper available from the authors, Department of Community Medicine, University of California at San Diego.

Sanders BS: Measuring community health levels. AJPH 54:1063-1070, 1964.

Sellin T, Wolfgang ME: The Measurement of Delinquency. New York, John Wiley, 1964.

Sheldon EB, Freeman HE: Notes on social indicators: promises and potential. Policy Sci 1:97-111, 1970.

Sullivan DF: A single index of mortality and morbidity, HSMHA Health Rep 86:347-354, 1971.

Sullivan DF: Conceptual problems in developing an index of health. U.S. Department of Health, Education, and Welfare, PHS Publication No. 1000, Series 2, No. 17, 1966.

Whitmore GA: Health state preferences: a mathematical characterization and a look at the common assumptions. Faculty of Management, McGill University, Working Paper. Presented at the 41st National Meeting of the ORSA, New Orleans, 1972.

World Health Organization: Statistical Indicators for the Planning and Evaluation of Public Health Programmes. 14th Report of WHO Export Committee on Health Statistics, WHO Technical Report Series No. 472, 1971, pp. 19-20.

AN APPROACH TO CONCEPTUALIZING THE LEVEL OF COMMUNITY HEALTH

Monroe Lerner

The Johns Hopkins University

I-Introduction

Apparently by common consent, the "old" ways of measuring health appear to be strongly in need of re-examination today, largely because of changing social conditions and because many of the formerly most "salient" health problems have to some considerable extent been dealt with. In this context the "old" ways included the use of mortality rates and a variant of these rates, life expectancy, and more particularly infant mortality rates, as indicators of the total health of some sort of a population aggregate, e.g., a "community".

While the shortcomings of mortality rates in whatever form as indicators of health were widely recognized, use of these indicators did have at least two virtues: 1. Mortality statistics were relatively readily and widely available, largely because they were collected as a by-product of death registration, a necessary procedure in a well-ordered society, especially a Western industrial society; and 2. In fact "high" mortality rates, i.e., large numbers of "premature" deaths, were widely recognized as the "most important" health problems of the day, in the sense that it may have appeared to many that these high rates could be reduced, if only because mortality rates were substantially lower in more affluent population groups and/or in groups where public health measures, as symbolized by pure water and sanitation, had been introduced and were being efficiently carried out. In addition, morbidity and impairments must have been highly correlated with mortality, so that in fact mortality may have been a good indicator of the total spectrum of health as it was defined at that time.

Later, during the second third of the 20th Century and as mortality was at first gradually and later rapidly reduced to levels even substantially lower than those prevailing at the beginning of the century, the morbidity and impairment components of health began to come to the fore. This was the time when chronic illness became a recognizable entity in its own right and when it was recognized that advances of medicine and surgery in prolonging life often resulted in the creation of large numbers of chronically ill and impaired individuals who had many years of life before them. At this point the correlation of mortality rates with morbidity and impairment rates was probably substantially lowered and, in addition and perhaps simultaneously, premature mortality did in fact become a less important problem than it previously had been.

However, measurement of morbidity and impairment is far more complicated and difficult than the measurement of mortality, largely because death is a unique, clearly-defined event while often the former is not. Also, morbidity and impairment often have social and emotional, as well as bio-physical, antecedents, and this adds an element of complexity to the measurement process. Morbidity and impairment have several dimensions--e.g., duration, intensity, the severity of the resulting disability, etc.--which complicate the measurement process no end. In addition, the most difficult problem is to devise means of combining mortality with morbidity and impairment into a single index, and to do this in a manner which is other than purely arbitrary.

If these problems were not enough, further and more complicated problems have developed today, and to find the solutions to these problems is very much on the present agenda. Thus morbidity and impairment are today more than ever recognized as encompassing a number of "socio-medical" conditions, i.e., conditions which may be as much "social" as "medical" pathologies, often including some elements of both in some kind of a mix, e.g., drug addiction and other forms of drug abuse, alcoholism, sexual pathology, etc. In this sense both morbidity and impairment may be thought of as part of that elusive concept, the "quality of life", a condition which is evidently unmeasurable and perhaps even undefinable. An even greater complication which appears in this context is that "socio-medical" conditions often involve a moral dimension, i.e., we are not sure whether to designate people who are afflicted with these conditions as ill or as merely immoral rather than ill.

The World Health Organization some years ago took what, in retrospect, appears to be a "giant step" in the right direction when it defined health as not merely the absence of disease, but rather as a state of complete physical, mental, and social well-being. While the moral dimension is not mentioned here, this definition has the merit of at least pointing to the artificiality of the distinctions among the physical, mental, and social dimensions of health. Thus we come full circle. In an earlier day, before extensive specialization of occupations and the corresponding division of labor, the "medicine man" or "witch doctor" of preliterate communities treated the "whole man" for all of his ailments, whether physical, emotional, social, moral, or some combination of all of these. Subsequently, as society developed its "helping professions", men were, at least for purposes of treatment, fragmented, split into separate components of being in perhaps a very unreal manner, since the human being remained a whole human being. Today, at a time of great looseness in societal structure, and in the face of breakdown in many formerly potent social controls, we must define health in these earlier terms, even taking the WHO definition several steps further. This should be done despite the very obvious difficulties in operationalizing these concepts and devising an "index" of health, one which will consist of all of these components combined into some sort of a valid mixture.

II-Some Theoretical Considerations and Issues in the Development of a Definition

The approach taken here is this: Health must be thought of as a qualitative, multi-dimensional characteristic, one which must be inferred since it cannot be observed or measured directly. The problem, which both manifests itself and must be solved at several levels, is: What are the components of this multi-dimensional characteristic, and how many levels do these components have? Given the fact that there are several levels (a statement to be explained shortly), how can the components at each of these levels be treated as indicators and put together into a mix, or index, i.e., how to assign weights, which are not purely arbitrary, among the components at the various levels? These weights should have some theoretical relevance, or underpinning, yet permit the index which is to be developed to be operational in character.

The first level (either of analysis or health) is this: Man as a species consists of men as individuals who, although individuals, nevertheless live in groups of various kinds and sizes, and these groups are crucial, in various degree, to the existence of man. Thus the starting point is, or must be, man as an individual. Man as an individual is a biological organism, characterized by life and therefore having a beginning and an end. The sheer duration or quantity of life can readily and easily be counted along the dimension of time; this poses no really formidable conceptual problem. However, this uni-dimensional characteristic, the quantity or duration of life, must be thought of as a pre-condition for health rather than as a component of it; considering man as an individual, biological organism, it is a necessary although not a sufficient condition for health. The approach taken here, following the WHO definition, is that health should be defined in terms of the quality and quantity of life, rather than merely in terms of its quantity. Further, the quantity of life cannot be readily and easily combined into an index of health with various components of the quality of life. This is precisely the point at which all index construction, at least up until now, has foundered.

Before turning to this, however, several points should be made:

1. Even though the quantity of life can readily be measured on a uni-dimensional contiuum along an axis of time, and therefore objectively each unit has the same weight (as is the case on an interval scale), <u>subjectively</u> human beings attribute different weights to various points along that continuum, i.e., life at different ages or at its various stages appears to have a different "meaning", and therefore a different value or "weight"; further, that meaning, value, or weight varies according to cultural factors and value-systems, and probably, therefore should be considered as "socially" defined. We informally weight it differently in different historical settings and times. Thus even time as the dimension along which duration of life is measured is not as uni-dimensional as it initially appeared to be. Perhaps it is best treated as a qualitative rather than as a quantitative variable. One implication of accepting this statement is that no single index of health will be applicable at all times and places, or even for different segments of a population in the same society.

2. We may be able to avoid this difficulty, at least when considering the health of groups or population aggregates of various sorts, by assuming that on a statistical basis, i.e., where comparisons of populations are involved, mortality rates at various ages will be highly correlated with one another and with overall (age-adjusted) mortality rates and/or life expectancy. A possible exception to this assumption may be the "saving" of "high-risk" or "impaired" lives during infancy and the younger years, a widespread consequence of improved modern medical technology, with a resulting carrying over of these lives to the middle years, where even modern medicine can no longer (as yet) assure their survivorship. The strength of the overall correlation is a matter which could be determined by empirical research.

3. The "health status" of an individual at any given moment of life should also presumably be measured by an "index" of some sort. However, for an individual this index should clearly not contain a mortality component, except perhaps a prognostication, e.g., at his present stage of vitality (or health?), how many years of life can this individual expect, or how many is he likely to have remaining to him? This seems to be something along actuarial lines of a life table for "impaired" lives. A prognostication of this type should perhaps be thought of as a "proxy" indicator of health status, or of some index of health status, rather than as the actual component of an index.

4. Even if we consider the quantity of life in simple, objective terms as uni-dimensional, can we assume it to be highly correlated with the various components of the quality of life (however we define those)? This also is a matter for empirical investigation.

How deal with, i.e., define and measure, the quality of life? Even the concept of "disability-free days" doesn't do this adequately, except at the grossest level. That is, it addresses merely the performance of <u>major</u> social roles in dichotomous terms--yes or <u>no</u>. (Even here, as Parsons and Sullivan say, we have no adequate delineation of roles for the aged.) The subjective aspect of the performance of roles--their meaning and/or satisfaction to the individual and the manner in which they are performed--is omitted.

Perhaps this concept--quality of life--can be understood in the following terms: The individual is a biological, social, moral, and "emotional" being. (The WHO definition which had included "social well-being" as part of health had in fact begun to approach this idea, but it was never operationalized.) Thus health must be considered as having at least these four components, <u>all related to functioning</u>, and for which life is a pre-condition. The assumption must be made that, for the individual, any life is healthier than no life at all. Thus an individual is healthier without a limb, or with impaired vision, so long as he is alive, than he would be if he were dead. The same principle holds for a group or population aggregate, i.e., anything or anybody or any institution that threatens its survivorship is ipso facto unhealthy.

Perhaps the measurement of "quality of life" can be approached from this point of view. The quality of life is high where an individual "functions" at a high level--where he lives and is free of organic impairment or illness, or at least disability due to organic impairment or illness; where he lives and fulfills his major social role obligations satisfactorily according to his own values and those of his group (the potential contradiction between these two is worth extended discussion); where he lives and has a high moral self-evaluation and also receives a high moral evaluation by his group (again the potential contradiction is worth extended discussion); and finally, where he lives and is emotionally healthy.

Can we have healthy individuals in a sick society? Or vice-versa? Yes, I think. However health is defined, the health of a group or population aggregate (society, community, etc.) must be considered as an emergent phenomenon, more than simply the health of the individuals comprising the group or aggregate. The group or aggregate has its own "needs" for suvivorship, i.e., for the continuation of its functioning, so that its health can be defined in these terms. As regards the individual, a "sick" society would impinge on the health of the individuals comprising it when it provides little in the way of "public health" (preventive) measures or facilities for personal health; when it sets up expectations for the performance of major social roles which are impossible of fulfillment; when the average score of its population is low on a scale of socially defined moral approval, again perhaps because the standards are impossible of attainment; and when the average level of "emotional" health is low. Tensions and strains which at the extreme interfere with the adequate functioning of individuals in the society would thus be created.

Ideally a "tight" society, with a "rigid" social structure, is healthiest for the aggregate of individuals comprising it. However, a loosely structured society might be best for some individuals. In these terms there appears to be almost a contradiction between an individual's health and the health of the group.

Essentially a "model" is implicit in these statements, something along these lines: A "healthy" society (in another context, a "good" society) is one which permits the average individual--once he enters the state of life--to do the following:

1. Live a relatively "long" life. How long is long? Perhaps "long" is the best that can be, or has been, done for any large population aggregate, anywhere, in the state of the arts current at the moment of measurement. Clearly, a standard is implied, and it must be a changing standard as the state of the arts improves. This implies an upper limit, at least for this component of an index. It assumes that there are individual differences among men, that even in the "ideal" society, where social conditions are "perfect", length of life will vary among men due to individual differences in innate biological or genetic endowment. Also, subjective elements related to the socially defined value of life at different stages of the life cycle have to be included here.

2. Live a life relatively free of disability (as expressed in behavior rather than in subjective "feeling states") due to "illness" or physical or organic "impairment". Clearly a standard is implied here also, i.e., the best that can be, or has been, done for any large population aggregate, anywhere, in the state of the arts current at the moment of measurement. Here too individual differences and subjective evaluations of the social importance of freedom from disability at various stages of the life cycle are relevant also.

How can we combine these components into an index, i.e., what weight can we give to each? Very arbitrarily, a healthy society is one in which, at an absolute minimum, the number of individuals who are born, achieve maturity, and reproduce is sufficient to preserve the continuity of the most essential cultural components of the society from generation to generation. So by this definition alone, a definition which emphasizes societal continuity, what happens to people during their childhood and reproductive years, i.e., whether they survive sufficiently long to reproduce, is more important than what happens at other stages of their life cycles. However, one problem here is that while the upper limit to the reproductive period can be fixed for females at least within a relatively narrow age-range, this is not true at all, or perhaps not as true (i.e., there may also be a range, but it may be much wider) for males. Another problem is that, at least in industrial and post-industrial society, with its relatively high degree of control over mortality, survivorship through the reproductive period is no longer the major problem it once may have been. The pendulum has swung all the way over to the other side, and the shoe is now on the other foot. At this moment it looks as though there may be another swing of the pendulum, although perhaps of another kind, if fertility declines to below replacement levels.

These comments should not be understood as implying that zero or negative weight should be given in the construction of an index to the components of the index represented by survivorship and freedom from disability during the non-reproductive years. However, they do imply that less weight should be given to survivorship and freedom from disability during these stages of life than during the reproductive years.

3. Live a life in which each individual's major social role obligations are performed "satisfactorily" according to the values of that individual and the consensus of values of the "groups" to which he relates. Here the word "group" must be considered at many levels. For example, it may be considered as one's own immediate reference group--perhaps coincidental, or congruent, with one's primary groups, as the term was used by Cooley, one's family (either of orientation or procreation or both, either im-

mediate or extended), one's peers, etc.; it may be considered as some combination of the vast number of "secondary" groups, again in Cooley's sense, which impinge on each individual; or it may be considered as the larger community, however defined, or society, of which one is a member. The individual receives social reinforcement or reward from all of these groups, although at varying levels of intensity, and the outcome of this process has a varying degree of salience or importance to him; as a consequence, however, they are all of some importance to him in terms of his self-evaluation (an important element in Jahoda's first criterion of emotional health, self-perception).

What if the individual disagrees with the group, e.g., about the roles that he should play or the way in which he should play these roles, his life style, etc.? Is there room for the completely autonomous individual (in the sense of Riesman, Maslow, etc.) in society? Or, should we consider only the "adjusted" person (in Riesman's sense, adjusted to the social character type predominant in the society in which he lives), as healthy? More important for present purposes, is the autonomous individual healthy or unhealthy? My own bias along these lines is that the healthiest situation occurs when there is some disagreement or contradiction between the individual and his groups, and when as a consequence there is a level of tension along these lines sufficient to stimulate in the individual what Jahoda calls "positive mental health", i.e., some kind of a "creative" existence which in the end improves rather than detracts from the total level of the group, however that may be defined. However, this means that the disagreement cannot be too great, i.e., the disagreements cannot involve too many individuals, so that the essential functions of the group are not threatened. This returns us to our model, mentioned earlier, of the "healthy" society which, at an absolute minimum, produces a situation sufficient to preserve the continuity of the most essential cultural components of the society from generation to generation.

One implication of the preceding statements is that societies, like individuals, may have a hierarchy of "needs" along the lines elucidated by Maslow and which require satisfaction, and that these needs may be different at various stages of societal and technological development. For societies at low levels of technological development --hunting and gathering, pastoral, etc.--the most important requirements or needs are those of sheer survival, as indicated earlier; later this becomes less important, paralleling the situation in which mere subsistence, since it represents a problem essentially "solved", becomes less important to the individual. To take this a step further, this means that although we can, perhaps, devise an index of health which will have universal applicability to all societies at all stages of historical development, everywhere, in the sense that for each of these it will contain the same set of components, the relative weights accorded to each component must vary in accordance with the "problems" or "challenges" (in Toynbee's terms) facing that society at a given "moment" (a historical era, or epoch) in time and the extent to which any society has ever solved these problems or met these challenges. The same reasoning is applicable to an individual, in society, with regard to constructing an index of health status. The components must be the same for everyone, but the weights given to each component must vary according to the problems facing that individual and the best that any individual, in similar circumstances, has done in solving those problems. The same reasoning is also applicable to the construction of an index of health levels for population aggregates of various types. Again, the components must be identical, but the weights accorded these components must vary in accordance with the nature of the challenge to the aggregate and the "best" responses that have been made.

Now the problem of combination of components into an index is once again far more complicated than it was when this problem was confronted in the earlier phase or our discussion. We now have to suggest ways in which to combine survivorship and freedom from disability, considered as individual components of an index of health, with still another component, satisfactory role performance. To reiterate, the society requires at least a minimum level of satisfactory role performance, while for each individual, role performance which is satisfactory in both the individual's and his various groups' terms presumably results (or there is at least a correlation) in a satisfactory self-evaluation on the part of the individual. What numeric weights can we give to each component?

Here is what I suggest that we do, at least thus far: From the point of view of constructing an index of health levels for an aggregate, we start with the concept "expectation of disability-free years of life" or "disability-free survivorship" along the lines of suggestions made by Sanders and Sullivan. However, we have to modify this in several ways:

1. The expectation of years of disability-free survivorship at each age has to be summed for all ages, however with different weights to be given to the figures at each age. For example, the weights would be highest under 50 for females, perhaps under 65 for males, with the weights tapering off at older ages. The figure to be included in an index of health status of an individual could also be approached from this point of view: A prognosis could be made for the expected disability-free survivorship of a single individual at a given moment, based on his general state of physical health, perhaps as judged by a physician and based on the presence or absence of illness or impairment of various kinds and at various levels of severity, and the negative probability of being afflicted with any of these on the basis of one's life style. Each individual should be scored differentially, along the lines suggested above for population aggregates (it is most important to survive disability-free up to the ages specified above and less important thereafter, etc.), but this derived score for an individual would be modified by relating it to the best scores obtained by anyone in his age-and-sex group, etc.

2. Role performance can be judged behaviorally at the simplest and grossest level merely in terms of whether the individual is actually performing the major social roles expected of him and "appropriate" according to "social" definition,

i.e., deemed appropriate by both the individual and his various groups. These major social roles, I believe (clearly, this is my own bias) are of two general types, perhaps for the moment at least to be weighted equally in the construction of a sub-index--occupational and familial. However, here too some ambiguity arises. For example, what is the appropriate occupational role for a retired person? How do we classify unmarried adults, married couples without children, etc.? In some way people have to be classified in terms of their "social adjustment" as expressed in the performance of major social roles, and scores computed, so that these scores can be put into an index.

Having said this, the next question is, "How important is this social adjustment relative to disability-free years of life?" The answer to this question will determine the relative weights to be accorded to each of these components of an index of health. Again, sheer physical health (disability-free years) may be more important to one's overall or general health in the years prior to adulthood, while "social adjustment" may be more important to one's overall health during the adult years and through to the end of the child-bearing period, and once again physical health (disability-free years) may be more important later.

Physical and/or emotional impairment deserves mention in this context; these types of impairment set limits on what can be expected of an individual. For example, a blind person surely cannot be expected to perform a job for which sight is a requirement; nevertheless, there are other jobs which he can perform and which do not require sight. Should a blind person be considered less healthy than a sighted person?

In terms of the physical health component of an index, "yes" is the correct answer because the individual does have a physical handicap; he may have some disability, even if minor, resulting from it, and his sheer life expectancy, especially his expectancy of disability-free years, may in fact be substantially less (the latter because of accidental deaths, etc., the former because he may require relatively specialized care as he becomes older, etc.).

In terms of the "social adjustment" component of an index, however, it is primarily the subjective considerations which determine whether an individual with an impairment, e.g., a physical handicap such as blindness, should be considered less healthy than others. The question is, "How do the impaired individual himself and society in general, including the primary and secondary groups significant to him, define his impairment?" What expectations should he live up to? The social adjustment of the impaired person must be deemed satisfactory if he indeed lives up to the expectations for him. The expectations themselves are likely to be a blend of what is possible and what is desirable. If the impaired person can carry on an occupation and/or hold a job (even in a "sheltered workshop" type of situation), this is all to the good and his social adjustment must be rated as high. Similarly, if the impaired individual can carry on normal family relationships, the same must be said, i.e., his rating must be high.

A distinction must be made here between impairments which have no presumed "voluntary" component in them and those which have. This attribute of impairments must be seen as a continuum, with an infinite number of gradations along the line. Some impairments are essentially self-inflicted, i.e., they occur because the individual engaged in behavior with a high probability of becoming impaired as a consequence, while others may be thought of as purely accidental. Much less stigma attaches to the latter, and individuals in this category are much more likely to make a proper "social adjustment" to their impairment and to be defined as "healthy" in spite of their impairment.

At the societal level, societies with large proportions of impaired individuals--impaired in the former sense, with a large component of presumed voluntariness involved--should be rated as less healthy societies. This is particularly true of societies with large numbers of alcoholics, drug abusers of various kinds, sexual deviants, psychopaths, etc. But societies as the unit of measurement should also be rated as healthy or less healthy in terms of the amount of family disorganization, crime, and unemployment which characterize them.

4. <u>Live a life in which each individual has a high moral self-evaluation but also receives a high moral evaluation from his group</u>. This ties in very closely with the preceding discussion. Individuals who meet their major social role obligations are likely to receive a high moral evaluation, both from themselves and from their groups. They are likely to be considered worthy persons and their self-concepts are likely to be good (in Jahoda's terms).

In Western society, at least, a negative moral evaluation is likely to result from "copping out", i.e., abdication of moral responsibility for work and satisfactory family relationships, and particularly when these occur in the absence of an "achievement-orientation" on the part of the individual, i.e., an orientation to control and transform nature and the environment to suit man's ends. Conformity to Riesman's inner-directed social character type remains the moral ideal. Although some changes may be occurring in this ideal, it does not appear that basic concepts have been in any way altered. The alienated commune-oriented individuals constitute but a small minority, an unimportant segment of the total.

Elsewhere, however, i.e., in other civilizations, values differ. Thus in societies where Hindu and/or Buddhist values predominate (e.g., India, S.E. Asia, and elsewhere), an "escapist" orientation is perceived much less negatively than is the case in Western societies. Moral approval is often conferred upon individuals who retire from the "active" life to a life of meditation and contemplation. But whatever the activities or life styles which result in moral approval is everywhere required as an essential component of health.

Here also an important methodological question is, "How important is moral approval relative to social adjustment and disability-free years of life?" The answer to this question, again, will determine the relative weights to be accorded to each of these components of an index of health. However, moral approval unlike social adjustment does not appear to have an age-referent; it probably is equally important at all ages and stages of life.

In any community, disability-free years of life can be aggregated for all individuals comprising the community and divided by the number of individuals, and an average number of disability-free years per individual in the community thus obtained. Communities can be compared on the basis of these averages. The same may not be true, however, for social adjustment and moral approval, except as indicated below. Both social adjustment and moral approval are essentially socially conferred attributes, and probably their distribution in any community is described by a normal curve. That is, some individuals in any community will be very well-adjusted, some will be socially deviant, and probably most will fall somewhere in the middle. Probably the mean scores will be similar for most societies within the industrial society category and similar also among traditional societies which are culturally similar and similar in terms of their major value-orientations. The same situation probably characterizes moral approval. The implications of these statements for the measurement of health levels and the construction of an index of health merit further exploration and discussion.

III-The Health of Individuals Versus the Health of Groups or Aggregates

However health may be defined, a major theoretical issue which does not appear to have been treated adequately in the literature is whether health should be defined differently for individuals than for groups. In practice the concept of health status is used to refer to the health of individuals, while the concept of health level is used to refer to the health of a group, a community, or some sort of a population aggregate.

Another way of asking this question is to ask whether, once health has been defined, the measurement process should be that the sum total of the health statuses of all individuals in the group should be simply cumulated and averaged to derive a measure of the health of the group as a whole, or is the health of the group an emergent phenomenon, over and above the health of all the individuals comprising it, and therefore to be measured in some other fashion? Some logic arguing that there may actually be different definitions for the health of individuals and for the health of groups stems from the differing "functional requirements" of individuals versus those of groups. (Groups in the present usage may be societies, families, communities, etc.) "Functional requirements" should be understood here as the requirements for survival of the individual as a system--biological or otherwise--and clearly the requirements for survival of a group qua group, i.e., as a system--cultural or otherwise.

Three definitions of health, and therefore three levels of analysis, appear to be involved here. One of these relates to the individual and his functional needs for survival as a biological organism. He needs food, clothing, shelter, and perhaps many other goods and services. But Maslow's conceptualization of a hierarchy of needs among individuals is relevant here, and at the simplest level an individual may require only the most necessary ingredient for survival, i.e., food. The health of a group, in the second level of analysis, may be thought of as simply the sum total of the health of all the individuals comprising that group. Here the group or community is considered as an aggregate, no more or less, and without emergent properties.

But there is also a third meaning to the notion of health, especially with reference to the health of a group. This has to do with its sheer survival. In other words, for survival the group itself must be considered as having certain functional needs qua group. This is important because all individual human beings are members of groups. They derive their distinctive humanity from group membership. Therefore it is incumbent on the group to survive if human beings are to remain human.

From this point of view it may be argued that, since human societies and communities are comprised of human populations, their survival depends at least in part upon the provision of an adequate physiological relationship to the setting in which they exist, including a considerable degree of control over fertility and at least a minimum degree of control over mortality. Related to these, human society also depends for survival upon such essential functions as socialization, language and communication, economic production, the preservation of order, maintenance of motivation, and the establishment and maintenance of integrated values. Thus, fertility and mortality are sociological phenomena interrelated with other essential features of human societies. Population is an endogenous variable in the analysis of social systems. Each society has structural patterns with consequences for fertility and mortality, i.e., its structures have to be suitable for survival in demographic terms.

Not surprisingly, therefore, preservation of life is a universal value in all human societies, except under those circumstances in which the taking of one's life, or perhaps of a small number of lives, rather than decreasing the society's chances for survival reaffirms the values of the group and stresses its solidarity, or in some other way improves its chances for survival in the face of a hostile environment. All societies institutionalize patterns of behavior intended to preserve health, or to restore and maintain it, and to prevent death. Often these patterns of behavior may actually be inefficient and inconsequential in achieving this end; probably, prior to the modern era, such mortality control as was achieved in primitive and pre-modern communities stemmed more from advances in technology, including improvement in the food supply, and from the maintenance of political order and protection than from behavior explicity intended to maintain health.

Measurement implications

a. If we consider any individual separately for purposes of measuring his health status, we must ask: How well does he function at a given moment? That is, is he well enough as a biological system to carry on his major social roles; is he well enough emotionally (as a member of a society and as a member of the various social groups to which he belongs) to accomplish this end; and does he have a sufficiently high moral evaluation? These are all questions which relate to the quality of his life, as distinct from its quantity. But we must also ask, how long has this individual lived up to this point and what is the duration of his life likely to be, given the quality of his life at this moment? We thus bring in some estimate of the quantitative dimension of health and add it to the qualitative.

b. If we consider any group or aggregate of people for purposes of measuring its health level, we simply aggregate, in some manner, the various measures of the quality and quantity of life of all individuals comprising it, as indicated above. This says nothing about the health level of the group qua group; it relates only to the health level of the aggregate of individuals comprising it. For this purpose the quantity of life is readily measured by conventional mortality rates or life expectancy.

c. If we consider any group as a cultural, societal, or civilizational collectivity, or as a community, for the purposes of measuring its health quotient, we have to consider measures of the quality of life in that collectivity as lived by individuals at any given time, and measures of the quantity or duration of existence of the collectivity itself as a collectivity and independent of the lifetimes of any individuals comprising it. From this point of view, the health status of individuals and the health levels of groups may be independent of the societal or civilizational health quotient; a healthy individual may exist in a sick society or civilization and a sick individual may similarly exist in a healthy society or civilization.

IV-The Quantity Versus the Quality of Life as Components of an Index of Health

Another major issue in defining the concept of health stems from its multi-dimensionality. In the broadest terms, health may be considered as related in some way to both the quantity and the quality of life. This is true regardless of whether we are considering the health status of individuals, the health levels of aggregates, or the health of collectivities. Within each of these, the problems are:

a. What indicators shall we use to measure the quantity and what indicators shall we use to measure the quality of life?

b. How shall we add these indicators into an index, i.e., what weight do we give to each?

The quantity of life may be thought of as uni-dimensional since it is measured along a continuum of time. However, various indicators have been used to measure the force of mortality on population aggregates; perhaps the best-known of these are the crude mortality rate, the age-adjusted mortality rate, the infant mortality rate, the proportional mortality ratio, and the expectation of life.

A major problem underlying all of these measures (except the infant mortality rate) is that the weights assigned to each year of life are implicit rather than explicit; clearly, they should be explicit before indicators of the force of mortality are selected to be included in an index of health. Also, some measures give more weight to certain ages than to others; for example, the expectation of life is affected much more strongly by changes in mortality during infancy than at the older years. This may be in accord with the values of our society, or it may not, and empirical research along these lines is sorely needed. Also needed is research on the statistical relationship or degree of association of the various measures to each other, so that we may know whether any one may be validly taken to represent the others.

The quality of life, by any analysis, is best thought of itself -- like health in this respect -- as a multi-dimensional concept. But in accordance with the framework offered here, the quality of life may be thought of as consisting of the following components:

a. A bio-medical component.

b. A socio-emotional component.

The bio-medical component of quality relates to the degree to which an individual is able to function free of bio-medical illness or impairment. The extent to which bio-medical illness or impairment results in some form of disability or less-than-perfect functioning is a key factor here, as is also the extent to which it results in a shortening of life. Asymptomatic illnesses or conditions, or illnesses or conditions in a pre-symptomatic stage, may not result in disability or less-than-perfect functioning; nevertheless, they may result in shortening the affected individual's life. The bio-medical component of the quality of life is affected by all of these. But a complication of measurement of the bio-medical component is that an individual in perfect health along the bio-medical continuum may die instantly due to some external cause (violence, poisonings, accidents, etc.). This individual moves instantly from one end of the continuum to the other, in the process skipping all intervening stages; this poses a measurement problem.

The socio-emotional component of the quality of life involves at least three first-level sub-components: emotional health, social functioning, and moral worth. However, the emotional health of an individual is to a considerable extent (certainly not entirely) conditioned by his perception of his own social functioning and moral worth. To some degree the individual's perceptions along these lines are correlated with the perceptions of others; to the extent that they are not, a difficult measurement problem emerges. However, this may be solved by using the incongruity of self-perceptions with the perceptions of others as one among other items (a second-level sub-component) measuring emotional health.

How shall we mix indicators of the quantity of life (e.g., life expectancy for an aggregate of population) with indicators of the quality of life (e.g., the number of disability-free days at various levels of disability)? That is, what weights shall we give each indicator used as a component of an index? Unfortunately, all judgments along these lines must necessarily be arbitrary since no empirically validated criteria have ever been established.

BIBLIOGRAPHY

1. Baumann, Barbara, "Diversities in Conceptions of Health and Physical Fitness", J. of Health and Human Behavior 2,1:39-46 (Spring) 1961.

2. Brockington, Fraser, World Health. (Balto.: Penguin, 1958.)

3. Cooley, Charles Horton, Human Nature and The Social Order. N.Y.: Shocken, 1964.

4. Dubos, Rene, Mirage of Health: Utopias, Progress, and Biological Change. (N.Y.: Doubleday, 1959.)

5. Dunn, Halbert L., "High-Level Wellness for Man and Society", Am. J. of Pub. Health 49,6:786-792 (June) 1959.

6. Jahoda, Marie, "Current Concepts of Positive Mental Health". Joint Commission on Mental Illness and Health, Monograph Series, No. 1. (N.Y.: Basic Books, 1958.)

7. Maslow, Abraham H., Toward A Psychology of Being. (N.Y.: Van Nostrand-Reinhold, 1968.) Second Edition.

8. Mead, Margaret, Editor, Cultural Patterns and Technical Change. (N.Y.: New American Library, 1955.)

9. Parsons, Talcott, "Definitions of Health and Illness in the Light of American Values and Social Structure", in: E. Gartly Jaco, Editor, Patients, Physicians, and Illness. (Glencoe: Free Press, 1958), pp. 165-187.

10. Parsons, Talcott, "Toward a Healthy Maturity", J. of Health and Human Behav. 1,3:163-173 (Fall) 1960.

11. Riesman, David, et.al., The Lonely Crowd. (New Haven: Yale, 1961.)

12. Sanders, Barkev S., "Measuring Community Health Levels", Am. J. of Pub. Health 54,7: 1063-1070 (July) 1964.

13. Sullivan, Daniel F., "Conceptual Problems in Developing an Index of Health". National Center for Health Statistics, Vital and Health Statistics, Series 2, No. 17. (Wash. D.C.: U.S. G.P.O., May 1966.)

14. Toynbee, Arnold, A Study of History. (N.Y.: Dell, 1965.) Paper.

DISCUSSION

Daniel B. Tunstall, Office of Management & Budget

I would like to discuss these papers in quite a different order from that proposed by the chairman. Let me begin with the paper on theoretical considerations in the development of a definition of health by Munroe Lerner; follow with my thoughts about the Bush, Chen, and Patrick paper on social indicators for health based on functional status and prognosis; and end with an examination of Karl Yordy's views on why we need health indicators.

From my point of view there is a paper missing from the list. What is needed is an exploration of the relationship between indicators of health and a system of social indicators, the development of new statistics using social indicators as a framework, and the development of social analysis and social policy formulation using social indicators as a primary tool for measurement and judgment. In fact Abe Aidenoff, Acting Director of the U.N. Statistical Office, was expected to present his ideas on a system of social and demographic statistics that would have provided one link between the development of measures of health status and the development of new health statistics. To some extent I can substitute my interest in the research and development of social indicators and thus represent both the person interested in the development of statistics from the point of view of a central statistical office and the user of health statistics in the development of social indicators of health.

Another general point: I wish we had the opportunity to continue a discussion of these papers this afternoon. I have a great deal to talk about and comment on and I wish I had the chance to discuss these ideas with the participants.

As I read Dr. Lerner's paper I had the feeling that he is talking about two types of health: the health of the individual and the health of society. The health of an individual, he suggests, is made up of four components or states: biological man, social man, moral man, and emotional man, and each should be thought of in terms of function or behavior. The health of society, he indicates, has two dimensions: the individual and something greater than the individual. A healthful society is composed of men who live long lives, free of disability, and perform their social role satisfactorily. That something greater than the individual is that the race be preserved. While I agree with Dr. Lerner that it is best to think of the health of the individual in broad terms, up to now too narrowly focused on the biological needs of man, I do not agree that we should combine our judgments of an individual's health, no matter how extensively conceptualized with judgments about society. I definitely prefer separating the two. In effect, let's not end up saying that the ultimate goal of society is to be healthy, or to be educated, or to be fully employed and satisfactorily employed, or fully recreated. It is so clear that individuals want all these states of well-being and that society can be judged in terms of the number or percent of individuals who attain various states of health, education, and recreation. But society is greater than the sum of these measured states of man. Using human attributes or state of well-being to describe society is too limiting and, I think, misleading.

I have another criticism. I would drop the use of the phrase "quality of life" to express the health condition of the individual and return to the more limited term, health status of the individual. And, I prefer the use of the term social concern to reflect the issues and goals we as members of society have for the well-being of individuals and the term societal concern to reflect issues and goals of society that are beyond the scope of merely an individual's well-being. Thus, my social concerns for health would be for the individual to live a long life and that his life be as free from disability as possible; the societal concern for health may be expressed in terms of preservation or survival of the species, or other species, and/or of a prescribed ecological habitat.

Throughout his discussion, Lerner draws heavily on ideas expressed by the World Health Organization. However, when the time comes to develop an index of health, he relies on the work done by Sullivan with a number of modifications.

Lerner begins with Sullivan's "expectation of disability-free years of life" which he would weight according to the following biological, social, moral, and emotional factors: age; prognosis based on life style using occupation and/or family status; extent of permanent disability; moral self-evaluation and group evaluation; and, although not explicitly included, I would expect he would want to add a weight for "emotional" health. Just making out the list leads me to react against such an index. Clearly, if it were available and meaningful, it would be useful. But, notwithstanding the problems of data availability, wouldn't indicators of each component be sufficient to make a number of important judgments about health status? Do we need to quantify all 4 components and combine them into one measure of health?

I think Lerner is correct in challenging those of us who will use health indicators to go beyond only those concepts which we can quantify. But that should not mean that we force quantification of aspects that are not easily quantified. The judgments that allow men to make decisions about health policy surely do not come about through quantification alone. In fact, one of the tenets of the social indicator movement, which many social scientists believe, is that much more can be quantified than is currently being quantified. I think Lerner realizes that he reached the limits of quantification as a means to communicate health status of individuals because he did not suggest examples for the emotional or moral states in which individuals exist.

His examples of the social state are age, family status, and occupation, which I believe are quite useful additions to the Sullivan model.

Since Bush and his associates give extensive documentation for a health status index, I would like to turn to their paper next. Let me add, however, that Lerner's contribution at this point in the development of health indicators makes me listen. The publication of Lerner and Anderson's Health Progress in the United States, 1900-1960, was surely one of the most important steps in the progress of improved health reporting during the past decade, and perhaps his new approach will now challenge the rest of us to think quite differently about the need for health indicators in the 1980's.

The Bush, Chen, and Patrick Paper

There are three ways to expose one's research to a group such as the American Statistical Association. A researcher can write a think-piece outlining the need for study in a given area; he can prepare a working paper about a project in progress; or he can present his findings upon completion. The Bush, Chen, and Patrick paper is an example of the best kind of paper reporting on work in progress. The authors have a strategic view of the role of the health analyst ranging from the task of identifying indicators of health conditions to suggesting how health indicators can be used in many aspects of health planning.

Quite simply, the authors propose the development and use of two health models and related health indexes. The first model is Functional Status, defined in terms of the activities an individual can perform (social, mobile, physical), with a related index that takes into account measures of these activities and measures of people's evaluation and rating of these activities. The second model is Health Status, defined in terms of health criteria such as years of life, with a Health Status Index that is defined in terms of years of life adjusted for time spent disabled. Thus, the Health Status Index contains a measure of prognosis.

Before I explore various aspects of these two models and their statistical indicators, let me list the nine criteria that a health index should meet, suggested at the beginning of the Bush paper. Throughout, I will criticize their models and indicators in terms of these criteria. An index of health should be of normative interest, thus reflecting perceived values of society; useful for planning; sensitive enough to detect significant changes in health status; reflective of multiple states of health; available in time series; comprised of an aggregation of components related to policy and thus open to manipulation by policy intervention; easily reproduced; and available without extensive new data collection.

A basic component of the Bush Function Status model is a ranking of health states according to functional levels. Three functional categories are used: a social activity scale, a mobility scale, and a physical activity scale. The authors intend each scale to be exclusive of the others and, when taken together, comprehensive in their coverage of functional activities. They construct a ranking based on the three scales, ranging from L1 (Death) to L30 (Optimum Function). These "objective"(my term) states of health are then evaluated through a survey of people's attitudes and ranked on a scale ranging from 0.0 (Death) to 1.0 (Well). Obviously, everyone in the country can be given a measured value (0.0-1.0) that corresponds to a functional level (L1-L30) and these values can be summed and averaged for the entire population. Changes in value over time would thus reflect changes in health status of a given population. Bush's Table 3 shows how the Functional Status Index is computed.

To develop a Health Status Index, the authors attempt, through evaluation of research and existing health conditions in the U.S., to determine the probability of an individual moving from one level on the Functional Status scale to another over a given period of time. Applying existing probabilities to the Functional Status Index gives us a Health Status Index for the population. The Bush Health Status Index, as Bush notes, is very similar to the index developed by Sullivan--expected years of life free of disability. However, the Bush measure includes a social value component.

Bush's Health Status Index is presented in terms of what is commonly called a criterion indicator--years of life--but weighted according to an "objective" measure of disability and a "subjective" (my term) evaluation of that disability. Thus the unit of measure for the Health Status Index is expected quality-adjusted years of life, a term which, with enough coaxing and increased use, could be understood by the general public.

The major difference between the two models is that the first, Functional Status, contains significant detail to show variety of health status and is thus sensitive to change in health status in the population. Its major use would be to accurately monitor the health status of the population as a whole or any select group within the population, such as an age, sex, or racial group. The second model, Health Status, is more useful for planning and decisionmaking because it takes into account prognosis, that is, it contains a measure of the probability of an individual moving from one health state to another.

Let me quickly sketch out a number of problems and criticisms I have with the two models and indexes. The Functional Status Index is based on a data base that doesn't really exist. Data for the Social Activity Scale are taken from the Health Interview Survey of the National Center for Health Statistics and data for the Mobility and Physical Activity Scales would have to come from other governmental surveys, such as the Social Security surveys, which do not cover all groups within the population.

None of the government surveys are coordinated to give data on a timely basis. In fact, only death is counted on an annual basis, and it is reported two years late. There are no national surveys of attitudes about functional status that can be used to assign values for the operational measures of function. This will remain a major problem with the Bush model and other social models that attempt to combine "subjective" attitudes about social conditions with existing "objective" data.

Clearly, there are reasons why these data have not been collected. Should death be included in a functional index of health? In the Bush model there is no way to include death in the index because a person, in terms of health, ceases to be a person at death. Do the three scales cover all forms of health function? I don't think so. We need more research to define social functions. The clearer the specification of function, the more likely the Index will be used, and that is the goal of this effort.

The authors also suggest that they have conquered the controversial definitional problem of positive and negative health, but I don't think so. Being "well" is not the same as being "physically or mentally fit"; it just means that the individual does not have any symptom. The Bush values do not distinguish between being well and being fit. This may suggest that a fourth scale is needed. Nor is the problem of pain adequately dealt with. If pain is associated with disability in some form then it is taken into account in the Bush model; otherwise, it is not. Thus the person suffering pain and restricted in his activity is ranked the same as the person with equal restriction but no pain.

Looking more closely at the actual scales used, it appears to me that it may be more practical to have 5 or 6 levels rather than 30, and, in fact, the measured values seem to clump at 5 levels: well (1.00), well but with some condition (.65-.80), some disability (.50-.59), serious disability (.39-.49), and extremely limiting disability (.27-.34).

Clearly, the problem with the Health Status Index is that probability data are not available. For now, Bush has gone beyond Sullivan's model conceptually, but practically speaking he can go no further than Sullivan because the data have not been organized in a fashion to meet the needs of social planning. I think this point is crucial. We do not necessarily need a massive new data-gathering effort to develop social indicators of health; we need instead the conceptual framework to organize existing data and pinpoint weaknesses and gaps in our information. This is a major contribution of this paper and the research of Bush and his associates. Their framework clearly highlights the weaknesses in developing and using health indicators with the current data base. Let me give an example. By clearly specifying the need for data on the probability of an individual moving from one functional status to another, the authors have provided a framework for organizing the vast amount of information that exists on medical treatment. One can measure individual (and group) movement into and out of a functional status. The more homogeneous the group (age, sex, race, occupations, etc.) and the more observations of treatment (social experimentation), the more likely one can identify the effects of medical and/or preventive action on functional status; and, thus, the clearer the link between health indicators and health programs. This linkage provides a sound science base for health policy. And because peoples' evaluation of health status is included, the framework provides a sound political base for health policy as well.

Bush and his associates now need to influence the development and organization of new health data to meet the requirements of their Functional Status model and influence the organization of medical research findings to meet requirements of the Health Status model. From past experience, they will be successful if they can show data users how their models can be used in health planning. This paper and their ongoing research is a major contribution in this field.

<u>The Yordy Paper</u>

To start with, I believe Yordy makes four main points: (1) health indicators are important for policymakers and for policymaking; (2) policymakers will be the ones to create the demand for indicators; (3) health should be viewed as a desirable end result and not just as a means to improve other social goals such as economic growth, education, etc.; and (4) there is urgent need for health indicators to quantify health goals, to monitor change in health over time, and, to evaluate programs designed to improve health in our society.

As support for these points, Yordy describes some of the major health programs begun during the 1960's (community mental health centers program; medicare and medicaid; establishment of neighborhood health centers for the poor, etc.) and points out that during this time a number of important health problems were uncovered (the poor have higher infant mortality rates and morbidity rates; there was an increasing number of deaths due to heart disease and cancer; health care costs were a serious financial burden for the elderly, etc.). Yordy does not say directly that the indicators influenced the new health policies but the implication is there. I would like to know more specifics.

What was the role of health indicators in gaining support for a community mental health center program rather than an expanded state institutional system? Was the development due to technology that would allow for more extensive home and hospital outpatient treatment or was the critical change merely the realization of high costs of institutionalization, which forced administrators to look for economic alternatives? Did health indicators (that is, measures of individual health status) play any role?

Let us take the new cancer research program, as another example. Was this policy decision made on an examination of the rapidly increasing death rates for malignant neoplasms for males and particularly for Negro males? Or was the program and the $100 million appropriation a culmination of the efforts of a cancer lobby that happened to push hard enough just at the time when the Administration and the Congress were jockeying for control over the direction of the health care system and over research and development budgets for the next 25 years?

Did the knowledge that the infant mortality rate was leveling off during the early 1960s trigger an effort during the past Administration to influence pre- and post-natal care and were those programs successful? We know that the infant mortality rate decreased from 26 infant deaths per thousand live births in 1960 to 25 per thousand in 1965, but, from 1965 to 1971, the rate dropped from 25 to 19.2 per thousand. Truly a remarkable turnaround--a 22% reduction in the infant mortality rate for whites and a 25% reduction for Negro and others.

While it is difficult to know just how information is used to determine policy shifts, Yordy could have told us how various types of information are elevated to indicator status. And while the political scientists may call this a problem of an information base for political decisionmaking, we who work with information may call it the politics of information. The President's Health Care Message in February, 1971, for example, used only one major health indicator--days of disability--and he noted that days of disability had been reduced by 6 percent from 1960 to 1970. In 1972, the Health Message contained three indicators of health status: life expectancy, which increased 3.4% from 1950 to 1970; maternal mortality, which declined 66% during the same period; and days lost from school and from work, which decreased 7.5 % and 3.5% respectively during the same period.

Why were these indicators selected? Did those writing the document realize that in 1971 they had neglected to include a measure of the institutional population and that, therefore, change in overall days of disability for the entire population may have been inaccurately reported? Is this why they changed to more specific but less health-oriented indicators, like days lost from school and work? Why change from a ten-year period to a twenty-year time period? Why focus on maternal mortality--an insignicicant cause of death, but possibly the only cause of death that showed such tremendous decrease in recent decades?

The selection and use of health indicators is part of a larger problem in government--that of sharing information about society with those who need it and want it. With all the money spent collecting health statistics, why can't HEW produce a publication that focuses attention on major health status indicators, thus giving the American public a statistical picture of health in the nation? And, why is there no publication of government sponsorship that analyzes health conditions in the U.S. and identifies health problems?

I was hoping Karl Yordy would explore this larger problem because it has a direct bearing on the use of health indicators to affect health policy during the next decades. If the 1960s was a time of "identifying deficits in the health of the nation," which, as Yordy suggests, was the role of health indicators and health policy, what will be the role of health information in the 1970s and the 1980s?

Yordy suggests that "the competition for public resource inputs leads to continued pressures to define output measures and to relate the input to the output" and "resource allocations should reward those programs which over time are more successful in defining measurable outputs and the relationship of input to output." The author has stumbled onto cost benefit analysis as a modus operandi for health indicators. In part he is right. There are health care deficits left in the U.S. (that's a sterile way of saying some people don't have enough food to eat, don't get emergency care when they need it, are not immunized when epidemics are expected, etc), but the critical problem for those in authority and for citizens as a whole is making a choice between health and medical benefits. Both citizens and authorities must judge expenditures for health care relative to needs for transportation, education, housing, environment, and other public and private needs.

Do we want to live longer? Do we want to eradicate cancer as a cause of death and disability, and pursue a massive research effort to do the same for heart disease? Do we want to emphasize preventive care to reduce the amount of disability among the middle aged? Do we want to make the effort to reduce death amoung the young (5-24), the causes of which are almost all related to violence and social behavior? Which decisions can be made by those in authority and which come from subtle but significant changes going on within society that social scientists and politicians alike are unable to explain.

The choices are great; they affect different groups in society, and benefits will not be equally distributed. To make these decisions about health goals and health care, it is imperative that government, the primary collector of health data, inform the public about health conditions so that everyone has the opportunity to participate in the decisions. The critical point is that health care and health goals are social issues, and there is no way to deal with social issues equitably except with social and political participation by the public. This can only take place if the public has information. The selection and communication of health status indicators to the public on a timely basis meets a definite social and political need.

DISCUSSION

Martin K. Chen

National Center for Health Services Research and Development

Dr. Bush's paper entitled "An Index of Functional Status and Prognosis," describes one approach for measuring health status. He classifies functional status into 31 levels, with optimum functioning as the highest level and death the lowest level. The functional status of a population is simply the weighted average of the functional levels given values between 1 and 0. The index of health is simply the function index with the addition of the prognostic factor in the form of transitional probabilities of a cohort of subjects remaining in the same functional level or migrating to different functional levels over time, using the first-order discrete-state Markovian process. The probabilities are to be estimated from survey data.

Undoubtedly this approach has its value in that it can be used in monitoring the health status of a population, provided that the functional levels can be accurately differentiated and that accurate data are obtainable. An examination of Table 1 reveals, however, that there are four or five inversions in the values of the function levels; that is, some lower function levels have higher values and vice versa. Furthermore, it does not appear that the three components of social activity, mobility and physical activity are truly mutually exclusive, as the author claims. I would think that physical activity in the form of walking is a subset of mobility: I find it hard to believe that a person who cannot walk freely can travel freely.

My basic objection to the approach is that it is oriented toward the provider or the third-party payor rather than the consumer. The indices deal with expected values and averages rather than individuals. This approach is useful from the actuarial point of view, but it does not help the individual consumer who is interested in his own health status. The functional status index and the health status index are based on expected values of the distributions of cohorts. If the variances of the distributions are large, the expected value or mean has little meaning for an individual who is three standard deviations from it. It may be argued that homogeneity within cohorts can be achieved by subdividing a cohort into sub-groups, but if this process is kept up, the branching can become so unwieldly that even the largest computer cannot handle the immense transitional probability matrix.

The paper by Dr. Yordy, "Why Health Indicators?", provides a comprehensive picture of the interplay of the political process and design of health indicators. Dr. Yordy traces the intricacies of political decision-making in relation to the requirement of evaluation of health intervention programs. His prophecy that policy-makers will in the future prefer outcome indicators to input indicators in the evaluation of health programs is both reasonable and encouraging-encouraging in the sense that such a shift in emphasis will mean the inception of the maturing process of politicians concerned with the health of the American people. In the final analysis, what really counts is the improved health status of the people, not the number of hospital beds and/or clinics that are made available to them.

An insightful observation made by Dr. Yordy relates to the establishment of causality between program intervention and outcome. This is an experimental design problem that is perhaps as difficult as the sesign of health indicators itself. This is an area in which the greatest contribution has come from Dr. Donald T. Campbell of Northwestern University. It is possible that, through his work and that of others in the quantitative sciences a solution or a series of solutions will be found to the problem of establishing a logical nexus between program input and program outcome in an open system full of unknown and therefore uncontrolled extraneous influences.

Dr. Lerner, in his paper entitled, "An Approach to conceptualizing Levels of Health," addresses himself to two key issues: the advisability or inadvisability of aggregating quantitative and qualitative indicators in a single index of health and the need for precision in health meansurement. There is no doubt that problems exist in trying to combine quantitative and qualitative indicators of health in a single quantitative entity. First of all, it is difficult to quantify qualitative information. But even if a way were found to quantity the qualitative information, the problem of comparability of units of measurement would still remain. One could not aggregate quantitatives that were in different units of measurement.

In spite of these difficulties, there will be political pressures on health administrators to develop a single index of health that can be used in monitoring the health of the people. This single index, when developed, can be used in the same way that the GNP is used in monitoring the nation's economic health.

I do not agree with Dr. Lerner that precision is not an important consideration in the development of health indicators. I believe that scholars who are working in this area are trying to make a science out of the art, and without precision we cannot have a science. A health indicator that is imprecise serves little useful function. For example, if we compared two communities in terms of the health indicator, any true difference that exists would be masked because of the large components of error in the indicator. Such an indicator would not be worth the time and energy devoted to its development

Dr. Lerner appears to endorse the WHO definition of health whole-heartedly. This definition, that "health is not only the absence of disease, but a complete state of physical, mental and social wellbeing" is intuitively appealing because it is comprehensive and covers all the essential aspects of what may be called "the quality of life." But how does one define complete physical, mental and social wellbeing? And if one could define these terms satisfactorily, what kind of data would one collect? There is no doubt that conceptually the WHO definition is superior to many other definitions of health, but there does not appear to be any solution, at least for a few years to come, to the problem of operationalizing the concept to the satisfaction of health measurement specialists.

STATISTICS AND POLITICS

Philip M. Hauser, University of Chicago

Statistics had its origin in politics--in the broad sense of the term. Historically considered they were "state-istics" and statisticians were once "statists." But "politics" in the title of this essay is used in the narrow sense--in the narrow usage associated with the way in which this nation conducts its system of governance on the federal, state and local level. That is, examined here is the relationship between the politician in office and the collection, compilation and analyses of quantative data which are derived from government records or censuses and sample surveys conducted by government. In the broad sense this involves the consideration of the role of statistics and of the statistician in government and use of the statistical product by policy officials in government, elected and appointed.

More specifically, what is examined is the extent to which statistics and statisticians are subjected to pressures to make a given administration, administrator or agency "look good," "make a case" or support a decision already taken on other than factual grounds.

First, it is well to state at least some of the basic assumptions on which this discussion rests. Among them are:

1. Statistics are quantitative facts collected, aggregated and analyzed to provide intelligence, to facilitate understanding and to serve as a foundation for formulation of policy, development and administration of programs and evaluation of the impact of programs.

2. The statistician is the professional specialist whose function it is to design, produce and analyze statistics and to present his findings in an objective manner with probity and integrity for use by policy makers, administrators, researchers and consumers in general.

3. The consumers of statistics constitute the audience for whom statistical intelligence is produced and who, it is assumed, want to "know the facts."

4. Although the relationship between consumers and producers of statistics must be a close one so that the information produced is relevant to the problems which confront consumers, this relationship should not in any way have the effect of impairing the integrity of the statistical undertaking as a fact finding enterprise or requiring the equivalent of a "directed verdict."

The Present Situation

Developments over the past several years have raised serious questions about the extent to which political pressures are being exerted on government statisticians so as to compromise the integrity of their product; and about the way in which the statistical product is being used in the interest of the Administration or the administrator of an agency, rather than in the interest of the general public. The political smoke which has risen from developments in the Bureau of Labor Statistics and the Bureau of the Census, and smolderings in a number of other statistical agencies, including those in the Department of Agriculture, the Department of Health, Education and Welfare, and the Statistical Policy Division of the Office of Management and Budget, have incited many, including statisticians, politicians, businessmen, labor leaders, reporters and editorial writers, to look for statistical fires. After all, it is part of our folk wisdom to assume that "where there is smoke there must be fire."

The widespread suspicion that unwholesome pressures have been placed by the present Administration on statistics and statisticians and much public airing of the matter in the mass media have, thus far, produced no clear-cut evidence <u>pro</u> or <u>con</u>. Even to those of us close to the production and analysis of government statistics the situation is a complex one with, at best, mixed or conflicting information about even the most publicized cases.

As a statistician, with now more than 42 years of continuing contact with government statisticians and statistics, I have, of course, become aware of the statistical "smoke." Specific instances which have occasioned suspicions of undue political pressure include:

1. The sequence of events which led to the premature retirement of the Deputy Director of the Census, the Chief of the Population Division, the Chief of the Construction Statistics Division, and a number of other senior statisticians in that Bureau.

2. The sequence of events which led to termination of the monthly press conference by the Assistant Commissioner of the Bureau of Labor Statistics in re the <u>Monthly Report on the Labor Force</u>, his reassignment and, eventually, his premature retirement.

3. The contemplated reorganization of statistical activities within the Office of Management and Budget, fortunately rescinded, which would have made the old Division of Statistical Standards an arm of that Bureau's management personnel and which did remove the title of Assistant Director of the Office from the head of statistical activities in that Office.

4. The reported political pressures from

the Secretary's office in the Department of Agriculture on the statistical activities within that Department.

5. The conflict within the Department of Health, Education and Welfare in regard to the conduct of a survey on nutritional deficiencies of children in this nation.

6. The order to the Bureau of the Census to stop using "Poverty" in the title of its annual report on the poor in the United States. (This was almost an ingenious and quick way to abolish poverty in this nation.)

7. The reprimand of the Census staff associated with the release of the 1971 poverty report.

8. The cancellation of the Urban Employment Survey during this election year 1972, which since 1969 had been providing labor force and other information about residents in "poverty areas" in much of urban America.

9. The imposition of political clearance procedures for members of statistical as well as other advisory committees. This has initiated a widespread search for Republican statisticians, demographers, and other scientists. (Perhaps there can now be developed new fields of Republican mathematical statistics, Republican demography, Republican physics, Republican medicine and so on.)

10. The placement within the Bureau of the Census of five persons who are not inaccurately described as "political commissars" whose function it was to oversee statistical operations and analyses. In one flagrant situation the Assistant Chief of one of the Census divisions was peremptorily removed from for the convenience of the political functionary who was then provided with amenities not previously afforded assistant division chiefs. (I have the names of the five political functionaries now reduced to two by reason of pressures brought to bear.)

11. The collapse of morale among statisticians in a number of agencies by reason of the "reign of terror" generated by the presence of political functionaries placed at the statistical operating and analytical levels. This, of course, was one reason for the premature retirement of many able career service statisticians.

12. The tendency to delay or withhold statistical reports deemed adverse to the interests of the Administration.

I am sure that many other instances could be added to this list which would further bolster the suspicion that statistics and statisticians are being subjected to unprecedented political pressures. I say unprecedented having in mind the scale of events generating suspicion. On the basis of my own experience with government statistics I know of no administration in which some zealous politician or politically minded press relations "eager beaver" did not, at some point, try to impair the integrity of statistical reports; but never have I witnessed as widespread and insistent efforts to politicize the statistical enterprise.

Furthermore, in the midst of the clamor about possible political contamination of statistics, the President's Commission on Statistics in the Federal Government issued its report without even a reference to the importance of safeguarding the integrity of the Government's statistical output. The Commission included statisticians of unquestionable competence. I do not challenge either their integrity or their competence. But I do think that they can legitimately be faulted for their collective insensitivity to the political issue.

Another piece of possible evidence of protecting the Administration, even if this were not the explicit intent, is to be found in the Commission's position on the proposed quinquennial census of population. Despite the fact that this issue has been before the nation for a century, during which time a quinquennial census has been recommended repeatedly and, most recently, by the Decennial Census Review Committee of the Secretary of Commerce (of this Administration) in the Spring of 1971, the Commission stated that the evidence on behalf of a quinquennial census was neither positive nor negative and suggested that the Bureau of the Census, which has been on record for it for many years, make still another study of the matter to see if one was needed in 1975. The fact that the Administration opposed such an undertaking as an element in its Budget policy may well have had an impact on the Commission--deliberate or unconscious.

In most of the instances cited above there is not as yet clear-cut evidence that the Administration acted only and entirely out of political motives. In placing political appointees in statistical agencies below the level of Presidential appointees the political motive is clear and perhaps can be defended, at best (or worst?), as examples of patronage not unknown on the American scene. Similarly, the insistence on political clearance of members of advisory committees is certainly politically motivated and may similarly be defended as another instance, even if carried to an extreme, of political patronage. There is reason to conclude, however, that these actions constitute a deliberate effort to place into statistical agencies an ideological point of view comparable to the placement of "conservatives" and "strict constructionists" on the Supreme Court. Some Administration supporters have explicitly stated that in their judgment too many "New Dealers" and "liberals" have been frozen into civil service status in the statistical agencies; and that, consequently, this Administration must protect its own ideology and political interests. It would seem reasonable to take the position that the burden of proof for this assertion rests upon this Administration, a burden which certainly it has not yet assumed.

In the other instances listed there is no case in which the Administration has not given reasons, other than political, for the actions taken. For example, in the case of the Deputy Director of the Census it was true that the time schedule on the production and release of 1970 Census information was not kept; and that the Deputy Director was not always responsive to the wishes of his political superiors. Furthermore, it is also true that he was offered a position of the same civil service grade in another agency which he declined. In the case of the reprimands to census staff associated with the 1971 Poverty report it was alleged that release procedures were violated and that some elements of the press were given a priority over others. In the case of the Assistant Commissioner of Labor, it had been previously decided that a civil servant should not be asked to hold a press conference on the employment and unemployment statistics because he would necessarily be subjected to political questions and placed "on the spot." Furthermore, his reassignment was the result of a reorganization of the entire statistical system set in motion before the conflict in the interpretations of the labor force report of the Assistant Commissioner and the Secretary of Labor. The proposed reorganization of the Office of Management and Budget in which the Division of Statistical Standards would be abolished and merged with a management function could be defended in terms of some prevailing concepts of management structure and function. The abolition of the word "poverty" and substitution of "low income" for the term can be defended as insistence on a more neutral designation than that adopted by the previous administration for its political purposes. The abolition of the Urban Employment Survey can be justified by the need to revise the sample of "poverty areas" in the light of the 1970 Census returns.

Examples of pure political pressure uncontaminated by other factors are difficult, if not impossible, to find. Yet the total picture, if not any one specific instance, certainly justifies suspicion and action to assure that the integrity of Federal Statistics is not impaired.

Reaction and Outlook

Fortunately, the statistical and social science fraternity, both within and without government, is endowed with professional elan and political independence. In their professional capacities statisticians can often be browbeaten but never subdued. Fortunately, also, the Government is not a monolithic structure and not all Republicans, nor all Democrats, are subservient to political pressures from their respective political parties. In consequence, it should not be surprising to learn that forces within as well as without this Administration are resisting and are determined to resist any political contamination of the government statistical product. Resistance and reaction to the types of pressures described are mounting even among the ranks of those in high office appointed by this Administration. In general, it may be said that both statistical and non-statistical personnel with professional backgrounds are resisting attacks on the probity of statistics and exerting pressures of their own against their Republican colleagues playing the pro-Administration political and public relations game. Evidence of such resistance is to be found even within the Executive Office of the President, for the Chief Statistician and head of the new Division of Statistical Policy within the Office of Management and Budget, supported at least by the former Director of that Office (now the Secretary of the Treasury) have pursued courses to maintain the probity of Federal statistics. (The Chief Statistician, before reorganization, was an Assistant Director of the Bureau of the Budget and head of the Division of Statistical Standards.)

For example, a significant step has been taken to protect the integrity of government statistics in Circular No. A-91, originally issued in February, 1969 and revised on April 26, 1972, which calls for the "prompt compilation and release of statistical information." This Circular, addressed to "Heads of Executive Departments and Establishments," requires that "the shortest practicable interval should exist between the date or period to which the data refer and the date when compilation is completed." Moreover, it states that "prompt public release of the figures should be made after compilation." The Circular applies to "the principal statistical series...issued by agencies to the public annually or more frequently...and that the publication dates...are made publicly available in advance." In accordance with the objectives of the Circular, release dates for principal economic indicators appear each month in the OMB Statistical Reporter and agencies are required to submit reports to OMB on the release of statistical series.

Although Circular A-91 constitutes a significant step in the right direction, its enforcement is beset with problems, especially in respect to statistical reports that are intermittent and not usually on a rigid time schedule. Obviously, the action taken by OMB is designed to prevent deferring or advancing the release of statistics better to serve political interests. It cannot but help to protect statistics against, at least, one form of political interference.

Perhaps the most significant feature of this Circular is that requiring a "one-hour separation between the issuance of the release by the statistical agency and related commentary." This provision clearly separates the professional statistical product from subsequent "interpretations" which may be politicized. Each administration certainly has the prerogative of interpreting the data as it sees fit at its own risk; but it is important to distinguish such interpretations from the original statistical report.

Another example of efforts to protect the integrity of statistics within the Administration is afforded by the pressures brought to bear by OMB upon the Department of Commerce and the Director of the Census to eliminate three of the five political functionaries placed within the Census Bureau. Of the two that remain one, I understand, achieved civil service status and in this new role, hopefully, this person will serve the Government as a professional rather than as a political overseer.

The Committee on Statistical Policy, which is advisory to the OMB and which is appointed by the OMB in consultation with the ASA, has also been active on behalf of the integrity of statistics. This Advisory Committee comprises mainly past presidents of the American Statistical Association. It exerted pressure upon the OMB to restore the independence of the Division of Statistical Policy, an action that was taken; and it recommended to the Association the creation of a Committee to investigate any possible political interference with statistics or statisticians. Other organizations joined in this recommendation. In consequence, as reported in the American Statistician (Vol. 26, No. 3, June 1972, p. 2) the Board and Council of the Association on January 28th authorized the President to appoint a Committee on Integrity of Federal Statistics to work with a similar committee from the Federal Statistical Users' Conference, which had expressed similar concerns. This Committee, which is holding its first meeting at this Annual Meeting of the Association, includes: Daniel H. Brill, Commercial Credit Corp.; A. Ross Eckler, a former President of the Association and former Director of the Census; Robert S. Schultz, New York State Council of Economic Advisers; and De Ver Sholes, Chicago Association of Commerce and Industry. The Population Association of America and the American Sociological Association have been invited to send observers to the meeting. A Committee of the Conference on Income and Wealth has also asked to be kept informed of the new Committee's work. Other professional societies will undoubtedly follow the work of the Committee with interest and may join in its work.

Another instrumentality that may be expected to exert increasing influence on behalf of statistical integrity is the newly created Committee on National Statistics of the National Academy of Science. This Committee, of which William Kruskal, a member of the now defunct President's Commission on Statistics, is Chairman, has as its terms of reference a continuing review and evaluation of the national statistical output. It is, in effect, a semi-private mechanism for, among other things, monitoring the probity of government statistics.

Pressures in behalf of preserving the integrity of government statistics are to be found in many other significant places; but the ability to resist political interference is not always successful even when such resistance is offered by Bureau heads who, in the last analysis, must serve their political superiors or exercise their option to resign.

Concluding Observations

The American Statistical Association and related interest professional associations can, of course, expect an impartial and competent report from its Committee on the Integrity of Federal Statistics. I am sure that its major work will come after the confusion of this election year and that it will have implications and impact transcending the tenure of this Administration. In the meantime, it is essential that all statisticians be alerted to the possibility of political pressures and make such information as they may acquire available to the Association's Committee.

With the increasing complexity and interdependence of our society requiring increasing government interventionism, statistics are becoming ever more important in providing a basis for policy and action, both in the government and private sectors. In consequence, the temptation to use statistics for administration, agency or other interest, as distinguished from public interest, will increase.

Although the politician may be able adversely to influence statistics in the short run, there can be no doubt that his cause is a hopeless one in the long run. This Association and the other professional associations are aware that just as "eternal vigilance is the price of liberty," similar vigilance is required to defend statistics and statisticians from political contamination. If necessary, the professional fraternity can appoint statistical "truth squads" that can hold their own press conferences to counter political distortions or falsifications, delays or withholdings of the data. In defending the probity and integrity of statistics, statisticians and related professional personnel are not only exercising an important professional and citizenship right and obligation but, also, in the long run they are defending the politician from himself; for nothing could undermine the politician as much as accumulated and intense public distrust and the generation of both a credibility and incredibility gap.

INCOME AS A MEASURE OF WELFARE: THE ISSUES AND A DISSENT

Wm. C. Birdsall, Social Security Administration

When I finished my homework on the topic of income as a measure of welfare, I saw my task as a relatively simple one: itemizing and organizing the technical problems which everyone runs into in building up from income to something like welfare, problems such as price indices and non-market consumption; how to include assets, leisure, and public services in the index; how to take into consideration need, uncertainty, and future expectations. As I proceeded through these problems, I found it impossible to confine my thoughts to those technical questions. The conviction grew in my mind that the important issue is what is this all about: Is the problem essentially a scientific one, similar to comparing temperatures? Is the problem one of discovering the actual norms used by society and government? Or is the problem one of social reform? That is, is it describing the world in such a way as to convince the government or the populace that it ought to be changed, either in general or along particular lines? I think that the basic question is social reform.

As a framework for explaining why I arrived at that conclusion and what it leads me to, I will briefly review some of the technical issues in the economist's progress toward a measure of welfare. My device is to set up a dialogue between two economists who share the conviction that a measure should and can be found. The dialogue I describe is apt to arise in the context of attempting to decide on the appropriate base for taxation on the assumption that somehow welfare of "well-offness" should be the norm. The catalogue of issues is not meant to be exhaustive, but I hope it is representative. The practical solutions that I mention are not universally accepted, but they are at least typical of actual solutions. One economist is a practical man, trying to get the job done. The other is a stickler or a devil's advocate. For my purpose it will be sufficient to limit the dialogue to the question of "equal treatment of equals", or horizontal equity. Extra problems arise in putting families or individuals on some scale to quantify inequality in order to reach the goal of a measure for vertical equity. The dialogue opens:

Pragmatist(tentively): Annual money income seems a reasonable norm for equal treatment.

Stickler: If two people are subject to different sets of prices, they cannot both buy the same set of goods if they have equal money incomes. Furthermore, by a theorem of Samuelson [2], in this circumstance there is no price index which will tell us unambiguously who is better off, even if both people have identical needs and tastes.

Pragmatist: I grant the point. But we must get some measure; let's either accept money income or estimate real income by some agreed upon price index and get on with it.

Stickler: How about such things as home grown food and owner-occupied housing, shouldn't they somehow be included?

Pragmatist (condescendingly): Using known market prices we will impute the value of these types of consumption and add that to income.

Stickler (slyly): How about the fact that some sets of people, such as families, pool their incomes?

Pragmatist: We will simply use these groups as our basic units.

Stickler (pouncing): Then how about the fact that families differ in size, age composition, etc. and thus differ in need?

Pragmatist (triumphantly): I have two possible solutions: We scientifically determine need via diets and build up to needed income from these; thus two families with equal real income are only to be treated as equals if they have equal need; or alternatively we assume that two families that spend the same percentage of their income on, say, food are equals, independent of what their real income is.

[The Stickler has problems with the scientific character of estimates of nutritional needs and he knows that the second solution assumes any differences in consumption tastes between families is due solely to the number of persons in the family; but he lets it go since he has no alternative solution.]

Stickler: Earlier you were happy to impute consumption of home-grown food and housing to income. How about leisure: what if one man earns twice as much per hour and only works half as many hours, loafing the rest of the time; isn't he better off even though both have the same income?

Pragmatist: Based on utility maximization theory we know that leisure is a good valued by each at his or her marginal wage rate; thus I advocate adding the product of leisure hours times the wage rate to income.

[The Stickler knows that this solution assumes that individuals are free to adjust hours-worked as they please. He sees a dilemma: at best the criterion must discriminate either against families who work more hours or against families involuntarily working fewer hours. But again, he has no better solution.]

Stickler (tiring): How about assets and wealth and the fact that they vary by family, shouldn't that be somehow incorporated? In fact, shouldn't our norm for equal treatment take into consideration the welfare potentiality of the family over a much longer span than just a year?

Pragmatist: I agree that in theory we should be looking for a lifetime norm. Practically we must settle for the following: we have techniques based on life expectancies, rates of return, etc. of translating these stocks of assets into a constant annual flow. We just add that to income.

Stickler: How about public housing? Shouldn't that subsidy be added to income?

Pragmatist: Yes, by the same logic as owner-occupied housing.

Stickler (somewhat exhaustedly): How about local public services and the environment? Certainly, you won't consider two families equal that had the same income and paid the same amount in taxes if they differed by the amount of smog or public safety each consumed?

Pragmatist (defensively): I agree with you, but we just haven't yet come up with reasonable coefficients to translate these differences into money terms to add them to or subtract them from income.

[The Stickler is disappointed that science is finally defeated in its quest and he is silent, not raising the additional problem of the uncertainty about income and/or the variability of income, the possible interdependence of welfare beyond the basic family, and so forth.]

CURTAIN

In the light of all these problems and solutions, I ask the question, "What is this all about?" "What are we really trying to do or where are we trying to get?"

One possible answer is that we are trying to measure actual relative well-offness, that is, to find an indicator which will truly tell us when two persons or families are equally well off, like a thermometer will tell us when two persons have the same temperature. I, and I think virtually all economists, admit that this is not what is really being done. We are too aware that welfare depends on many more things than we even dream of taking into consideration.

Is the goal to find the actual norms which society or government uses for equality and inequality of "treatment"? Certainly not; to do this we would simply look to our laws to find the actual norms used. For direct norms we would examine such laws as our tax and welfare provisions; for the indirect norms we would examine the rest of our legislation, and find such norms as acres of cotton planted--and acres of cotton not planted. Important as this task is, it is not what the dialogue is all about.

We are certainly somehow in the realm of ethics, a rule for equality of treatment is the goal. And we seem to be in that realm in a very practical sense, in the social reform sense that we are proposing a norm by which to develop social policy.

If we grant that the essential forum of the dialogue is social reform, what society ought to do, it will probably amaze anyone but an economist that there is little overt ethical discussion in the dialogue. We do not hear the pragmatist say that public housing tenants should be taxed more or given less or that leisure should be taxed.

I think that the reason why the dialogue takes place in something of an ethical no man's land is that the actual course of the discussion is determined by economists' natural desire to take into consideration the actual or potential consumption of all goods; that is, to deal with the general constraint on consumption. Because of the presumed need for a one dimensional measure of equality and inequality, he simply translates this constraint on consumption into a single dimension. To him this is a technical task and his rules do not need examination, except by the science of economics.

I take strong issue with this. My starting point is the earlier-mentioned theorem by Samuelson. It says essentially that even linear constraints are only comparable if the slopes in every dimension are identical. This theorem applies not only to the initial question of the price index; it applies at each decision in the dialogue. Another way of stating the theorem is this: if one insists on translating contraints into a single dimension all important differences between contraints must be reconciled by scientifically arbitrary decisions. At each step in construction unless we agree with what is an essentially ethical decision we must conclude that the index is simply a "compounding of arbitraries" which is certainly itself arbitrary. Thus we must examine the rule or goal governing the pragmatist's decisions.

Especially since the pragmatist tries to translate everything into dollars, his goal may seem to be to find the point at which the family's constraint cuts the money income axis, the point we can call maximum potential money income. Attaining this goal is a near impossibility. We seldom have any data about the shape of the constraint in the vicinity of any axis and especially near the point of maximum potential money income. For example, what on earth is the marginal wage rate at maximum work hours, where "maximum" here means at the limit of one's physical capability? Certainly the pragmatist is not striving for this goal since he evaluates leisure at the actual wage.

I think that two norms govern most of his decisions; one is applied when differences in the constraints faced by various groups arise due to differences in how government already treats people; the other is used for differences due to (apparently free) consumer choice.

Decisions made in the face of present differences in treatment by government seem to be based on the assumption that where treatments differ, cost differences are equivalent to income differences. Except in the case of money transfers this is doubtful. Take public housing for example. To attribute the subsidy to the dweller is to assume a remarkable efficiency on the part of the housing authority, and it assumes that public and non-public housing is otherwise identical. Unless there is a long waiting line for public housing and a low vacancy rate it is difficult to argue that any net subsidy is received by the tenants. Thus this rule seems very arbitrary to me.

The pragmatist decisions in confronting differences due to apparently free consumer choice, such as in the leisure/work hours question, seem to be made in accordance with the goal of making the index neutral with respect to that choice, at least in the vicinity of present equilibrium. That is, if the index were used as a tax base, its existence would not affect your marginal choice between leisure and work. While there are ceteris paribus arguments for neutrality in government treatment rules, I think that the ceteris are seldom paribus. That is, none of us is really indifferent about all the marginal choices made by others, both because we judge various choices as good or bad for the chooser and his or her family and because choices will affect costs, both dollar costs of programs and externality costs. Neutrality should be put in its proper perspective, not as a norm given by God or Moses to be followed in all circumstances, but as a good characteristic of a policy if it is feasible and if society does not care about individual choices in that area.

In summary, I judge any welfare index, the translation of the complexities of the consumption constraint into a single index, as arbitrary; I see such indices as based on questionable assumptions and inadequate acknowledgement or examination of the hard ethical decisions which must be made in building a treatment norm, and more importantly, that must be made in determining good social policy in general.

In rushing to an index I think we rush right by what is most effective for social reform, data about the concrete consumptions constraints that families, especially poor families face. That data, not an index, is the most useful thing for determining the need for social reform, for convincing others of its need, and for discovering solutions. (I do not mean to overstate the givenness of data or the availability of important data. I would, in fact, emphasize the money, the time, and the inventiveness needed to develop good data and to squeeze the meaning out of data.) Especially if our impetus is toward social reform, we should not underestimate the ability of facts to change opinions and to change preferences about social policy based on erroneous opinions. I wish we had a national opinion poll from 1965 to tell us who were thought to be poor by age, sex, race, and working status to compare with the data in "Counting the Poor [1]". Gathering more and better data on income by sources and analyzing it better have been the breakthroughs of the late sixties. I hope that doing the same for consumption of market and non-market goods will be the breakthrough of the seventies. Besides the problems our pragmatist had to deal with, I would include mental illness, disease, malnutrition, and crime as topics of such consumption studies.

While gathering and analyzing the facts are, in my opinion, the most important contributions economists can make to social reform, our contribution must not end there. The task of trying to predict the direct and indirect cost of alternative policies is certainly next in priority. If errors about characteristics of the poor stymied reform in the 60's, our fears about the direct costs of ameliorative policies and about indirect costs due to the recipients' and taxpayers' reactions to these policies may stymie reform in the 70's. Unfortunately, excuses may be able to stay ahead of refutations, but excuses unrefuted never seem to die.

REFERENCES

1. Orshansky, Mollie, "Counting the Poor: Another Look at the Poverty Profile", Social Security Bulletin, January, 1965 (28, No. 1), 3-27.

2. Samuelson, Paul, "Evaluation of Real National Income", Oxford Economic Papers (New Series), Vol. II, No. 1, 1-29.

INCOME INSTABILITY AS A DIMENSION OF WELFARE[a]

Jacob Benus and James Morgan
Institute for Social Research, University of Michigan

Economists have long argued that classifying individuals according to income level in any given year may result in an erroneous representation of economic well-being.[1] That is, since some low-income individuals are temporarily below their "permanent" income level and some high-income individuals are temporarily above theirs, single year income data may misclassify those with income fluctuations. To guard against this, many have favored the use of income averaged over several periods as a measure of "permanent" income level. With a longer accounting period, income level is less likely to reflect temporary vagaries, thereby reducing the likelihood of misclassification.

Income level, even when it is taken over a long accounting period, however, presents only a partial picture of economic well-being. That is, two persons may experience vastly different income patterns yet have the same income level. To capture such differences in income patterns, additional parameters such as trend and instability are necessary. These parameters not only help describe the behavior of income over a period of time, but also serve as additional dimensions of economic welfare, distinct from income level. Clearly, an individual with constant income is in a different welfare position from another individual whose income fluctuates unexpectedly - even if their average incomes are the same. With a stable income, for example, an individual may make long-run plans and commitments with confidence that his income level will continue at a steady rate. On the other hand, an individual who experiences substantial fluctuation in his income is likely to refrain from committing himself to any long-run obligations.

Since it is easier to adjust to unexpected income increases than to unexpected decreases, direction of income change (i.e., trend) is another important dimension of welfare. That is, an individual who experiences sporadic income increases is certainly better-off than another individual who experiences sporadic decreases, even if their level and instability are identical. Since instability measures treat income increases and decreases identically, one cannot distinguish between two such cases using level and instability alone. To fully describe an individual's welfare position, therefore, one must combine three dimensions: level, trend and instability.

In this paper we analyze these three dimensions of economic welfare and examine whether a trade-off exists between level, trend and instability. In addition, we isolate those subgroups in the population who face substantial income instability and examine the relationship of these parameters within each of the subgroups.

The data used in this analysis are from the OEO Panel Study of Income Dynamics.[2] The Panel is currently composed of approximately five thousand households, many of which have been interviewed annually since 1968. However, since we focus on head's labor income we restrict our analysis to those households with the same family head throughout the period, 1968-1971. Furthermore, since some family heads voluntarily entered or left the labor force during the analysis period, we eliminate those who were in the labor force less than 1500 hours (i.e., hours worked plus hours missed due to unemployment and/or illness) during any year. The resulting sample following these restrictions is composed of 2326 individuals.

Measures of Level, Trend, and Instability

The availability of four years of income data leaves several options for measuring level, trend, and instability. The most obvious measure of income level is a simple four-year average of the annual incomes. Since we are measuring income level over a four-year period rather than "permanent" income as of year four, the simple average seems superior to a weighted average of past incomes.

An appropriate measure for income trend is less obvious. One measure, however, which is recommended by its straightforward interpretation is the least-squares slope of the regression of income on "time." If we set the origin at the mid-point of the period (i.e., T = -1.5, -.5, .5, 1.5) the equation for the slope of the time trend is:

$$b = \frac{\sum_{t=1}^{4} Y_t T_t}{\sum_{t=1}^{4} T_t^2} = \frac{1.5Y_4 + .5Y_3 - .5Y_2 - 1.5Y_1}{5} \qquad (1)$$

where Y_t = head's labor income in year t (henceforth we omit the index of summation when it is t=1, ..., 4).

Still less obvious is an appropriate measure for income instability. One possibility is the proportion of the variance in income around the mean unexplained by the time trend:

$$u_1 = \frac{\Sigma(Y-\hat{Y})^2}{\Sigma(Y-\bar{Y})^2} = 1- \frac{\Sigma(\hat{Y}-\bar{Y})^2}{\Sigma(Y-\bar{Y})^2} = 1- \frac{\frac{(\Sigma YT)^2}{\Sigma T^2}}{\Sigma Y^2 - \frac{(\Sigma Y)^2}{N}} \qquad (2)$$

A problem with using u_1 as the instability measure arises when the slope of the trend line equals zero. That is, since $\frac{(\Sigma YT)^2}{\Sigma T^2}$ in equation (2) may also be written as $b^2 \Sigma T^2$, u_1=1 whenever b=0. The problem arises since the slope may equal zero under very different circumstances. For example, an individual with constant income over the analysis period would have a zero trend slope. At the same time another individual whose income varied substantially over the period may also end up with a zero slope. In both cases, equation (2) yields the same instability measure (i.e., u_1=1). Actually this happens very infrequently in practice.

To avoid the problem of assigning the same instability level to all individuals with zero time trends, one may use the following instability measure:

$$u_2 = \frac{\Sigma(Y-\hat{Y})^2}{\Sigma Y^2} = \frac{\Sigma Y^2 + \Sigma(\hat{Y}-\bar{Y})^2}{\Sigma Y^2} =$$

$$1 - \frac{\frac{(\Sigma Y)^2}{N} + \frac{(\Sigma YT)^2}{\Sigma T^2}}{\Sigma Y^2} . \qquad (3)$$

This measure represents the proportion of the total sum of squares around zero unexplained by the regression. Using this measure, an individual with a constant four-year income would have u_2=0, whereas an individual whose income fluctuated but with b=0 would have 0 <u_2 <1.

While u_2 may solve the problem created by zero slopes, it is not necessarily a superior measure. The chief difficulty with u_2 as an instability measure is that it is highly correlated (negatively) with the income level. On the other hand, u_1 is unaffected by the income level since it involves only deviations from the mean. These points may become clearer with an example. Suppose two families with initial incomes of $5,000 and $10,000, respectively, have a $1,000 increase between years 1 and 2 and $500 increases thereafter. Using u_2, the instability level of the $5,000 family exceeds the instability level of the $10,000 family by over three fold; whereas, using u_1, the instability levels are the same for the two families. That is, since four-year income of one family is nearly twice the income of the second family, squaring income in the denominator of u_2 results in a substantially lower instability level for the higher income family. Since the denominator of u_1 involves only deviations from the mean, both families exhibit the same instability.

Another possibility for an instability measure is the coefficient of variation:

$$u_3 = \frac{\sigma}{Y} = \sqrt{\frac{\Sigma(Y-\bar{Y})^2}{T-1}} \Big/ \bar{Y} . \qquad (4)$$

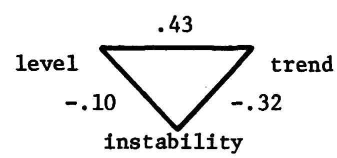
It may be thought of as a measure of relative income variation since it involves a ratio of the standard deviation of income to its mean. Thus, if the standard deviation were one-quarter the level of income, u_3 would equal .25. This measure, like the previous one, is strongly negatively correlated with income level. In computing the coefficient of variation (or any of the other measures of level, trend, and instability for any group, we first calculate the parameter for each individual in the group and then average over the entire group.

One deficiency with each of these instability measures is that they do not differentiate between increases and decreases in income. Since only unexpected decreases are likely to cause difficulties we present in addition to the instability measures, measures which consider only income decreases. Several such measures suggest themselves. If one assumes, for example, that individuals expect income to continue at the previous year's level, then an appropriate measure for unexpected decreases is the relative decline from the previous year's income. On the other hand, it may be argued that a more appropriate measure is the relative decline from previous-peak income since individuals adjust slowly to income decreases. The list of potential measures may be extended ad infinitum if we consider various possible weights for declines with different time lags. To avoid this difficulty we present only the first two measures, i.e., the sum of annual declines from previous-year income and the sum of annual declines from previous-peak income

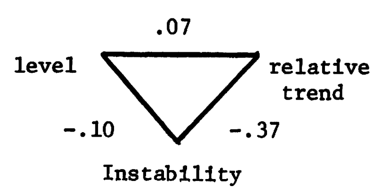
Empirical Results

Since one objective of this paper is to examine the trade-off between level and instability for various subgroups, one important criterion in choosing an instability measure is that it be realtively uncorrelated with income level. Otherwise, the trade-off that we observe may be dominated by the subgroup income level, thereby hiding any behavioral differences that may exist. While each of the proposed instability measures is inversely correlated with income level (see Table 1), the measure least correlated with income level is u_1, the unexplained variation around the trend. In fact, the fraction of the total sum of squares explained by income (i.e., ETA^2) is .020 for u_1 as compared with .084 and .107 for u_2 and u_3, respectively. As a result, we choose u_1 as our instability measure.

The average instability level for our sample of 2326 households is .42, the four-year average of head's labor income and the annual income trend are $9,250 and $734, respectively. The relationship among these parameters may be summarized by the following correlation coefficients for each pair of parameters:



As expected, the correlation between level and instability is negative and relatively weak. The correlation between level and trend, on the other hand, is positive and quite strong. However,since absolute income trend is automatically correlated with income level (since large annual increases in income lead to high income), we replace absolute trend by relative trend (i.e., absolute trend divided by level):



The effect of this substitution is to reduce the correlation between trend and level to .07 from .43. Thus, both the correlation between level and relative trend as well as the correlation be-

Table 1

Measures of Income Instability by Income Level
(same head and in the labor force at least 1500 hours each year, N = 2326)

Head's 4-year Average Labor Income	%	$u_1 = \frac{\Sigma(Y-\hat{Y})^2}{\Sigma(Y-\bar{Y})^2}$	$u_2 = \frac{\Sigma(Y-\hat{Y})^2}{\Sigma Y^2}$	$u_3 = \frac{\sigma}{\bar{Y}}$
Under $2000	1.2	.50	.08	.48
$2000 - 3999	7.9	.52	.04	.28
$4000 - 5999	16.4	.45	.02	.23
$6000 - 7999	21.9	.43	.02	.20
$8000 - 9999	18.4	.39	.01	.17
$10,000 - 11,999	13.2	.37	.01	.16
$12,000 - 14,999	11.6	.40	.01	.16
$15,000 - 19,999	5.7	.34	.01	.19
Over $20,000	3.6	.34	.02	.23
TOTAL OR AVERAGE	100.0	.41	.02	.20
ETA^2		.020	.084	.107

tween level and instability are relatively weak. The relationships, however, are in opposite directions; high (low) income level is associated with high (low) relative trend and low (high) instability level. The third pair of parameters, relative trend and instability, exhibits a relatively strong, negative correlation. Thus the higher the relative trend, the lower the instability level.

In order to examine whether the same relationships among the parameters persist for subgroups with various levels of instability, we employ an AID analysis to isolate groups with different instability levels. The results, presented in Figure 1, indicate that the most important determinant of income stability is occupation. The self-employed and the farmers have the highest level of instability; the white collar occupation, the lowest. For the remainder of the population other important determinants of income instability are size of largest city, race, and education. In highly urbanized areas those with less than a high school education have substantial instability. In less urbanized areas it is the blacks who face substantial instability.[3]

In Table 2 we present for each of the six final subgroups of the AID tree the estimates of the parameters discussed above. The results suggest an inverse relationship between income level and instability across groups; for example, the white collar group (#2) has the highest income level and the lowest instability level. The relationship, however, is not monotonic. The self-employed and farmer group (#5) has the highest instability level, yet its income level is higher than several other subgroups. Thus, we may conclude that low-income subgroups do not necessarily suffer greater instability, but the tendency is in that direction.

The relationship between level and absolute trend, on the other hand, is positive. Those who experience large annual income increases tend to have high income levels. Again this relationship is not monotonic with group 5 (self-employed and farmers), for example, receiving the lowest income increases of any group, but with one of the highest income levels. Substituting relative trend for absolute trend nearly eliminates the positive relationship between trend and level found above. For example, groups 2 and 8 (with the highest and lowest income levels, respectively) have the two highest relative trends.

The relation between trend and instability completes the matrix of inter-relationships. We observe a strong inverse relationship between absolute trend and instability; however, between relative trend and instability the relationship is weak. The latter result is surprising inasmuch as the micro correlation between relative trend and instability is stronger than the correlation between absolute trend and instability. This finding suggests that while some subgroups exhibit both a high relative trend and high instability, the micro correlations within these groups is negative (i.e., high relative trend is associated with low instability). For example, group 8 has a high average relative trend as well as a high average instability level, yet within the group the correlation between relative trend and instability is negative.

Further examination of the correlations within the AID subgroups reveals substantial differences across groups. For example, for group 5 (i.e., self-employed and farmers) the correlation between relative trend and instability is only -.11; for the remaining subgroups the correlation is at least three times as large. We may, therefore, conclude that for most of the sample high relative trend is associated with income stability. For the farmer and self-employed group, however, this association is rather weak.

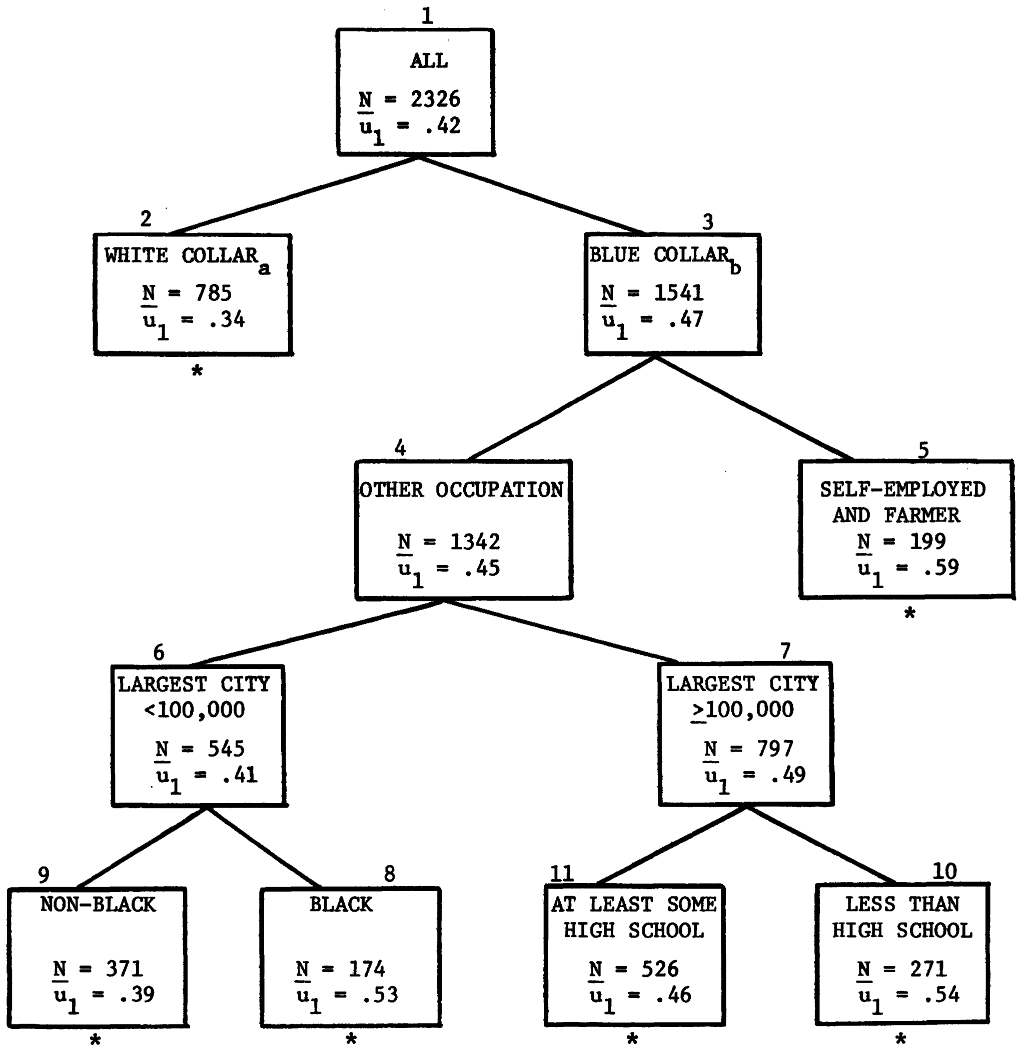
Another interesting result that may be drawn from Table 2 is that, while the correlation between level and relative trend is positive for the sample as a whole, it is negative for several subgroups (i.e, groups 8, 9, and 10). An exam-

Figure 1

AID on Instability (u_1) of Head's Labor Income

(same head and in the labor force at least 1500 hours each year)

Variation explained = 7.1%



[a] White collar group includes: professionals, managers, clerical and sales workers and miscellaneous occupations.

[b] Blue collar group includes: self-employed, craftsmen, operatives, laborers and farmers.

---

ination of the income level reveals that these three groups represent the lowest income groups in the sample. As a result, we suspect that the negative correlations reflect the fact that small absolute changes are likely to represent large relative changes for those with very low incomes. Thus, our choice of relative rather than absolute measure of trend leads to negative correlations for the lowest income groups. Since the results depend so heavily on the measure used, care must be exercised in their interpretations.

A comparison of our instability measure (u_1) with a measure which corresponds closer to economic difficulties (i.e., sum of relative declines in income) results in the same ordering of the six subgroups. The self-employed and farmer groups, for example, have both the highest instability and the highest relative decline from previous-peak income. The latter measure, however, indicates that this group suffered twice the declines of any other subgroup; whereas the instability measure suggests only slightly higher instability for the group. Thus, while the two measures yield a similar ordering of the groups, the relative decline measure may more accurately describe the severity of the economic difficulties.

Conclusion

In this paper we analyzed the relationship among three dimensions of economic welfare: level, trend, and instability. We observed, for example, a negative correlation between level and instability and between relative trend and instability; between level and relative trend, on the other hand, we observed a positive correlation.

However, an examination of these correlations within selected subgroups revealed substantial differences across groups. Since our conclusions depend heavily on the subgroups selected for analysis, as well as the measures chosen to represent the welfare dimensions, further research in this area is required.

Footnotes

[a]This analysis as well as the data collection for the Panel Study of Income Dynamics was funded by the Office of Economic Opportunity.

[1]See Milton Friedman, A Theory of the Consumption Function, Princeton University Press, 1957.

[2]For a complete description of the study see: James Morgan, et. al., A Panel Study of Income Dynamics, Study Design, Procedures, Available Data, 1968-1971. Institute for Social Research, 1971.

[3]Variables which proved unimportant include: age, sex, region, and size of family.

Table 2

Level, Trend, Instability, and Their Correlations for Various Subgroups

			Group Average				Correlation Of:		
roup umber	Description	N	Level	Absolute Trend	Relative Trend	Instability	Level & Relative Trend	Level & Instability	Relative Trend & Instability
2	A. White collar	785	$11,602	$1,086	.092	.34	.02	-.01	-.40
	B. Blue collar								
5	1. Self-employed and farmers	199	7,643	341	.025	.59	.16	-.07	-.11
	2. Other occupations								
	a. Largest city <100,000								
8	1) Black	174	4,450	427	.105	.53	-.15	-.17	-.51
9	2) Non-black	371	7,149	549	.079	.39	-.05	-.10	-.37
	b. Largest city ≥100,000								
10	1) Less than high school	271	6,681	398	.060	.54	-.04	.17	-.48
11	2) Some high school or more	526	8,494	513	.058	.46	.06	-.07	-.39
1	ALL	2326	9,250	734	.074	.42	.07	-.10	-.37

Jotes: Level = 4-year average of head's labor income
Trend = least-squares regression of income on "time."
Instability = proportion of variance unexplained by regression

Table 3

Instability and Sum of Relative Declines for Selected Subgroups

			Sum of Relative Declines	
Group Number	Description	Instability (u_1)	Previous Year	Previous High
2	A. White collar	.34	9.7	12.1
	B. Blue collar			
5	1. Self-employed and farmers	.59	36.4	49.9
	2. Other occupations			
	a. Largest city <100,000			
8	1) Black	.53	20.1	24.5
9	2) Non-black	.39	16.6	18.2
	b. Largest city ≥100,000			
10.	1) Less than high school	.54	20.1	24.6
11	2) Some high school or more	.46	14.2	20.7
1	ALL	.42	15.5	19.8

PREPARING ECONOMIC AND SOCIAL PERSPECTIVES FOR FAMILIES AND UNRELATED INDIVIDUALS FROM THE CURRENT POPULATION SURVEY

Murray S. Weitzman
Social and Economic Statistics Administration, Bureau of the Census

I. Introduction

There is still to this day and I imagine there will be for a long time to come, a not insignificant amount of useful and meaningful income information locked up in the considerable pile of data amassed each year. This material can be released and used profitably in stimulating and furthering income research through examining different functional uses of the data, coupled with innovative formulations of strategies, quantitative applications, and table designs.

In this paper I want to describe and discuss two sets of analytic income tables. One set of tables is now published annually but has not been featured in any analytic presentation and may have gone unnoticed in the plethora of detailed tables that the Bureau of the Census publishes. The other set of tables is part of a larger goal that revolves around casting income and other economic statistics into modules or entities. Both sets of tables can be used directly and effectively in economic and social analyses and also can be integrated into a system of economic and social accounts. These two sets of tables represent efforts to complement and expand the information and analytic potential of the main series of income tables pertaining to families, unrelated individuals, and persons and which are presented in the Bureau's annual P-60 Current Population Report on Consumer Income. The income data in the P-60 report are presented primarily as income size distributions cross-classified by a wide-ranging variety of variables covering differing geographic, demographic, social, and economic elements. For the set of published analytic tables the initial objective was to bring together as many variables as practical that are scattered among different tables in the P-60 report in order to describe the characteristics of families at different income class intervals. Presenting data in this kind of format has been instrumental in germinating the idea of preparing an illustrative set of economic and social account type tables. The current goal is to organize economic and social data around important and sufficiently different groups of persons who display significantly different responses and performances within the context of economic and social conditions, emphasizing at the same time the additive features associated with the data.

II. Published Analytic Tables

The first six detailed tables in the Current Population Reports, "Income in 1969 of Families and Persons in the United States," Series P-60, No. 75, pages 19 to 22, some of which are reproduced here, are designated as published analytic tables. Income in 1969 is chosen to discuss their content because the data available for the experimental tables are for that year.

The first table in the series (table 1 and shown here as table A) sharply delineates the direct relationship on the average between increases in family income in terms of mean family income or per capita income and increases in the number of persons per family, earners per family, and median school years completed by the head.

The ratio of nonearners to earners statistic in table A is an inverse variant of earners per family and the ratio varies inversely with income. The median age of head by income class is a reversed jay-shape curve "ᒐ", with higher older ages, about 62 years of age, at the lower part of the income scale and lower older ages, 50 years of age, beginning at the $25,000 income level. Within a broad band of income, ranging from $7,000 to $14,999, the median age of head approximates 42 years.

The three following tables in the published set of analytic tables (tables 2, 3, and 4) are in the form of cumulative percent distributions of aggregate family income, number of families, and some of the variables used in table 1, plus the number of years of school that the head completed and the number of weeks that the head worked the previous year. Table 2 is reproduced here as table B. Some of the interesting observations that can be gleaned from the three tables are summarized in table C. Table C presents ratios based on the cumulative percent distributions between aggregate family income of specific universes and selected family characteristics within these universes at four income plateaus—$5,000, $10,000, $15,000, and $25,000. The information in the table suggests that where there is some indication within a selected universe of greater response and uniformity of effort of family heads and members of the family to the economic product of the country than the average, the aforementioned ratios are lower and also tend to increase at a slower rate at the different income plateaus (in all cases for the first three plateaus) than the ratios for all families. This observation applies to all earners in families, family head with 4 or more years of college, and family head who worked year round full time. The situation is reversed for families in which the head has no more than an elementary school education or did not work last year.

Estimates of Gini ratios, a summary measure of income concentration, by selected family characteristics tend to support the observation that the greater the homogeneity in family characteristics the more likely that the ratios of percents of aggregate family income to percents of family characteristics will be lower than for family characteristics which exhibit greater heterogenity. As shown in the Bureau of the Census Technical Paper No. 17, "Trends in the Income of Families and Persons in the United States: 1947-1964," table 32, the Gini ratio for families tends to be lowest (less income concentration) for families with heads who worked year round full time the previous year and highest (more income concentration) for families with heads who did not work the previous year. The Gini ratio for 1964 for all families was .352, for families with

Table A.—SUMMARY MEASURES OF FAMILY CHARACTERISTICS, BY TOTAL MONEY INCOME IN 1969

(Social and demographic estimates as of March 1970. They exclude inmates of institutions but include members of the Armed Forces in the United States living off post or with their families on post. Dollar estimates relate to income received in 1969)

Total money income	All families (thousands)	Mean family income (dollars)	Per capita family income (dollars)	Average number of--			Earners per family	Ratio of non-earners to earners 1/	Median age of head	Median school years completed by head 2/
				Persons per family	Related children					
					Per family	Per family with children				
Total	51,237	10,577	2,919	3.62	1.36	2.34	1.69	1.14	45.6	12.2
Under $1,000	804	51	17	3.02	1.30	2.27	0.71	3.27	47.6	9.8
$1,000 to $1,499	670	1,255	434	2.89	1.07	2.37	0.71	3.05	58.2	8.2
$1,500 to $1,999	930	1,750	632	2.77	0.87	2.28	0.66	3.17	62.3	8.3
$2,000 to $2,499	1,231	2,234	805	2.78	0.89	2.28	0.69	3.00	61.8	8.6
$2,500 to $2,999	1,140	2,736	946	2.89	0.92	2.39	0.79	2.66	61.9	8.5
$3,000 to $3,499	1,328	3,224	1,014	3.18	1.16	2.59	1.01	2.15	56.7	8.7
$3,500 to $3,999	1,377	3,739	1,176	3.18	1.15	2.72	1.02	2.11	54.2	8.9
$4,000 to $4,999	2,752	4,475	1,354	3.30	1.24	2.48	1.23	1.69	50.5	9.3
$5,000 to $5,999	3,033	5,457	1,586	3.44	1.33	2.44	1.43	1.41	44.9	10.4
$6,000 to $6,999	3,281	6,436	1,814	3.55	1.41	2.39	1.50	1.36	43.3	11.4
$7,000 to $7,999	3,726	7,453	2,077	3.59	1.42	2.33	1.55	1.32	42.5	12.0
$8,000 to $8,999	3,787	8,443	2,323	3.63	1.46	2.32	1.63	1.23	41.4	12.2
$9,000 to $9,999	3,602	9,447	2,559	3.69	1.45	2.28	1.76	1.10	41.9	12.2
$10,000 to $11,999	6,662	10,876	2,907	3.74	1.49	2.31	1.83	1.04	41.9	12.4
$12,000 to $14,999	7,020	13,280	3,434	3.87	1.49	2.32	2.05	0.89	43.7	12.5
$15,000 to $24,999	8,005	18,284	4,602	3.97	1.38	2.27	2.34	0.70	46.8	12.8
$25,000 to $49,999	1,665	35,786	8,819	4.06	1.25	2.30	2.37	0.71	50.4	15.6
$50,000 and over	224			4.00	1.47	2.39	1.77	1.26	50.0	16.5

1/ The number of all persons without earnings in families divided by the number of persons 14 years old and over with earnings in families.

2/ Restricted to families with head 25 years old and over.

Source: U.S. Bureau of the Census, Current Population Reports, Series P-60, No. 75, "Income in 1969 of Families and Persons in the United States," U.S. Government Printing Office, Washington, D.C., 1970, table 1, page 19.

Table B.—CUMULATIVE PERCENT DISTRIBUTIONS OF SELECTED FAMILY CHARACTERISTICS, BY TOTAL MONEY INCOME IN 1969

(Numbers in thousands)

Total money income	All families	Aggregate family income	Total persons in families 1/	Total related children in families 1/	All earners in families 1/
Total	51,237	$541,934,000	185,396	69,786	86,711
CUMULATIVE PERCENT DISTRIBUTION					
Under $1,000	1.6	(Z)	1.3	1.5	0.7
Under $1,500	2.9	0.2	2.3	2.5	1.2
Under $2,000	4.7	0.5	3.7	3.7	1.9
Under $2,500	7.1	1.0	5.5	5.3	2.9
Under $3,000	9.3	1.5	7.3	6.8	3.9
Under $3,500	11.9	2.3	9.6	9.0	5.5
Under $4,000	14.6	3.3	12.0	11.3	7.1
Under $5,000	20.0	5.6	16.9	16.2	11.0
Under $6,000	25.9	8.6	22.5	22.0	16.0
Under $7,000	32.3	12.5	28.8	28.6	21.7
Under $8,000	39.6	17.7	36.0	36.2	28.3
Under $9,000	47.0	23.6	43.4	44.1	35.4
Under $10,000	54.0	29.9	50.6	51.6	42.7
Under $12,000	67.0	43.2	64.0	65.8	56.8
Under $15,000	80.7	60.5	78.6	80.8	73.4
Under $25,000	96.3	87.5	95.8	96.6	95.0

Z Less than 0.05 percent. 1/ Distributed by income levels of their families.

Source: U.S. Bureau of the Census, Current Population Reports, Series P-60, No. 75, "Income in 1969 of families and Persons in the United States," U.S. Government Printing Office, Washington, D.C., 1970, table 2, page 19.

Table C.—RATIOS AND INDEXES OF THE PERCENTS OF AGGREGATE FAMILY INCOME TO SPECIFIED CHARACTERISTICS ABOVE $5,000, $10,000, $15,000, AND $25,000 INCOME LEVELS

Characteristics of families and sources of data	Percent of characteristics in brackets at $5,000 income level	Ratios of the percents of aggregate family income to specified characteristics above $5,000 income level	Percent of characteristics in brackets at $10,000 income level	Ratios of the percents of aggregate family income to specified characteristics above $10,000 income level	Percent of characteristics in brackets at $15,000 income level	Ratios of the percents of aggregate family income to specified characteristics above $15,000 income level	Percent of characteristics in brackets at $25,000 income level	Ratios of the percents of aggregate family income to specified characteristics above $25,000 income level	Index of ratios (Ratios in table 2 of aggregate family income to all income=100)			
									Above $5,000 income level	Above $10,000 income level	Above $15,000 income level	Above $25,000 income level
Table 2—ALL FAMILIES												
Aggregate family income..............	100.0— (5.6)1/	(X)	100.0—(29.9)	(X)	100.0—(60.5)	(X)	100.0—(87.5)	(X)	(X)	(X)	(X)	(X)
All families........................	100.0—(20.0)1/	1.18:1 1/	100.0—(54.0)	1.52:1	100.0—(80.7)	2.05:1	100.0—(93.3)	3.38:1	100.0	100.0	100.0	100.0
Total persons in families..........	100.0—(16.9)	1.14:1	100.0—(50.6)	1.42:1	100.0—(78.6)	1.85:1	100.0—(95.8)	2.98:1	96.6	93.4	90.2	88.2
Total related children in families.	100.0—(16.2)	1.13:1	100.0—(51.6)	1.45:1	100.0—(80.8)	2.06:1	100.0—(96.6)	3.68:1	95.8	95.4	100.5	108.9
All earners in families............	100.0—(11.0)	1.06:1	100.0—(42.7)	1.22:1	100.0—(73.4)	1.48:1	100.0—(95.0)	2.50:1	89.8	80.3	72.2	74.0
Table 3—HEADS 25 YEARS OLD AND OVER												
Aggregate family income..............	100.0— (5.2)	(X)	100.0—(28.2)	(X)	100.0—(58.9)	(X)	100.0—(86.9)	(X)	(X)	(X)	(X)	(X)
All families........................	100.0—(19.2)	1.17:1	100.0—(52.1)	1.50:1	100.0—(79.5)	2.00:1	100.0—(96.1)	3.36:1	99.2	98.7	97.6	99.4
Elementary school, 8 years or less												
Aggregate family income..............	100.0—(15.1)	(X)	100.0—(51.2)	(X)	100.0—(79.3)	(X)	100.0—(95.9)	(X)	(X)	(X)	(X)	(X)
All families........................	100.0—(38.7)	1.38:1	100.0—(75.3)	1.98:1	100.0—(92.6)	2.80:1	100.0—(99.3)	5.85:1	116.9	130.3	136.6	173.1
High school, 4 years												
Aggregate family income..............	100.0— (3.0)	(X)	100.0—(27.5)	(X)	100.0—(63.3)	(X)	100.0—(91.5)	(X)	(X)	(X)	(X)	(X)
All families........................	100.0—(11.2)	1.09:1	100.0—(46.8)	1.36:1	100.0—(80.1)	1.84:1	100.0—(97.4)	3.27:1	92.4	89.5	89.8	96.7
College, 4 or more years												
Aggregate family income..............	100.0— (0.9)	(X)	100.0— (9.2)	(X)	100.0—(31.3)	(X)	100.0—(68.5)	(X)	(X)	(X)	(X)	(X)
All families........................	100.0— (5.0)	1.04:1	100.0—(22.1)	1.17:1	100.0—(51.9)	1.43:1	100.0—(85.4)	2.16:1	88.1	77.0	69.8	63.9
Table 4—FAMILIES WITH CIVILIAN HEADS												
Head worked year round full time												
Aggregate family income..............	100.0— (2.0)	(X)	100.0—(22.8)	(X)	100.0—(55.3)	(X)	100.0—(85.8)	(X)	(X)	(X)	(X)	(X)
All families........................	100.0— (7.3)	1.06:1	100.0—(40.9)	1.31:1	100.0—(74.3)	1.74:1	100.0—(95.2)	2.96.1	89.8	86.2	84.9	87.6
Head did not work in 1969												
Aggregate family income..............	100.0—(31.3)	(X)	100.0—(65.6)	(X)	100.0—(82.8)	(X)	100.0—(94.7)	(X)	(X)	(X)	(X)	(X)
All families........................	100.0—(62.3)	1.82:1	100.0—(88.3)	2.94:1	100.0—(95.9)	4.20:1	100.0—(99.4)	8.83:1	154.2	193.4	204.9	261.2

1/ Computation of ratio: 100.0—(5.6) = 94.4 and 100.0—(20.0) = 80.0; $\frac{94.4}{80.0}$ = 1.18.

X Not applicable.

Source: Table B in text and published analytic tables 3 and 4 included in Appendix A.

heads who worked year round full time—.301, and .452 for families with heads who did not work in 1964.

The fifth table in the published set combines the cumulative distribution of aggregate family income and the component sources of such income. Aggregate family income is presented as a cumulative percent distribution and the sources of income as a percent distribution by income class intervals. Although there are many interesting relationships that can be drawn from this table as it now stands, the task of interpretation can benefit from some calculations to obtain the absolute dollar estimates on which the table is based. These estimates are shown here in table D to assist people who may find table 5 in the published report of some interest. This table points up the very substantial role of the Government in contributing to money income for families with incomes under $5,000 in 1969 and its rapidly diminishing role in the higher income intervals as the earnings component of income becomes increasingly more important.

The last of the six tables in the published series supplements the previous tables by showing specified characteristics of families as a percent of all families, by total money income.

III. Economic and Social Perspective Tables

The income and related 1969 data from the March 1970 Current Population Survey have been rearranged to produce a set of tables that in my view have the attributes required for the establishment of a network of annual economic and social accounts linking the activities and contributions of significant and identifiable groups of persons to the economy and to the economic position of the family. The tables try to systematically describe the commitment and efforts put forth by heads of different groups of families, unrelated individuals, and persons in producing the economy's output in terms of current labor force status and for such earnings and income-generating factors as weeks worked last year, worked full time or part time last year, education, and age. This information for heads, for example, is linked to the numbers of persons in the specific family group and aggregate family income. From these relationships, comparisons are derived indicating the economic contributions of the family head, average family size, per capita income, and other measures.

There is an extensive variety of possible combinations of family, unrelated individual, and persons groups for which individual tables can be prepared and summed to larger aggregates since the base file for this experimental work has been tabulated by two race breaks (Negro, and white and other), three income breaks, and the aforementioned variables. The package of experimental tables, which is available by writing to the author, is intended to illustrate the potential of this type of presentation for economic and social accounts. Included in the package are all families, families headed by females, and two sets of tables for male heads, one emphasizing weeks worked and the other worked full time last year, education, and age by three family income breaks—less than the low economy standard budget, more than the low economy standard budget but less than $10,000, and more than $10,000.

Table D.--AGGREGATE MONEY INCOME OF FAMILIES BY TYPE OF INCOME, BY TOTAL MONEY INCOME IN 1969
(In millions of dollars)

Total money income	Total aggregate income	Aggregate family money income									
		Earnings				Income other than earnings					
		Total	Wage or salary income	Nonfarm self employment income	Farm self-employment income	Total	Social Security and Government Railroad Retirement	Dividends, interest, net rental income, income from estates or trusts	Public Assistance and Welfare payments	Unemployment and workmen's compensation, Government employee pensions, and Veterans' payments	Private pensions, annuities, alimony, royalties, etc.
Total 1/..	542,174	483,296	434,295	40,916	8,085	58,879	16,333	20,688	3,765	9,100	8,993
CUMULATIVE DISTRIBUTION											
Under $1,000..	370	176	176	—	—	194	92	21	49	5	27
Under $1,500..	1,210	469	402	33	34	741	382	53	208	30	68
Under $2,000..	2,836	931	799	52	80	1,905	1,058	103	517	88	139
Under $2,500..	5,582	1,864	1,556	128	180	3,718	2,056	192	939	235	296
Under $3,000..	8,700	2,994	2,514	186	294	5,706	3,130	351	1,378	416	431
Under $3,500..	12,983	5,172	4,306	424	442	7,811	4,200	524	1,798	619	670
Under $4,000..	18,134	7,943	6,776	577	590	10,191	5,365	749	2,140	903	1,034
Under $5,000..	30,448	16,026	13,837	1,134	1,055	14,422	7,310	1,260	2,629	1,507	1,716
Under $6,000..	47,004	28,934	25,303	2,087	1,544	18,070	8,908	1,817	2,863	2,182	2,300
Under $7,000..	68,119	46,810	41,591	3,216	2,003	21,309	10,064	2,457	3,109	2,703	2,976
Under $8,000..	95,891	71,343	64,299	4,581	2,463	24,548	11,107	3,257	3,264	3,388	3,532
Under $9,000..	127,863	100,426	91,394	5,993	3,039	27,437	11,894	4,013	3,380	4,017	4,133
Under $10,000.	161,892	131,756	121,023	7,267	3,466	30,136	12,663	4,768	3,475	4,584	4,646
Under $12,000.	234,339	199,828	184,490	10,872	4,466	34,511	13,667	6,164	3,547	5,704	5,429
Under $15,000.	327,561	287,416	266,502	15,340	5,594	40,145	14,793	8,503	3,643	6,938	6,268
Under $25,000.	473,916	424,275	392,151	25,041	7,083	49,641	16,005	13,656	3,745	8,587	7,648

— Represents zero.

1/ Excludes net losses, therefore the data shown here are not strictly comparable with the aggregate income shown in table 2 which includes net losses. Moreover aggregate negative amounts in nonfarm and farm self-employment income were changed to zero values in all computations.

Source: U.S. Bureau of the Census, Current Population Reports, Series P-60, No. 75, "Income in 1969 of Families and Persons in the United States," U.S. Government Printing Office, Washington, D.C., 1970, based on table 5, page 21.

Family income below the low economy standard budget is based on the 1969 family income cutoffs established for use in connection with the Bureau of the Census statistics on the characteristics of the low income population. Family income below the low economy standard budget is a variable measure depending primarily upon size of family, presence and number of family members under 18 years of age, sex of head, and farm and nonfarm residence and a nutritionally adequate food plan that could be implemented under temporary or emergency conditions when funds are low. References to the work in this area of the Department of Agriculture, of Mollie Orshansky of the Social Security Administration, and a summary of the 1969 income cutoffs are presented in the Current Population Report on Consumer Income, P-60, No. 76, pages 17 to 20.

The estimates of average income and earnings of persons 14 years old and over in the economic and social perspective tables relate to all persons in a given group irrespective of whether they have income or earnings from a particular income source. This procedure differs from the one used for the persons tables that appear in the P-60 Consumer Income Reports. In these reports income distributions, and median and mean income estimates are restricted to persons with income or particular types of earnings. Both procedures are acceptable considering the respective objectives of the two kinds of tables. In presenting income size distributions by type of income, it is inappropriate, for example, to include persons who are proprietors and receive zero wages and salaries with persons who do receive wages and salaries. In contrast the all persons concept is used in the economic and social perspective tables because of the reliance on a modular or entity approach in which the different attributes and performances of diverse population groups are summed to totals.

Tables E and E' illustrate the presentation of the economic and social perspective for all families with male heads utilizing the stub that emphasizes worked full time, education, and age.

At this methodological stage of the project in which the emphasis is on the skeletal structure of the system, the table presentation has been kept relatively unencumbered. The modular approach to the tables, however, permits a great deal of flexibility in adding or subtracting modules. For example, the column on number of persons in families can be decomposed into different family relationships and sex and age characteristics. Similarly, the income data detail can be expanded by different sources of income. The occupational variable has not been tabulated in the present file used in this work in an effort to keep the volume of printouts from ballooning to unmanageable proportions. Also the limited size of the CPS sample, 50,000 households, would result in a still smaller number of observations in a cell than the file now contains. Of course, it is possible to collapse the detail for other than occupation variables, to obtain more observations per cell. The 1970 Census of Population and Housing, through its associated package of basic tapes and the various sample tapes makes available many if not all of the variables used in the CPS and thus represents an important data source for the preparation of tables for national and less than national universes.

Table E.—ECONOMIC AND SOCIAL PERSPECTIVE OF FAMILIES WITH MALE HEADS WITH SPECIAL EMPHASIS ON MALE HEADS WHO WORKED FULL TIME LAST YEAR: MARCH 1970

NUMBER OF PERSONS IN FAMILIES, AGGREGATE FAMILY INCOME, AND EARNINGS OF HEADS BY SELECTED LABOR FORCE CHARACTERISTICS, AGE, AND EDUCATION

(Social and demographic estimates as of March 1970. They exclude inmates of institutions but include members of the Armed Forces in the United States living off post or with their families on post. Dollar estimates relate to income received in 1969. Numbers may not add to totals because of rounding)

Characteristics of male family heads	Number of families (in thousands)	Number of persons in families (in thousands)	Aggregate family income (in millions of dollars)	Aggregate income of male family heads (in millions of dollars)	Aggregate earnings of male family heads (in millions of dollars)	Percent distribution: Number of families	Number of persons in families	Aggregate family income	Aggregate income of male family heads	Aggregate earnings of male family heads
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Total population	45,658.5	167,100	508,910	398,700	360,600	100.0	100.0	100.0	100.0	100.0
Civilian population	44,647.5	163,369	500,469	391,710	353,741	97.8	97.8	98.3	98.2	98.1
Civilian labor force	38,516.4	147,200	462,775	367,600	348,300	84.4	88.1	90.9	92.2	96.6
Employed	37,511.8	143,400	453,743	360,900	342,300	82.2	85.8	89.1	90.5	94.9
Worked full time last year	35,814.1	138,600	441,480	352,900	337,400	78.4	82.9	86.7	88.5	93.6
Under 25 years	2,256.0	6,625	17,996	13,820	13,620	4.9	4.0	3.5	3.4	3.8
Elementary—less than 8 years	70.0	252	430	340	338	0.2	0.2	0.1	0.1	0.1
Elementary—8 years	104.8	347	658	539	533	0.2	0.2	0.1	0.1	0.1
High school—1 to 3 years	454.8	1,415	3,195	2,601	2,570	1.0	0.8	0.6	0.7	0.7
High school—4 years	1,121.3	3,286	9,235	7,148	7,080	2.5	2.0	1.8	1.8	2.0
College—1 to 3 years	361.0	981	3,095	2,231	2,161	0.8	0.6	0.6	0.6	0.6
College—4 or more years	144.2	343	1,383	956	941	0.3	0.2	0.3	0.2	0.3
College—4 years	119.4	280	1,185	810	799	0.3	0.2	0.2	0.2	0.2
College—5 or more years	24.8	63	199	146	142	0.1	(Z)	(Z)	(Z)	(Z)
25 to 44 years	17,762.1	78,453	211,405	177,300	172,600	38.9	46.9	41.5	44.5	47.9
Elementary—less than 8 years	1,106.5	5,737	8,889	6,881	6,744	2.4	3.4	1.7	1.7	1.9
Elementary—8 years	1,205.9	5,695	11,269	8,952	8,768	2.6	3.4	2.2	2.2	2.4
High school—1 to 3 years	2,853.7	13,190	28,751	23,080	22,540	6.3	7.9	5.6	5.8	6.3
High school—4 years	6,875.1	29,908	77,317	64,470	63,180	15.1	17.9	15.2	16.2	17.5
College—1 to 3 years	2,462.4	10,269	31,691	26,600	25,940	5.4	6.1	6.2	6.7	7.2
College—4 or more years	3,258.4	13,653	53,483	47,300	45,370	7.1	8.2	10.5	11.9	12.6
College—4 years	1,799.6	7,546	27,991	24,640	23,630	3.9	4.5	5.5	6.2	6.6
College—5 or more years	1,458.9	6,107	25,494	22,660	21,740	3.2	3.7	5.0	5.7	6.0
45 to 64 years	14,804.8	51,108	201,271	153,300	144,600	32.4	30.6	39.5	38.4	40.1
Elementary—less than 8 years	1,786.8	6,667	16,062	11,000	10,600	3.9	4.0	3.2	2.8	2.9
Elementary—8 years	2,179.3	7,175	22,748	16,770	16,060	4.8	4.3	4.5	4.2	4.5
High school—1 to 3 years	2,656.4	8,821	31,917	23,180	22,090	5.8	5.3	6.3	5.8	6.1
High school—4 years	4,678.5	15,826	63,487	47,640	45,020	10.2	9.5	12.5	11.9	12.5
College—1 to 3 years	1,599.2	5,486	26,176	20,620	19,090	3.5	3.3	5.1	5.2	5.3
College—4 or more years	1,904.5	7,133	40,874	34,130	31,710	4.2	4.3	8.0	8.6	8.8
College—4 years	1,046.8	3,946	21,026	17,560	16,280	2.3	2.4	4.1	4.4	4.5
College—5 or more years	857.7	3,187	19,848	16,570	15,430	1.9	1.9	3.9	4.2	4.3
65 years old and over	991.2	2,375	10,802	8,429	6,594	2.2	1.4	2.1	2.1	1.8
Elementary—less than 8 years	215.4	559	1,511	1,066	873	0.5	0.3	0.3	0.3	0.2
Elementary—8 years	228.3	537	1,970	1,510	1,210	0.5	0.3	0.4	0.4	0.3
High school—1 to 3 years	150.6	353	1,394	1,076	868	0.3	0.2	0.3	0.3	0.2
High school—4 years	164.9	403	1,819	1,379	1,088	0.4	0.2	0.4	0.3	0.3
College—1 to 3 years	72.6	156	995	844	677	0.2	0.1	0.2	0.2	0.2
College—4 or more years	159.4	367	3,114	2,553	1,878	0.3	0.2	0.6	0.6	0.5
College—4 years	80.5	193	1,307	953	697	0.2	0.1	0.3	0.2	0.2
College—5 or more years	78.9	174	1,808	1,601	1,181	0.2	0.1	0.4	0.4	0.3
Worked part time last year	1,465.2	4,173	11,070	7,357	4,634	3.2	2.5	2.2	1.8	1.3
Did not work last year	232.5	679	1,201	690	254	0.5	0.4	0.2	0.2	0.1
Unemployed	1,004.6	3,760	9,029	6,649	5,997	2.2	2.3	1.8	1.7	1.7
Seeking full-time employment	936.5	3,564	8,612	6,394	5,832	2.1	2.1	1.7	1.6	1.6
Worked full-time last year	873.8	3,361	8,145	6,160	5,688	1.9	2.0	1.6	1.5	1.6
Worked part time last year	26.9	94	144	63	35	0.1	0.1	(Z)	(Z)	(Z)
Did not work last year	35.8	109	323	171	109	0.1	0.1	0.1	(Z)	(Z)
Seeking part-time employment	68.1	197	417	255	164	0.1	0.1	0.1	0.1	(Z)
Not in civilian labor force	6,131.1	16,169	37,694	24,110	5,441	13.4	9.7	7.4	6.0	1.5
Worked full time last year	1,012.0	2,933	9,032	6,422	4,691	2.2	1.8	1.8	1.6	1.3
Under 25 years	82.4	220	393	228	190	0.2	0.1	0.1	0.1	0.1
25 to 44 years	175.2	730	1,501	1,020	721	0.4	0.4	0.3	0.3	0.2
45 to 64 years	407.1	1,180	4,039	2,836	2,278	0.9	0.7	0.8	0.7	0.6
65 years old and over	347.2	803	3,098	2,338	1,503	0.8	0.5	0.6	0.6	0.4
Worked part time last year	555.1	1,535	3,109	1,992	665	1.2	0.9	0.6	0.5	0.2
Under 25 years	29.2	73	132	59	38	0.1	(Z)	(Z)	(Z)	(Z)
25 to 44 years	35.6	178	200	117	84	0.1	0.1	(Z)	(Z)	(Z)
45 to 64 years	113.6	395	706	358	130	0.2	0.2	0.1	0.1	(Z)
65 years old and over	376.7	888	2,071	1,458	413	0.8	0.5	0.4	0.4	0.1
Did not work last year	4,564.1	11,702	25,554	15,690	85	10.0	7.0	5.0	3.9	(Z)
Under 25 years	44.2	103	211	76	20	0.1	0.1	(Z)	(Z)	(Z)
25 to 44 years	141.3	670	787	402	8	0.3	0.4	0.2	0.1	(Z)
45 to 64 years	930.6	2,863	5,955	3,231	47	2.0	1.7	1.2	0.8	(Z)
65 years old and over	3,448.1	8,066	18,602	11,990	9	7.6	4.8	3.7	3.0	(Z)
In Armed Forces	1,011.0	3,731	8,441	6,990	6,859	2.2	2.2	1.7	1.8	1.9

Z Less than 0.05.

Note: Summary measure statistics from the Current Population Survey are usually not shown for a population base of less than 75,000 because of low reliability of the estimates. Such statistics are shown here primarily as components contributing to the aggregate estimates and generally should not be employed for between-cell comparisons. This note applies to all tables in the Economic and Social Perspective series.

Estimates for the Armed Forces were obtained by subtracting estimates for the civilian population from the total population. The relatively small size of the Armed Forces in comparison with the total population, the rounding procedures used to conserve space in the tabulation printout (estimates in the hundreds of billions of dollars were rounded to the nearest hundred million) and the residual procedure used to estimate the parameters for the Armed Forces may result in some substantial errors for some of the Armed Forces estimates involving income and earnings.

Source: Current Population Survey, special tabulation.

Table E'.--ECONOMIC AND SOCIAL PERSPECTIVE OF FAMILIES WITH MALE HEADS WITH SPECIAL EMPHASIS ON MALE HEADS WHO WORKED FULL TIME LAST YEAR: MARCH 1970

DERIVED AVERAGE ECONOMIC AND DEMOGRAPHIC MEASURES AND RELATIONSHIPS FROM AGGREGATE DATA

(Social and demographic estimates as of March 1970. They exclude inmates of institutions but include members of the Armed Forces in the United States living off post or with their families on post. Dollar estimates relate to income received in 1969. Numbers may not add to totals because of rounding.

Characteristics of male family heads	Average family income (in dollars)	Average income of male family heads (in dollars)	Average earnings of male family heads (in dollars)				Income of male family head as a percent of family income	Average family size	Per capita family income (in dollars)	Per capita family income excluding average earnings of male family heads	
			Total	Wages and salaries	Nonfarm self-employment	Farm self-employment				Dollars	Percent of per capita family income
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Total population	11,146	8,732	7,898	6,936	803	159	78.3	3.66	3,045	887	29.1
Civilian population	11,209	8,773	7,923	6,939	821	163	78.3	3.66	3,063	898	29.3
Civilian labor force	12,015	9,543	9,043	7,924	939	180	79.4	3.82	3,145	778	24.7
Employed	12,096	9,621	9,125	7,983	958	184	79.5	3.82	3,166	778	24.6
Worked full time last year	12,327	9,853	9,421	8,276	959	186	79.9	3.87	3,185	751	23.6
Under 25 years	7,977	6,124	6,039	5,884	106	49	76.8	2.94	2,713	659	24.3
Elementary--less than 8 years	6,139	4,854	4,833	4,725	84	24	79.1	3.61	1,701	356	20.9
Elementary--8 years	6,282	5,149	5,088	5,103	—	-15	82.0	3.31	1,898	361	19.0
High school--1 to 3 years	7,024	5,721	5,651	5,569	72	10	81.4	3.11	2,259	441	19.5
High school--4 years	8,236	6,375	6,314	6,101	149	64	77.4	2.93	2,811	656	23.3
College--1 to 3 years	8,573	6,180	5,987	5,885	74	28	72.1	2.72	3,152	951	30.2
College--4 or more years	9,594	6,632	6,528	6,318	49	161	69.1	2.38	4,031	1,288	32.0
College--4 years	9,921	6,783	6,692	6,438	59	195	68.4	2.35	4,222	1,374	32.5
College--5 or more years	8,016	5,903	5,736	5,736	—	—	73.6	2.53	3,168	901	28.4
25 to 44 years	11,902	9,981	9,720	8,700	883	136	83.9	4.42	2,693	494	18.3
Elementary--less than 8 years	8,033	6,218	6,095	5,621	377	97	77.4	5.18	1,551	374	24.1
Elementary--8 years	9,345	7,424	7,271	6,280	657	334	79.4	4.72	1,980	439	22.2
High school--1 to 3 years	10,075	8,087	7,933	7,315	500	117	80.3	4.62	2,181	464	21.3
High school--4 years	11,246	9,377	9,189	8,397	606	186	83.4	4.35	2,585	473	18.3
College--1 to 3 years	12,870	10,802	10,536	9,714	746	75	83.9	4.17	3,086	496	16.1
College--4 or more years	16,414	14,516	13,924	11,727	2,163	34	88.4	4.19	3,917	594	15.2
College--4 years	15,554	13,691	13,132	12,342	731	59	88.0	4.19	3,712	578	15.6
College--5 or more years	17,475	15,534	14,900	10,968	3,930	2	88.9	4.19	4,171	615	14.7
45 to 64 years	13,595	10,357	9,765	8,353	1,156	256	76.2	3.45	3,941	1,110	28.2
Elementary--less than 8 years	8,989	6,155	5,932	5,262	438	232	68.5	3.73	2,410	820	34.0
Elementary--8 years	10,438	7,693	7,368	6,336	515	516	73.7	3.29	3,173	933	29.4
High school--1 to 3 years	12,015	8,726	8,316	7,390	694	232	72.6	3.32	3,619	1,114	30.8
High school--4 years	13,570	10,183	9,622	8,450	920	252	75.0	3.38	4,015	1,168	29.1
College--1 to 3 years	16,368	12,896	11,937	10,285	1,427	225	78.8	3.43	4,772	1,292	27.1
College--4 or more years	21,462	17,921	16,650	13,041	3,561	48	83.5	3.75	5,723	1,283	22.4
College--4 years	20,086	16,777	15,550	13,390	2,080	81	83.5	3.77	5,328	1,203	22.6
College--5 or more years	23,141	19,318	17,991	12,616	5,368	7	83.5	3.72	6,221	1,384	22.2
65 years old and over	10,898	8,504	6,653	4,984	1,324	345	78.0	2.40	4,541	1,769	39.0
Elementary--less than 8 years	7,016	4,952	4,056	2,915	882	258	70.6	2.60	2,698	1,138	42.2
Elementary--8 years	8,627	6,613	5,300	4,174	576	550	76.7	2.35	3,671	1,416	38.6
High school--1 to 3 years	9,254	7,146	5,761	4,794	622	345	77.2	2.35	3,938	1,486	37.7
High school--4 years	11,029	8,364	6,600	5,289	981	330	75.8	2.44	4,520	1,815	40.2
College--1 to 3 years	13,706	11,624	9,326	6,731	2,354	241	84.8	2.14	6,405	2,047	32.0
College--4 or more years	19,537	16,021	11,781	8,009	3,539	233	82.0	2.30	8,494	3,372	39.7
College--4 years	16,234	11,831	8,655	6,648	1,743	264	72.9	2.40	6,764	3,158	46.7
College--5 or more years	22,910	20,299	14,974	9,399	5,374	201	88.6	2.21	10,367	3,591	34.6
Worked part time last year	7,555	5,021	3,163	1,906	1,080	176	66.5	2.85	2,651	1,541	58.1
Did not work last year	5,165	2,967	1,092	1,077	2	13	57.4	2.92	1,769	1,395	78.9
Unemployed	8,988	6,619	5,969	5,739	216	14	73.6	3.74	2,403	807	33.6
Seeking full-time employment	9,196	6,828	6,228	5,984	230	14	74.2	3.81	2,414	779	32.3
Worked full time last year	9,321	7,049	6,509	6,256	238	15	75.6	3.85	2,421	730	30.2
Worked part time last year	5,355	2,345	1,292	1,033	259	—	43.8	3.51	1,526	1,158	75.9
Did not work last year	9,024	4,788	3,054	3,054	—	—	53.1	3.04	2,968	1,964	66.2
Seeking part-time employment	6,129	3,741	2,413	2,375	33	5	61.0	2.89	2,121	1,286	60.6
Not in civilian labor force	6,148	3,932	887	755	79	53	64.0	2.64	2,329	1,993	85.6
Worked full time last year	8,925	6,346	4,636	4,045	369	222	71.1	2.90	3,078	1,479	48.1
Under 25 years	4,774	2,761	2,300	2,251	49	—	57.8	2.67	1,788	927	51.8
25 to 44 years	8,568	5,821	4,112	3,866	206	40	67.9	4.16	2,060	1,071	52.0
45 to 64 years	9,921	6,966	5,595	4,788	535	272	70.2	2.90	3,421	1,492	43.6
65 years old and over	8,922	6,734	4,329	3,690	331	308	75.5	2.31	3,862	1,988	51.5
Worked part time last year	5,600	3,589	1,198	835	203	160	64.1	2.76	2,029	1,595	78.6
Under 25 years	4,525	2,027	1,295	1,270	25	—	44.8	2.50	1,810	1,292	71.4
25 to 44 years	5,613	3,285	2,372	1,984	369	19	58.5	5.00	1,123	648	57.7
45 to 64 years	6,212	3,153	1,148	833	301	14	50.8	3.48	1,785	1,455	81.5
65 years and over	5,498	3,870	1,095	694	172	229	70.4	2.36	2,330	1,866	80.1
Did not work last year	5,599	3,439	19	16	—	2	61.4	2.56	2,187	2,180	99.7
Under 25 years	4,765	1,730	457	457	—	—	36.3	2.33	2,045	1,849	90.4
25 to 44 years	5,571	2,844	59	57	1	—	51.1	4.74	1,175	1,163	99.0
45 to 64 years	6,399	3,472	50	48	—	2	54.3	3.08	2,078	2,061	99.2
65 years and over	5,395	3,476	3	—	—	2	64.4	2.34	2,306	2,304	99.9
In Armed Forces	8,349	6,914	6,784	6,795	3	5	82.8	3.69	2,262	424	18.7

— Represents zero.
Column 7: Column 2 divided by column 1 times 100.
Column 9: Column 1 divided by column 8.
Column 10: Column 1 less column 3 divided by column 8.
Column 11: Column 10 divided by column 9 times 100.

Note: Summary measure statistics from the Current Population Survey are usually not shown for a population base of less than 75,000 because of low reliability of the estimates. Such statistics are shown here primarily as components contributing to the aggregate estimates and generally should not be employed for between-cell comparisons. This note applies to all tables in the Economic and Social Perspective series.

Estimates of the Armed Forces were obtained by subtracting estimates for the civilian population from the total population. The relatively small size of the Armed Forces in comparison with the total population, the rounding procedures used to conserve space in the tabulation printout (estimates in the hundreds of billions of dollars were rounded to the nearest hundred million) and the residual procedure used to estimate the parameters for the Armed Forces may result in some substantial errors for some of the Armed Forces estimates involving income and earnings.

Source: Current Population Survey, special tabulation.

There are innumerable patterns and relationships that emerge readily from the economic and social perspective tables. The discussion here will concentrate primarily on some of the differences in group performance that show up among the three income categories for male family heads. Some of the highlights are summarized in tables F and G.

Table F.--SELECTED STATISTICS FOR FAMILIES WITH MALE HEADS BY THREE FAMILY INCOME GROUPS: MARCH 1970

(Social and demographic estimates as of March 1970. They exclude inmates of institutions. Dollar estimates relate to income received in 1969. Percents may not add to totals because of rounding)

Characteristics of family and head	Family income Below Low Economy Standard Budget			Family income more than Low Economy Standard Budget but less than $10,000			Income more than $10,000		
	Heads (in percent)	Average income of head (in dollars)	Income of male head as a percent of family income or average family size	Heads (in percent)	Average income of head (in dollars)	Income of male head as a percent of family income or average family size	Heads (in percent)	Average income of head (in dollars)	Income of male head as a percent of family income or average family size
All heads in civilian population (in thousands)	3,084.9	(X)	(X)	19,135.3	(X)	(X)	22,427.3	(X)	(X)
CIVILIAN POPULATION	100.0	1,740	(X)	100.0	5,553	(X)	100.0	12,486	(X)
Income of male heads as a percent of family income	(X)	(X)	78.7	(X)	(X)	82.0	(X)	(X)	76.9
Average family size	X)	(X)	3.94	(X)	(X)	3.32	(X)	(X)	3.90
CIVILIAN LABOR FORCE	56.9	1,969	(X)	79.7	6,085	(X)	95.9	12,613	(X)
Income of M/H as a % of family income	(X)	(X)	80.7	(X)	(X)	84.8	(X)	(X)	77.7
Average family size	(X)	(X)	4.70	(X)	(X)	3.56	(X)	(X)	3.94
EMPLOYED--WORKED FULL TIME LAST YEAR	42.1	2,086	(X)	71.5	6,305	(X)	92.9	12,665	(X)
Income of M/H as a % of family income	(X)	(X)	82.2	(X)	(X)	86.2	(X)	(X)	78.0
Average family size	(X)	(X)	4.92	(X)	(X)	3.65	(X)	(X)	3.95
Under 25 years	4.0	1,867	(X)	8.2	5,564	(X)	2.5	8,599	(X)
Income of M/H as a % of family income	(X)	(X)	91.1	(X)	(X)	81.8	(X)	(X)	68.8
Average family size	(X)	(X)	3.40	(X)	(X)	2.96	(X)	(X)	2.77
25 to 44 years	21.4	2,597	(X)	35.1	6,755	(X)	46.3	12,535	(X)
Income of M/H as a % of family income	(X)	(X)	85.5	(X)	(X)	89.0	(X)	(X)	82.2
Average family size	(X)	(X)	5.73	(X)	(X)	4.31	(X)	(X)	4.40
Elementary--less than 8 years	6.7	2,838	(X)	3.1	5,727	(X)	1.4	9,478	(X)
Income of M/H as a % of family income	(X)	(X)	86.6	(X)	(X)	83.3	(X)	(X)	70.0
Average family size	(X)	(X)	6.30	(X)	(X)	4.84	(X)	(X)	5.11
Elementary--8 years	3.2	2,639	(X)	3.4	6,370	(X)	2.0	9,974	(X)
Income of M/H as a % of family income	(X)	(X)	84.6	(X)	(X)	87.3	(X)	(X)	73.1
Average family size	(X)	(X)	5.81	(X)	(X)	4.56	(X)	(X)	4.71
High school--1 to 3 years	3.9	2,584	(X)	7.5	6,463	(X)	5.8	10,376	(X)
Income of M/H as a % of family income	(X)	(X)	85.6	(X)	(X)	87.4	(X)	(X)	76.0
Average family size	(X)	(X)	5.71	(X)	(X)	4.47	(X)	(X)	4.68
High school--4 years	5.3	2,532	(X)	14.5	7,010	(X)	17.6	11,322	(X)
Income of M/H as a % of family income	(X)	(X)	87.6	(X)	(X)	90.2	(X)	(X)	80.7
Average family size	(X)	(X)	5.45	(X)	(X)	4.24	(X)	(X)	4.38
College--1 to 3 years	1.4	2,211	(X)	3.8	7,134	(X)	7.6	12,578	(X)
Income of M/H as a % of family income	(X)	(X)	73.9	(X)	(X)	91.4	(X)	(X)	82.4
Average family size	(X)	(X)	5.32	(X)	(X)	3.93	(X)	(X)	4.24
College--4 or more years	0.9	1,712	(X)	2.8	7,301	(X)	12.0	16,117	(X)
Income of M/H as a % of family income	(X)	(X)	87.6	(X)	(X)	90.9	(X)	(X)	88.2
Average family size	(X)	(X)	3.77	(X)	(X)	3.90	(X)	(X)	4.25
College--4 years	0.6	1,416	(X)	1.6	7,329	(X)	6.6	15,195	(X)
Income of M/H as a % of family income	(X)	(X)	86.9	(X)	(X)	92.1	(X)	(X)	87.6
Average family size	(X)	(X)	3.45	(X)	(X)	3.94	(X)	(X)	4.26
College--5 or more years	0.3	2,224	(X)	1.2	7,263	(X)	5.4	17,232	(X)
Income of M/H as a % of family income	(X)	(X)	88.3	(X)	(X)	89.2	(X)	(X)	88.9
Average family size	(X)	(X)	4.31	(X)	(X)	3.83	(X)	(X)	4.25
45 to 64 years	14.7	1,594	(X)	25.6	6,065	(X)	42.2	13,001	(X)
Income of M/H as a % of family income	(X)	(X)	73.6	(X)	(X)	83.9	(X)	(X)	74.6
Average family size	(X)	(X)	4.49	(X)	(X)	3.10	(X)	(X)	3.59
65 years old and over	1.9	626	(X)	2.6	4,963	(X)	1.9	13,777	(X)
Income of M/H as a % of family income	(X)	(X)	73.4	(X)	(X)	83.2	(X)	(X)	76.1
Average family size	(X)	(X)	2.21	(X)	(X)	2.29	(X)	(X)	2.54
EMPLOYED--WORKED PART TIME LAST YEAR	8.6	1,430	(X)	4.8	3,764	(X)	1.3	12,250	(X)
Income of M/H as a % of family income	(X)	(X)	74.2	(X)	(X)	66.1	(X)	(X)	66.1
Average family size	(X)	(X)	3.80	(X)	(X)	2.51	(X)	(X)	3.05
EMPLOYED--DID NOT WORK LAST YEAR	2.1	1,070	(X)	0.8	3,289	(X)	0.1	6,441	(X)
Income of M/H as a % of family income	(X)	(X)	64.0	(X)	(X)	58.4	(X)	(X)	52.2
Average family size	(X)	(X)	3.33	(X)	(X)	2.66	(X)	(X)	3.37
UNEMPLOYED	4.1	2,349	(X)	2.7	5,126	(X)	1.6	10,295	(X)
Income of M/H as a % of family income	(X)	(X)	80.5	(X)	(X)	77.6	(X)	(X)	70.5
Average family size	(X)	(X)	5.05	(X)	(X)	3.28	(X)	(X)	3.96
NOT IN CIVILIAN LABOR FORCE	43.1	1,466	(X)	20.3	3,466	(X)	4.1	9,511	(X)
Income of M/H as a % of family income	(X)	(X)	75.7	(X)	(X)	66.8	(X)	(X)	58.2
Average family size	(X)	(X)	2.94	(X)	(X)	2.41	(X)	(X)	3.15
WORKED FULL TIME LAST YEAR	3.9	1,582	(X)	3.0	4,443	(X)	1.4	11,712	(X)
Income of M/H as a % of family income	(X)	(X)	75.1	(X)	(X)	72.3	(X)	(X)	70.1
Average family size	(X)	(X)	3.91	(X)	(X)	2.70	(X)	(X)	2.87
WORKED PART TIME LAST YEAR	4.3	1,470	(X)	1.9	3,518	(X)	0.2	9,570	(X)
Income of M/H as a % of family income	(X)	(X)	74.8	(X)	(X)	66.7	(X)	(X)	55.3
Average family size	(X)	(X)	3.19	(X)	(X)	2.48	(X)	(X)	3.73
DID NOT WORK LAST YEAR	34.8	1,427	(X)	15.4	3,268	(X)	2.5	8,259	(X)
Income of M/H as a % of family income	(X)	(X)	75.9	(X)	(X)	65.4	(X)	(X)	51.5
Average family size	(X)	(X)	2.80	(X)	(X)	2.35	(X)	(X)	3.26
45 to 64 years old	8.4	1,425	(X)	2.7	3,179	(X)	0.7	7,737	(X)
Income of M/H as a % of family income	(X)	(X)	71.0	(X)	(X)	55.6	(X)	(X)	49.2
Average family size	(X)	(X)	3.28	(X)	(X)	2.79	(X)	(X)	3.67
65 years old and over	23.7	1,350	(X)	12.2	3,317	(X)	1.7	8,599	(X)
Income of M/H as a % of family income	(X)	(X)	78.4	(X)	(X)	69.4	(X)	(X)	52.5
Average family size	(X)	(X)	2.36	(X)	(X)	2.22	(X)	(X)	3.02

X Not applicable.

Source: Based on tables for the Economic and Social Perspective of Families With Male Heads: March 1970.

Table G.--NUMBER, AGGREGATE INCOME, AND AGGREGATE EARNINGS OF CIVILIAN POPULATION 14 YEARS OLD AND OVER AND MALE FAMILY HEADS BY FAMILY INCOME CATEGORIES: MARCH 1970

(Dollar estimates relate to income received in 1969. Numbers may not add to totals because of rounding)

Civilian population of male family heads	Civilian population 14 years old and over			Worked full time in 1969 and employed in March 1970			Worked full time in 1969 and employed in March 1970 divided by civilian population 14 years old and over x 100		
	Number (in thousands)	Aggregate income of persons (in millions of dollars)	Aggregate earnings of persons (in millions of dollars)	Number (in thousands)	Aggregate income of persons (in millions of dollars)	Aggregate earnings of persons (in millions of dollars)	Number (in thousands)	Aggregate income of persons (in millions of dollars)	Aggregate earnings of persons (in millions of dollars)
	NUMBER						PERCENT		
Total....................	144,143.4	600,430	523,630	64,137.2	493,600	472,400	44.5	82.2	90.2
All male family heads.........	44,647.5	391,710	353,741	35,814.1	352,900	337,400	80.2	90.1	95.4
FAMILY INCOME CATEGORIES									
Less than Low Economy Standard Budget...	3,084.9	5,378	3,217	1,299.7	2,712	2,521	42.1	50.4	78.4
More than "LTLESB" and less than $10,000	19,135.3	106,260	90,344	13,674.1	86,220	83,550	71.5	81.1	92.5
More than $10,000......................	22,427.3	280,019	260,214	20,840.3	263,900	251,300	92.9	94.2	96.6
	PERCENT DISTRIBUTION								
Total....................	100.0	100.0	100.0	100.0	100.0	100.0			
All male family heads.........	31.0	65.2	67.6	55.8	71.5	71.4			
FAMILY INCOME CATEGORIES									
Less than Low Economy Standard Budget...	2.1	0.9	0.6	2.0	0.5	0.5			
More than "LTLESB" and less than $10,000	13.3	17.7	17.3	21.3	17.5	17.7			
More than $10,000......................	15.6	46.6	49.7	32.5	53.5	53.2			

Source: Based on tables for the Economic and Social Perspective of Families With Male Heads: March 1970.

There is an appreciable difference in economic effort exhibited by each group of male heads. The progression of effort beginning with male heads in families with family income below the low economy standard and extending to male heads in families with income above $10,000 may be illustrated by referring to the percent of civilian male heads in an income category who were employed full time in 1969 and were employed in March 1970, the survey month. This statistic climbs rapidly from 42 percent for male heads in families below the low economy standard budget to 72 percent for the intermediate group, and 93 percent for males in families with income above $10,000. It should also be noted that for the same group of heads, those in the highest family income classification are associated with family members who contribute proportionately more to the family than do family members in the other two income categories. Family members other than the head contributed 22 percent to family income for the above-$10,000 category, 14 percent for the intermediate category, and 18 percent for the lowest income category.

With each higher income category, the proportion of male heads with less than 8 years of education and who were employed full time in 1969 and were employed in March 1970 declines from 23 percent for the lowest income category to 19 percent for the intermediate one, and 11 percent for the highest one; the latter percent represents some 2.5 million families. The modal educational level of male heads for the two highest income categories is 4 years of high school, 25 percent for males in families with income more than the low economy standard budget but less than $10,000, and 34 percent for those in families with income above $10,000.

For the 25- to 44-year age group and for the various income categories shown in table F, the income of male heads as a percent of family income tends to increase while average family size tends to decrease as education of head increases. At each educational level, the 25 to 44 age group in the $10,000 and over income category has a larger family size than does the intermediate income category. In the less than college level, the average size of family at each educational level is largest for families with income less than the low economy standard budget.

To place the importance of the statistics on male heads who were employed full time in 1969 and were employed in March 1970 in sharper perspective with respect to their contributions to the economic activity of the economy and within their income categories, several percentages that appear in table G are cited. All male family heads comprised 31 percent of all persons in the civilian population, 14 years old and over, and accounted for 65 percent of all income of persons and 68 percent of their earnings. Male family heads who were employed full time in 1969 and were employed in March 1970 represented 25 percent of all persons in the civilian population and generated 59 percent of the aggregate income of persons and 65 percent of total earnings of $524 billion by the civilian population. For male heads in families with income of $10,000 or more the comparable estimates were 14.5 percent of the civilian population and 48 percent of aggregate earnings of persons. Within the highest income group, $10,000 or more of family income, the 93 percent of male heads who worked full time had 94 percent of the aggregate income and 97 percent of the earnings. This very uniform performance and reward record in the economy for a group which produces 44 percent of all income and 48 percent of all earnings is considerably different than for the other two income categories. As shown in table G, male heads who worked full time last year made up 42 percent of all male heads in families with income less than the low economy standard budget and received 50 percent of all income for the group. For male heads in the intermediate family group the comparable figures were 71.5 percent and 81 percent.

IV. Improving and Expanding the Data Source and Scope of the Economic and Social Perspective Tables

The Current Population Survey generates each month a wealth of data across a broad spectrum of subjects. As a beneficiary of this largess, I have never quite ceased being amazed by the rational statistical network of social and economic relationships that the survey unfailingly produces. This does not mean however that the data are without limitations or that the money income concept is not losing some of its dominance as the principal indicator of welfare. The inaccuracy in reporting, particularly underreporting, interest and dividends, social security benefits and public assistance, and the difficulties in collecting reliable self-employment farm and nonfarm income is well recognized and has appeared in the income literature. The problem becomes more critical as more and more Government programs in the social sphere rely on these income statistics for planning and evaluation purposes.

Improvements in CPS income estimates by correcting for dificient income responses can be accomplished by various adjustment procedures. These may range from arbitrary allocations of income to respondents to make the survey income estimates equal independent control totals to discretionary assignments based on special surveys and associated information. For some sources of income, for example public assistance, a particularly effective method of appraising the quality of individual responses and improving coverage is to conduct surveys sampled from an established list of income recipients. This procedure permits analyses of respondents who incorrectly report no income or who correctly report the income source but incorrectly report by either overstating or understating the amount of income. The list sample approach cannot of course detect respondents who incorrectly report receiving income from a given source. Reducing this type of error can only proceed from the CPS vehicle and is expensive to conduct and the results are subject to a large element of uncertainty because of the difficulty of locating respondents on a given list.

Two hundred years ago in this country nonmoney income was an important and pervasive characteristic of an economy in which a large proportion of the population was engaged in semisubsistence-type agriculture. The currently expanding nonmoney income phenomenon seems to be considerably less anchored in the technological characteristics of the economy than 200 years ago and more in man's social perception of the needs and requirements of men and women in securing stipulated social objectives. Employers increase their financial participation in private pensions and health plans and contributions to the Social Security Trust Fund with respect to their wage bill and Government assumes, develops, and finances more supportive social and health services. The growing list and value of nonmoney income sources and personal benefits derived from labor/employer contracts and practices and Government programs portend increased interest in collecting information on the subject in household surveys that are compatible with the Current Population Survey that can serve as a basic core vehicle. The precise vehicles and methods to estimate and distribute defined nonmoney income sources and personal benefits will have to be determined. The options that come to mind are possible supplements to CPS, collateral surveys, some of which could be of the list sample variety, administrative statistics, and at a minimum the heroic efforts of intelligent, imaginative, and experienced analysts in developing estimations and adjustment procedures to modify money income estimates in order to take account of nonmoney income and personal benefits derived from Government programs.

In an informative paper delivered before this Association in 1971 at Fort Collins, Colorado, "Changes in the Distribution of Taxes Among Income Groups: 1962 to 1968," Roger A. Herriot and Herman P. Miller broadly followed the methodological roads travelled earlier by other analysts to merge money and nonmoney income sources. Neither endorsing the estimates presented in table 9 of that paper nor the conclusions drawn at its finish, I am presenting them here as table H to indicate in broad terms what we can ultimately attain in bringing together more elements than money income in an overall presentation of the assignment of this Nation's output of goods and services to its population.

The flexibility of the modular format used to present the economic and social perspective tables as indicated in Section III of this paper with respect to adding additional variables or level of detail from the CPS or 1970 Census of Population and Housing can be extended to embrace the tax and Government expenditure activities discussed in the Herriot and Miller paper. In addition the modular format can accommodate consumption and wealth elements associated with identifiable groups in the population.

Table H.--PERCENT DISTRIBUTIONS OF FAMILIES AND UNRELATED INDIVIDUALS AND OF TOTAL INCOME BEFORE AND AFTER TAXES, TRANSFERS, AND GOVERNMENT EXPENDITURES, FOR 1968

Adjusted money income intervals	Families and unrelated individuals	Total income before taxes, transfers, and Government expenditures	Total income after taxes, transfers, and Government expenditures		
			A 1/	B 2/	C 3/
Total....................	100.0	100.0	100.0	100.0	100.0
Under $2,000..................	9.9	0.8	2.1	2.6	3.6
$2,000 to $4,000..............	12.4	2.9	5.1	5.4	6.6
$4,000 to $6,000..............	12.8	5.9	7.5	7.3	8.6
$6,000 to $8,000..............	13.9	9.7	10.5	10.1	11.1
$8,000 to $10,000.............	13.0	11.6	11.9	11.1	12.1
$10,000 to $15,000............	22.2	27.3	26.4	25.1	25.6
$15,000 to $25,000............	12.3	23.6	21.9	21.0	20.1
$25,000 to $50,000............	3.0	11.3	9.8	10.4	8.4
$50,000 plus..................	0.5	6.9	4.8	7.0	3.9

1/ A = Unallocable expenditures distributed by total income.
2/ B = Unallocable expenditures distributed by total wealth.
3/ C = Unallocable expenditures distributed by number of families and individuals.

Source: American Statistical Association, "Changes in the Distribution of Taxes Among Income Groups: 1962 to 1968," 1971 Proceedings of the Business and Economics Statistics Section, by Roger A. Herriot and Herman P. Miller, table 9, page 113.

V. Concluding Comments

The work reported in this paper is geared to synthesizing a large amount of economic and social statistics produced from the CPS and molding it into a format suitable for presenting information and establishing a structure for economic and social accounts. A major theme that emerges from the data is the connecting link in the United States economy between economic effort and reward in terms of money income before taxes but including Government transfer payments. This finding is compatible with at least two hard facts of life that persistently reappear in history with boring regularity. First an economic system must in essence, if it is to achieve a modicum of success other than perpetuating itself by healthy doses of repression, be able to distribute rewards to productive forces in the economy in proportions that are approximately commensurate with their economic performances and also have their acquiescence with respect to Government finance and expenditure programs. As a corollary, words, phrases, and exhortations cannot paper over the first observation for any extended period of time.

Looking at the different cumulative income size distributions in Section II, table C, prompts me to point out that the ratios of percents of aggregate family income to all families is a very gross relationship measure that can stand a great deal of refinement in terms of standardized universes in order to make it a more acceptable statistic for use in economic and social analysis. This comment applies equally as well to other generalized measures such as Gini ratios to describe the income inequality of all families or persons. One can pick up the 1971 Statistical Abstract of the United States, turn to table 72, page 55, and find that for 1967 the death rates per 1,000 population were identical for white and Negro and other at 9.4 per 1,000. I do not think we would leap from this statistic to the conclusion that this indicated that the health experiences of both population groups, which are different in age and sex composition, were the same. Examination of age-specific survival rates by sex shows that the populations on the average are not equally at risk with respect to death. Similarly the data as summarized in table C and other tables in this report most certainly reflect measurable differential dollar and time investment commitments and current family operating costs, again in terms of dollars and time expenditure, in securing given income levels, exclusive of any differences in innate qualities, functioning abilities, and cultural aspirations of family heads and members.

USING SOCIAL STATISTICS FOR POLICY AND PROGRAM DEVELOPMENT

Robert B. Hill, National Urban League
Dorothy K. Newman, Washington Center for Metropolitan Studies

Introduction

Most social analysts and policymakers agree that a prerequisite for developing sound social policies and programs is the systematic collection and analysis of social and economic data. But many national, state and local programs are still being developed on the basis of inadequate or minimal pre-program research, misuse of statistical data, and, sometimes, conscious rejection of known facts which contradict the conventional wisdom.

For example, despite the large amount of data revealing a strong work orientation among the poor, most income maintenance programs today contain provisions that coerce poor people to work. Such misconceptions are crucial in the failure of many statistical and social programs to function effectively.

Since the mid-1960's, when a number of works on minority populations were published, such as the useful compendium on the social and economic status of blacks by the Bureau of Labor Statistics in 1966; and the poverty index was developed by the Social Security Administration, the movement toward developing social indicators for policy and programs has accelerated. Unfortunately, the pressure to derive key social indicators for major policymakers in the way that key economic indicators are being used, has often led to misinterpretation and misapplication of social statistics. Sometimes, indicators are the synthesis of a network of unproved assumptions or stereotypes. The effort to derive key social indicators of trends in the society was never intended to be a substitute for careful, systematic analysis of the social and economic circumstances relating to the summary data being used.

In this paper, we attempt to illustrate briefly some of the problems encountered when using social statistics for social policy in such fields as occupations, housing, income and family structure and class patterns. We also try to suggest some ways by which the statistical data can be used more effectively for developing social policies and programs.

Income

A recurrent, but difficult, question for social analysts and policymakers is, "How do we properly assess the extent to which minority groups are achieving income parity with the majority?" A basic issue is the need to redefine or to better measure income.

Income alone, however, does not measure "wealth" or the capacity to consume and prosper. It does not take account of assets, fixed taxes on wages or sales, tax benefits (sometimes called "loopholes"), deferred investment income, lines of credit from which blacks are often barred, and pension and welfare benefits to which only persons in particular industries, occupations or unions are entitled. While black workers are much less likely than whites to have assets, they tend to pay a higher percentage of their income in taxes. They also can afford less for health care and thus have shorter life spans. If all assets were taken into account--both the tangibles and the as yet unmeasured intangibles--that contribute to a more accurate concept of total "wealth," there is little doubt that white families, even at similar levels of income, would be much more well-off than black families.

The problem of assessing income parity between minority and majority groups is compounded when the differential impact of multiple earners is not taken into account. Most analyses of income differences between black and white families, for example, are in fact comparing many one-earner white families with multiple-earner black families. This is primarily because black wives are much more likely than white wives to be in the paid labor force. In 1970, it took 2 earners among black families to achieve an (median) income (of $8,430) almost as high as 1-earner white families (with a median income of $8,713). Black families with 3 or more earners had less income than 2-earner white families.

Thus, a more accurate way of assessing the extent of income parity is to compare the incomes of those black and white families having similar numbers of earners. When the Census Bureau takes account of multiple earners among husband-wife families, it finds, contrary to the usual analysis, that the income gap between black and white families with only one earner (i.e., the male head of household) has not significantly closed over the past decade.

Occupation

Many analysts of occupational mobility continue to equate "white collar" with "high" status and earnings. But the fact is that each group contains a heterogeneous mix of jobs. Many "blue collar" occupations, particularly in the crafts, pay substantially more than many white collar positions such as cashiers, bank tellers, dispatchers and office machine operators. But the services, also a low status, "blue collar" group, includes such disparate occupations as policemen, firemen, barbers, hair dressers, dishwashers, night watchmen, waiters, cooks and attendants.

So-called "women's jobs," such as secretaries and sales persons in retail stores average less pay than those jobs which men tend to occupy almost exclusively in heavy industry. For example, full-time, year-round women workers in sales and clerical jobs averaged (at the median) less than $5,500 in 1970, whereas men in manufacturing and even so-called "unskilled" laborers earned substantially more than that. Craftsmen averaged well over $9,000.

It is clear that occupational status and the entire classification scheme of occupations require substantial analysis and reassessment. Some work in this area is already going on within the government. It should be accelerated and more attention paid to behavioral science resources outside the government. Empathy and insight into people's perceptions are needed. Required also is a careful analysis of many kinds of data that disaggregate totals and relate details to industry earnings, wage rates, the nature and extent of collective bargaining agreements, location and economic base and indicators of social and economic mobility.

Current data are being collected by the Bureau of the Census, the Bureau of Labor Statistics and the Equal Employment Opportunity Commission and should not be ignored. They should be used, however, with a basic understanding of the society they reflect.

A related policy issue that is generating much discussion is, "What proportion of any job category should be occupied by a particular ethnic minority (or by women) in order for them to achieve equal employment opportunity." Many economists and statisticians advocate using the participation rate of that group in the labor force as the yardstick.

The ratios in each broad occupation among Polish, Italian and Irish-origin men correspond roughly to the distribution of all American men --about 14 percent in professional or technical work, 14 percent managerial, and 20 percent craftsmen or operatives. Each of these white ethnic groups constitutes less than ten percent of the male labor force; therefore, according to the conventional rule of thumb, they are heavily overrepresented in manual compared to professional work. On the other hand, Chinese and Japanese Americans, a minute fraction of the labor force, are much more heavily represented in the professions than any of the white ethnic groups mentioned, and underrepresented in the industrial occupations.

Actually, the economy's employment requirements dictate an occupational distribution which changes with the demands of industry, trade and the professions. Therefore we should really be considering first, what talents the economy needs and how those needs are changing, and second, whether preferences and group customs affect occupational distribution when discrimination is not present.

Among black American men, who are 11 percent of the total male labor force, 8 percent were in professional or technical types of jobs in 1971, 5 percent were managers, 13 percent were craftsmen, and 26 percent were operatives. This picture suggests that substantial numbers of black men have not yet been permitted to contribute their talents to facets of the American economy requiring professional, technical, and managerial competence or craftsmanship in the skilled trades.

Housing

Housing policies and programs require much more intensive analysis using social statistics. One policy issue often debated today relates to housing subsidies--that is how much of their income should the poor pay for shelter if they live in subsidized housing.

At present, most proposals are designed to insure that the poor (and families) pay a higher proportion of their income for shelter than the non-poor (or the elderly). A recent analysis of consumer expenditures, BLS normative family budgets, and the criteria used by mortgage lenders and the FHA for loans to middle-income families, do not support as high a ratio as the one assigned to the poor under present and proposed law.

What is more, in every case of a subsidy or supplement to the poor for housing, income and the elements included in housing costs are itemized in such a way as to maximize the cost to the poor, but provide a cushion for the more well-off. Maintenance, insurance, and utility costs are often excluded in estimating the shelter-to-income ratio for poor families, and income is established as gross income from all sources from all adult members before taxes and withholding.

Higher income renters or buyers, however, are considered on the basis of dependable income over a span of years, after federal taxes; shelter includes an estimate for the normal repair, utilities, maintenance, and insurance costs.

A more effective and equitable housing policy would assess family shelter costs and income alike for the poor and the well-to-do. Shelter costs and family income should be related to the sex and age of the family head and by family size. The data should be assembled by income class and family size from consumer expenditure studies, family budget analyses and area and local transaction records.

Family Structure

Policy issues relating to family patterns are hotly debated today. As indicated earlier, many national policies and programs are based upon stereotypes or conventional wisdom about family instability, disorganization or pathology among minority and low-income groups. This orientation is best exemplified in the widespread practice of equating family structure with family functioning.

Many social scientists, like the man on the street, continue to treat the statistics relating to the proportion of families headed by women as if it were a valid measure of the functioning of those families. One-parent families are morally prejudged to be unstable and disorganized, while two-parent families are prejudged to be stable and organized--without any empirical investigation of the actual functioning patterns of these families.

In fact, some analyses of family functioning have revealed that two-parent families are not necessarily "intact" and one-parent families are not necessarily "broken." Moreover, some of these investigations have indicated that many one-parent families function more effectively than many two-parent families. A serious deficiency in the literature on low-income and black families has been the lack of systematic investigation of the many strengths and resources existing in both two-parent and one-parent families.

Social Class

Among the most misleading of all social statistics are the correlates of social class. Earlier, we discussed the ambiguity in using "income" as a measure of "wealth." But an equally serious problem is posed by the widespread practice of equating income groups with class categories, values and life-styles.

For example, many analysts continue to equate "middle-income" with "middle-class" and "low-income" with "lower-class." They forget that many "middle-income" families can have "lower-class" values and life-styles while many (if not most) "low-income" families have "middle-class" values, and live according to that model.

Moreover, as a recent analysis of middle-income black families revealed, it is not clear whether "middle-income" black families should be compared with "middle-income" white families at the white "norm" (or median) or at the black norm (or median). For this analysis revealed some striking differences in life-styles between black and white families at similar income levels.

Although income is a strong correlate of socio-economic status, it is not as valid a measure of class life-styles and values. For income only describes how much money a family has, but class determines how that money is spent--or saved.

More systematic research is needed about the variety of class values and life-styles exhibited by persons and families at similar and different economic levels. A number of studies have already revealed that there is more homogeneity in goals, values and aspirations between persons or families at different economic levels than is commonly believed.

Methodological Biases

Finally, we feel it is necessary to indicate that our own methodological biases may in fact impede the effective use of social statistics for policy and program development.

One of these biases is the widespread practice among social scientists of making generalizations about groups based on unrepresentative proportions. For example, in most discussions about black families, greatest attention is usually given to the 30 percent of black families headed by women. The fact that husband-wife families represent the overwhelming majority of black families has usually not deterred many references to a "matriarchy" as being the "typical" pattern of black families.

In fact, data from the recent Census Employment Surveys in 60 low-income (or "poverty") areas in 51 cities, reveal that many of the patterns that are said to be "typical" of ghetto residents are not. For example, only 1 out of 4 residents of urban "poverty areas" are below the official poverty level. More than 3 of 5 black families in the low-income areas are husband-wife families, and fewer than 3 out of 10 of the black families in urban "poverty areas" receive public assistance. Thus, the major thrust of the nation's policies and programs for low-income areas are directed toward a distinct minority, for which the majority of ghetto residents are often not eligible. Clearly, a more varied approach is needed in order for social policies and programs to be of substantial benefit to all of the families in need--both poor and non-poor.

SELECTED BIBLIOGRAPHY

Aug, Robert G. and Bright, Thomas, "A Study of Wed and Unwed Motherhood in Adolescents and Young Adults," Journal of the American Academy of Child Psychiatry, 9 (October 1970), 647-655.

Goodwin, Leonard, Do the Poor Want to Work? Washington, D.C.: Brookings Institution, 1972.

Hill, Robert B., The Strengths of Black Families, (A National Urban League Research Study), New York: Emerson Hall Publishers, Inc., 1972.

National Urban League Research Department, "Black Unemployment: A Crisis Situation," (A report released at the National Urban League's 62nd Annual Conference in St. Louis Missouri), August 1, 1972. (Mimeo.)

Newman, Dorothy K., "The Black Middle-Income Family," Tuesday at Home, October 24, 1971.

______, "Housing the Poor and Shelter to Income Ratio," in Papers Submitted to Subcommittee on Housing Panels on Housing Production, Housing Demand and Developing a Suitable Living Environment, Part II, Committee on Banking and Currency, House of Representatives, 92nd Congress, First Session, 1971.

______, "Are Sociologists Flunking Social Policy?" (A paper presented to the Southern Sociological Society, 35th Annual Meeting, New Orleans, Louisiana), April 6, 1972.

______, "The Black American Worker," The American Negro Reference Book, New York: Prentice Hall, to be published by early 1973.

Newman, Dorothy K., and Steffes, Robert B., "Occupational Attainment of Ethnic Groups and Women in 15 Industries," (Unpublished study by the National Urban League Research Department for U.S. Equal Employment Opportunity Commission), 1971.

Tausky, Curt, and Wilson, William J., "Work Attachment Among Black Men," Phylon, (Spring 1971), 23-30.

U.S. Bureau of the Census, "Differences Between Incomes of White and Negro Husband-Wife Families are Relatively Small Outside the South," U.S. Department of Commerce News, February 18, 1971.

______, "Differences Between Incomes of White and Negro Families by Work Experience of Wife and Region: 1970, 1968, and 1959," Current Population Reports, Special Studies, Series P-23, No. 39, 1971.

______, "Income in 1970 of Families and Persons in the United States," Current Population Reports, Consumer Income, Series P-60, No. 80, 1971.

U.S. Bureau of the Census, "The Social and Economic Status of the Black Population in the United States, 1971," Current Population Reports, Special Studies, Series P-23, No. 42, 1972.

______, "Employment Profiles of Selected Low-Income Areas, United States Summary--Urban Areas," 1970 Census of Population and Housing, PHC(3)-1, 1972.

U.S. Bureau of the Census and U.S. Bureau of Labor Statistics, "The Social and Economic Status of Negroes in the United States, 1969," Current Population Reports, Special Studies, Series P-23, No. 29 and BLS Report No. 375, 1970.

______, "The Social and Economic Status of Negroes in the United States, 1970," Current Population Reports, Special Studies, Series P-23, No. 38 and BLS Report No. 394, 1971.

U.S. Bureau of Labor Statistics, The Negroes in the United States: Their Economic and Social Situation, Bulletin No. 1511, 1966.

U.S. Department of Health, Education, and Welfare, Toward a Social Report, Ann Arbor Paperbacks, The University of Michigan Press, 1970.

POLICY USES OF URBAN INDICATORS

Robert K. Yin
The New York City-Rand Institute

From the point of view of the urban policy-maker, research on social indicators has still failed to provide systematic reports on the social and physical condition of the metropolis. In fact, there are at least four distinct and major phases of indicator development that are in need of improvement, and I would like to take this opportunity to review each of them briefly. The first has to do with the development of certain basic information about the city; the second deals with the development of indicators that are usable in the framework of policy decisions; the third involves a topic not usually considered part of social indicators, i.e., the exploration of new policy options; and the fourth raises the question of the second-order effects of indicator systems.

Basic Information About The City

Despite the considerable attention given to social indicators over the last decade and the pioneering efforts of people like Eleanor Sheldon, Raymond Bauer, and Bertram Gross [1], we still have a little information about the quality of life in the contemporary American city. To be sure, no one expected that by this time there would be routine reports for value laden indicators like residential satisfaction. The difficulties of making precise measurements of social factors, added to the problem of making value judgments about what to measure and what constitutes an "improvement," both create a formidable barrier for any attempt at developing urban social indicators of this sort. However, any review of the existing information about cities will reveal that little is yet reported even about certain basic urban conditions. For instance, no matter what their value orientation, most people would agree that cities ought to have systematic reports on the following conditions:

- the total number of people
- the number of people by age, sex, and race
- the rate of population mobility
- the health status of the population
- the educational status of the population
- the environmental safety of the neighborhood (including accidents, fires, and crimes)
- family income
- the employment (and unemployment) of the population.

The reports should be available at least on an annual if not a monthly basis, and should be broken down to the level of small areas like census tracts. Such reports would be invaluable to citizens, researchers, and policy-makers, and yet this type of information is virtually non-existent.

One possible reason for the failure to develop indicators of these conditions is that researchers have tended to lean too heavily on residential surveys as sources of information. Surveys include the national census, periodic urban surveys like the Detroit Area Survey (University of Michigan) or the Boston Area Survey (Harvard--M.I.T. Joint Center for Urban Studies), and special surveys. In general, however, surveys are too costly to be carried out frequently enough or with large enough samples of respondents to provide even annual small-area data for an entire city. As a remedy, what may be needed is a large dose of imagination and more exploration of two other sources of indicators: direct observations of street events, and municipal records. Both of these sources should be examined for direct measures of the urban conditions described as well as for indirect measures, or proxies of those conditions.

Direct observations of the street would involve the identification of certain events that may be unobtrusive signs of urban social conditions. For instance, abandoned cars, the number of people on the street, and drug and other retail stores are all easily observed and enumerated. Yet few investigators have examined the validity of these street indicators as proxies, say, for poverty, population, and health. Although considerable field work might be needed to assess these street indicators, once identified and validated they could easily be monitored by mobile research teams, policy-makers, or even residents themselves [2]. The final result would be a routine series of reports, produced at frequent intervals, and according to neighborhood locations.

Municipal records are those records maintained by municipal agencies. The records are often unreliable and in difficult-to-manage form. However, where these problems can be overcome, the records can serve as a potentially good resource for indicators because the information is usually recorded in sufficiently fine spatial and temporal detail. As a first step, several basic criteria might be used in determining the usefulness of the events in these records [3]:

- the events should be consistently defined;
- the events should be discrete, with a known time of occurrence;
- the events should be reported by very small geographic area;

-the reported incidence of events should approximate the actual incidence; and
-the events should be reported routinely and with as little delay as possible.

Researchers have tended to pay little attention to municipal records, and particularly to one indicator that most eminently satisfies all of these criteria, urban fire alarms [4]. For instance, Figure 1 shows the different alarm types for New York City. Each separate type of alarm tends to follow different statistical patterns, suggesting that each one could potentially be used as an indicator of different urban conditions.

Developing Usable Indicators

Assuming that basic information about the city were available through surveys, street indicators, or municipal records, a second major concern has to do with the shaping of the information into usable form. Usability, in this case, is defined strictly in terms of the needs of the urban policy-maker.

Generally, this means that the information should be accurate, timely, and brief. One prototype (and one of the few examples of an existing urban indicator) of the report needed is the Temperature-Humidity Index devised by the U.S. Weather Bureau. The Index is not intended as a measurement of weather conditions. Instead, it purports to record the amount of personal discomfort due to heat and humidity on any given day. Unfortunately, this Index is not reported for all urban areas, but where it is reported, the daily report provides highly relevant and timely information. In some cities, air pollution indices have also been devised and are now reported routinely [5].

Simple reporting, however, is not quite sufficient. The analyst must also give the policy-maker some sense of the degree to which an indicator deviates from normal. This means that, just as with the national system of economic indicators, the urban indicators must be seasonally adjusted or otherwise normalized. For fire alarms, for instance, the analyst would have to point out the expected seasonal patterns as shown in Figure 2. Interestingly, such seasonal patterns may underlie survey results as well, and replication studies as suggested in a recent indicator document will have to be designed with this possibility in mind [6]. As another example, the alarms also have persistent geographic patterns; Figure 3 shows the variation in small areas for the borough of the Bronx for a one-month period.

In addition, the analyst should also provide the policy-maker with substantive interpretations of the indicator. For instance, there has been a considerable amount of research on the topic of neighborhood change, and particularly residential turnover from white to non-white populations [7]. If such research showed the existence of critical points in the turnover process, then such points should be identified along with the indicator report. Another substantive contribution would be the identification of early warning signs of significant urban change, whether involving an incipient riot, neighborhood abandonment, economic shifts, or urban renovation [8]. Many have hypothesized, for instance, that a rise in false fire alarms in a neighborhood is an early sign of large population turnover, primarily from middle-income to lower-income families. In short, policy-makers would benefit most from an indicator system that not only provides the essential information but also identifies the potential repercussions of indicator change.

The final key to a usable set of indicators is a well-designed, computer-based information system. Much has already been written about urban information systems, most of it belaboring the obvious points concerning system design and programming talent [9]. The main reason for a computer-based system is simply that, for the large American city, the number of relevant incidents is usually large and the calculations required for various indices are complex, but the data must be reported as quickly as possible. Hence there is ample justification for creating a computer-based urban indicator system.

Exploration Of New Policy Options

The third major concern in the use of urban indicators is not one that is normally regarded as part of the development of social indicators. This concern does not involve the creation of new and usable information as much as it involves the development of new types of program options that are available to the policy-maker. In short, urban social indicators must be created hand-in-hand with new urban programs and responses.

Why should this be a part of the indicator job? The answer is related to the problem of making value judgments in indicator research. As Moynihan has written, social indicators in general are likely:

> ...to be developed by professors and government executives who will be far more concerned with what is bad about cities than with what is good about them. These men will judge good and bad in terms of their own rather special values acquired in the course of family, religious, educational, and occupational experiences that, by and large, are quite different from those of the urban masses whose condition they will seek to measure [10].

Thus much care must be exercised in deciding what should be measured and how changes in indicators should be interpreted. By now, most people have been sensitized to the problem, although adequate means for dealing with implicit value judgments have not emerged.

However, the very creation of an indicator series is also a prejudicial step if only one response action is available to the policy-maker. For instance, supposing an investigator were asked to develop an early warning system for urban riots, and such an indicator actually came into existence. At the present time, most mayors have only one recourse in a riot situation: summoning control forces and preventing or combating the riot. Under such conditions, the early warning system has therefore become not only a riot indicator, but it has also become an instrument of riot suppression. Similar cases can be made for crime indicators, or any indicator concerning the formation of youth gangs. In each case, if there is only one response available to the decision-maker, the indicator becomes an unwitting partner of a fixed urban policy action.

In order to create a usable and less biased series of urban indicators, analysts must therefore also develop new policy options for the decision-maker. In the case of urban riots, other options could involve peaceful street negotiations, perhaps by the mayor himself, or the development of other emergency contingencies to help rectify whatever immediate situation is leading to the riot. Perhaps even more important, longer-term follow-up programs could be considered to provide more meaningful and permanent improvements. It should be pointed out that the development of new programs and options is a very complex affair, in that the programs often have unintended consequences [11], but that in spite of these difficulties the development of new options should be an integral part of the development of new indicators.

Second-Order Effects Of Indicator Systems

The need to develop new policy options is actually but one example of a whole class of broader second-order concerns [12]. Where reliable indicator systems are developed, in other words, policy-makers and citizens may both pursue new courses of action as a result of the new indicator information. The policy-maker, as has been pointed out, can use his existing program options more effectively.

The citizen, however, can also change his preferences and activities because of the indicator. An obvious example involving the citizen would derive from a crime indicator system. If crimes and perceived safety were routinely reported for different neighborhoods, citizens might then be better forewarned to protect themselves and their property. On the other hand, the routine crime reports might also serve to accelerate residential relocation from high-crime to low-crime neighborhoods, and such movement (since it would be selectively limited to the families that could move) might leave the high-crime neighborhood vulnerable to even more crime. The indicator reports themselves could thus have both positive and negative effects.

The development of social indicators therefore entails the serious risk that conditions can be aggravated simply because of the fact that indicators have been created and reported. These second-order effects could be entirely undesirable, and could be avoided if the indicators were not reported in the first place. The researcher engaged in developing social indicators must consider the possible second-order effects of his indicator system, and weigh the likely advantages and disadvantages of the system. Naturally, this decision-making process will involve another set of value judgments, and more research is needed to determine the appropriate criteria for recommending which indicators on a certain topic be developed or not be developed.

Summary

In summary, the researcher who attempts to develop meaningful and useful urban indicators is faced with several problems beyond the normal methodological concerns of statistical validity and reliability. First, he must locate a good information source for indicators, knowing that at present there is very little known about basic urban conditions, even on such essential matters as the overall population or the population's mobility, health, or income. The main suggestion here has been that researchers give more attention to the use of street indicators and to municipal records, and less attention to the use of residential surveys, which tend to be cumbersome and unresponsive to the policy-maker's needs for timely decisions on a small-area basis. Second, the good analyst must not only provide indicator data in a compact and usable form, but he must also carry out relevant analysis on indicator trends so that he can tell the policy-maker when indicator changes represent significant deviations, and what changes may be important as critical points or as early warning signs. Finally, the analyst must also keep in mind the potential second-order effects of the creation of indicator systems. These include the development of new policy options to go along with the creation of new indicators, so that the indicators do not automatically become instruments of single public policies. The second-order effects also include the judgment that urban conditions will not be aggravated simply as a result of the existence of the indicator information.

These requirements perhaps pose a formidable agenda for the researcher. However, the requirements must be met if urban indicators are to share any part of the reality of the American city in the nineteen-seventies.

NOTES

1. Eleanor Bernert Sheldon and Wilbert E. Moore (eds.), Indicators of Social Change: Concepts and Measurements (New York: Russell Sage, 1968); Raymond A. Bauer (ed.), Social Indicators (Cambridge: M.I.T. Press, 1966); and Bertram M. Gross (ed.), "Social Goals and Indicators for American Society," The Annals, 371 (May 1967) and 373 (September 1967).

2. For further discussion on the development of street indicators, see Robert K. Yin, Participant-Observation and the Development of Urban Neighborhood Policy, The New York City-Rand Institute, R-962, April 1972.

3. One can use these (and other) criteria to compare the potential usefulness of available records, like reported crime, school enrollment, the incidence of mental illness, etc. See Robert K. Yin, "The Development of Social Indicators: The Case of Fire Alarms," paper presented at the Eastern Sociological Society annual meeting, April 1972, Boston, Massachusetts.

4. For a description of the potential interpretations of fire alarms, see Ibid.

5. It is interesting to note that the few urban indicators that have been attempted (e.g., the Temperature-Humidity Index, air pollution indicators, and noise pollution indicators) are all psychological, and not social indicators. In other words, the information relates the urban environment to the needs of the individual citizen, but no assessment has been made of his social activities.

6. Otis Dudley Duncan, Toward Social Reporting: Next Steps (New York: Russell Sage, 1969). For one of the few studies that has dealt with the seasonal factor on a social issue, see William Michelson, "Some Like It Hot: Social Participation and Environmental Use as Functions of the Season," American Journal of Sociology, 76 (May 1971), 1072-1083.

7. For instance, see Morton Grodzins, The Metropolitan Area as a Racial Problem (Pittsburgh: University of Pittsburgh Press, 1958); Davis McEntire, Residence and Race (Berkeley: University of California Press, 1960); Eleanor P. Wolf, "The Tipping-Point in Racially Changing Neighborhoods," Journal of the American Institute of Planners, 29 (August 1963), 217-222; and Harvey Molotch, "Racial Change in a Stable Community," American Journal of Sociology, 75 (September 1969), 226-237.

8. For example, see Joan E. Jacoby, "The Neighborhood Early Warning System--Design and Development," Socio-Economic Planning Sciences, 4 (1970), 123-129.

9. Two discussions that not only cover some of the technical requirements but also raise important social and political issues are: Edward M. Goldberg, "Urban Information Systems and Invasions of Privacy," Urban Affairs Quarterly, 5 (March 1970), 249-264; and Edward H. Blum, "The Community Information Utility and Municipal Services," The New York City-Rand Institute, P-4781, February 1972.

10. Daniel P. Moynihan, "Urban Conditions: General," The Annals, 371 (May 1967), 159-177.

11. One of the few analytic models of the longer-term effects of urban programs can be found in Jay W. Forrester, Urban Dynamics (Cambridge: M.I.T. Press, 1969).

12. The question of second-order consequences has been explicitly raised primarily in relation to new technological changes. See Raymond A. Bauer, Second-Order Consequences: A Methodological Essay on the Impact of Technology (Cambridge: M.I.T. Press, 1969).

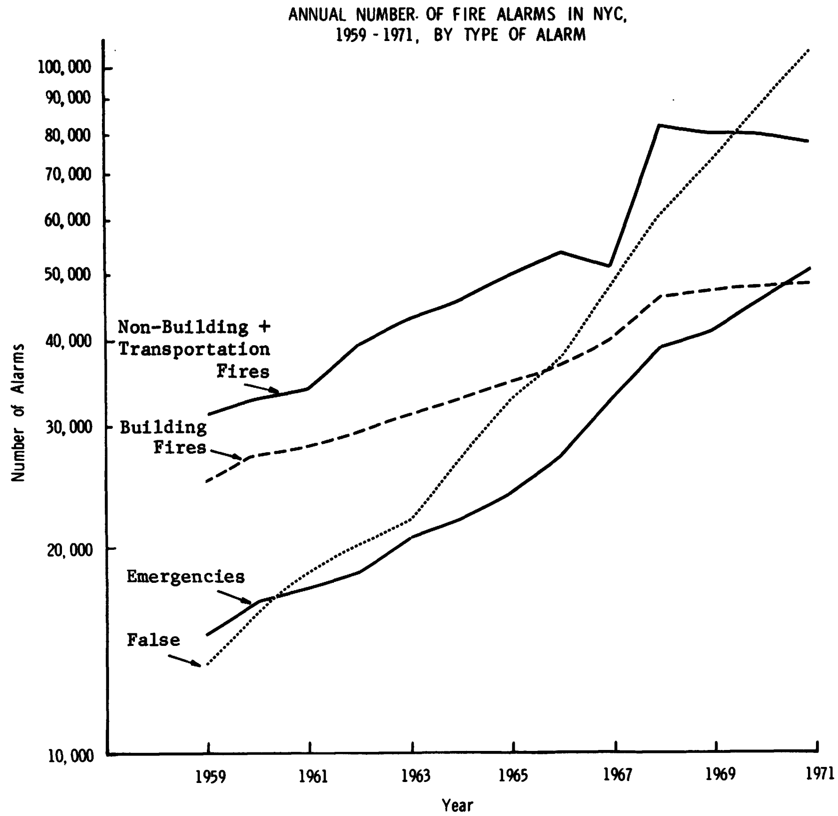


Figure 1

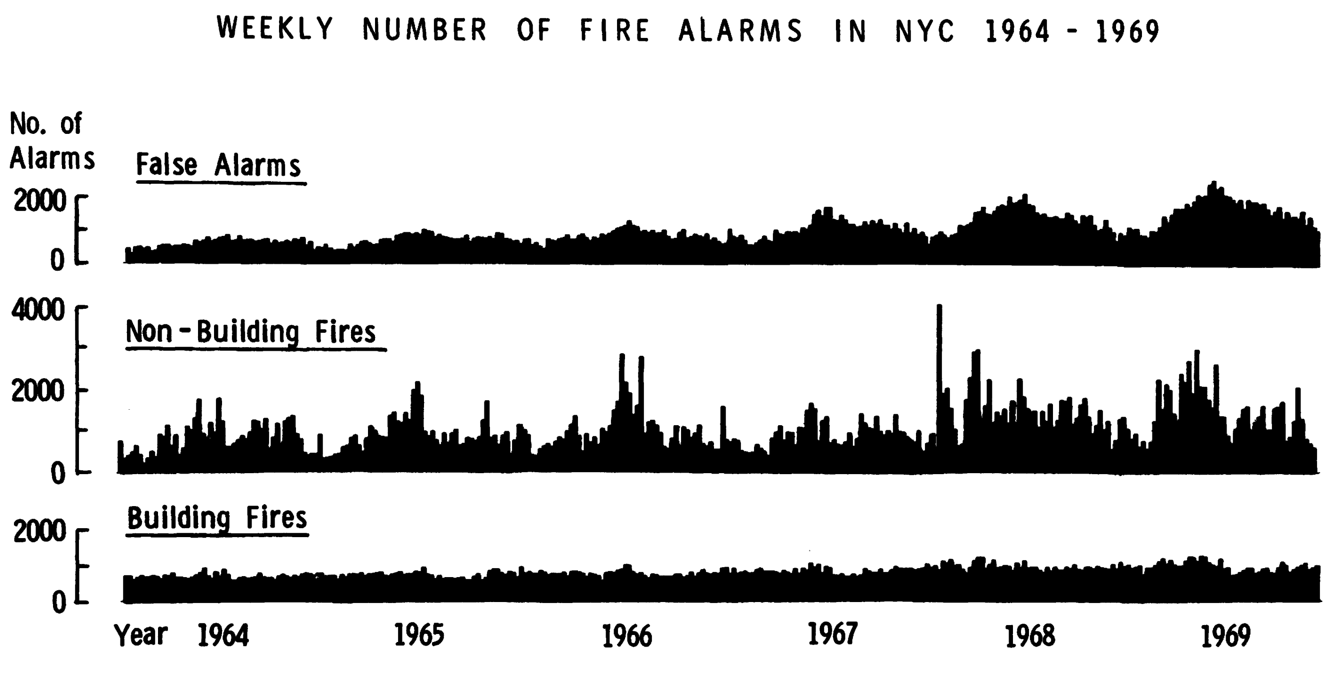


Figure 2

FALSE ALARMS BY HEALTH AREAS

THE BRONX

JANUARY 1970

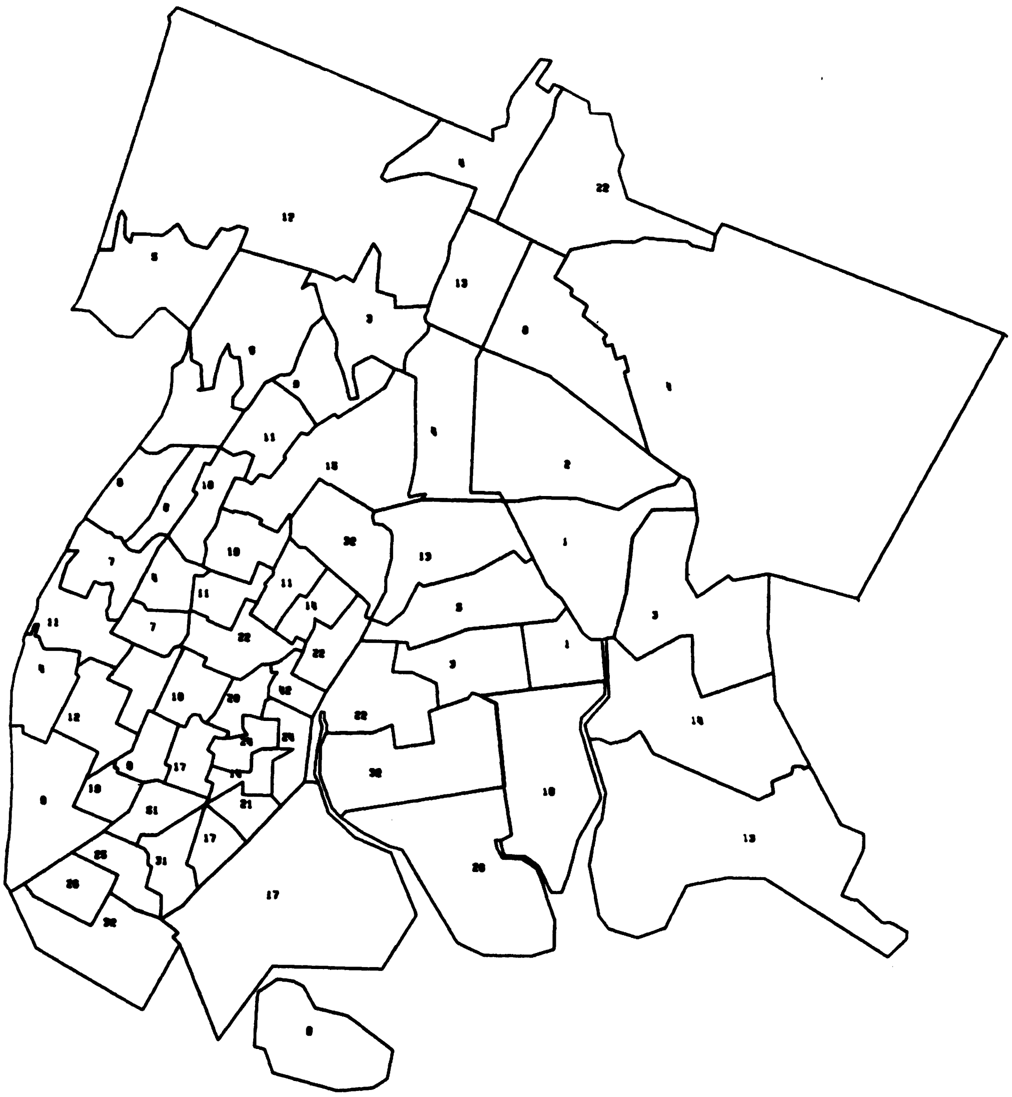


Figure 3

POLICY IMPLICATIONS OF CURRENT DRUG ABUSE ESTIMATES FOR LOCAL AREAS IN NEW YORK CITY

Mary G. Powers, Ruth M. Cullen, and John M. Martin
Fordham University

A recent article in the New York Times referred to attempts to estimate the number of addicts in the United States as a "numbers game."[1] Many who have made or used such estimates for program purposes would agree. Ignoring for the moment the interesting question, "Who is an addict?" and relying on operational definitions based on existing data, the number estimated varies widely depending on what sources and methods are actually used.

This paper examines and evaluates Narcotics Abuse rates derived from New York City Narcotics Register data, New York State Narcotic Addiction Control Commission (NACC) estimates, and from New York City police arrest figures. These are the three basic sources of data for most estimates of the magnitude of the drug problem in New York City. Data taken from the Narcotics Register and from estimates made by NACC were used to describe rates of opiate use in New York City and the various subareas or neighborhoods of the City for 1964 through 1967. These subareas were then ranked according to rate of opiate use. (Graph A).

This ranking was made from estimates of drug use per 1,000 population 15-44 years old, based on Register data and tabulated by Health Center Districts. There are thirty Health Center Districts in New York City. These may be defined as sub-communities or neighborhoods within the city.

The ranking provides a basis for viewing the problem of drug abuse in New York City as a whole, and comparatively, among local communities within the city. Because of problems with the base data, however, extreme care must be taken in interpreting these and other available data - especially in using them to derive estimates of numbers of addicts. Yet one constantly hears in radio, press, and T.V. coverage of the drug problem references to the magnitude of the problem in terms of large numbers - estimates of from 100,000 to 600,000 addicts in New York City. The accuracy of numbers such as these cannot be evaluated precisely because of the sources of the estimates. The data and method used to derive the numbers are described as follows.

Register Data

According to the 1969 report of the Narcotics Register, there were 94,699 drug abusers reported in New York City for the year ending December 31, 1969.[2]

Manhattan ranked first among the five boroughs in reported cases of opiate abuse. Of the seven health districts in Manhattan, Central Harlem ranks first, East Harlem had the second largest number of cases, and the Lower East Side ranked third. Among the 30 Health Center Districts in New York City as a whole, the Lower East Side ranked fifth in newly reported cases of narcotics abuse in 1969.

Central and East Harlem are relatively homogeneous with respect to ethnic make-up. The Lower East Side is not. It is still a heterogeneous area. This evaluation of available data focuses on figures for the Lower East Side relative to all other areas. It was performed as part of an overall evaluation of Horizon Project, a drug treatment program in the area operated by the New York City Addiction Services Agency and funded by the National Institute of Mental Health. The results have implications for neighborhood based programs in all areas, however.

Table A describes the total number of narcotics abusers reported to the Register from the Lower East Side, Manhattan, and New York City. The table suggests a dramatic increase in drug use over the period shown. The increase probably results from both an increase in the number of narcotics abusers in the City and increased reporting to the Register by the agencies involved.

What are the addicts like in terms of social characteristics? More detailed tabulations have been made of the social and demographic characteristics of heroin abusers than of all narcotics users. The 1969 data from the Register indicates that 87 percent of all drug abuse is heroin use. Thus, the material describing heroin abusers has been used interchangeably with that for total narcotics abusers.

Narcotics Register data on heroin abusers residing in the Lower East Side indicate that the ethnic breakdown of users was similar to the ethnic make-up of the community. Approximately 50 percent of the cases reported to the Register were Puerto Rican, 27 percent were other whites and 19 percent were blacks. This ethnic breakdown remained relatively stable over the period for which data were available. Most users reported to the Register were between 20-30 years of age. The proportion of blacks was highest in the thirty and over age category, however, 31.1 percent compared to 19 percent of the total. This may indicate 1) that drug use started earlier among blacks than others, and therefore, some users are older, or 2) that drug use may be a temporary or "youth" phenomena among Puerto Ricans and other whites and a long-term behavior pattern among blacks.

If these figures reflect the actual addict population in the Lower East Side, Puerto Ricans are over-represented, particularly in the younger age categories, while blacks are proportionately greater in the thirty and over age category. As will be seen shortly, however, Register data must be interpreted with extreme caution.

Evaluation of Data and Estimates

The figures cited indicate a massive and increasing drug abuse problem in New York City and in the Lower East Side. It is, however, difficult to evaluate the completeness and accuracy of these data because of 1) the method of reporting to the Register and 2) the method of correcting for under-reporting used by NACC in making estimates of rates per 1,000 population for each Health Center District.

The Narcotics Register Project

The Narcotics Register Project of the New York City Department of Health is an example of the "cumulative case register" approach to disease surveillance used in many areas of public health practice.[3]

A case register is a system for collecting information on a population with a particular disease or problem based on cases reported by agencies or individuals (e.g., physicians, therapists) within a specified area. The Narcotics Register was designed to develop methods of establishing a reliable up-to-date, unduplicated count of narcotics addicts in New York City.[4]

The Narcotics Register has been viewed as a research project with no anticipated case intervention by either public health practitioners or law enforcement officials. Under the New York City Health Code, the reports supplied by the various social service agencies working with addicts are confidential and not subject to subpoena. Some agencies have been reluctant to report to the Register, however, because they fear confidentiality will not be maintained. In fact, even with all the publicity given to drug problems in public schools, the Board of Education has only recently agreed to report. Part of the increase in the incidence of drug abuse thus may be attributable to increased and better reporting to the Register.

Definition of Addiction

Definitions of addiction have been formulated by a number of organizations in the addiction field. They usually incorporate the three related phenomena of tolerance, physical dependency, and habituation to opiates or opiate-like drugs. The Narcotics Register has had considerable difficulty in developing suitable operational definitions. Cases are filed as narcotics addicts if

1. the individual had been accepted by a hospital or clinic for in-patient treatment,

2. the individual was reported by a medical source, and,

3. the individual was accepted for treatment and reported by an established social service agency.[5]

The above represents an acceptance of the "addict" definition used by the agency of report. Cases with only a single report, therefore, may be very different with respect to their actual drug involvement. Very little is really known about patterns of drug use over time. We do not know how many temporary or even one-time users exist. Thus, and most significantly, it may be quite inaccurate to list a one-time user, who happens to have been arrested on the occasion of use, in a Register which is designed to be primarily a cumulative list of active users. In addition to the fact that each reporting agency's definition of "addict" is accepted, inaccuracies in the data occur because some addicts may be reported more than once during the year, sometimes under different names and/or addresses. However, Register personnel feel they have minimized the duplication of reporting since 1969. Another possible source of error is suggested by the findings of another paper in the evaluation program which indicates that there is good reason to believe that a small number of persons in various treatment programs are not addicts at all.[6] Youths apprehended for a crime may plead addiction and opt to go to a treatment center in lieu of prison.

Problems of Estimating Drug Use

In addition to the problems of definition of "addict" or "user" and duplicate reporting to the Register, another difficulty may arise from the method of correcting for under-reporting to the Register. The rates of opiate use in the 30 Health Center Districts in New York City were computed by NACC's Division of Research, based on data from the Register. In order to correct for the acknowledged under-reporting to the Register, NACC research personnel devised a novel method for estimating the number of actual users from the Register cases. The method is described as follows:

"In the period January 1, 1964 through December 31, 1967, 46,400 different drug abusers were reported to the Register. Of this group, 38,751 were heroin users reported one or more times in this period. A recent check by the New York State Narcotic Addiction Control Commission of known addicts living on a slum block in Manhattan disclosed that 78 percent were listed in the Register. A check by the City Department of Health on persons whose deaths were certified by the Medical Examiner as due to narcotism revealed that 60 percent of the death certificates ascribing death to evidence of narcotics were listed in the Register. Thus, it seems reasonable from the aforementioned clues, to estimate that the Register is about 65 percent complete as a list of regular heroin users in New York City. From this, one can estimate that there are 58,500 heroin users in the City. Since 87 percent of the opiate users reported to the Register were heroin users, the 58,500 heroin users were readjusted for this proportion, and it is estimated that there were approximately 65,000 regular users of opiates in New York City at the end of 1967.

This estimate is based on a four-year period, 1964-67. In addition, a recent NACC Benchmark Survey Report indicated under-reporting in low drug density areas. Where observations and the assumption that those reported are more than occasional users and a large number who may have tried opiates have not been included, a higher figure for opiate users may be justified.[7]"

The above estimate of 65,000 opiate users was then used to calculate the rates per 1,000 population 15-44 years of age for the city as a whole, for each of the five boroughs and for the thirty Health Center Distrists into which New York City is divided. Construction of a rate based on the population 15 through 44 years of age appears reasonable in view of the fact that 97 percent of all heroin users listed in the Register from 1964-1967 were in this age group. To the extent that there is any difficulty with the rates, it arises primarily from two sources: 1) the correction factor for under-reporting and, 2) the use of "known addicts" in a four-year period as the numerator with an estimated population for one year, 1965, as the denominator.

The correction for under-reporting estimates that the Register contained only 65 percent of the opiate addicts for 1964-1967. The NACC research personnel recognized that the correction factor would have to be recomputed each year and that there is more under-reporting in low drug use areas than in high use areas such as the block from which the 78 percent figure was derived. It is the only attempt to correct for the known under-reporting, and probably as good as any especially when used for comparative purposes. There is, however, no way of knowing whether the absolute numbers derived from the rates for any single area are correct. They may be higher in some areas and lower in others.

A second source of difficulty may contribute to an over-estimation of the number of addicts. A purist might argue about the logic of the mathematics of rate construction using a 4-year numerator and a 1-year denominator. On the other hand, because of the nature of narcotics addiction, it might be argued that "known addicts" in 1964 are still addicts in 1965, 1966, and 1967 and therefore, legitimately part of the numerator, whether reported to the Register each year or not. There would still be a question of residence, however. A "known addict" with an address in the Lower East Side in 1964 may still be an addict in 1966, but with an address in the Bronx. And, it is possible that he may not even be an addict. So little is known about the process of "cure," especially if it is not in a program, that it is impossible to estimate the number of former addicts who are cured or "clean." Also, the original report to the Register may have involved minimal and/or temporary involvement with drug use, because of the nature of reporting to the Register. One would have to examine each case reported by the police, for example, to determine the range of involvement with drugs used by the police to define "user" for reporting purposes. The same is true for other reporting agencies.

The point of all this is that a rate is usually expressed for a single period of time. Thus, for example, to describe the crude birth rate for 1965, one would need to know the number of births in 1965, and the total population in 1965. When it comes to a rate of opiate use, however, the existing NACC estimate for New York City utilizes a count of users obtained over a four-year period. This is then corrected upwards for known under-reporting, and related to the population base for a single year.

To compute a meaningful rate per 1,000 of opiate users, one needs: 1) the number of known users in a given year and, 2) the estimated population for that same year. At present, there is no provision for removing drug names from the Register. Once listed, a name remains on the Register. Thus, at any time, the Register includes some "cured" addicts and some non-users who experimented briefly or on a single occasion in the past. A better "user" population is needed to construct a rate. If a report is made each time the user comes in contact with the law, social service agencies, and medical sources, it might be possible to get a somewhat more accurate unduplicated list of addicts, for a single year (regardless of whether they appear on the Register in previous years), correct that for under-reporting, and use it as the numerator in the calculation of a rate instead of the cumulative number in the Register. The research cited suggesting that non-addicts are sometimes categorized as addicts and placed in programs, may also make it necessary to correct for "over-reporting."

An alternative method for estimating rates might utilize New York City Police Department drug arrest figures for supplementary information. Narcotics arrests have increased dramatically over the last five years in New York City. The increase is due in part, to what appears to be a real increase in the drug abuse problem, but in part also to a redeployment of manpower and an increase in enforcement efforts by federal and city police agencies. The arrest data generated by the police department indicate the number of arrests, not the number of people arrested during the period covered. At present, data are not available indicating how many different people are involved in the total arrests report. Furthermore, as indicated previously, not all people arrested for narcotics offenses are addicts; and many addicts are arrested for non-narcotic offenses including theft, assault, and disorderly conduct.

Narcotics arrests of adults in the 89 police precincts in the city totaled 52,479 in 1970 compared to 35,178 in 1969 and 22,428 in 1968. The increase in felonies is especially notable in 1970 compared to earlier years. (Felonies represent arrests for hard drug abuse.)

The number of narcotics arrests each year represents a sizeable proportion of the number on the Register for the same years. The Police Department reports all narcotics arrests to the Register. Yet arrest reports are only one of several sources of case data for the Register. It would be valuable to determine how many of the arrests reported are duplicate cases (also reported by drug treatment agencies for example). Such a clarification is not yet available. Yet given a file of unduplicated names which contain all of the drug use reports on each case (whether from police, hospital, etc.), it might be possible to design a method for estimating how much activity is generated by each reported addict each year, and thereby improve somewhat the accuracy of current estimates of the size of the addict population in the City. This would not, of course, correct for all sources of error in the Register.

It appears that the present method of estimating narcotics abuse rates, while useful for comparative purposes such as ranking the various sub-communities within the City, most likely results in an over-estimation of the total <u>number</u> of known addicts in New York City when the absolute figures are utilized to specify the extensiveness of drug addiction. Whatever the numbers of addicts are, they are large and increasing. Whether they are the same numbers as those derived from existing data is open to question, and it is suggested that existing estimates be used cautiously.

One policy implication of this evaluation concerns the funding of neighborhood based drug treatment programs. Any increase or decrease in the number of drug treatment programs in a community such as New York City is likely to influence the number of addicts reported to any central register. Moreover, the rapid expansion of drug programs as alternatives to overcrowded criminal justice systems is also likely to increase the number of "addicts" as youngsters opt for drug treatment instead of jail.

When central registers are used to compare the number of "addicts" in the sub-communities of the City with a number of drug treatment programs in operation, or districts subject to intensive police surveillance, they may be expected to report more individuals to central registers than districts with few such programs or pressures. These same districts may then make a case for receiving more funds and programs than areas reporting fewer cases of drug abuse. Thus, to reiterate an earlier point, current estimates must be used with caution.

<u>Footnotes</u>

1. <u>New York Times,</u> Tuesday, June 6, 1972, p. 18.

2. Narcotics Register Statistical Report--1969. New York City Department of Health, Office of Research. (No date)

3. Narcotics Register Project Report, 1967-1968.

4. Zili Amsel, Carl L. Erhardt, Donald C. Krug and Conald P. Conwell in "The Narcotics Register: Development of a Case Register," paper presented at the 31st Annual Meeting of the Committee on Problems of Drug Dependence; Division of Medical Sciences, National Academy of Sciences--National Research Council. February 25, 1969.

5. <u>Ibid</u>., p. 2.

6. Emile J. Pin, Madeline H. Engel, Gerald Chasin, "Preliminary Statistical Analysis of Horizon Project Population in Treatment," Fordham University, Institute for Social Research, Horizon Project Evaluation, Internal Document, October 29, 1971, p. 15... For further elaboration, see Thomas M. Quinn, "The Impact of Law on the Heroin Distribution System, The Addict and the Community," Fordham University, Institute for Social Research, Horizon Project Evaluation, Internal Document, July 15, 1971.

7. New York State Narcotic Addiction Control Commission and New York City Narcotics Register, <u>Opiate Use in New York City</u>, pp. 3-4, November, 1969.

Bibliography

Amsel, Zili, Carl L. Erhardt, Donald C. Krug and Donald P. Conwell, "The Narcotics Register: Development of a Case Register," paper presented at the 31st Annual Meeting of the Committee on Problems of Drug Dependence; Division of Medical Sciences, National Academy of Sciences--National Research Council. February 25, 1969.

Establishing a Narcotics Register--Progress Report, 1967-1968. New York City Department of Health, Office of Research. (No date)

Kovaler, Florence, Paul M. Densen, Zili Amsel and Donald C. Krug, "Narcotics Register Project: Demographic Data, 1964-1966," paper presented at the Meeting of the Board of Health of New York City, February 15, 1968.

Narcotics Register Statistical Report--1969. New York City Department of Health, Office of Research. (No date)

Opiate Use in New York City. New York State Narcotics Addiction Control Commission and New York City Narcotics Register. Prepared by Mary Koval, NACC Division of Research. November, 1969.

TABLE A. NUMBER OF NARCOTICS ABUSERS REPORTED TO THE REGISTER FROM THE LOWER EAST SIDE, MANHATTAN AND NEW YORK CITY, BY YEAR OF FIRST REPORT

Year	Lower East Side (LES)	Manhattan	New York City	LES as a Percent Manhattan	LES as a Percent New York City
1964-1969	4,632	37,475	94,699	12.4	4.9
1964-1968	3,043	25,337	58,095	12.0	5.2
1969	1,589	12,138	36,604	13.1	4.3

Source: Narcotics Register Project, New York City Department of Health, 1969 Statistical Report.

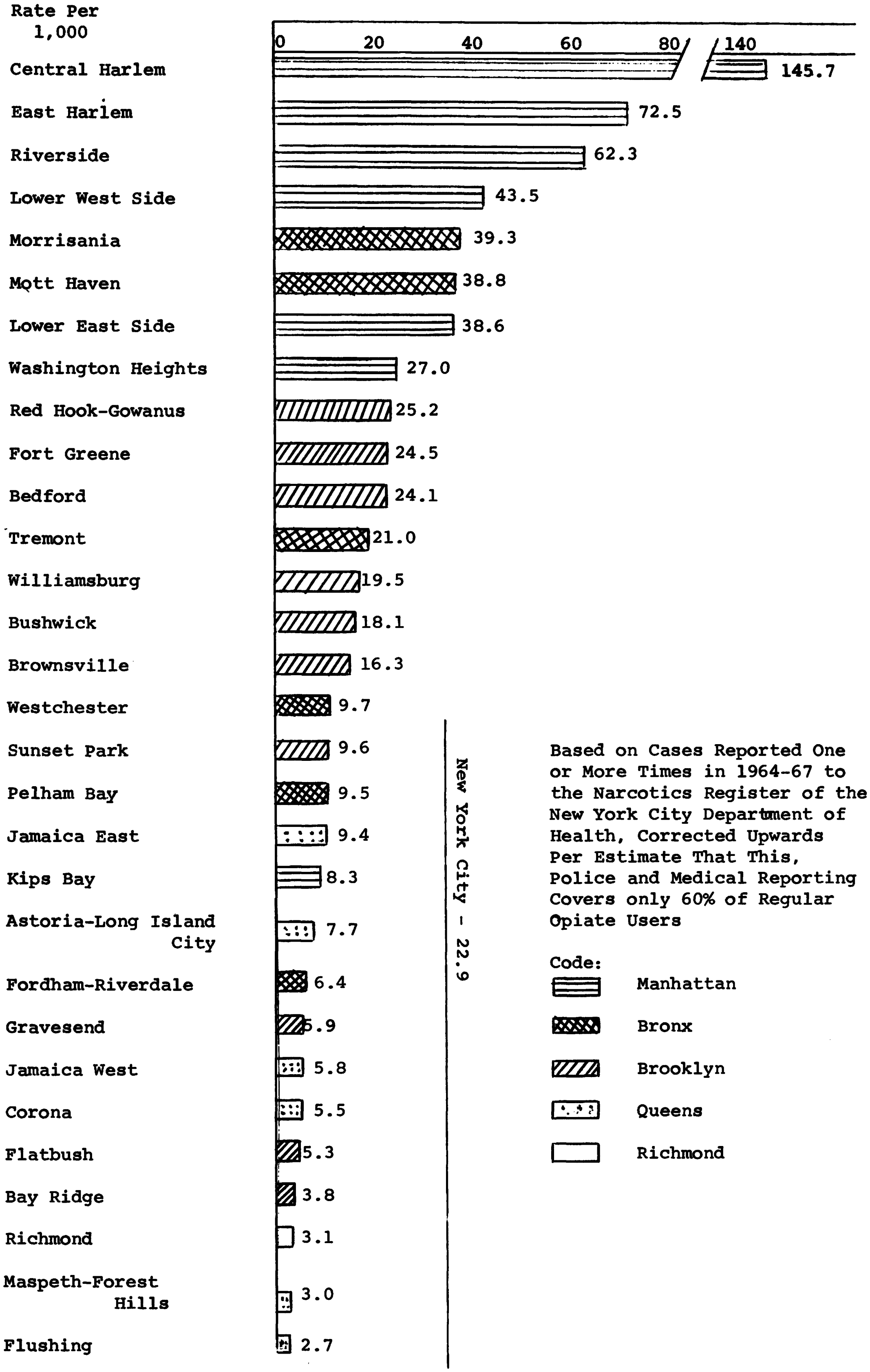
TABLE B. HEROIN ABUSERS REPORTED TO THE NARCOTICS REGISTER FROM THE HORIZON PROJECT AREA, 1964-1968 by ETHNICITY

	YEAR											
	1964 - 1968		1968		1967		1966		1965		1964	
Ethnicity	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Total	1,729	100.0	394	100.0	334	100.0	371	100.0	342	100.0	288	100.0
Puerto Rican	863	49.9	202	51.3	163	48.8	185	49.9	161	47.1	152	52.8
Other White	470	27.2	93	23.6	90	26.9	101	27.2	120	35.1	66	22.9
Negro	329	19.0	74	18.8	73	21.9	72	19.4	51	14.9	59	20.5
Other/Unk.	67	3.9	25	6.3	8	2.4	13	3.5	10	2.9	11	3.8

Source: Narcotics Register Project, City of New York, Department of Health, 1968.

GRAPH A

NEW YORK CITY RATES OF OPIATE USE PER 1,000 POPULATION, 15-44 YEARS OLD - RANKED BY HEALTH CENTER DISTRICTS



HORIZON PROJECT AREA

TABLE C. TOTAL NUMBER OF INDIVIDUALS REPORTED TO NARCOTICS REGISTER BY AGE AND ETHNICITY 1964-1968

Age and Ethnicity	1964 No.	1964 %	1965 No.	1965 %	1966 No.	1966 %	1967 No.	1967 %	1968 No.	1968 %	Total 1964-1968 No.	Total 1964-1968 %
Under 20												
Total	25	100.0	36	100.0	40	100.0	51	100.0	67	100.0	219	100.0
White	7	28.0	20	55.5	13	32.5	16	31.4	18	26.9	74	33.8
Negro	2	8.0	0	0.0	4	10.0	5	9.8	4	6.0	15	6.8
Puerto Rican	16	64.0	16	44.4	23	57.5	28	54.9	41	61.2	124	56.6
Other/Unknown	0	0.0	0	0.0	0	0.0	2	3.9	4	6.0	6	2.7
20 - 29												
Total	171	100.0	213	100.0	217	100.0	190	100.0	222	100.0	1,013	100.0
White	39	22.8	71	33.3	69	31.8	55	28.9	55	24.8	289	28.5
Negro	30	17.5	28	13.1	34	15.7	31	16.3	45	20.3	168	16.6
Puerto Rican	101	59.1	112	52.6	113	52.1	102	53.7	115	51.8	543	53.6
Other/Unknown	1	0.6	2	0.9	1	0.5	2	1.1	7	3.1	13	1.3
30+												
Total	73	100.0	83	100.0	109	100.0	86	100.0	80	100.0	431	100.0
White	15	20.5	24	28.9	18	16.5	19	22.1	13	16.3	89	20.6
Negro	24	32.9	22	26.5	32	29.4	35	40.7	21	26.3	134	31.1
Puerto Rican	26	35.6	30	36.1	47	43.1	28	32.5	43	53.7	174	40.4
Other/Unknown	8	11.0	7	8.4	12	11.0	4	4.7	3	3.7	34	7.9
Age Unknown												
Total	19	100.0	10	100.0	5	100.0	7	100.0	25	100.0	66	100.0
White	5	26.3	5	50.0	1	20.0	0	0.0	7	28.0	18	27.3
Negro	3	15.8	1	10.0	2	40.0	2	28.6	4	16.0	12	18.2
Puerto Rican	9	47.4	3	30.0	2	40.0	5	71.4	3	12.0	22	33.3
Other/Unknown	2	10.5	1	10.0	0	0.0	0	0.0	11	44.0	14	21.2

SOURCE: Narcotics Register Project, City of New York, Department of Health, 1968

TABLE D. NARCOTICS ARRESTS BY THE NEW YORK CITY POLICE DEPARTMENT IN NEW YORK CITY AND IN THE PRECINCTS IN THE HORIZON PROJECT AREA, 1968, 1969, 1970

Precinct and Type of Arrest	ARRESTS 1968	1969	1970	Percent Change 1968-69	Percent Change 1969-70
New York City					
Felonies	9,626	15,431	26,799	+60.3	+73.7
Misdemeanors	12,802	19,747	25,680	+54.2	+30.0
Total	22,428	35,178	52,479	+56.8	+49.2
5th Precinct					
Felonies	198	189	211	-4.5	+11.6
Misdemeanors	95	104	125	+9.5	+20.2
Total	293	293	336	0.0	+14.7
7th Precinct					
Felonies	208	177	322	-14.9	+81.9
Misdemeanors	277	271	286	-2.2	+5.5
Total	485	448	608	-7.6	+35.7
9th Precinct					
Felonies	642	938	1,286	+46.1	+37.1
Misdemeanors	299	699	574	+133.8	-17.9
Total	941	1,637	1,860	+74.0	+13.6
13th Precinct					
Felonies	80	126	193	+57.5	+53.2
Misdemeanors	86	154	151	+79.1	-1.9
Total	166	280	344	+68.7	+22.9

Source: Special tabulation prepared from New York City Police Department data.

DISCUSSION

Robert Parke, Social Science Research Council

In an essay entitled Toward Social Reporting: Next Steps, published by the Russell Sage Foundation, Dudley Duncan outlined two approaches to the study of social change. One approach is to think through the question of what ought to be measured and then figure out how to measure it. The other is to measure something we know how to measure and then figure out what the measures mean. Broadly, and with some injustice, I would classify the paper by Powers, Cullen and Martin on drug abuse estimates as an instance of the first approach and the papers by Yin and by Hill and Newman as instances of the second approach.

Yin settles upon fire alarms as an indicator series on the basis of purely statistical criteria, that is to say, the consistency of definition of observations, consistency and promptness of reporting, coverage of the universe, etc. On such grounds he regards fire alarms as worthy of extended statistical analysis simply because fire alarms make good statistics, not because they have any settled meaning for the social statistician, as yet.

By contrast, the work of Powers, Cullen and Martin is clearly in response to a pre-established necessity to come up with decent data on drug abuse. They claim, justly I believe, that the statistics describe a massive and increasing drug abuse problem in New York City and the lower East Side, and that, for comparative purposes as between subareas of New York City, their data are probably quite adequate. However, they note many problems with the data, some of which we recognize in other contexts as well. For example, Powers, Cullen and Martin note that the data from the narcotics register rise as a consequence of improved reporting by agencies as well as the result of the increase in use. The confounding of trend data by such processes is something we are familiar with in crime statistics. Second, the duplication and subsequent unduplication of reporting for individuals has the effect of making the revised data look like wholly new series. Third, the authors note the difficulties with the statistical definition of addiction. Addiction has many dimensions--sociological, psychological, physiological, and pharmacological--which make it impossible to establish a single statistical classification scheme for addiction. Here, it seems to me, the investigators are imposing an unreasonable requirement on their own work. It is not clear the development of summary indices presupposes unidimensionality in the basic statistics. Fourth, the authors note the difficulty of knowing whether the absolute rates of addiction are correct, while acknowledging the usefulness of the data they present for comparing addiction rates for different areas. Finally, they suggest that the need to use a multi-year numerator poses a problem. This is not necessarily the case. In vital statistics we are familiar with the practice of preparing rates by consolidating numerator data for several years, especially in the case of rare events.

As a partial approach to the resolution of some of these problems, Powers, Cullen and Martin suggest in passing the possibility of using New York City Police Department figures on drug arrests. One must wonder whether this is really a viable alternative. In effect, it amounts to taking a measure of the treatment of a condition as the operational definition of the prevalence of the condition. Biderman has pointed out the extent to which conventional crime statistics are measures of the performance of the police rather than the prevalence of crime. There is an inflationary dynamic built into statistics of this character which impairs confidence in estimates of the prevalence of the condition. Just as a marketing director distinguishes between sales and the degree of market penetration, we must maintain the distinction between treatment and potential treatment, which here means prevalence.

These considerations suggest two general comments. One has to do with the problems associated with data production for an undefined audience in a field where the measures are slippery. The other has to do with the nature of the task in social indicators. In determining what is an adequate measure, the statistician must ask "What difference does it make?" That is, what would anyone do differently if the number given out were a rather than b. How good does a number have to be? There is, in general-purpose data programs, a perfectionist imperative which frequently results in the statistician's imposing higher standards on the data than are necessary for all but a very few uses. The result--and I have written my share of statistical reports this way--is to label the data "Use with caution," and in effect to say, here are the numbers but don't believe them. The real question is what does the policy maker need to know?

I think it is fair to say, given the disappointment of Powers and her colleagues with the results of what appear to have been heroic efforts, that drug abuse is an area in which there are no general-purpose data. This would seem to be one area where the most important number the policy maker needs to know is the telephone number of the analyst who prepared the estimate. He may be unable to use the numbers intelligently otherwise. While I emphatically dissent from Mr. Yin's suggestion that it is the function of the statistician to devise new policy options, I do believe that in the drug abuse area it is likely that the policy maker needs the opportunity to test the meanings he imputes to the statistics against the judgment of the statistician. As a statistician, one must be concerned with the policy maker's possible misapprehension of the data. This is different, however, from taking responsibility for the policy maker's misuse of the data in the sense which Mr. Yin discusses. I refer to Mr. Yin's concern about the statistician serving as an instigator of repression in the instance where the policy maker's options are limited, and the indicators, and therefore the analyst who produced them, serving as "an unwitting partner of a fixed urban policy action."

These considerations lead me to a further suggestion about the relationship between social indicators and social reporting. In many of those interested in what is called the social indicators movement have thought of indicators development as an essential prologue in a process the outcome of which is social reporting on the state of the nation or the community with respect to some aspect of well-being. The problems in using highly imperfect data on highly important topics suggest grounds for selectively reversing the relationship here. The efforts of Powers, Cullen and Martin suggest that in the area of narcotics addiction we are dealing with a subject where indicators for general use are not likely to be forthcoming soon. It is equally clear that this is an area where a limited range of propositions can be supported by the available data. What the statistician can do in this instance is to present those propositions as the chief outcome of his endeavor, supported, of course, by the estimates suitably qualified. That this exegesis requires care goes without saying. I think Hill and Newman provide us with a good view of some of the potential pitfalls, in the area they are concerned with.

ON PROBLEMS OF ACCURACY AND COMPARABILITY OF INTERNATIONAL FERTILITY DATA

Arjun Adlakha, Food Research Institute, Stanford University

Introduction

When one consults the United Nations Demographic Yearbook, the major source of international population data, he finds a large supply of data on human fertility. One will find birth rates given for a large majority of the countries of the world. However, after going through the footnotes to tables, which often run to several hundreds for a single table, one finds that the apparent existence of reliable fertility data is an illusion. To anyone more than a casual observer, it becomes clear that for a majority of the world's population, data from vital registration statistics are so deficient that they cannot be used for statistical purposes.

For many countries parameters of fertility are estimates from sources other than vital registration systems. Most often these are derived from censuses or surveys with the aid of demographic techniques developed in recent years. In a few instances information on fertility is available from dual systems, i.e., data are collected through two independent sources and corrected for events missed in both sources.

This article briefly evaluates the current availability of fertility data and discusses some major problems of data comparability and adequacy. A complete analysis of the present availability of data should be based on a tabulation showing for each country of the world the kind of information available and how it is obtained. This would be a monumental task and is not undertaken here. Further, problems of comparability of data of less developed countries differ in nature from those of more advanced countries. In general, satisfactory basic data on fertility for more advanced countries are available, though a skilled demographer may not find for all countries the detailed classification he needs for sophisticated analysis. On the other hand, data on fertility for less developed countries are usually estimates derived indirectly and suffer from numerous deficiencies. The present paper gives more attention to the latter.

The paper is divided into two sections, one dealing with vital registration systems and the other with information on fertility obtained for several countries from sources other than the vital registration system. Discussion in this paper is confined to availability of data on the national level.

Vital or Civil Registration Systems

In the latest United Nations Demographic Yearbook, 1970, which lists more than 250 countries and territories, there are some data on births for a great majority of the areas, and for certain years. However, for a large number of countries, under-registration of births is on the order of 30 per cent or more. In 1970, with the admittedly lenient criterion of completeness of birth reporting (90 per cent or better is considered satisfactory), 66 per cent of the estimated world population lacks reliable birth registration. This is shown in Table 1. Around 1947 usable vital statistics were available for about 30 per cent of the world's population.[1] There has been only a little improvement in availability of birth data during the last two decades.

Deficiency in vital registration statistics is by no means randomly distributed among different geographical regions. As might be expected, availability of birth statistics and social and material development are correlated. In Table 1 data are presented by major divisions of the world. Almost 100 per cent of the population of Europe (including the USSR) and Northern America have reliable statistics on live births while only a small fraction of the population in other regions are so classified. Even such rudimentary information as total number of births is missing for about 2.4 billion people living in developing areas.

The availability of fertility data of satisfactory quality from civil registration systems by type of tabulation is given in Table 2. The table lists the types of tabulation and the aggregate population as well as per cent of total population for which each type of tabulation is available in various regions of the world. The contrast between the more developed and the less developed regions is startling with respect to the availability of satisfactory data for birth tabulations. While satisfactory data on total births are available for only a small percentage of the populations of developing nations, more detailed tabulations are even less widely available. In Africa, while about 6 per cent of the population is classified as having virtually complete birth statistics, only one-half of one per cent has a classification of births by age of mother. The situation in Asia, excluding Japan, is not very different from that of Africa. In Latin America the situation is only slightly better than in Asia and Africa.

This general lack of reliable data on births from vital registration systems for a large portion of the world population, however, is not the only problem. Even in countries classified as having virtually complete data there are problems of comparability. These arise from differences in procedure for evaluation of birth data and differences in recording procedure.

As stated above, the United Nations classifies data virtually complete for a country if coverage of births is at least 90 per cent complete. The basis for the evaluation is information accumulated from questionnaires, from direct correspondence, and from relevant official publications. However, there is considerable subjectivity involved, as in several instances the basis for a country's own evaluation is not provided. Even if one takes the evaluation results at face value, countries classified as having virtually complete data may vary in coverage between 90 and 100 per cent among themselves, as well as in the same country over time.

Despite the recommendation of the United Nations to adopt a uniform statistical definition of births and procedures for tabulation of births there exist differences among countries.[2] In many countries infants who are born alive but die before registration are not included in the live-birth registers, e.g. Algeria, Belgium, and France. In other countries the law specifies that an infant must survive for at least 24 hours before it can be considered for inclusion in the live-birth register, e.g. Ecuador and Spain. The effect of these factors is to underestimate total births.

In addition, many countries, due not to obduracy but to necessity, report data on births by date of registration,[3] rather than date of occurrence, as recommended by the United Nations. A lag between date of occurrence and date of registration of 20-25 years is not uncommon, though the majority of births are recorded within two to four years. Unless registration is prompt, such data are neither comparable internationally nor in the same country over time. For example, in Costa Rica, where data are tabulated by date of registration, it was observed that in recent years when registration improved the difference between births classified by date of occurrence and date of registration could be as high as 10 per cent. It is estimated that about 62,000 births occurred during 1961, while registered births in 1961 were 68,377.[4] The official birth rate, published in the United Nations' compilations for the same year, was derived from the latter figure and was therefore substantially exaggerated.

In brief, while vital registration systems provide satisfactory fertility data for most advanced countries, such data are missing for a majority of developing populations where problems of population are most salient and where fertility plays a prominent role in the population problems.

Other sources of Data on Fertility

In the absence of usable data from vital registration systems, the basic information used to estimate fertility parameters for most developing countries comes from censuses or surveys. The estimation of fertility from census or survey data has been feasible due to development of ingenious demographic techniques in recent years. There are a large variety of techniques, among them those known as the stable population technique, the survival ratio technique, the Brass technique, etc. On the basis of these techniques the Economic Commission for Africa and the Demographic Center for Latin America have provided estimates of fertility for most African and Latin American countries. Even more recently the use of such techniques has made it possible for the United Nations to provide plausible estimates of birth rates for countries having no usable official data on births.

The latest United Nations Population and Vital Statistics Report, which contains data available as of April 1, 1972, provides estimates of the birth rate for some 90 countries with no reliable registration data on births. For most of these countries the estimates are made by the United Nations. Details of computation for each country are not available; however, as explained in the Report these rates have been estimated from retrospective information obtained from sample surveys and censuses, by application of the "reverse survival" method and by other methods of analysis.[5] Earlier, the Population Division of the United Nations provided data on birth rates for 1950-55, 1955-60, and 1960-65 for a large number of countries, mainly using data from censuses and surveys with the aid of the reverse survival method and the stable population technique.[6]

Though these techniques have been useful tools in providing information which otherwise would not have been available, the information obtained in this manner is of limited utility. Several of the deficiencies in estimates obtained by the use of these techniques are illustrated below by detailing the two commonly-used techniques, stable population and survival ratio. The purpose of the discussion is merely to point out some problems in the application of the techniques and some limitations in the parameters obtained, not to describe the various techniques in use.[7]

Stable Population Technique

The stable population technique has been useful for obtaining estimates of the crude birth rates and--less often of other parameters of fertility for several countries without satisfactory vital statistics. For the proper application of the technique, the conditions required are that fertility and mortality in the population have remained unchanged in the long past and the population did not experience migration. If these conditions are met, estimates of fertility parameters for such a population can be obtained by comparing the age distribution of the population with a stable population which has a growth rate or mortality level identical to that of the population under consideration. Such comparison has been facilitated by the work of Coale and Demeny, who have constructed a series of stable population models[8] and also extended their use in situations of changing mortality but constant fertility in a population.[9]

There are several limitations of the estimates obtained in this manner. These are as follows:

First, as stated above, the technique is valid only if fertility in the long past has remained unchanged in the population under consideration. In several instances where the technique is used, it is difficult to ascertain directly whether this requirement is met since satisfactory information on fertility levels is not available. It is, therefore, generally assumed, rather than ascertained, that fertility in the population has been constant in the past.

Second, because of the problems of age misreporting and of age-selective coverage in the census or survey enumeration, the technique often yields a wide range of birth rate estimates within which the true value for the population probably lies. This can be seen from the data in Table 3, taken from the work of Zachariah on estimation of birth rates of Arab countries using the stable population technique.[10] The birth rate corresponding to the cumulative proportion of population at different ages and growth rates differs widely within each country shown in Table 3. A difference of more than 10 points between the lowest and highest value of birth rate for a country is not uncommon, while in several instances it is more than 15 points and in a few instances even greater than 20 points. When such large difference in estimates of the birth rate for a given country exist, any averaging procedure is likely to provide an estimate which could be far from the true value. Estimates obtained in this manner, therefore, provide only a crude picture of the level of fertility of a country as well as of differences in levels of fertility among countries.

Finally, the technique is useful for measuring levels but not for trends. If the objective is to study the change in the level of fertility of a population the technique cannot be used since such a population is not, by definition, stable.

To illustrate these problems explicitly, stable population estimates of birth rates for Pakistan as obtained by various scholars are shown in Table 4.[11] The data in Table 4 show that using the same technique and even the same data different persons arrive at different results. For example, using 1951 census data, Krotki estimates a birth rate of 55 for Pakistan. Using the same data, Ahmed's estimates of the birth rate for East Pakistan (now Bangladesh) and West Pakistan are 59 and 62 respectively, both higher than Krotki's estimate. When 1951 census results are used the estimated birth rate is higher in East Pakistan than in West Pakistan. Results from 1961 census data are in just the opposite direction. The only conclusion one can derive from these data is that fertility in Pakistan is high. The data fail to provide any precise information on how high it is or on the direction and magnitude of differences in level of fertility between East Pakistan (now Bangladesh) and West Pakistan.

Use of Census Survival Ratios in the Estimation of Fertility

When a country has had two censuses the birth rates can be obtained from the census survival ratios, that is the proportion of age-group cohorts surviving from one census to the next, plus additional information from censuses. Details may vary, but the commonly-used techniques involve obtaining a life table representative of observed survival ratios. A birth rate can be obtained by two alternative procedures. The one commonly used by the United Nations is the "reverse survival" method, i.e. increasing the number of children of a given age group, usually 0-4 years, recorded in a census, by the life table survival coefficient, in order to estimate the number of births from which these children are survivors. An alternative procedure applies the age-specific death rate from the life table to the observed population by age to get a crude death rate. The birth rate is obtained by adding this death rate to the rate of natural increase usually determined as the rate of population growth between the two censuses.

The use of survival ratios in obtaining birth rates has an advantage over the stable population technique because it is applicable even if fertility is changing. However, estimates of fertility obtained by this technique have limitations of a similar nature to those already described. Some of these are detailed below by illustrative data from Turkish and Brazilian populations which have had regular censuses and are representative of several Asian and Latin American populations.

The construction of life tables, or even the estimation of some level of overall mortality, by this procedure requires that two censuses should be accurate, or at least that errors in the two censuses are of the same magnitude, and that ages are correctly reported. Generally, data from developing countries do not satisfy either of these requirements. Coverage probably varies from census to census, and there are certainly problems of age misreporting. Survival ratios obtained from census data are usually deficient and require some adjustment or smoothing. For example, census survival ratios obtained from the Turkish censuses conducted in 1960 and 1965, shown in Table 5, provide results which are improbable. In certain ages the survival ratios are greater than unity, while at other ages they are surely too low. This pattern is not only common in Turkish populations but also is found in other Asian populations, notably India, Pakistan, and Indonesia. Improbable results are also noted in the case of females in Brazil; however, survival ratios based on data for males in Brazil are much more regular in pattern. Similar results are observed in other Latin American countries, i.e. problems of age-reporting are much less serious in the case of males than females.

Recently Demeny and Shorter and Coale and Demeny have suggested a new procedure which minimizes the effect of age misreporting. They suggest the use of cumulative survival ratios instead of survival ratios for single cohorts, i.e. computing the proportion of the population surviving from age "x and over" in one census to age "x + 10 and over" in a census taken 5 or

10 years later. This should be done for each age x from 0 to 40. Computations involve projecting the initial population by various model life tables to conform to the population enumerated in the succeeding census and locating a model life table corresponding to each age "x and over" where x takes values from 0 to 40. Various life tables so selected provide a range of mortality estimates which is considered as the range of errors introduced by age misreporting. The median life table, it is suggested, should be selected and considered as the best estimate of the mortality level, relatively free of age misreporting.[12]

Through the use of the procedure outlined above, values of e_0^o have been obtained for males and females from the Turkish census conducted in 1960 and 1965 and the Brazilian census conducted between 1960 and 1970 and are shown in Table 6. In spite of the fact that this method is better than using individual cohort survival ratios, the range of mortality level is still too wide to provide an unbiased estimate of the true level of mortality. When values of e_0^o show variations of about 15 years, as in Turkey and for females in Brazil, the median may not be representative of the true value.

Since estimation of birth rates depends upon the selected life table, errors in the latter are introduced in the former. In Table 7 we give estimated birth rates for Turkey and Brazil using the lowest, median, and highest levels of mortality in these countries as obtained in Table 6 by two different methods outlined at the beginning of this section. The detailed interpretation of this table is left to the reader; however, a few points are worth noting. When basic data are of poor quality, as in the case of Turkey, the possible range for the birth rate is very wide and estimated values differ widely when one uses different procedures. Even when basic census data are of reasonably good quality, as in Brazil, estimated values are quite different when one uses different assumptions.

It should be further noted that birth rates estimated by the reverse survival technique, with a few exceptions, are lower than birth rates obtained by growth rate and death rate. This finding is not unexpected. The estimated value of birth rate by the reverse survival method depends upon the accuracy of the census count of children. Serious underreporting of children in censuses is a common feature of less as well as more developed populations. Thus not only is comparability limited by the grossness of the estimates: it is also systematically biased by the methods used, some of which by this inherent weakness produce lower estimates than others.

The limitations of the techniques we have cited are known to demographers who use these data; however, they are explicitly brought out here. We have described only two of the many techniques, but other techniques have similar limitations. We have also limited our discussion to estimation of birth rate, though techniques have been and are used to extract more detailed information on fertility. This requires additional assumptions, which may introduce additional biases.

Concluding Remarks

Reliable data on fertility are available for only a small per cent of the world's population, mainly living in the more developed countries of the world. A source of non-comparability for this group of countries arises from differences in the availability of detailed data. But in so far as basic data on births and births by age of mother are concerned, these are available for most of these countries and can be considered as of reasonable quality for statistical purposes.

For a large per cent of the populations of developing countries reliable information on fertility is non-existent. The available information for these countries for the most part comes from retrospective data with the use of demographic techniques. The burden of this paper has been that information obtained in this manner is suspect and should be used with care for cross-cultural comparison or studying changes in fertility over time in one country. The techniques used have several limitations:

First, the techniques usually provide a range rather than a unique value of parameters to be estimated. Obtaining a precise figure requires some sort of averaging, dependent upon the choice of the researcher and sometimes on his preconceived notion of the plausible level of characteristics for the population under study. In fact, different scholars do arrive at different results nominally using the same data and the same methods for a population.

Second, the use of the techniques often requires making some general assumptions, e.g., population is stable, censuses have identical coverage, which may or may not be true for the population under study. The techniques also make use of models derived mainly from the experience of Western nations, which may or may not be valid for populations of developing countries.

Third, though techniques do provide some plausible estimates of levels they often fail to provide information on trends.

The Population Division of the United Nations has been doing an excellent job in publishing data for a large number of countries in a single volume. Yet there is a need for some modifications in the presentation of data so that they are not misused. First,the criterion for classification of data as "virtually complete" should be more rigorous than presently used. It is suggested that for countries where strong evidence exists that data are reasonably accurate the data should be given in a separate table. For the remaining countries data should be published without any code for completeness. This should include those countries for which there is evidence that data are probably accurate but evidence is either fragmentary or of dubious nature. These limitations may be explained in footnotes to the table. Second, estimated rates as published in the United Nations' volumes give an impression of false precision as they are given in decimal points. In addition to the average, the range of estimated value should be included so that the user of the volume becomes aware of the approximate nature of the estimates and uses them with care.

Demographers are usually aware of deficiencies in available data and have been careful in using them and arriving at conclusions from such data. However, on occasion the data are used indiscriminately. In the case of less developed countries users of the data should recognize (a) that for most countries birth rates are obtained indirectly from census data; (b) that the methods employed may build in systematic biases; and (c) that model life tables and stable population models are elegant methodologies that may substitute a spurious plausibility and regularity for the more complex reality that careful analyses can reveal. In addition the user should carefully go through the footnotes to tables to become acquainted with the deficiencies of the data.

Techniques, no matter how sophisticated, cannot substitute for good data. There is a need for improving basic sources of data as the existing information on fertility is inadequate for identifying changes in fertility in developing countries and for making valid projections of future changes in fertility--and hence for evaluating the achievement of family planning programs now under way in several countries. Responsibility for improving the vital registration system to make it capable of providing comparable and reliable information on fertility lies mainly with the governments. Progress in recent years has been disappointing, understandably due to real problems of organization and funds. But these difficulties can be minimized by proper organization of the system and by financial and technical assistance.

In the meantime, for several developing countries we will have to rely on data obtained from sources other than vital registration. Well-designed sample surveys, closely supervised with a skilled staff, can provide indispensable data on current levels and current trends in fertility. Finally, there is need to improve the analytical methods, perhaps most important, methods of estimating mortality for developing countries. Currently the procedure for estimating mortality in less advanced countries uses a model based mainly on the experience of European and more advanced countries. There is evidence that such models may not be applicable to currently developing countries[13] and a new model based on the experience of these countries should be constructed.

ACKNOWLEDGMENTS

The research on which this paper is based was carried out pursuant to Contract No. NIH 70-2189 with the National Institutes of Health, Department of Health, Education, and Welfare as part of a larger study at Stanford University on "Socio-economic Factors in the Reduction of Natality in the Less Developed Areas," under the direction of Professor Dudley Kirk.

FOOTNOTES

[1]Dudley Kirk, "Problems of Collection and Comparability of International Population Statistics," in book of same name, Milbank Memorial Fund, 1949, pp. 20-39.

[2]For recommendation of the United Nations, reference is made to United Nations, Principles for Vital Statistics Systems, Statistical Papers Series M, No. 19, New York, 1953. For greater detail on deficiencies in registration data see United Nations, Report of the Seminar on Civil Registration and Vital Statistics for Asia and the Far East, Statistical Papers, Series M, No. 50, New York, 1970, and Demographic Yearbook 1969.

[3]Tunisia, Antigua, Barbados, British Honduras, British Virgin Islands, Costa Rica, Dominica, Grenada, Jamaica, Mexico, Montserrat, St. Kitt's-Nevis-Anguila, St. Vincent, Guyana, Ceylon, China (Taiwan), Hong Kong, Jordan, Ireland, Isle of Man, Northern Ireland, Scotland, Australia, Cook Islands, Fiji, New Zealand, Norfolk Island. All these territories are characterized as having "virtually complete" data.

[4]Direccion General de Estadistica y Censos, Principales, Hechos, Vitales Occurridos en Costa Rica, Depto. de Estadisticas Sociales, Seccion de Estadistica Vital, No. 28, 1963, p. 18.

[5]United Nations, Population and Vital Statistics Reports, Data available as of 1 April 1972, ST/STAT/Ser.A/100, V. XXIV, N. 2, 1972.

[6]United Nations, Estimates of Crude Birth Rates, Crude Death Rates, and Expectation of Life at Birth, Regions and Countries, 1950-1965, ESA/P/WP/38, Feb. 22, 1971.

[7]A succinct summary of the techniques in use is provided by Ansley Coale, "The Determination of Vital Rates in the Absence of Registration Data," The Milbank Memorial Fund Quarterly,XLIX, 4:2, Oct. 1971; for detailed description and application of technique refer to Ansley Coale and Paul Demeny,Methods of Estimating Basic Demographic Measures from Incomplete Data, United Nations, Manual IV, New York, 1967.

[8]Ansley J. Coale and Paul Demeny, Regional Model Life Tables and Stable Populations, Princeton, New Jersey, Princeton University Press, 1966.

[9]Ansley J. Coale and Paul Demeny, "Estimating Vital Rates for Populations in Process of Destabilization," Demography 2, 1965, 516-530.

[10]K.C. Zachariah, "The Demographic Measures of Arab Countries, A Comparative Analysis," in Demographic Measures and Population Growth in Arab Countries, Cairo Demographic Center, Research Monograph Series, No. 1, Cairo, 1970, pp. 279-326.

[11]These are cited in Sultan S. Hashmi, "Fertility Studies in the Pakistan Institute of Development Economics, Karachi," in Minoru Tachi and Minoru Muramatsu (eds.), Population Problems in the Pacific: New Dimensions in Pacific Demography, 1971 (Proceedings of the Congress Symposium No. 1 and Divisional Meeting of Section VIII, No. 5, 11th Pacific Science Congress, Tokyo, August 22-September 10, 1966).

[12]Paul Demeny and Frederic Shorter, Estimating Turkish Mortality, Fertility, and Age Structure: Application of Some New Techniques, Istanbul Faculty of Economics, University of Istanbul, Publication No. 218, 1968.

[13]Arjun Adlakha, "Model Life Tables: An Empirical Test of their Applicability to Less Developed Countries," to be published in Demography, November 1972 issue.

Table 1.--Estimated Total Population, Population With and Without Reliable[a/] Birth Date, and Per Cent of Population Without Reliable Birth Data as of 1970 by the Major Regions of the World

Population in Millions

	Total	With reliable birth data	Without reliable birth data	Percentage without reliable birth data
World	3638	1230	2408	66.3
Africa	344	21	323	93.9
Western	101	0.3	100	99.7
Eastern	98	1.6	96.4	98.4
Northern	87	19.3	67.7	77.8
Middle	36	0.6	35.4	98.3
Southern	23	0	23	100.0
Northern America	227	227	0	0.0
Latin America	285	105.4	180	63.1
Tropical	153	1.2	151.5	99.2
Middle	67	62.5	4.7	7.0
Temperate	39	34.1	5.3	13.5
Caribbean	25	7.8	17.6	69.2
Asia	2056	152	1904	92.6
East Asia	930	1221	817.9	88.0
South Asia	1126	303	1095.7	97.3
Europe	462	462	0	0.0
Oceania	19.4	18.6	0.8	4.1
USSR	243	243	0	0.1

Source: United Nations, Demographic Yearbooks, 1969 and 1970.

a/ Complete or virtually complete data as reported to the Statistical Agency of the United Nations. Some countries in Africa have complete data for certain ethnic groups. In these calculations such countries are considered as lacking complete data. Inclusion of these countries, however, does not make any substantial difference in the overall situation.

Table 2.--Aggregate Population and Per Cent of Total Population for which Different Types of "Virtually Complete" Data from Vital Registration Systems are Available, by Major Regions of the World

Population in Millions

Type of information	Africa Population	Africa %	Asia Population	Asia %	Latin America Population	Latin America %	North America Population	North America %	Europe Population	Europe %
Number of live births	21	6.1	152	7.4	105	36.9	227	100.0	462	100.0
Birth rate[a/]	21	6.1	152	7.4	105	36.9	227	100.0	462	100.0
Live births by age of mother	2	0.5	151	7.3	104	36.7	227	100.0	460	99.6
Live birth ratesby age of mother[b/]	2	0.5	151	7.3	74	26.1	227	100.0	427	92.4
Live births by age of mother & birth order	2	0.5	119	5.8	29	10.2	227	100.0	426	92.2
Live birth rates by age of mother & birth order[c/]	2	0.5	110	5.4	23	5.1	227	100.0	164	35.5

Source: United Nations, Demographic Yearbooks 1969 and 1970

a/ Birth rate is the number of live births reported for a calendar year per 1,000 persons in the same area at the midpoint of the year.

b/ Rates specific for age of mother are the number of live births in each age group per 1,000 female population in specified age group.

c/ Rates are the number of live births in each birth order and age group per 1,000 total female or married female population in the specific age group.

Table 3.--Estimates of Crude Birth Rate by Stable Population Method Using Cumulative Age Distribution and Growth Rate*

		Growth Rate	Age 5	10	15	20	25	30	35	40	Range	Difference
Morocco	M	27	48	59	54	42	40	41	42	43	40-59	19
	F	27	52	58	45	36	37	43	47	48	36-58	22
Algeria	M	26.9	55	57	65	61	54	51	51	49	49-65	16
	F	27.2	52	53	55	53	50	49	51	50	49-55	6
Tunisia	M	27	48	53	58	52	45	43	43	43	43-58	15
	F	27	48	51	54	49	46	46	48	50	46-54	8
Libya	M	27[a/]	45	48	47	39	39	41	42	42	39-48	9
	F	27	49	54	48	43	43	48	48	50	43-54	11
UAR	M	24.7	38	45	50	46	43	41	42	44	38-50	12
	F	22.5	38	43	45	41	40	41	42	46	38-46	8
Syria	M	27	57	63	66	55	50	50	51	50	50-63	13
	F	27	52	53	51	47	47	49	48	49	47-53	6
Jordan	M	27	48	50	60	64	61	60	56	52	48-64	16
	F	27	43	43	46	52	52	54	53	54	43-54	11
Iraq	M	27.5	54	60	54	46	42	40	39	39	39-60	21
	F	26.2	51	53	47	44	41	42	43	42	41-53	12

* K.C. Zachariah, "The Demographic Measures of Arab Countries: A Comparative Analysis," in Demographic Measures and Population Growth in Arab Countries, Cairo Demographic Center, 1970, p. 302.

a/ Assumed growth rate.

Table 4.--Estimate of Crude Birth Rate for Pakistan Using Stable Population Technique

Source	Data used	East Pakistan	West Pakistan	Total
Krotki	1951 census			55
Ahmed	1951 census	59	62	
Krotki	1961 census	58	51	
Planning Commission	--	57	52	
Khan	1962 PGE			49
Zelnik and Khan	1962 PGE	38-68	37-51	
Robinson	1962 PGE	54	53	

Source: Sultan S. Hashmi, "Fertility Studies in the Pakistan Institute of Development Economics, Karachi," in Minoru Tachi and Minoru Muramatsu (eds.), Population Problems in the Pacific, New Dimensons in Pacific Demography, Proceedings of the Congress Symposium No. 1 and Divisional Meeting of Section VIII, No. 5, 11th Pacific Science Congress, Tokyo, August 22-September 10, 1966, 1971, pp. 113-114.

Table 5.--Survival Ratios of Males and Females in Turkey from 1960 and 1965 Censuses and in Brazil from 1960 and 1970 Censuses

	Turkey 1960-65[a/]		Brazil 1960-70[b/]	
Age	Males	Females	Males	Females
0-4	1.109	1.086	1.024	1.061
5-9	0.993	0.941	0.957	1.054
10-14	0.916	0.919	0.948	1.017
15-19	0.984	1.079	0.930	0.905
20-24	0.938	1.040	0.964	0.927
25-29	1.027	0.970	0.976	0.931
30-34	0.994	0.952	0.978	0.934
35-39	0.954	0.978	0.963	0.932
40-44	0.889	0.785	0.906	0.922
45-49	1.006	1.183	0.862	0.900
50-54	0.820	0.714	0.822	0.848
55-59	1.067	1.358	0.759	0.801
60-64	0.700	0.677	0.477	0.545

a/ Five years survival ratios
b/ Ten years survival ratios

Table 6.--Estimated Value of Expectation of Life (e_0^o) for Turkey 1960-65 and for Brazil 1960-70 by the Use of Census Survival Ratio

Age (x & over)	Turkey 1960-65 Males	Turkey 1960-65 Females	Brazil 1960-70 Males	Brazil 1960-70 Females
0	64.3	64.3	61.0	68.0
5	47.2	50.0	56.5	62.5
10	45.9	55.2	58.0	57.5
15	51.6	63.3	59.6	54.3
20	51.6	55.5	62.0	57.5
25	58.4	50.0	62.4	59.6
30	53.4	50.3	62.0	61.2
35	51.0	52.5	60.9	62.5
40	52.3	51.8	59.5	63.5
Median	51.6	52.5	60.9	61.2
Range	45.9-64.3	50.0-64.3	56.5-62.4	54.3-68.0

Estimates of e_0^o are obtained from the census data by the use of the Demeny and Shorter techniques discussed in the text: South and West Regional Model Life Tables of Coale and Demeny were used in Turkey and Brazil respectively.

Table 7.--Birth Rates in Turkey 1960-65 and Brazil 1960-70 Obtained from Census Survival Ratios Using Lowest, Median, and Highest Estimates of Mortality, Intercensal Growth Rate and Reverse Survival Method

Mortality Estimates	Turkey 1960-65: Birth rate based on estimates of mortality derived from data of Males[a/]	Females[a/]	Both sexes	Brazil 1960-70: Birth rate based on estimates of mortality derived from data of Males[b/]	Females[b/]	Both sexes
			With the Use of Intercensal Growth Rate			
Lowest	33.8	34.8	34.2	35.0	35.8	35.4
Median	39.4	41.2	40.3	36.1	38.8	37.5
Highest	44.2	42.6	43.4	37.9	42.4	40.2
			Reverse Survival Technique			
Lowest	34.0	34.8	34.3	35.4	35.5	35.5
Median	36.2	37.3	36.8	36.0	37.1	36.5
Highest	38.5	38.0	38.2	37.0	39.0	38.0

a/ Assumes a sex ratio at birth of 106.

b/ Assumes a sex ratio at birth of 105.

PROBLEMS OF INTERNATIONAL COMPARABILITY
IN INTERNAL MIGRATION RESEARCH

Barbara A. Anderson, Princeton University

Any discussion of problems of comparability of demographic data can concentrate on one of three types of problems: 1) problems of data collection, 2) analytic data manipulation to achieve comparable data, or 3) types of desirable comparisons. This paper concentrates on the second of these problems. Analytic method of obtaining comparable estimates of migration streams is described and discussed.

A major problem in the international comparison of internal migration is obtaining sufficiently detailed and accurate estimates of migration streams from data which is readily available for a large number of countries. The addition of a special question, even if the question is very good not only is very expensive but immediately raises serious questions of comparability with estimates from countries not using that question. Of course, no question can be added to a census which was taken in the past, and thus additional questions are useless in obtaining historical estimates.

In this paper I examine one method which shows some promise of providing such data. Focussing on a method also provides an opportunity to discuss some generally applicable questions of comparability and other methodological considerations.

Rogers and von Rabenau[1] have suggested a method of estimating internal migration streams which is appealing in its simplicity and in the availability of the data which it requires. Place of birth by sex by age by place of residence data from two censuses are used. Although data often are not tabulated in this form, such information practically always is collected and usually could be tabulated more easily than a special question could be added. Also, such tabulations might, unlike special questions, be obtainable for earlier censuses. This method is somewhat similar to the Grabill-Cho method of fertility estimation from questions about own children.[2] Both methods involve a retabulation of data which are often available, and both are of little use in areas where people know neither their age or date of birth. Unless there have been recent major boundary changes, misreporting is likely to be a more minor problem with this data than with responses to a question about place of residence a certain number of years earlier.

The basic assumption is that migration between regions which partition a country is a Markov process, for which, within an age-sex group, the probability distribution of destinations in the next time period depends solely on location in the present time period and is independent of location at any earlier time, especially place of birth. The validity of this assumption is discussed later.

Streams are estimated from the simultaneous equations shown in Table 1. Ignoring mortality and international migration, these equations are based on the observation that a person in some region at time (t+1) had to be in some region at time t. Ideally, for considering migration between n regions of a country, persons in that country of a given age and sex would be randomly assigned an integer ranging from 1 to n, and a person's assigned integer could not change. It would be a label which would stay with him and would be independent of migration probability. This would give the equations in Table 1 where k is the assigned integer. Since there are n^2 unknowns, the n^2 streams to be estimated, one needs n^2 equations. This means that there must be n categories of any labelling variable. Thus, using the same n regions for place of birth regions as for migration regions makes place of birth a natural label to use. If only two regions were used, both sexes could be studied together and sex could be the labelling variable. Both place of birth and sex are appropriate labels since they do not change throughout an individual's life. Unfortunately, neither is likely to be independent of migration probability. This problem will be discussed later.

In this paper, for simplicity, only two regions are considered. This can be extended to any number of regions. For two regions, disregarding for the moment mortality and international migration, the estimation procedure is shown in Table 2. The censuses are assumed to be ten years apart, and ten-year age groups are used, but any time interval and any age span which divided that interval could be used. Usually only one sex at a time is studied. Any other unchanging characteristic, such as race could be used to further divide the population.

Mortality and international migration would affect internal migration estimates using this or most other techniques. I recommend following Thomas Burch's suggestion and assuming that a person's mortality risk is dependent on his place of birth rather than on his place of residence. Besides being conceptually neat, since a person's place of birth cannot change, there is evidence that a person's mortality risk is strongly influenced by where he spent his first fifteen years.[3]

One method of handling mortality and international migration is to consider the conditional probability of migrating, given that a person is alive and in the country at time 2. To do this, the native population of each region at Time 1 is multiplied by a scale factor to make its size equal the size of the population native to that region at Time 2. This scale factor, then, is an estimate of the survival ratio. This procedure is shown in Table 3.

The other way to proceed is to include the probability of being dead or out of the country at Time 2 along with the migration probabilities. In this case, an estimated migration probability is the product of the scale factor for the appropriate region of birth and the conditional migration probability for the appropriate origin and destination. This procedure is shown in Table 4.

Naturally the unconditional and conditional probabilities differ. An example is shown in Table 5. The example is for Canadian females native to Canada who were aged 15-24 in 1951.[4,5] As expected, the unconditional migration probabilities are smaller than the conditional probabilities and among the unconditional probabilities, the smaller the scale factor, the smaller the migration probabilities.

International immigrants who were born in the country present a different problem than international emigrants native to the country, since immigrants were not in the base population at Time 1. Usually there are few such persons, and they are concentrated in the older age groups. Often they are people who have returned to their native land to retire. However, in countries such as Spain or Turkey, where persons frequently go to another country to work only to return in a few years, the problem is more serious. The major effect of native born persons present at Time 2 but not at Time 1 is to increase the scale factors. It is also important when comparing migration rates for two countries to know whether the censuses asked de jure or de facto residence.

Different considerations determine whether the conditional or unconditional probabilities should be used. Generally one should use the conditional probabilities when comparing two countries. Otherwise, one might believe that one country had greater internal migration simply because it had lower mortality.

For comparisons between regions of a country the unconditional rates should also be studied. The comparison of the scale factors with the values predicted from appropriate life tables gives an indication of the extent of international migration.[6] In the example in Table 5, the comparison of the scale factors with the life table values indicates that Ontario experienced a lower rate of international emigration than the rest of Canada. If it is known that emigration is light, then this comparison gives an indication of the quality of the data.

Within one country, migration probabilities from various regions may be used as an indicator of some attitude, such as susceptibility to change. In that case, international migration should also be considered. Otherwise, one might erroneously believe that the people in a particular region were very geographically immobile, when actually a large proportion was moving out of the country. This would especially be true of regions including large port or other border cities. This touches on a major problem in comparing migration probabilities both within and between countries when regions differ in size, shape, topography, and centrality in the country. Due to time limitations, that problem will not be pursued further here.

The essential difference between the two approaches is while conditional probabilities are concerned with the location at Time 1 of those who are in the country at Time 2, the unconditional probabilities concentrate on the distribution at Time 2 of those in the country at Time 1. This is somewhat similar to the difference between the total fertility rate and the crude birth rate. Like the total fertility rate, the conditional probabilities reflect the level and over time, the change in a basic process, uncontaminated by mortality. The unconditional probabilities, like the crude birth rate, may not have as great a theoretical significance as the former measure, but they may be more useful in economic models not concerned with the theories of demographic dynamics.

There are problems with both ways of handling mortality and international migration. In each case, the two are assumed to depend on the place of birth. This is fine if the loss between the censuses is entirely due to mortality. However, it is unlikely that international migration depends on place of birth, while internal migration depends on place of residence. The extent of the problem depends on the size of international migration compared to internal migration and loss due to mortality. This is one factor which introduces error into estimates.

The usefulness of this method rests on its robustness under violation of its basic assumption and its insensitivity to misreporting.

There are strong arguments that the assumption that migration probability is dependent on place of residence and independent of place of birth is not true. An extensive literature presents evidence that people may be divided into movers and stayers.[7] This means that a person who moves in one time period is more likely to move in the following time period than a person who did not move in the first time period.

A person's place of birth by residence category is unfortunately highly relevant to his mover-stayer status, since a person who does not live in his region of birth has moved at least once and thus is more likely to be a "mover" than a person who lives in his region of birth.

On the other hand, there may be regions in a country, perhaps highly urban regions, where the socialization process makes people highly susceptible to becoming movers, perhaps due to an awareness of the importance of taking advantage of economic opportunity. In that case, people who were born in such regions and spent their childhood there would be unusually likely to migrate, regardless of where they lived later.

Table 6 shows the robustness of this method to the violation of the basic assumption of independence of place of birth and migration probability but not of misreporting of age, mortality or international migration. T_{ij} is the average of the probability that a person born in region 1 and living in region i at Time 1 will be in region j at Time 2 and that probability for persons born in region 2. Thus T_{ij} is the average of the "true P_{ij}" values for persons born in the two regions. For example, $(T_{11}+u)$ is the probability of a person in group A being in region 1 at Time 2, while $(T_{11}-u)$ is the probability that a person in group B will be in region 1 at Time 2. If both u and v are positive, then we have a mover-stayer model. If they are opposite in sign, then people born in one region are far more likely to migrate than persons born in the other region. If they are both negative, there is a reversed mover-stayer model.

It would be nice if the absolute value of P_{ij} fell between the absolute value of T_{ij} plus or minus the absolute value of its deviation. That is, $(|T_{12}| - |u|) < P_{12} < (|T_{12}| + |u|)$.

This can be true only if u and v are opposite in sign. However, if u and v have the same sign, then the estimation of net migration will be more accurate than if they had opposite signs. This can be seen from the examples in Table 7. Notice that the size of the discrepancy between T_{ij} and P_{ij} depends only on the distribution of the population between regions at the first time and does not depend on the pattern of migration. In general, the more equal the two regions are on size and the greater the difference between the product of those living in their region of birth and of those not living in their region of birth, that is (AD-BC), the less the discrepancy.

It is also apparent from the examples that the error can be quite large, making the estimated P_{ij} greater than one or less than zero. This may cause doubts about the method as a whole. However, something about the magnitude of the discrepancy can be computed from the data used. For example using the Canadian data in Table 5 to substitute into Eq. 21 yields: $P_{OC} - T_{OC} = -(1.004u + .085v)$ and $P_{CC} - T_{CC} = (.097u + 1.004v)$ If in the case of the data used the coefficients of u and v were very large, then one would be reluctant to use the P_{ij} estimates. If the coefficients were small, though, one would know that the estimates for that particular data set were not strongly affected by small deviations from the basic assumption. Also if one knew there were not serious problems of coverage in the data, and if the coefficients were small, then unreasonable P_{ij} estimates would provide evidence that the basic assumption of this method was severely violated. This would be of theoretical interest.

It is also important how sensitive this method is to enumeration errors. It is easy to show that if two censuses differ in completeness of coverage, but if groups are not differentially covered according to region of birth or residence, then the conditional probabilities are unaffected. However, the unconditional probabilities would be affected, since the scale factors would be incorrect. If the two regions differ in completeness of coverage, but if within regions the completeness is the same in both censuses, then P_{ij} estimates will be affected primarily by the completeness of enumeration in region j. It is likely that both types of misenumeration will occur. Since the group studied is ten years older at the second census, and enumeration completeness usually differs by age group, the first type of problem is likely to occur. The second problem is also common. Usually recent migrants are underenumerated since they have lived in their community for a relatively short time and they may be transients who have no permanent residence.

In summary, this method is useful where age is known, where there have not been recent major boundary changes, and where regions do not differ greatly in population. It is especially useful where censuses differ in completeness since it is not a subtraction method. As for use, those concerned with economic projection will probably use the unconditional probabilities, while theorists concerned with the causes of migration would be more likely to study the conditional probabilities.

References

[1] Rogers, Andrei and von Rabenau, Burkhard, "Estimation of Interregional Migration Streams from Place-of-Birth-by-Residence Data," Demography, 8 (May 1971), 185-194

[2] Grabill, Wilson H. and Cho, Lee Jay, "Methodology for the Measurement of Current Fertility from Population Data on Young Children," Demography, 2 (February 1965), 50-73

[3] Burch, Thomas K., "Internal Migration in Venezuela: A Methodological Study," Unpublished Ph.D. dissertation, Princeton University 1962, 59-60

[4] Dominion Bureau of Statistics, Census of Canada 1951, Volume II, Cross-Classification of Characteristics, Ottawa 1953, Table 11

[5] Dominion Bureau of Statistics, Census of Canada 1961, Volume I, Part 3, Birthplace and Citizenship by Age Groups, Ottawa 1963, Table 90

[6] George, M.V., Internal Migration in Canada, 1961 Census Monograph Series, Dominion Bureau of Statistics, Ottawa 1970, 41

[7] Goodman, Leo, "Statistical Methods for the Mover-Stayer Model," Journal of the American Statistical Association, 56 (December 1961), 841-868 This is a discussion of some of the statistical aspects of the model.

Table 1
Basic Estimating Equations

$$\sum_{i=1}^{n} P_{i1}\,{}^{1}A_{it} = {}^{1}A_{1(t+1)},\ \sum_{i=1}^{n} P_{i1}\,{}^{2}A_{it} = {}^{2}A_{1(t+1)},\ \ldots\ \sum_{i=1}^{n} P_{i1}\,{}^{n}A_{it} = {}^{n}A_{1(t+1)} \qquad \text{Eq. 1}$$

$$\vdots \qquad\qquad\qquad\qquad \vdots$$

$$\sum_{i=1}^{n} P_{in}\,{}^{1}A_{it} = {}^{1}A_{n(t+1)},\ \ldots\ \sum_{i=1}^{n} P_{in}\,{}^{n}A_{it} = {}^{n}A_{n(t+1)}$$

Where

P_{ij} = probability that a person in region i at time t will be in region j at time (t+1)

${}^{k}A_{it}$ = the number of persons born in region k who are in region i at time t

n = the number of regions

Table 2
Solution of Basic Estimating Equations

	In Region 1		In Region 2	
	Born Reg 1	Born Reg 2	Born Reg 1	Born Reg 2
Time 1 Aged x to x+10	A	B	C	D
Time 2 Aged x+10 to x+20	E	F	G	H

Capital letters represent the number of persons in each category.

From those born in Region 1, $P_{12}A + P_{22}C = G$ Eq. 2

From those born in Region 2, $P_{12}B + P_{22}D = H$ Eq. 3

Since destination probabilities at time 2 sum to 1 for each location at time 1,

$P_{11} = 1 - P_{12}$ and Eq. 4

$P_{21} = 1 - P_{22}$ Eq. 5

Solving,

$P_{12} = (GD-HC)/(AD-BC)$ Eq. 6

$P_{22} = (HA-GB)/(AD-BC)$ Eq. 7

$P_{11} = 1-(GD-HC)/(AD-BC) = (ED-FC)/(AD-BC)$ Eq. 8

$P_{21} = 1-(HA-GB)/(AD-BC) = (FA-EB)/(AD-BC)$ Eq. 9

Table 3
Conditional Migration Probabilities

	In Region 1		In Region 2	
	Born Reg 1	Born Reg 2	Born Reg 1	Born Reg 2
Time 1 Aged x to x+10	A(E+G)/(A+C)	B(F+H)/(B+D)	C(E+G)/(A+C)	D(F+H)/(B+D)
Time 2 Aged x+10 to x+20	E	F	G	H

S_1 = Scale factor for those born in Region 1 = (E+G)/(A+C)

S_2 = Scale factor for those born in Region 2 = (F+H)/(B+D)

Solving using Eq. 6&7,

$P_{12} = [GD(F+H)(A+C)-HC(E+G)(B+D)]/[(F+H)(E+G)(AD-BC)]$ Eq. 10

$P_{22} = [HA(B+D)(E+G)-GB(F+H)(A+C)]/[(F+H)(E+G)(AD-BC)]$ Eq. 11

Then $P_{11} = 1-P_{12}$, $P_{21} = 1-P_{22}$

Table 4
Unconditional Migration Probabilities

Let ${}_{k}P_{ij}$ = the unconditional probability that a person born in Region k who is in Region i at Time 1 is in Region j at Time 2

Let P_{ij} = the conditional probability that a person who is in Region i at Time 1 is in Region j at Time 2, conditioned on him being alive and in the country at Time 2

Then,

${}_{1}P_{12} = S_1P_{12} = [(E+G)/(A+C)]P_{12}$ Eq. 12

${}_{1}P_{22} = S_1P_{22} = [(E+G)/(A+C)]P_{22}$ Eq. 13

${}_{2}P_{12} = S_2P_{12} = [(F+H)/(B+D)]P_{12}$ Eq. 14

${}_{2}P_{22} = S_2P_{22} = [(F+H)/(B+D)]P_{22}$ Eq. 15

${}_{1}P_{11} = 1-[(1-S_1)+S_1P_{12}] = S_1P_{11}$ Eq. 16

${}_{1}P_{21} = 1-[(1-S_1)+S_1P_{22}] = S_1P_{21}$ Eq. 17

${}_{2}P_{11} = 1-[(1-S_2)+S_2P_{12}] = S_2P_{11}$ Eq. 18

${}_{2}P_{22} = 1-[(1-S_2)+S_2P_{21}] = S_2P_{22}$ Eq. 19

Table 5
Example Showing Conditional and Unconditional
Migration Probabilities
Canadian Females Aged 15-24 in 1951

	In Ontario		In Other Canada	
	Born Ontario	Born Other Canada	Born Ontario	Born Other Canada
1951 Aged]5-24	287,568	36,212	12,176	750,366
1961 Aged 25-34	274,889	63,375	23,850	713,190

Total born in Ontario 1951=299,744
Total born in Ontario 1961=298,739
Thus S_O=.99664

Total born in rest of Canada 1951=786,578
Total born in rest of Canada 1961=776,565
Thus S_C=.98727

Life table survival value for Ontario=.99313
Life table survival value for rest of Canada=.99030

Conditional Migration Probabilities
P_{OC}=.04254
P_{OO}=.95746
P_{CO}=.03934
P_{CC}=.96070

Unconditional Migration Probabilities
${}_OP_{OC}$=.04225
${}_CP_{OC}$=.04213
${}_OP_{OO}$=.95088
${}_CP_{OO}$=.94817
${}_OP_{CO}$=.03907
${}_CP_{CO}$=.03896
${}_OP_{CC}$=.95410
${}_CP_{CC}$=.95138

Table 6
Estimation of Robustness of Method

	In Region 1		In Region 2	
	Born Reg 1	Born Reg 2	Born Reg 1	Born Reg 2
Time 1	A	B	C	D
Time 2	$(T_{11}+u)A+$ $(T_{21}+v)C$	$(T_{11}-u)B+$ $(T_{21}-v)D$	$(T_{12}-u)A+$ $(T_{22}-v)C$	$(T_{12}+u)B+$ $(T_{22}+v)D$

Note:The population at Time 1 has already been adjusted to account for mortality and international migration and the probabilities computed are conditional probabilities.

Using Eq. 6,

$P_{12}=[((T_{12}-u)A+(T_{22}-v)C)D-((T_{12}+u)B+(T_{22}+v)D)C)]/(AD-BC)$ Eq. 20

Then, $P_{12}-T_{12}= -[u(AD+BC)+2vCD]/(AD-BC)$ Eq. 21

Similarly,

$P_{22}-T_{22}=[2uAB+v(AD+BC)]/(AD-BC)$ Eq. 22

$P_{21}-T_{21}=[2uAB+v(AD+BC)]/(AD-BC)$ Eq. 23

$P_{11}-T_{11}=[u(AD+BC)+2vCD]/(AD-BC)$ Eq. 24

But,

$(P_{21}-P_{12})-(T_{21}-T_{12})=[(u-v)(AD+BC)+2(vCD-uAB)]/(AD-BC)$ Eq. 25

Table 7
Examples of Robustness of Method

A	B	C	D	u	v	T_{12}	P_{12}	T_{22}	P_{22}	$T_{21}-T_{12}$	$P_{21}-P_{12}$	$P_{12}-T_{12}$	$P_{22}-T_{22}$
1000	100	100	1000	.1	.1	.3	.178	.6	.222	.1	.1	-.122	.122
1000	100	100	1000	.1	-.1	.3	.219	.6	.519	.1	.272	-.081	-.081
1000	100	100	1000	.01	.01	.3	.288	.6	.612	.1	.1	-.012	.012
1000	100	100	1000	.1	.01	.1	.080	.8	.830	.1	.090	-.020	.030
1000	800	100	200	.1	.1	.3	.034	.62	.160	.1	-1.194	-.266	1.560

FEDERAL STATISTICS, NON-FEDERAL STATISTICIANS

Otis Dudley Duncan, University of Michigan

The first of "three basic questions" in the President's charge to his Commission on Federal Statistics was, "What are the present and future requirements for quantitative information about our society?" The Commission was enjoined to "identify the major problems of today and tomorrow for which information is or will be needed." As Eckler (American Statistician, April 1972) has noted, however, "the content [emphasis added] of the federal statistical system" is an "important subject receiving only a limited amount of attention" in the Report of the Commission proper (Volume I).

I, for one, do not know how the Commission might best have responded to the directive to "indicate the important gaps in economic and social statistics," given a period of less than a year to accomplish this and other tasks. In the Commissioners' shoes, many of us might have proceeded in the ways that they did: pointing out how the "convulsive nature of political events rules out orderly specification of statistical requirements"; incorporating in Volume II, with appropriate disclaimer of endorsement, a consultant's review of developments and desiderata in the field of social indicators; exhorting the Statistical Policy Division (Office of Management and Budget) and the National Science Foundation to continue work that will "insure the long range development of social statistics to serve the needs of the nation"; and recommending the establishment of a continuous outside review committee by the National Research Council.

In his evaluation of the Commission report, to which reference has been made, Eckler looked particularly at the report's treatment of issues relating to activities of the federal statistical agencies. I, too, adopted a particular perspective in scanning the text of the report: that of an academic statistical social scientist who generates some statistics, is a user of others, and from time to time gains a mere inkling of small parts of the "Federal statistical system." I wondered if the Commission was as aware as it should have been of the importance, actual and potential, of non-federal sources of "federal statistics." I was satisfied on this score and pleased that the Commission noted some of the special difficulties under which non-federal producers operate (see pp. 55-59, Vol. I).

The Commission's typology of "user groups" is instructive, highlighting as it does the diversity of interests in and claims upon the products of a statistical system. Too often the uses of statistics are assumed to be solely those of the "program manager." In this report, however, his somewhat circumscribed function is distinguished from that of the "policy maker," one of whose functions is the definition of problems, which involves translating public perceptions of situations into political terms. The policy-maker's plight, vis-a-vis statistics, is depicted in terms of the necessity to react to crises without time to plan for the collection of relevant information. But as to the public, whose role is crucial in this recurrent dilemma, the Commission finds: "Statistics ... enter, in diverse ways, into the formation of the citizen's perception of the world he lives in."

The aim of the whole game, I suppose, is that through improvements in statistics we make the citizen's perceptions more nearly true, the policy-maker's definitions more enlightened, and the program managers' and other users' actions more efficacious. The Commission did not fail to understand the open-ended nature of the process, as the reader realizes (among other ways) upon encountering the perceptive discussion of the notion of statistical "gaps." A paradoxical conclusion is implicit in the tenor of this discussion: the number of gaps you perceive varies directly, and not inversely, with the amount you know already. Thus, in recommending the upgrading of exploratory social science research to produce (with federal funding) more and/or better primary data, the Commission surely sensed that a major by-product will be more numerous and more prominent "gaps" in our knowledge of what is going on in the society.

There is a great deal that policy-makers need to know, but the report quotes a perceptive official to this effect: "... the government is simply not good at defining what it wants to do in terms of needed social science research. ...the government, in general, can only articulate the area in which it needs information..." I would go even further and argue (if pressed to do so) that there is more that government needs to know than it necessarily wants to know. This brings us back to the citizen and his "perception of the world he lives in." For it is the changes in both the world and the perceptions of the world, as communicated by the citizenry to politicians and bureaucrats, that drive the processes of democratic government.

Responsive as the federal policy-maker must be to changes in the citizen's perceptions, for reliable knowledge of which he must have statistics, it does not follow that the needed "federal statistics" are necessarily federally-collected statistics. Indeed, precisely insofar as perceptions of the citizenry are at stake, one has to be a little, or more than a little, suspicious of the potential bias of auspices.

Consider the following statistics that ought to be of the highest interest to at least some policy-makers. My colleagues at the Institute for Social Research report (ISR Newsletter, Winter 1972) "a massive erosion in the trust the American people have in their government," as evidenced, among other things, by the decrease from 62 to 35 per cent in the proportion of the public expressing a "high" level of trust in government between 1964 and 1970.

This grim intelligence raises at least two questions for statisticians: If the populace is losing its trust in government in general, is it also losing its trust in that part of the government that carries out population enumerations and sample surveys? Moreover, if policy makers should come to acknowledge the importance of this kind of data, would they do better to ask the Bureau of the Census to collect them, or to rely upon a third party, such as ISR or another reputable survey or polling organization?

Let me cite another example. In a current project we have responses to the following question, asked of probability samples of the metropolitan Detroit population in 1958 and 1971: "Do you feel that someone who doesn't believe in God can be a good American, or not?" In 1958, 57 per cent said he can; in 1971, 77 per cent. If not the "policy-maker," then someone in government needs to understand that the American public increasingly dissociates Christian theology from American democracy, public political ceremonies to the contrary notwithstanding. To me, at least, it seems evident that this is one small piece of statistical information--admittedly a small one indeed, adduced only for illustrative purposes--that ought to be in the hands of the Congress as it debates a bill for federal aid to parochial schools or the President as he decided whether to sign or veto such a bill.

The issue for federal statistics, of course, is that no federal statistical agency has any business whatever inquiring about anyone's religious beliefs, even though information about the distribution of beliefs in the population is pertinent to various federal policies. Again we must concur with the Commission that the scope of "federal statistics"--meaning, the domain of statistical intelligence useful to federal officials--exceeds the scope of statistics produced by federal statisticians. To the reasons given by the Commission for this state of affairs, I have tried to all one, to wit, that some kinds of subject matter are, or surely should be, beyond the competence of the government to include in questions put to, or records compiled about, the individual citizen.

If we should happen to become serious about the matter of social indicators in the manner suggested by the Commission's consultant (Vol. II, Ch. 7) or in some other substantial way, we can only anticipate that the relative importance of these subjects, off limits to the federal statistician, will increase. The personal life of the citizen and his "perception of the world he lives in" are precisely what some of the most informative social indicators will be about. Social indicators are needed, both by the policy maker and by the citizen in his several capacities, because of the increasing complexity and interdependence of the society and the accelerated pace of social change. Our recognition of problems for policy often depends on how this change impinges upon the personal lives of individuals, both in those respects that government has an acknowledged right to look into--criminal behavior, for instance--and in respects that are no business of the government vis-à-vis the individual citizen but are relevant to government's relations with the aggregate of the citizenry.

I feel, therefore, that the Commission was well advised in acknowledging the role of the non-federal producers of some statistics useful in governance. The Commission might well have done even more to delineate appropriate roles for such producers. The report does appropriately touch on the matter of statistical standards in the work these producers carry out under contracts or grants. As the Commission notes, statistics so procured are supposed to be subject to review by the Statistical Policy Division, but under present conditions such review does not provide adequate control. It is not clear to me that increasing the number of SPD personnel assigned to this function and extending its authority to scrutinize "research" as well as "information collection" will, by itself, resolve the problem of variable and substandard quality of non-federal statistical work undertaken with federal funds.

What is really at stake is the development and diffusion of suitable standards --standards that would apply to all serious research for public consumption and not only that part of it undertaken at federal behest. Just as a suggestion for further consideration, it seems to me that the Committee on National Statistics established by the National Research Council in response to a Commission recommendation might put this issue on its continuing agenda. We need, first of all, some good documentation of the nature of the problem and estimates of its current and forseeable magnitude. Then, I suppose, there would be a basis for discussions and studies involving selected non-federal producers of statistics, the SPD, and CNS.

If the times appear suitable for a cooperative effort, we might envision the establishment under CNS or other appropriate auspices a National Conference for Standards of Social Measurement. I would envision this as an organization of individuals and non-federal research units engaged in producing social statistics. Their purpose would be to develop and disseminate recommendations for standard procedures to be used routinely in carrying out and reporting statistical studies, in the absence of sufficient and explicit reason to the contrary.

The Conference would have no police powers and should be chary of the exercise of even informal sanctions. Its authority would derive from the careful staff work and thorough discussions directed toward defining the "state of the art" with respect to each facet of study design. In particular, the Conference should move toward recommendations for standard definitions of those social categories and variables which are used in a routine way in a variety of studies. At the present time one of the most annoying obstacles to the cumulation and comparability of statistical intelligence is the capricious variation from one survey to another in definitions of such elemental items as educational attainment, religious affiliation, or political party preference. In principle, it is not an insuperable problem to suggest criteria for the choice among alternative forms of these questions in the context of general-purpose surveys. The mere announcement that a competent national body recommends a particular alternative might go far toward reducing the current level of chaos.

There would, of course, be other kinds of tasks for a National Conference on Standards of Social Measurement. What I am suggesting, I suppose, is an effort directed to the broad non-federal statistical community so as to accomplish for it something of what the Commission recommendations are intended to accomplish for federal statistics in the narrow sense.

FEDERAL STATISTICS FOR WHAT?

Hal H. Winsborough and Karl E. Taeuber, University of Wisconsin

The authors of this paper are demographers and sociologists. We approached the task of discussing the Report of the President's Commission on Federal Statistics from the point of view of these two disciplines. Thus, as we considered what we wanted to talk about several months ago we were tempted to engage in special pleading; to describe why our disciplines desperately need data not called for in the report. To do that, however--to crank our squeaky wheels--would be to take a particular position as to what presidential commissions are all about. We would be agreeing with the report that the calling of a commission is a tactic used by a policy maker to gain time in the face of a crisis in order to collect information and carefully decide what to do. That sanguine view suggests that someone cares what the commission recommends, that some policy action is brewing, and that this would be a good time to get in our licks.

On the other hand, Mr. Feldman's very interesting paper in volume II of the report "Commissions on Statistics and Statistics on Commissions," presents rather compelling evidence that "systems overview" commissions have no effect at all. Specifically, his final hypothesis states, "If there is a problem calling for a creation of a commission, the solution is recognized when the commission begins rather than when it presents its findings." Mr. Feldman's position suggests that all the action occurred several years ago and that it makes very little difference what the report says or what we say about it. That thought led us to consider dropping the whole matter.

But, we wondered, if Mr. Feldman is right, why would anyone call a systems overview commission in the first place? Our knowledge of the sociological literature suggested an answer that puts the purpose of presidential commissions in a somewhat different light. Perhaps commissions are called, even when their problem is well on the way to being solved, in order to lend legitimacy to the existence of the problem and to its solution. When sociologists talk about legitimacy in this sense, they mean that "certain something" about an idea, a position, a premise, or a theory which says you have to take it seriously--you can't simply dismiss it as hairbrained or ridiculous or trivial.

In general, then, we think that one of the important effects of presidential commissions is to lend legitimacy to ideas, positions, premises, theories, policies, problems, and the like. This legitimating activity in no way assures that the recommendations of the commission will be carried out. It does, however, influence the political process because it makes the ideas and recommendations publicly debatable without the risk of the contestants losing face. An example of this legitimating activity is the recommendation by the Report of the Commission on Population Growth and the American Future that abortion laws be liberalized. That recommendation in no sense assures the liberalization of laws, but it does make the topic more open for public debate than it was before.

If we are right that one of the effects of a presidential commission is to lend legitimacy to ideas, positions, and policies, then the important thing to do here is not to plead a special case, nor to give up the whole thing as a paper which missed its time, but rather to investigate what things are being legitimated by the Commission and to see what we make of them.

The most obvious legitimating activities of the Commission have to do with the place of statistics (as a body of knowledge) and statisticians in the operation of the government. The report stresses the utility of statistical methods and ways of thinking in the formulation of policy, the management of governmental programs and in the evaluation of programs. The Commission also, by the example of its careful and judicious view of the problems of privacy and confidentiality, subtracts a certain amount of legitimacy from the extreme positions which have been stated about these important matters.

With these kinds of legitimating activities we have no quarrel at all. Indeed, we applaud the fresh stress this report places on statistical activities beyond the production of time series of descriptive indicators. However, throughout much of the report's discussion of the government as a producer of statistics--in the sense of numbers--a theory is made explicit about how the system works, a theory with which we must take some exception. We would like to call attention to its difficulties before it becomes so legitimate as to be an unchallengeable premise.

As demographers we are inclined to regard the questionable theory as one designed to explain the birth and death of statistical series. The authors of the report, however, couch the theory in a mold taken from the field of public finance within economics and regard it as a theory about the supply of numbers generated by the government. In this view, statistics are seen as public goods with the federal budgeting activity operating in place of the market to translate demand for statistics into an appropriate supply and to adjudicate between competing demands (pp. 77-78).

How does the budgeting process operate as a pseudo-market? Two principles are enuciated by the report which are thought to guide budgeting decisions. First is the "squeaky wheel" principle. Noisy and politically powerful interests get satisfied. Second is the persistence principle--once a data collection activity begins it tends to continue beyond its usefulness (p. 34).

The first principle is thought to be an appropriate political mechanism for uncovering those broadly-based national interests which should be satisfied:

> The Commission recognizes that the allocation of resources, at a particular time, to the gathering and analysis of various data reflects a political interpretation of the national importance of various user groups and this interpretation

is best made by representatives in the legislative and executive branch (p. 79).

The persistence principle is viewed as less benign. Data collection activities are thought to continue indefinitely because "bureaucrats who preside over the demise of an activity make many enemies but few friends and derive little personal satisfaction from the process" (p. 116).

This theory, then, posits two mechanisms controlling the supply of federal statistics: a political (squeaky wheel) mechanism controlling their creation and a bureaucratic (persistence) one controlling their demise. Both principles seem to us to have difficulties as tools for understanding what's going on. The persistence principle presents the most serious logical difficulties. Let us deal with it first.

It is a common assertion that a statistical series, or for that matter any federal program, tends to persist long after it is needed. What data the Commission used to support this serious accusation about federal statistics is not clear from the text of the report. Apparently hearings were held and people described superannuated data. A few examples are cited in the text.

But the principle that, on the average, data series outlive their usefulness cannot be asserted on the basis of such information. That would be like studying mortality by interviewing centenarians--a newspaperish trick we are sure is decried somewhere in Wallis and Roberts. What we need to know to have any confidence in the persistence principle is the mortality experience of birth cohorts of statistical series. We find no attempt to produce even exploratory data on this matter. Opportunities abound because any knowledgeable person can cite examples of dead statistics from the top of his head. There has not been a Census of Religious Bodies since 1936. A series on the Statistics of Cities was begun in 1894 and terminated in the 1920's. Real Property Inventories were conducted in 1934 and 1939, but not since. Social security number was asked in the Census of 1940, and not since. Religion was asked on a Current Population Survey in 1957, and not since. One could go on at considerable length citing only data about population and housing.

Had all these data outlived their time long before their demise? Perhaps so, but by no means certainly. It is just as likely that some of these series terminated prematurely. Indeed, if the resurrection of statistical series following a lapse of several years or even decades is an indication, a case can be made for premature death in a number of instances. The Real Property Inventory of the 1930's collected data on mode of transit to work. These data reappeared in the 1960 Census and are continued in the Census of 1970. The great interest of the 1960's in statistics on poverty calls to mind the turn-of-the-century census reports of paupers. A National Health Survey was collected in 1935 and 1936 which, after a lapse of twenty years, became a regular series in 1956.

In summary, then, our first complaint about the persistence principle is that appropriate data which would warrant its assertion are not presented nor, apparently, investigated.

A second difficulty with this principle lies in the mechanism presumed to produce it, if indeed it does exist. While the "good" (squeaky wheel) principle is seen as a result of the political process, the "bad" (persistence) one is seen as a result of bureaucracy. But, just as reasonable as leaning on "bureaucratic pathology" is the notion that, in a single specific case, the political costs of offending a small but vigorous interest group by cutting off its "subsidy" outweighs the political gain to be achieved by satisfying a broadly-based but diffuse interest in efficiency and lower taxes. It is, after all, the politicians who presumably have ultimate control over the budget.

We do not doubt that bureaucrats are disinclined to propose terminating many statistical series, nor do we doubt they are unhappy when they have to. But we can think of some cognitive as well as emotive reasons for the disinclination in many cases--reasons which are justifiable concerns. Administrations change, but the bureaucrats go on and on. This year's problem deserving benign neglect may be next year's crisis. Next year, or four years from now, the bureaucrat now asked to preside over the demise of a series may be called upon to yield up all manner of data for policy-making purposes on the history, present state, and causes and consequences of the old problem now newly defined as a crisis. Thus it may be that the bureaucrat's disinclination to drop a series too quickly serves (just as the greed of the speculator serves in the market) to smooth out radical booms and busts in the supply of statistics. It seems doubtful to us, however, that a bureaucrat so motivated will find a convenient place to indicate these reasonable concerns on the Assessment Checklist the Commission proposes should be returned at budget review time.

The assertion of the persistence principle for the operation of the pseudo-market for statistics, then, seems to us to be a classic case of poor social science analysis. An empirical generalization is put forward from improper data. A mechanism is asserted to explain the generalization which is inconsistent with other aspects of the proposed theory and which is chosen from among potentially competing mechanisms by no clear decision rule.

In spite of the errors involved in formulating the persistence principle, its assertion yields recommendations which have only the most modest direct costs to the operation of the system. Whatever may be the distribution of the observed time of termination of statistical series around their "optimum" stopping time, reducing the errors of overshooting the mark is desirable so long as it doesn't introduce excessive undershooting and thereby increase the amplitude in cycles of unfulfilled demand. We suspect the bureaucrats, pathological or rational, can deal with the proposed increase in paper work skillfully enough to moderate these difficulties.

Difficulties with the first of the two proposed principles of the pseudo-market, the squeaky wheel one, seem to us to produce greater difficulties for the Commission Report.

The President's letter establishing the Commission states their first charge as follows:

> In general terms, the Commission should identify the major problems of today and tomorrow for which information is or will be needed. Within this general perspective, it should indicate the important gaps in economic and social statistics, and in related management data and environmental statistics (pp. 37-38).

Given this charge, the report spends a good deal of space discussing proposed gaps in the statistical offerings of the Federal Government. The first principle is used as a touchstone to discern a gap. According to this principle, what constitutes a gap? The report is explicit:

> Although the nature of the budget process causes one to expect to find gaps where demand for new statistics is rising, a gap cannot be said to exist in the Federal Statistical System until a question or a problem is perceived and data relevant to the answer or analysis is unavailable. Thus, perception of the issue is a necessary condition for existence of a gap. As a practical matter, however, perception alone is not a sufficient condition. Perception must be accompanied by an effective desire on the part of an important member of the decision-making complex to have an answer to the question and an analysis of the problem (p. 109).

Thus, a gap is presumed to exist only when a "squeaky wheel" operated by important interests gets no grease. But how do you tell when that has happened? There are lots of squeaky wheels. The problem is deciding which interests are important. The theory proposes a touchstone for that issue as well: "The budget allocation is at once a reflection of the importance placed upon groups of users in the political process, and the means by which they are enabled to pursue their interests in gathering data" (p. 79). Thus, things proceed in a circle. A gap exists only when important interests want data badly. You can tell that the interest is important because the data collection activity they desire is included in the budget. Therefore, there are, by definition, no gaps without a budget allocation and those transient ones will be quickly filled. Pangloss rides again.

Our first point about the squeaky wheel principle, then, is that it deals with the question of persistent gaps by definition rather than by identifying "the major problems of today and tomorrow for which information is or will be needed."

Given this definitional rationalization, there seems little need to discuss the feelings of users. Nonetheless, the report proceeds to consider the classes of users to inquire if their interests are being served. The first group investigated are policy makers. A difficulty with the definitional solution promptly presents itself. Apparently policy makers feel there <u>are</u> gaps, ones which only become apparent in "crisis" situations. Faced with the need to provide policy decisions rapidly, the appropriate data are often lacking. Indeed, one can sense this concern for preparation for crisis management in the President's charge to the Commission.

Herein lies a serious difficulty. One may have plenty of clout but be unsure about which wheel to squeak or when to begin cranking. According to the theory, of course, this uncertainty cannot represent a "real" gap as defined by the first principle until the crisis is upon the policy maker. Then the gap will quickly disappear. (But not, perhaps, until the crisis is well over.)

Thus, although the theory shows why there can be no persistent gaps in the statistical offerings, there can, apparently, be lots of serious transient gaps. Why do these transient gaps appear? The report assures us that this difficulty arises, not from the statistical system, but from the nature of the demands placed upon the policy maker by the political system.

> The political system is conclusive; it acts when the electorate perceives that a crisis exists, whether the crisis is one of inflation, unemployment, pollution, the failure of the educational system, or crime in the streets. But the public perception of a crisis often antedates the presentation of statistical evidence that there is indeed a crisis. Hence, when the legislature or the executive is faced with an aroused public, time is not available to design a survey or experiment, gather the requisite data, and perform a careful analysis pointing toward an optimal policy recommendation. . . .

Are decisions in the face of a crisis, then, made without benefit of data? Apparently not. The report continues:

> When a crisis arises, some data are used to support action decisions. The data used are often a combination of existing benchmark data produced by census-type agencies, management data produced by agencies with related responsibilities, data presented by lobbyists who support a particular position, and particularly important, data on public opinion gathered <u>ad hoc</u> by specialized private polling organizations like those of Gallup, Roper, and Harris (p. 83).

Is there anything that can be done to improve the quality of data available to the policy maker in a crisis situation? The report proposes two tactical procedures the policy maker can use. He can propose a harmless and ineffectual solution, or establish a commission--both being devices to gain time in which to gather better data. Otherwise, "the foresight of the policy-maker in anticipating problems is the strongest determinant of the quality of the data used in solving policy problems."

This advice and this observation seem inadequate answers to the President's charge to the Commission that it "identify the major problems of today and tomorrow for which information is or will be needed."

Our second point about the squeaky wheel principle, then, is that it deals with the problem of transient gaps by transferring the blame for them to the political system. Further, the theory seems to regard the political system as exogenous, that is, influencing but not influenced by the statistical system. Hence, we presume no changes are recommended for the statistical system to extend the foresight of beleaguered policy makers.

Is this separation of the political and the statistical systems reasonable? We think not. Our view about how a crisis descends on a policy maker is somewhat different. Often, it seems to us, crises arise out of well-recognized arenas of difficulty. The economy, racial problems, poverty, education, health, criminal justice--all of these are longstanding arenas for the generation of crises. Some arenas may be particularly "hot" at one time and relatively "cool" in another. As problems are "hot" and actively producing crises, more data are collected pertaining to the problem. Sometimes these data are highly crisis-specific. Sometimes they are of the benchmark kind which become useful in making the next crisis less critical. As problems are ameliorated (and we might hope that improved data lead to more ameliorative policy), interest in keeping the data current may wane. Alternatively, a change in the political climate may make other issues seem momentarily more critical. We can think of two ways this might happen. On the one hand, the development of a national crisis such as a war may distract both public and policy making attention from a persistent problem such as race relations, leaving it, like a pot boiling unnoticed on the back of the stove, likely to boil over into a crisis at any time. On the other hand, the periodic changes of the party in power--each having its own constituency of problems--may again lead to a lessened interest in data within a particular arena. Indeed, there may be a positive disinclination to produce data which would permit the "outs" to claim and try to rally support for the existence of a crisis in an area affecting their constituency.

All of this is to say that it seems to us unreasonable to insist that the political system which produces crises is exogenous to statistical system. The statistical system of the government influences the political system. On the one hand its adequacies may make policy more effective in ameliorating problems and thus reduce the incidence of crises. On the other hand, it may fuel the fires of the opposition in mounting an attack of crisis proportions upon the administration. Perhaps most importantly, the statistical information produced by the government sets some empirical constants in the political debate.

In summary, then, we feel that the squeaky wheel principle of the operation of the pseudo-market for statistics is not very helpful. Although it manages to define away gaps in the offering of statistics, it does so by transferring the blame for problems to the political system. The theory then regards the political system as exogenous to the statistical one and, interpreting its charge to mean not thinking about politics, calls it quits.

Taken over all, then, we regard the theory presented in the Commission's report about how the statistical system works as most unsatisfactory. Its principle to account for the birth of statistics, the squeaky wheel principle, serves only to transfer any responsibility for gaps in the offerings of the system out of the range of the system or the Commission. The second principle dealing with the mortality--or the lack of mortality--of statistics is technically incompetent.

In spite of these difficulties, this theory seems the major basis for the second of four major recommendations of the Commission, that "more systematic efforts should be made to eliminate obsolete statistical programs." The brief text following this recommendation goes on to assert: "The filling of statistical gaps is nearly automatic and causes us relatively little concern, but the elimination of unproductive programs is important and, at the same time, one of the most difficult problems facing the government."

In thus a manner is considerable legitimacy added to dubious propositions. While these assertions may be true enough, the report has given us only the most meager reasons for believing them.

DATA BASES FOR DATA USERS

Edmund M. Murphy, Chief, Development, Regional Statistics Staff Statistics Canada

This paper will concentrate upon the output side, the dissemination side of data base construction using examples drawn from the dissemination programme designed for the data base created from the 1971 Canadian Census of Population and Housing.

One may conceive of three major steps in constructing a data base:

a. collecting or assembling the data
b. processing the data to create the data base itself
c. drawing information from the base and placing it in the user's hands.

A dissemination programme must consider all three phases. The technology of the third phase, the publication phase, has the most apparent effect upon dissemination; however, unless the user is given adequate information about the two preceeding stages, the dissemination programme may become an exercize in user deception rather than user education.

At every step of the collection and processing phases assumptions are made, approximations used, shortcuts taken that can have an important effect upon the quality of the data. The user must be fully aware of the impact of these procedures lest he mislead himself and those who use his analyses.

1. The Collection Stage.

It is a truism that the first step in using a data set should be to read the instruction manual for the enumerators who collected the data. Apart from such utopian and perhaps unfair demands on the user, it is a responsibility of a dissemination programme to give the user a description of the collection procedures as they affect the data.

The presentation of such information should take two forms. First, a general discussion of methods and procedures, and second, a discussion, where necessary, of the impact of collection procedures on specific data sets.

For the 1971 Canadian Census the major methodological innovations were self-enumeration (drop-off, mail-back) and the extensive use of sampling. Preliminary processing procedures that have an impact on the data include acceptance standards for incomplete questionnaires, call back procedures, plus regional office coding of occupation, industry and such open ended questions as ethnic group and language.

Describing the impact of these procedures is no simple task. A general description of the procedures presents no problem, but, aside from estimates of sampling error, it is difficult to go beyond cautionary notes when discussing the effect of specific procedures upon specific data items. What is needed is a quantitative, item specific estimate of the reliability of the data.

2. The Processing Stage

Between the collection stage and the creation of the data bank, the information is subjected to a series of processing steps that have an additional effect upon the data. Again a dissemination programme must include a description of these procedures and an estimate of their impact upon specific data items.

In addition to the regional office processing described with the collection procedures, the major steps in processing the 1971 Canadian Census data were the editing and the imputation programmes and the weighting procedures applied to the sample data.

In general weighting procedures and their effect upon reliability are relatively easy to describe and to quantify. The weighting procedures flow from the sample design, and their impact can be presented as part of the estimates of sampling error.

The 1971 Canadian Census used a ratio ranking procedure for weighting. For private dwellings, enumeration areas were aggregated to weighting areas of roughly 4,000 population. Weights were assigned so that for each weighting area sample estimates and complete counts matched on certain key items common to both the sample and complete count questionnaires.

Presenting to the user a summary of the editing and imputation programmes and their impact upon the reliability of the data is not so simple. Stating the philosophy of these programmes is easy enough: to remove obvious errors and inconsistencies, on the one hand, and to make meaningful estimates of missing data on the other. However, to indicate the operational steps involved is a problem of another order of magnitude. A list of the edit specifications for the 1971 Census is itself an impressively thick volume.

Again the question is one of balance. The user must not be swamped with indigestable masses of information, but he must be given enough information about these procedures to enable him to properly interpret his data and the results of his analyses.

For the editing programme general lines of strategy could be indicated giving the level of priority for each type of data. For example, one could indicate, perhaps in a flow chart, that age is checked against other ages in the household, that marital status is checked against age as is educational attainment, etc.

For the imputation programme one can indicate the general rules. In the 1971 Canadian

Census there were two types of imputation procedures. The first used deterministic rules, a husband's age could be estimated from the age of his wife, for example. The second type was the well known "hot deck" procedure where imputation groups are defined by a set of basic variables and missing data for a record are imputed from the previous complete record falling in the same imputation group.

The user must have a firm idea of the rules of these procedures. Moreover, he must also have a quantitative, item specific estimate of their effect. This latter information is even more difficult to convey than the former. Not only does the impact of these procedures vary from data item to item but from area to area.

These variations by area are especially critical in Census data, which is a major source of data for small areas. The 1971 Census will publish a great deal of data for enumeration areas, which have a population of about 400 to 600. The impact of the editing and imputation programmes can vary significantly from EA to EA, and the user should be warned when he is using an EA that is largely hypothetical. The same problems exist to a lesser degree for larger areas.

3. The 1971 Census Master File and Data Dictionary.

The purpose of these collection and processing steps is to create the data base. For the 1971 Census data this base is called the Census master file; from the user's point of view this file _is_ the Census data. It is from these data as coded, edited, imputed and weighted that all 1971 Census tables must be drawn. To analyse a Census table or to design a special tabulation from the data, the user must understand the data on the file. He must know what data are recorded on the file, what categories are used to record the data, for what subpopulations the data are reported, and what are the strengths and weaknesses of the various data items.

For 1971 the Census is producing a document that contains much of this information — the Census Data Dictionary. In this document the user will find a definition of each data item recorded on the master file, the categories used in coding the responses, which of the last three Censuses contained this data item, differences in definition between these Censuses, the sampling fraction where applicable, the population for which the data are reported, and finally, a remarks section pointing out any problems in interpreting and using the data.

By reading this Dictionary the user can discover for example which ethnic groups are coded on the file and by implication which are grouped as "others"; he can see that the labor force status information is reported only for those 15 and over; that age data is available up to an open ended group 100 and over, etc. He can also discover that no information exists in the Census data base on health.

This Census dictionary is a major start towards providing the user with the information he needs about this data base. The introduction to the dictionary provides the needed overview of the Census collection methodology and processing procedures. It does lack, however, at this stage of development, two of the elements noted above as being essential to a full description of a data base. There is no discussion of the detailed strategy used in the editing and imputation steps, nor is there any indication in the dictionary itself of the reliability of the data for particular items or areas.

4. Root Mean Square Error

On another front, Statistics Canada has also made progress in indicating to the user the reliability of the data. After the evaluation studies of the 1971 Census are complete, the information obtained will be summarized in one measure, the Root Mean Square Error. This combines in one measure both sampling error and non-sampling error. The measure is constructed so that it can be used, in general, like a sampling variance to estimate confidence limits for a figure of a given magnitude.

When this information is added to the Census Data Dictionary as part of a section in each definition on "Reliability", the Census Data Dictionary will come close to being a model for a description of a data base. Even in its present form it is valuable aid to any user in understanding the 1971 Census data base and the tables that are drawn from it.

5. The Dissemination Programme

The first stage in the dissemination programme is to describe to the user what data are available, to give him an understanding of the data base. The next step is to give him some data.

5.1. The Media

Before discussing the various strategies for placing data in a user's hands, a brief discussion of the media of presentation is desirable. There are, in general, three types of media applicable to data transmission — printed publications, optical devices, such as microfilm, and computer readable storage such as magnetic tape. As one reads down this list the media become increasingly more expensive to read, to peruse, and increasingly less expensive to store and manipulate. Printed reports are bulky, difficult to store and the data must be later transferred to machine readable form for manipulation. On the other hand, one does not have to sell his patrimony to buy equipment to read a book.

At the other extreme, data on computer readable media can be introduced directly into a machine, but one needs a computer to actually see the material and documentation and compatability problems often exist. Optical devices such as film are easy to store and reading devices are at least cheaper than a computer, but again the data must be transferred to another medium for analysis.

The 1971 Census data will be available in all three types of media, the standard printed reports, microfilm rolls and microfiche, and user summary tape.

Since printed materials are very easy to peruse but difficult to store, the evident strategy is to provide summary information in printed form and to use film or tape for transmittal of detailed data. The 1971 Census tabulation programme takes a step in this direction. The preplanned tabulation programme contains some 3,500 tables, about 3.8 million pages of computer printout. A small proportion of these will be published in printed reports. The rest of the tables, except for those produced only for check-out purposes and internal analysis, will be available on microform or as part of the summary tape programme.

5.2. Data to the user

After having described the data base, and insured a range choice of publication medium. The next step is to decide what information is to be transmitted. A problem that often arises is that of confidentiality. For most data bases, it is usually contrary to public law, professional ethics, or common decency to release to the public data that can be related to an identifiable individual unit. One is thus left with three types of publication:

a. Publication of summary data as in tabulations
b. Publication of individual records adjusted so that individuals cannot be identified
c. In-house manipulation of individual records and publication of the results.

For the 1971 Census data, Statistics Canada is considering options b. and c., but the major thrust of the dissemination programme is upon a large and flexible tabulation programme. Such tabulation programmes have two major components, a set of preplanned tables plus systems for permitting the user to design and obtain tables to his own specifications.

5.2.1. The preplanned tabulations

No matter how flexible and efficient the special request tabulation system may be, it is still cheaper and quicker for a user to obtain an "off the shelf" table when he can find one that suits his needs. The twin problems for a preplanned tabulation programme are to design the tables so that a large proportion of user needs will be met and to inform the user as to what tables are available.

The proper design of the programme should involve at least two elements: professional judgement of what would be useful and desirable, plus the equivalent of market research to ascertain the efficiency of previous tabulation programmes and to estimate changing demand for tabulations.

The 1971 Census tabulation programme emphasized the first aspect, professional judgement. The market research consisted largely of an analysis of the special requests generated by the earlier Censuses. Requests for special tabulation are one important source or information about the effectiveness of a preplanned programme, and this type of research will be continued and expanded for the 1971 special requests. In addition, the Census Division is setting up a special unit to conduct market research in the broader sense as an input to the 1976 and 1981 tabulation programmes.

If the preplanned tabulation programme is large, as it is for the 1971 Census, the problem of informing the user of exactly what is available assumes major proportions. The 1971 Census approaches this problem in two ways. The published tables are described in the traditional Census catalogue, and the entire preplanned programme is listed in a new document, the Tabulation Directory.

This directory is the first step toward an automated search system. Each table in the preplanned programme is completely described on an entry in a computer file. This file could be made available on tape to a user who could then devise a programme to search it, or he could list the file or parts of it in a manner that suits him. In addition the Census Division will make available a listing of the file. This listing will be cross sorted for ease of access. For example, all tables containing data on a given variable will be listed in one place cross-referenced by the other variables in the table. A user interested in a table showing age by sex by occupation can turn at once to a given section of the directory and find a list of all tables containing these variables.

5.2.2. Special Request Tabulations

No preplanned tabulation programme could or should meet all requirements for data from a given base. If every need is being satisfied, then the preplanned programme is probably too large, some of the tables are being used by a minute number of people. Thus every dissemination programme must provide for the timely and flexible production of special request, user defined, tabulations.

The basic element of such a system is a description of what tables could be produced from the given data base. The description of the data base discussed above should serve this purpose. For the 1971 Census, the Data Dictionary does give the user the information he needs to design tables that can be produced from the data base.

The next step is to develop a system that will deliver the requested tables quickly and cheaply. Since internal priorities and staff shortages are responsible for many delays in servicing special requests, the system should minimize the necessity for in-house staff intervention.

The need for in-house experts to advise and assist users can never be completely eliminated; however, the system should permit the user who is so inclined to accomplish himself the work necessary to generate a special tabulation. Ideally, there would be direct interaction between the user and the data base; the user types in a request on a terminal in his office and a second later the table appears. (The same kind of interaction is also a goal of the 3rd dissemination alternative mentioned above, in-house analysis of individual records and publication of results.)

There are several problems to be resolved before such direct user interaction is a reality, but the 1971 Census has made a major step in this direction with its STATPAK tabulation system. The TARELA language developed for the STATPAK programme is close enough to ordinary notation that the user can after some study essentially write the control cards for the programme himself. This system is new, and it has not yet been tested on real census data with real users, but it is a very promising step. In theory at least, the user can design the table, write the control cards, and send the request to the Census Division. His request will be checked, the control cards punched, the table run, and the results returned to the user with the obvious gains in overall turnaround time and costs.

A system such as this imposes an obvious burden upon the user. He must take the trouble to decide exactly what he wants, check its availability, write the control cards, and accept responsibility for the results. It is easier to write a note asking for "something by age and sex and occupation for the city of Sorel". However, with an interactive special request system, the user will have ample rewards for his efforts in cheap and timely answering of his request for special tabulations.

Another requirement for a special request programme is flexibility, especially in delineation of areas. The 1971 Census dissemination programme will give the user a great deal of flexibility in areal delination with its Geocoding programme. This is a computer system that codes all Census data to small units — block faces in 14 urban areas, Enumeration Areas for the rest of the country — and permits the user to construct his own areas as aggregations of these units simply by drawing his area on a map.

5.2.3. Confidentiality

Even with summary data, problems of confidentiality of individual records arise. It is often quite easy to pick out individuals in a tabulation and deduce information about them. One Census example is a table showing that the only Doctor in a certain small area has a certain income. In addition to these direct disclosures, there exists the problem of residual disclosure, obtaining information by combining two or more tables which themselves contain no direct disclosures. And finally, many data bases are plagued with problems of dominance, one unit's data dominating the data for a given area.

Confidentiality is a severe problem for any data base and especially one with direct user access. The procedures used to protect against disclosure must work, they must be automatic, and they must not distort the analytical meaning of the data.

For the 1971 Census data, which has problems with direct and residual disclosure but rarely with dominance, a random rounding procedure has been adopted. This procedure essentially shifts the unit of census tables from one to five.

The procedure gives good protection against disclosure; no number smaller than five appears in any table nor can a number smaller than five be derived by manipulating tables. It is automatic; a computer algorithm rounds each figure to one of the two nearest multiples of five using a random number generator and a schedule of probabilities. And it has little impact upon the analytical meaning of the data; only very small numbers are significantly changed, and these are very unreliable in census data in any case.

This, procedure will not be appropriate for, all data bases, but any data base which incorporates direct user interaction or rapid in-house preparation of special requests will need confidentiality procedures with similar properties: adequate safeguard, automatic application and minimal impact upon analytical meaning.

6. Summary

The construction of a data base has one goal: to place in the hands of analysts and policy makers the information they need to do their job. The major tasks of a dissemination programme are to provide the potential user with an adequate description of the data base including the procedures by which it was constructed, to plan a series of standard data packages to meet the needs of the general user, and to provide quick, accurate, and cheap special request data for users with specialized needs.

The 1971 Canadian Census will provide such a data base by means of its expanded summary tape and microfilm publication programme, its Statpack and Geocoding special request systems, and its expanded data documentation.

(Further information on the 1971 Canadian Census Data Dissemination programme may be obtained from:

Census User Enquiry Service, Census Division Statistics Canada, Ottawa K1A OT7)

EVALUATING THE QUALITY OF THE 1970 CENSUS

Joseph Waksberg
U. S. Bureau of the Census

Considerations in Determining Evaluation Program

Papers presented in earlier years, and published reports, have described in general terms the program to evaluate the quality of the 1970 United States Census of Population and Housing. This paper will provide further details of the program. However, before going on to a description of the various projects used in the evaluation process, it seems useful to discuss the considerations taken into account in determining the evaluation program, and why it took the particular form it did.

In a mathematical model extensively used by the United States Bureau of the Census and other statistical organizations, both measurement and sampling errors are identified as contributing to variance and bias which are then separately estimated in experimental studies. This model has been utilized effectively to guide census and survey design. In a sense, the variance can be considered to reflect the "noise" in the census-taking process and bias the misinformation in the system. Of course, what contributes to variance and what contributes to bias depends on what is being held fixed. A given questionnaire may have a certain consistent or average error effect, defined as bias, but if we consider this particular questionnaire as a sample of one from a population of possible questionnaires that might be considered for the survey, the differences among such questionnaires would contribute to variance.

A desire to measure both response bias and response variance of Census statistics had a major effect on the methods used to evaluate the Census. Frequently, the best method of measuring bias is to compare Census data with outside, presumably more accurate, sources. Although as much of this was done as possible, other devices had to be superimposed in order to measure response variance as well. Generally, the additional programs also had secondary values in providing breakdowns of the biases for sub-classes of the population. As a result of these and other considerations, projects were developed using the following methods for evaluation.

(1) checks of the reasonableness of Census totals and distributions by comparing them against outside data;
(2) examination of distributions for internal consistency;
(3) review of the effect of operating procedures, such as accuracy of coding, imputation for nonresponse, etc;
(4) re-interviews of samples of households or persons with case-by-case matching of results;
(5) record checks in which information from outside sources for a sample are compared case-by-case, with Census data;
(6) designed experiments; in 1970, primarily a randomization of work units among a set of enumerators.

The evaluation program also included experiments to test the effects of proposed changes in census-taking process in contrast to procedures in use, although such experiments obviously do not directly measure the quality of the Census.

Given that these were the measurement methods to be used, decisions were still needed on what aspects of census-taking should be studied and the emphasis to be given to each. The criteria used in allocating priorities involved determining that the evaluation program should satisfy the following broad classes of objectives:

(1) To determine the effectiveness of the new procedures adopted for the conduct of the 1970 Census. In a sense, one of the purposes of the 1970 evaluation program is to ascertain whether the Census moved in the right direction in adopting the procedures utilized for the first time in 1970 (and for that matter, in continuing the methods first introduced in 1960).

(2) To obtain another round of readings on the components of mean square error of important census statistics. (1970 would be the third point in a time series on evaluation.)

(3) To solve some problems of evaluation that were inadequately dealt with in 1960.

(4) To pursue more intensively the correlates and causes of error in census statistics.

Description of Evaluation Projects

A group of projects was developed to measure, as accurately as possible, the extent and cause of coverage and content errors. About half of the resources to be devoted to the evaluation program was assigned to each of these classes. The first ones we shall describe are those that involved checks of Census totals and distributions, analyses of internal consistency, and reviews of Census operating procedures:

1. Demographic analysis of coverage. This comprises techniques that check census statistics against data derived from combination and manipulation of birth and death statistics, previous census results, data on migration into and out of the country.

2. Comparisons of Census totals and distributions with other available aggregated statistics, e. g., comparing labor force distributions with the Current Population Survey figures, school enrollment with Office of Education data, etc. Of course, such comparisons can raise serious questions on the accuracy of different statistical sources and what are appropriate standards. This is a problem that permeates the entire evaluation program and one that is better dealt with on a project-by-project basis.

3. Analysis of the size and potential effect of nonresponses and the methods used to impute for them.

4. The effect of special procedures used for the first time in the 1970 Census to improve coverage in difficult-to-enumerate areas. Insofar as possible, we are obtaining counts of the numbers and some characteristics of persons added by these procedures. Unfortunately, this is only possible for some of the devices used. There is no practical method for isolating the effects of some coverage improvement techniques; for example, more intensive training of enumerators and crew leaders in selected areas.

5. A study of the composition of the 20 percent sample in the Census to ascertain whether any bias has occurred in the sample selection.

6. Analyses of the effectiveness of the various quality control devices (for both field and office operations) used in the Census to control error.

A greater part of the resources in the evaluation program was devoted to a series of projects in which samples of the population or housing units were selected and subjected to various operations to provide measures of error. The operations consisted of such techniques as record-checks with outside, and presumably more accurate, sources; re-interviews, sometimes with more probing questionnaires; randomizing the work units assigned to a set of enumerators; examining the effects of different stages of census processing, etc.

It should be noted that some of these processes seem to imply that there is always a correct answer to a question and that the difference between the correct reply and the one actually obtained is the error in the individual response or measurement. Without going into a discussion of the nature of truth, in some cases at least this notion is unproductive. The idea of "correct" or "incorrect" is uncertain for a number of items in the Census which are not defined with the precision necessary for an unambiguous reply to exist in all cases; for example, mother tongue, and vocational training. For these items, the re-interview focuses on providing insights on how the respondents interpreted the questions.

The specific projects in the evaluation program are listed in table 1. Some comments are in order.

For overall understanding of census counts, the Bureau is relying on demographic analysis with breaks by age, race, and sex. The detailed projects relating to coverage in table 1 complement the demographic analysis studies in a number of ways. Two studies are designed to improve these estimates. The Birth Registration Study examines the completeness of registration of births. The number of births is an important component of the estimates developed by demographic analysis, and they are obtained by adjusting registered births by estimates of nonregistration. The second study is the Coverage Check of Persons 65 Years and Over, the age group for which demographic estimates are probably weakest. This study has been performed by selecting a sample of Medicare registrants and determining whether they have been enumerated. The remaining studies are intended to explore various aspects of undercoverage. Several projects will measure various aspects of coverage of housing units, partly to shed light on the extent to which living quarters were omitted from the Census, partly to detect any weaknesses that might exist in the mail techniques introduced in 1970, and partly to evaluate the effectiveness of a number of procedures used in 1970 specifically to improve coverage. The remaining coverage projects explore the extent and reasons for missing selected classes of the population.

Measures of response errors, generally referred to as "content" errors, are mainly derived from record checks, re-enumerative studies, and the randomized enumerator variance study. The 1970 program has expanded the number of record checks, as compared to the 1960 and 1950 evaluations. The record checks appear to be the best devices to measure possible biases in Census statistics. The record checks were performed by matching census reports for a sample of households with independent record sources. In regard to the re-interview program, analyses of the results of the 1950 and 1960 programs indicated that, for most population items, re-enumeration does not provide satisfactory measures of bias, at least with the techniques used. Since adequate measures of simple response variance are coming from the CPS-Census match, most of these population items are omitted from the 1970 evaluation program. Instead it concentrates on: (a) housing statistics, (b) population items for which earlier experience indicates adequate measures of bias can be obtained, (c) new items not in CPS and for which no 1950 or 1960 evaluation data exist. The total response variance, including the correlated component produced by the tendency of some enumerators to make consistent errors, is measured by the randomization study. A more efficient randomization procedure was used than in 1960 or 1950 and more precise measures of variance are expected.

Two of the projects do not neatly fit into any of these categories. One is a study of the accuracy of geographic coding in the approximately 60 percent of the United States in which the codes were computer-assigned. The second, referred to as the National Edit Sample, traces a sample of questionnaires through various stages of census processing, to see how editing, follow-up, clean-up, etc., affected the quality.

Current Status of Program

Coverage: In a paper presented in the 1970 Annual Meeting of the American Statistical Association, Jacob Siegel reported that an analysis of a historical series of rates of net underenumeration for 1880-1970 supports the view that coverage of censuses has been improving, albeit irregularly, and that the 1970 Census had the lowest underenumeration rate over this period. The preliminary evidence is that the 1970 rate of underenumeration is lower than the 1960 rate, with the decrease in the neighborhood of 0.1 or 0.2 percent. (It must be emphasized that these are only preliminary figures and the Census Bureau will issue a full, final and more accurate report when it completes further work.) This decrease occurred in spite of the fact that changes in the age-sex-color composition of the population between 1960 and 1970 increased the proportion of the more-difficult-to-enumerate categories in 1970. If the same underenumeration rates for age-sex-color groups had occurred in 1970 as in 1960, the total underenumeration would have increased by about 0.2 percent.

Accuracy of Response: Comparisons are being made between Census tabulations and other sources of data for a number of statistics. As mentioned earlier, it is not always clear which source is "better." Furthermore, there are frequently small differences in the population covered, the definitions used, or the reference period which makes exact comparability impossible. However, such comparisons do indicate whether any serious problems exist and the general order of magnitude of errors.

Comparison of the Census with results of the Current Population Survey are shown in table 2 for a number of key subjects--labor force, income, educational attainment, and school enrollment. In general, CPS and Census data are quite close. The largest differences are in some school enrollment categories and these are probably due to the difference between the fall date for enrollment figures in CPS and the later date for the Census. (There are also some large differences in the proportions enrolled in public vs. private schools.) The compensating differences between the different levels of persons completing high school and college are, we believe, mostly due to the better reporting in the Census as a result of self-enumeration.

School enrollment figures were compared with data compiled by the Office of Education, although differences in timing do not permit exact comparability. The two sets of figures are generally similar, except for the division of private schools between parochial and others, where a major discrepancy exists.

In general, the quality of the statistics on items that have been traditionally collected in the Census and current surveys is similar to that in prior years. Some of the newer items in the Census were known to be difficult to collect, either because of lack of clarity in the concepts or because many of the respondents simply do not know the correct answer. Evaluations completed to date on two such items confirm the existence of high response errors. For example, the number reported as disabled in the Census appears to be low by about 30 percent. Such difficulties were anticipated and in the 1970 Census there was a conscious decision that the requirements for information for small areas were such that partial data would be useful.

As you can see, only scattered preliminary results are available at this time and these, of course, do not include the larger projects. These programs tend to be complex and slow, requiring a considerable amount of detailed work. In addition, there are always needs for priorities on work assignments and, of course, these are normally assigned to the production of the regular census publications.

We expect most of the projects to be completed within the next year. The results will be made available to the public in a series of 1970 Census publications which will be identified in the PHC-E series. The series will be announced in the usual manner of census publications.

The final evaluation of the quality of the 1970 Census, both in absolute terms, and as compared to earlier censuses must await these final reports.

Table 1. 1970 Census Evaluation Projects

Project	Sample Sizes
Coverage:	
Birth Registration Study	15,000 children under 5 years of age 7,500 white children 7,500 children of races other than white
Coverage Check of Persons with Motor Vehicle Operators' Licenses	1,000 licenses issued or renewed less than a year before the 1970 census to males in their 20's in District of Columbia tracts where the residents were low-income and predominantly of the Negro race.
Coverage Check of Persons 65 years and over	8,000 persons 65 years of age and over, registered for Medicare.
Housing Unit Coverage	A relisting of addresses in 9,000 small-area segments (usually about the size of a city block). In addition, a recheck of 20,000 addresses for completeness of coverage of all housing units at those addresses.
Misclassified "Vacant" Units	3,000 units enumerated as vacant in the Census.
Erroneously Deleted Housing Units	1,300 units originally listed, but later deleted from the Census
Coverage Errors based on the Definition of a Housing Unit	1,800 housing units enumerated in the Census, but with characteristics that indicate a potential undercount. 900 housing units enumerated in the Census but with characteristics that indicate a potential overcount.
Matching of Housing Units from the Current Population Survey Against the Census	56,000 housing units checked for inclusion in the Census. Persons in 10,000 housing units checked for inclusion in the Census.

Table 1 (continued)

Project	Sample Sizes
Content:	
Match of CPS and Census Data	Characteristics of housing units and persons in them compared with Census, for 10,000 housing units.
Accuracy of Income Reporting based on Match with the Internal Revenue Service	The 10,000 housing units cited above.
Accuracy of Occupation and Industry reporting Based on Match against Employer Records	7,500 employed persons (excluding self-employed, unpaid family workers and farmers).
Employment 5 years ago	4,000 persons 14 years and over reporting employment status in the Current Population Survey in July 1963.
Record Check on Expenditures for Gas and Electricity	5,000 rental housing units in the Census 20 percent sample selected from 5 metropolitan areas.
Record Check on the Value of Owned Homes	3,000 homes sold July-December 1971
Reinterview of Census Sample Households for Content Evaluation	10,000 housing units occupied in the Census. 1,000 vacant housing units.
Accuracy of Geographic Coding of Addresses to tract, Block, and Minor Civil Divisions	11,000 addresses.
Accuracy of Place of Work reporting and Coding	4,000 employed persons in the Census sample, in SMSA's
Disability Study	15,000 households with one or more persons reported as disabled. 25,000 households with no persons reported as disabled.
National Edit Sample (to measure effect of editing follow-up and clean-up on "Not Reported" rate)	7,500 questionnaires from the Census sample. 7,000 nonsample questionnaires.
Response Variance Study	1,100 enumerators in 35 Census District Offices.
Other:	
Experiment testing the feasibility of extending the mail census to areas of lower population density	Five matched pairs of District Offices for comparison of costs and administrative problems.
	For a coverage check, 13,000 housing units relisted in small areas selected from 420 enumeration districts.

Table 2. Comparisons of Census and CPS Statistics for Selected Items, 1960 and 1970

(CPS data for labor force are for April 1970; income and educational attainment are for March 1970; school enrollment data are for October 1969)

Item	1970			1960		
	Census	CPS	Difference	Census	CPS	Difference
Labor Force (%)						
In civilian labor force	58.4	60.1	- 1.7	55.5	57.0	- 1.5
Employed	55.8	57.5	- 1.7	52.6	54.0	- 1.4
Unemployed	2.6	2.6	-	2.9	3.0	- 0.1
Unemployment rate	4.4	4.3	0.1	5.1	5.2	- 0.1
Not in civilian labor force	41.6	39.9	+ 1.7	44.5	43.0	+ 1.5
Income						
Median family income	9590	9433	157	5660	5417	243
Aggregate money income (in billions)	635.5	603.3	32.2	332.3	304.5	27.8
Educational attainment for persons 25 years and over (%)						
Elementary						
0 to 7 years	15.5	14.3	1.2	n.a.	n.a.	n.a.
8 years	12.8	13.4	-0.6	n.a.	n.a.	n.a.
High school						
0 to 3 years	19.4	17.1	2.3	n.a.	n.a.	n.a.
4 years	31.1	34.0	-2.9	n.a.	n.a.	n.a.
College						
1 to 3 years	10.6	10.2	0.4	n.a.	n.a.	n.a.
4 years	6.1	6.8	-0.7	n.a.	n.a.	n.a.
5 years or more	4.6	4.3	0.3	n.a.	n.a.	n.a.
School enrollment 1/ (thousands of persons enrolled)						
Elementary school	33,210	33,788	-1.7 2/	28,988	29,382	-1.3
High School	14,481	14,553	-0.5	9,696	9,616	0.8
College	6,966	7,435	-6.3	2,935	3,340	-12.1

Note: Minus sign denotes higher C P S figure.

1/ 1970 data are for persons 3-34 years; 1960 data are for 5-34 years.

2/ Difference shown as percent of C P S totals.

ORGANIZING CENSUS RESULTS FOR USE BY SOCIAL SCIENTISTS*

Everett S. Lee, University of Georgia
Richard C. Taeuber, Oak Ridge National Laboratory

We stand at the beginning of an era in social research, unfortunately one which we are not prepared to enter. The Census Bureau has adopted a policy of releasing elaborate tabulations for small areas on magnetic tapes and has prepared a series of one-percent samples of individual records for 1970 (in which no person can be identified, of course), that permit the investigator to make whatever tabulations or correlations he wishes. Furthermore, the Census Bureau is providing comparable data for 1960 and has arranged for the release of Current Population Survey Data, after combining surveys in such manner that the statistical variation is not distressingly great. There is perhaps some hope that samples will be prepared from the records of earlier censuses so that much needed historical research can be performed with the same ease as that based on the 1970 Census.

The Census Bureau is not the only agency to have taken an advanced view of data dissemination. The Social Security Administration has released tapes which include one-percent samples of social security records, going back more than a decade, and there is experimentation with the cost and value of increasing the sample size, perhaps to ten percent. Data from the Census of Manufactures and the Census of Agriculture are also available on tape, as are the materials published in County Business Patterns. It may well be that the National Center for Health Statistics will follow suit. Already the amount of material available on magnetic tape is such that major efforts are necessary to catalog the available materials.

While we have long been handicapped more by lack of imagination than by lack of data, the situation of the social scientist and planner in this country is greatly improved over one that was already excellent. As Mary Jean Bowman has noted:

"The leadership of American economists in the development and empirical testing of models for analysis of investment in education rests firmly upon the imaginative and yet unwitting contributions of sociologists, demographers, and very practical market analysts in private business. It is these people who persuaded the U.S. Census to put earnings and educational attainments into the 1940 Census, along with breaks by sex, age, race, and region. These data were improved in the 1950 and especially the 1960 Census, and will be even better in 1970. They are the envy of economists in other lands concerned with human resource problems."[1]

This is a true statement but it deserves an addendum. Some of the contributions were not unwitting, and many of them were made by social scientists within the Census Bureau itself, who were led to them by their own analytical work. One of the great strengths of the Census Bureau has been that its leaders have distinguished themselves in the analysis of census and related materials. Consequently, they have been made aware of the pitfalls and limitations of the data, and they have had the easy access to other leaders in social science research that comes from intellectual communion and cooperation. While there were many people outside the Bureau who urged the development of public use samples and other magnetic tape releases, it was through the efforts of Census officials, themselves analytically oriented, that we have the current wealth of data.

We have moved so rapidly into this new era of data manipulation that none of us are prepared. The Census Bureau has understandably encountered difficulties and delays in making the tapes available in usable form. The spate of private research and data processing organizations that sought to use the census tapes in providing services to their clients has been decimated by inflation and recession. Only a few remain and probably all of these have channelled their efforts much more narrowly than originally intended. Universities and nonprofit research organizations have also had to curtail their plans, as the costs of assembling the census tapes, providing the necessary programming, and developing working tapes have become clear. There is not only a shortage of money, there is an even more disturbing lack of properly trained personnel. Furthermore, all of us have to be made aware of the multiplicity of opportunities we are offered, though as E. B. Wilson once remarked of mathematics, statistical technique under the spur of the computer is like a horse with so many and so varied gaits that the entranced rider is likely to ride off in all directions. Until we become better accustomed to both tools and data, we may fully expect to see regressions and multiple regressions without end or reason, and we will find certain variances explained by several hundred percent.

There are lessons to be learned from the Surveys of Economic Opportunity, imaginative undertakings sponsored by the Office of Economic Opportunity and conducted by the Census Bureau. After tapes from these surveys were released by the Census Bureau, it took the joint efforts of the Brookings Institute and the ASSIST Corporation to make the tapes usable to non-census people. Even so, important questions including sample weighting and variances, remain to plague investigators. Programming turned out to be no easy matter, and it is only now, through the combined efforts of social scientists from the University of Georgia, the University of Wisconsin, the Department of Agriculture and elsewhere, that data from these surveys can be brought to bear on problems of immediate importance.

Experience with these surveys and with a few 1960 census tapes, including the 1/1000 sample, has saved us many difficulties, but there are still many steps and costs involved in proceeding from the census tapes to research findings. In the first place, tapes from the Census Bureau are high in cost and sometines slow in delivery, because the copying and documentation of tapes has to be fitted into the Bureau's heavy operational work load. As received from the Census the tapes are not in the best form for processing. For the sake of economy in storage and handling they should be packed and converted for particular tape drives. Programming suitable for several types of computers and for varying materials (the Fourth Count as against the Public Use Samples, for example) is necessary.

Fortunately, these needs were foreseen by the Ford Foundation, the National Science Foundation, and the Center for Research Libraries,who took steps to facilitate the widest possible access to the Census data by providing grants to universities to obtain tapes and programs and by arranging with a non-profit organization, DUALabs, to provide compressed tapes and adequate programming at minimal costs. Arrangements were also made to provide extract or working tapes at the lowest possible costs to researchers. Perhaps just as important was the provision of an organization which could make copies of tapes without the long delays that are associated with governmental operations. To the researcher quick cost estimates and quick delivery of materials after they have been purchased are crucial concerns.

But even with these developments it was clear that there were alarming as well as desirable features in the new methods of data dissemination. It remains true that some of the most important papers will be written by researchers who, thumbing through a census volume, are struck with the relevance of a single table or combination of tables to an important concern, and thereafter with pencil and paper, and perhaps with desk calculator, perform the necessary recombinations and calculations. Nevertheless, the lone researchers or groups of researchers at smaller universities have been placed at a considerable disadvantage vis-a-vis those in places with better computers, more programming assistance, and money to amass tapes. What is being ushered into the social sciences is the development of Big Science, performed in large establishments with large research staffs and experienced support personnel, as is already the case in the physical sciences. Again, we emphasize that much can be done by the imaginative researcher who has nothing more than pencil, paper, and a census volume, but he may have his productivity crippled by the awareness that others have better access to data and can obtain their results more easily. We can only foresee an increase in the psychological blocks to productivity for researchers in small and poorly supported institutions.

Although the physical sciences have moved in the direction of Big Science, partly because of costs and partly because of the many advantages to assembling a group of scientists to work on a specific problem, ways have been worked out to facilitate cooperative efforts and increase the possibilities of faculty members and students at smaller or faraway institutions. The reactors at AEC National Laboratories have been used for experiments by scientists from many universities, and fellowships have permitted students from all over the country to use the complex and expensive equipment found in national laboratories. In such ways Big Science, as represented by large research organizations, has aided in research and training in many other institutions.

In the South an opportunity for cooperative arrangements among social scientists is being provided by the development of a data center at the Oak Ridge National Laboratory (ORNL) in connection with work being done for the Department of Housing and Urban Development. At this center all of the tapes from the 1970 Census are being amassed, along with available tape files from the 1960 Census, the Social Security One-Percent Sample, County Business Patterns tapes, and a number of valuable files made available by individual researchers, such as the Bowles-Tarver County Migration tapes for 1950-60 and the SEA by SEA migration flows for 1955-1960. Cooperating in this endeavor have been the Department of Agriculture, the Rand Corporation, the Tennessee Valley Authority, and a number of universities throughout the country.

The Oak Ridge Associated Universities, a corporation sponsored by 43 Southern colleges and universities with graduate schools, was established to further cooperation between the National Laboratory and educational institutions. Recognizing the possibilities afforded by the data bank and by the computer facilities and technology at Oak Ridge, a group of demographers asked to be allowed to utilize the Oak Ridge Associated Universities in somewhat the same way as physical and biological scientists. The Southern Regional Demographic Group resulted, a group which now has over 150 individual members from many colleges and universities. Its purpose is to ease access to ORNL data and capabilities for members of the group, to organize and participate in studies which were beyond the facilities of a single university, and to organize conferences and symposia addressed to research problems and needs.

Initial funding was received from the Ford Foundation, and a series of conferences have been held. In March, 1971, a conference on Research and the 1970 Census was sponsored by the National Center for Population Research of NICHD, the papers from which have been published. A meeting scheduled for September 28-29, 1972, deals with "Research Needs for a Southern Population and Urbanization Policy." A central office with a full-time executive secretary supplies needed materials to member researchers, arranges for the dissemination of information on research carried out by its members, and facilitates cooperative endeavors among universities.

At ORNL preliminary attention has been focused upon methodological and data processing questions. The public use samples are being converted into binary form as an aid in compression and in speed of processing. Available computer programming packages such as SPSS are being modified and extended for use with the public use samples, as is Howard Brunsman's CENTS program. In these endeavors we are cooperating closely with DUALabs and research organizations in NIH and the Department of Commerce. Tapes have been received, catalogued, combined, extracted, compressed, and otherwise processed for immediate research or for wide spread usage.

A number of pilot studies have been made which were designed to explore the possibilities of varied materials and to discover the difficulties in their use. One such study, carried out in cooperation of the Tennessee Valley Authority, dealt with migration into and out of the Tennessee Valley, and gave us valuable insight into the use of Social Security materials in assessing the flows of labor. We then secured the cooperation of David Hirschberg and the Bureau of Economic Analysis in making a study of changes in the labor force in metropolitan areas.

We have now moved to the construction of a special tape for the Atlanta Metropolitan Area which permits us to follow over a ten-year period anyone who was ever in covered employment in that city. We can determine where he was before he came to Atlanta, what his annual earnings were, what he earned while he was in Atlanta, and where he went when he left Atlanta, all that by age, sex, color, and industry. Once such a study is made we hope to encourage comparable studies by other researchers in other cities. Our aim is not to direct the course of research, but to permit replication and improvement by studies in other areas. Our hope is that the results of these initial efforts will be multiplied and improved upon so that large initial processing costs need not be borne by researchers with little equipment or limited funds.

Similar exploration has been done with census materials and a study of the usefulness of city directories as a supplement to census information is under way. Examinations of black suburbanization have been made in several metropolitan areas, while other studies deal with the shift of industry from central cities to the rings. Growth by size of place has been studied as it relates to past growth, to location in regard to metropolitan areas, or in regard to interstate highways. As an aid in understanding areal patterns we have experimented with computer graphing. For example, we have worked out ways of graphing migration rates for counties by age, sex, and race so that we can get graphs for all 3000 counties in a few minutes of computer time. When we have the basic data on hand we will make similar graphs for 1970 and offer sets for 1960 and 1970 for publication in the various states. Another program converts density by enumeration district or block group to density per kilometer, giving us density maps for 1960 and 1970 that can be compared to see which areas lost and which gained population. Once done for a single city, processing for a large number of cities is straightforward. Furthermore, the same technique can be extended to black population, the aged, broken families, or to any other quantity we choose to relate to area.

Although our efforts have been going on for more than two years, we feel we are just beginning. We have experienced enough irritation from members of our own group at our slowness and at our fumbling with unfamiliar materials that we have come to sympathize greatly with the Census Bureau and other organizations that are delayed in the processing of eagerly awaited data. Frustration, failure, procrastination, and sheer stupidity among our own ranks has taught us humility. We shall not soon master the necessary techniques for processing and analysis of such a flood of data but there is an encouraging acceleration in our efforts.

Furthermore, we are even more convinced that cooperation among institutions and the establishment of non-profit and semi-public data centers are essential to the proper exploitation of census and other materials. One of the great failures of social science research is the lack of replication of studies in other places and for different times. We believe that more research per dollar and per scientist-year can be obtained through the kind of cooperation we have initiated and that it can be done better and more quickly. We are moving to institute fellowships and summer research opportunities. In cooperation with two different universities we have now produced two Ph.D.'s and we have brought students from a number of instiutions into our ongoing research. We expect to cooperate with the graduate schools of our member universities in providing data and facilities for students who will work with our help but under the direction of their own faculty committees.

Finally, we are encouraged enough by results to date to recommend the establishment of regional centers for the receipt and processing of data tapes. The kinds of data we have described will multiply far faster than the persons who can process and analyze them, or even find them. All too often the establishment of a computer center and the turn toward magnetic tape has slowed down research and increased expenses, rather than the reverse. We have never been so much in need of guides to available materials, of methodological studies of particular sets of data, and of the development of quick and rapid processing. As social scientists, we should follow the lead of the physical scientists to combine the efforts of Big Science with the furtherance of individual research.

Footnotes

[*]Paper presented at the meeting of the American Statistical Association in Montreal, August 1972.

[1]M. J. Bowman in Mark Blaug, Economics of Education, Volume I. Cambridge, Cambirdge University Press, 1968, pp. 114-115.

DISCUSSION

A. Ross Eckler, Consultant

Dr. Taeuber, the organizer of this session, is to be commended for planning a well balanced program covering an important, but often neglected stage in the production of statistics, a modern data delivery system, and the creation of a mechanism outside the government for utilizing more of the products of a national census. If I may shift the order to proceed from production through final use, Mr. Waksberg has reported on the plans for evaluating the 1970 Census of Population in the United States, work which I regard as the final stage of the production of a census. Dr. Murphy has told us about the means for the delivery of data in various forms from the 1971 Census of Canada. And, third, Dr. Lee has described a cooperative effort to establish a mechanism for the fuller utilization of material from the United States Bureau of the Census as well as other agencies.

These papers relate to one another so well that it is easy to forget that they are based upon the work of two different countries. Indeed, the papers furnish a good example of the principle of interchangeable parts. If Waksberg's paper had been presented by his Canadian counterpart, and Murphy's by his U. S. counterpart, the story would have been much the same, illustrating the effect of the long-continued close working relationships and interchange between Canada and the United States. This would not be equally true for the paper presented by Lee. In this case, differences in the scale of operations and in the resources available to users favor the earlier development of such a cooperative effort south of our common border.

A further reason for starting with Mr. Waksberg's paper is to be found in my own professional interest in evaluation of the quality of statistical data, as expressed in meetings of the International Statistical Institute beginning in 1951. The subject is one, moreover, with theoretical ramifications of considerable interest to many statisticians.

Dr. Taeuber's reference to President Washington's concern about the quality of the 1790 enumeration may have erroneously suggested a continuing attention to this problem. The fact is that for most of the long history of census taking in the United States, there has been an official tendency to take for granted that "if it's in the Census, it's right." One break in the long period of neglect - unfortunately, not benign - was that provided by General Francis A. Walker, Superintendent of the Censuses of 1870 and 1880. About one hundred years ago, General Walker, in words that might well have been written today, referred to the duty of producers of statistics to be candid about their shortcomings, even though as a consequence, the results of a census might receive less credit. However, his point of view was an exception to the prevailing one, and it was not until the midforties that the first post-enumeration survey was undertaken by the Bureau of the Census. Since that time it has been the standard practice of the Census Bureau, and of many other statistical agencies, including Statistics Canada, to regard the provision of measures of error in a census or survey as part of the task of production.

I believe that Mr. Waksberg's agency deserve much credit for introducing and maintaining this tradition of conceding the existence of errors, and seeking measures of their size and nature. Nevertheless, I am disappointed that we were not given more information on the quality of the 1970 Census. While it is true that the results of evaluation of the censuses prior to 1970 were delayed until well after the completion of other work, I had hoped that by the time of this meeting, about two and one third years after the date of the census, there might be provided a substantial amount of information about quality in 1970. Let us recognize that a great deal of work has been going on at the Census and hope that the completion of the evaluation task will be regarded as a high-priority undertaking.

Mr. Waksberg's report on coverage, based on earlier work by Jacob Siegel, was that underenumeration in 1970 had declined fractionally from the level for the preceding census. The improvement would have been about twice as great if the age-sex-color composition had remained unchanged over the decade. I would note also that the enumeration in 1970 took place after a bitter conflict which could easily have led to higher rates of undercoverage in 1970, except for the changes introduced in that year.

From my observation of the series of evaluation programs beginning in 1945, I am impressed with the increasingly sophisticated character of these undertakings. They provide a great number of different measurements, intended, as Waksberg has pointed out, to determine the effectiveness of the new 1970 procedures, to obtain another round of readings on the components of mean square error of important census statistics, to solve some problems of evaluation left from the 1960 work, and to pursue more intensively the correlates and causes of error in census statistics.

Over the years in which the evaluation of censuses has been undertaken, there have been marked shifts in the relative emphasis given to the various parts of the evaluation program. For example, much less attention is now given to the reinterview survey as a means of determining the amount of error in a census. With the recognition that this device is subject to shortcomings similar to those of the original interview, dependence is now placed upon the analysis of demographic characteristics as a means of estimating overall undercoverage. The reinterview survey is retained, to be sure, but mainly to measure coverage of housing units and response variance.

Record matches are being employed to a greater extent than ever before, one of the major uses being to give a measure of reporting bias. The matches include many different files, such as Medicare registrants, motor vehicle registrations, Internal Revenue Service returns, expenditures for utility services, and prices received for houses sold. Each of these is used to provide evidence on the bias of one or more items in the census.

I was particularly interested in the report on the results of comparisons between the Current Population Survey and the Census. In the current comparison, as was the case ten years ago, the two sets of data were in quite close agreement. The situation was quite different in 1940 and 1950. In both those years, the census unemployment figures were significantly lower, even though the totals for the employed were in quite close agreement. The lower census figures for the unemployed were explained on the basis of the omission of considerable numbers of unemployed by the Census enumerators, who failed to ask all of the questions needed to identify certain types of unemployed. The most generally accepted explanation of the better agreement in 1960 and 1970 is that most of the population were given a chance to classify themselves, so that the numbers classified as unemployed more nearly equalled those obtained by the trained enumerators in the Current Population Survey.

Dr. Murphy's excellent paper describes an area of census work where changes have been unusually great in the past 20 years. After the 1951 Census in Canada, it was probably true that the great majority of all uses of the census depended upon the printed volumes, and relatively few upon other types of output. Even after the 1961 Census, if one may judge from experience in the United States, the amount of information provided in other than printed form was still not great. By 1971, however, the situation has changed sharply, with so many users having access to computers.

Nevertheless, it is at first surprising to hear that 90 percent of the prepared tabulations for the 1971 Census of Canada, the equivalent of 3,800,000 pages, will be in the form of micro-film, micro-fiche, or summary tapes. Even if only 10 percent is in conventional published form, the resulting 380,000 pages will be double that made available in the United States. The requirements for bilingual presentation in Canada would appear to explain only a small part of the difference.

I should like to commend Dr. Murphy for his statements on the responsibility of a dissemination program to give a description of the collection and processing procedures as they affect the data. He notes particularly that the imputation procedures may have a significant impact upon the results for very small areas, such as enumeration areas. I have wondered, therefore, whether his office plans to provide counts of the numbers of imputed cases for enumeration areas. Along similar lines, he has referred to plans to have the Dictionary eventually include measures of root mean square error. I raise the question as to how soon this will be done, and as in the case of the United States, would hope for all possible speed in getting out these measures, since providing information on the size of error should be regarded as an integral part of the production process.

Dr. Murphy has referred to three forms of publication: first, summary data; second, individual records, presumably with area detail deleted; and third, individual records manipulated so as to avoid disclosure. However, relatively little has been done apparently in the release of information under the second and third categories. In the United States, a great amount of information has been released in the form of individual records, so that we shall be much interested in the safeguards for such work adopted by the Canadians, who because of the smaller size of the country have a correspondingly greater problem in protecting confidentiality of individual records.

Dr. Murphy has told us of the great volume of preplanned publication tables which are listed in a new document called the Tabulation Directory. However, there will be some needs for tabulations not included in the Directory, and for these special work is required. I was particularly interested in his brief description of what might be called a "do-it-yourself" tabulation program in which the user willing to spend a little time in learning the system can use a special language called TARELA for a program designated as STATPAK. The need for this approach may be greater in

Canada than in the United States, where there are many summary tape centers to bridge the gap between the central office and the isolated user.

Finally, in Dr. Lee's interesting paper, we have a description of one kind of organization that has developed in the United States to serve needs that could be met only with considerable difficulty and delay by the Census Bureau. In Canada, where the provincial statistical systems are considerably stronger than their counterparts in the States, the corresponding decentralization of service to the public may take place in the provinces.

The potential services to be rendered by an organization such as Dr. Lee has described are quite great. It has been estimated that business has utilized no more than ten percent of the useful products of a decennial census, and there is little reason to believe that the academic and research users have done much better.

I recognize that I am less sensitive than I used to be to criticism of the Census Bureau, but it seems to me that the attitude toward the Bureau is more tolerant than it used to be. Indeed, the relationship between the organization described by Dr. Lee and the Census Bureau is described in a manner quite sympathetic to the Bureau. Occasional references are made to census errors and delays, but always in a context recognizing the size of the task of getting out census tabulations and tapes.

Dr. Lee's reference to the role of "Big Science" in the field of research and training seems to be well justified. The tapes at the data center of the Oak Ridge National Laboratory include all of those from the 1970 Census, as well as tapes from the 1960 Census, the Social Security 1% sample, County Business Patterns, the Census of Manufactures and a series of files from individual researchers. The development of appropriate software to use with these resources and of cooperative relationships with a number of social and physical scientists makes this new Center unusually promising for demographic research making use of new tools now available to social scientists.

One can wholly applaud Dr. Lee's statement that cooperation among institutions and the establishment of non-profit and semi-public data centers are needed for the proper exploitation of census materials. But the costs for many undertakings will be relatively high. There will be false starts and it will be necessary to write off some large outlays that fail to meet the expectations of their sponsors. But the potential yields are very great also, and we can be glad that organizations like the Oak Ridge National Laboratory have both the resources and the resilient attitude necessary to realize the research potentialities.

MANPOWER CHARACTERISTICS SYSTEM

Robert W. Cain, National Science Foundation

Introduction

The National Science Foundation continually examines the needs for information on scientists and engineers, the data being collected, and the means used to collect such information. On the basis of this continuous review the Foundation has designed data collection systems, which seek to obtain the most comprehensive and reliable information for the least cost. The National Register of Scientific and Technical Personnel was a part of this system. While it was always realized that the use of the data collected through the Register mechanism was limited because of the uncertainties introduced by a relatively large non-response rate, it became increasingly apparent that the operation of the National Register was too costly. Furthermore, it became evident that the corollary objective of the Register of acting as a locator of scientific and technical personnel in times of national emergency could no longer be justified.

Therefore, the President's budget for FY 1972[1/], as submitted to the Congress, recommended the "discontinuation of the National Register of Scientific and Technical Personnel in its present form." It was further recommended that funds be appropriated "to allow for the development of alternative mechanisms for obtaining required information on scientists and engineers, ---." In addition, the Foundation with the request for FY 1972 funds was to "examine new methods of making scientific and manpower analyses."

The National Register of Scientific and Technical Personnel had the following features: The information was collected from individual scientists and engineers; data could be secured directly on the personal and professional characteristics of these persons; the information on individuals could be processed with a flexibility impossible in systems collecting only aggregate data and; it provided a capability for longitudinal analysis of matched individuals.

The Register, though, was only a part of the scientific and engineering manpower data program within the Federal Government. Several other programs exist. Some of these are supported and/or directed by the Foundation; others are operated independently by other agencies with their output being used by NSF. These programs primarily involve collections of information from employers of scientists and engineers and produce summary information on the numbers employed and functions in which they engage. However, these types of collections do not provide a capability to collect detailed data on individual scientists and engineers, such as age, level of degree, work experience, or scientific specialty, primarily because in most instances, employers generally do not maintain records on such subjects and because they are usually not willing to provide anything but aggregated information. Thus, this method of collection does not provide a capability to obtain data on matched individuals for longitudinal analyses.

Determination of Needs for Data on Scientists and Engineers

The determination of needs for data must be based on the specific or stated and as well the anticipated requirements of policy or decisionmakers, both within and outside the Government. Such information also serves the requirements of persons desiring information for the analysis of technical manpower issues and problems. These needs for data vary by sub-group of the science and engineering population, such as level of educational attainment (doctorate and nondoctorates), sector of employment (education, Government, industry, etc.), and activity (R&D, etc.). Thus, various data elements are essential for the analysis of a given subject at a given point in time. Furthermore, a data collection system must incorporate the capability to supply needed information on a continuing basis. The mechanisms selected, however, need not provide for the continuous collection of any given data element. Rather, specific elements can be obtained at various time intervals from a given sub-group of the population with some elements collected more frequently than others.

NSF examined the data elements formerly incorporated in the Register operation from the standpoint of use and reliability. Other data elements, not previously collected through the Register, were also examined in terms of essentiality both for decisionmakers and other users. Also, reviewed were previously stated needs for information as expressed in requests of various bodies, such as the Congress, the National Science Board, the Executive Office, other NSF units, etc. These examinations indicate that two general classes of data elements can be set forth: Essential Needs--that is those elements which are needed on a more or less continuous basis from all subgroups of the science and engineering population; and Lesser Needs--those elements which might be collected less frequently or only from one or more subgroups. Essential Data Needs include: age, sex, level of degree, employment status, types of employer, program employment, work activity, Government support, field or specialty, annual salary, and postdoctorate status. Lesser Data Needs include: citizenship, student status, total professional income, work experience, and secondary school education.

Development of Alternative Systems

The development of alternative mechanisms to the National Register was undertaken by the National Science Foundation on the basis that these mechanisms should supply selected data on representative samples of individual scientists and engineers in the absence of the Register and be responsive to the development of an adequate, timely, and relatively comprehensive manpower data system.

Possible alternative systems were realistically circumscribed by such requirements as: (1) The information provided must be truly representative of a defined population (or populations); (2) The information collected must include at least the "Essential" Data elements required by policy makers; and (3) the information should be obtained and analyzed in the most efficient means possible, and preferably at a cost less than that of the National Register. The cost criterion also dictated that any proposed new systems be based on existing ones.

Manpower Characteristics System

These principles led to the identification and examination of several alternative approaches in terms of practicality of collection, coverage, costs, flexibility, and so forth. After consideration of these alternatives, the following one was selected as the appropriate approach. This approach now called the Manpower Characteristics System consists of three separate and specific mechanisms or systems: A Comprehensive Doctorate Roster, based on the existing National Research Council efforts; a continuous periodic collection from a National Sample of Scientists and Engineers based on the 1970 Decennial Census; and Surveys of Nondoctorate Entrants to Science and Engineering, based on the present surveys of undergraduates conducted by the American Council on Education. This combination of collections will permit detailed periodic coverage of doctorates, as well as relatively adequate coverage of the total science and engineering population and the critical new entrants. Though each of the subsystems consists of periodic collections of data from individual scientists and engineers, it will not be possible to match or integrate specific individual scientists and engineers across the subsystems, because of the confidential nature of the records. However, it will be possible to integrate and analyze data on separate cohorts or populations.

The three separate parts of the Manpower Characteristics System are presently in varying stages of development as follows:

Comprehensive Doctorate Roster. The Office of Scientific Personnel of the National Research Council is now proceeding under contract with the National Science Foundation to develop a Comprehensive Doctorate Roster. Over many years, the Council with the support of the Foundation, and several other Federal agencies, has maintained a Doctorate Record File. This file consist of information collected from each person receiving an earned doctorate, including data on personal characteristics, educational background, and career plans. One of the first stages of the development of the Roster will be to collate this file of information with the records on doctorates contained in the several biennial (1954 to 1970) registerations of the National Register. Data on persons earning a doctorate from other than U. S. institutions and presently employed in the United States will also be assembled. Using these sources, a universe of doctorates will thus be established. Under this plan, the Foundation hopes to have by the fall of 1972 a file containing the names, activity status, and current address of as many of the estimated 190,000 doctoral scientists and engineers as can be reached. Once established, the Comprehensive Doctorate Roster will be maintained on a current basis with respect to the inflow of recipients of newly awarded doctorate degrees from U. S. universities (to foreign and U.S. citizens). In addition, efforts will continue on including persons (U.S. and foreign citizens) receiving doctorates from foreign institutions and employed in the United States.

The Roster will then be available for general and special purpose data programs as needed. Without anticipating all uses of this resource, the operation will include routinely an annual survey of newly awarded doctorates, taken about a year after the award and covering the nature of post-award activity (other pertinent data will be available from the Doctorate Record File); a biennial collection of Essential Data from a cross-section sample of the entire science and engineering doctorate population, as well as a biennial updating of mailing addresses of the doctorate population. Additional data elements will be included as the need arises.

National Sample of Scientists and Engineers. The 1970 Census of Population collected information on the education, occupation, industry of employment, and other characteristics of a 20-percent sample of the entire U.S. population, including approximately 125,000 scientists, 240,000 engineers, and about 2.7 million persons with four or more years of college. This information will provide benchmark data on the science and engineering population, as it existed in 1970.

The Foundation is presently supporting a Postcensal Survey of Scientific and Professional Manpower directed at a 100,000-case subsample, drawn from the Census 20-percent sample, of persons in science and engineering occupations and persons with four or more years of college. The project is being carried out by the Bureau of the Census. The actual collection of detailed information took place in spring and summer 1972 and obtained 1970 and 1972 information to provide a more comprehensive picture of the 1970 science and engineering work force. (Mr. Seltzer will describe this project in greater detail).

In addition, the Foundation has supported the Bureau of the Census in obtaining an additional sample of 50,000 cases from the Decennial Census. These additional cases were surveyed in early summer of 1972 to obtain new addresses and selected items of data. Information collected from the two samples selected from the 1970 Census (100,000 from the Postcensal Survey and the additional 50,000 cases) will be examined using a set of screening criteria (age, level of education, and employment) to establish an adjusted science and engineering population. This population will then be used as the foundation for a series of periodic biennial examinations of the total science and engineering work force. Additional benefits accruing from the larger size sample include greater flexibility in cross tabulative data and an improvement in reliability, especially for smaller cells.

Survey of New Additions of Nondoctorates. An adequate collection of information on the scientific and engineering workforce must include data on persons entering the work force for the first time--primarily new graduates.The collection of such information is particularly important, as a supplement to the 1970 population of scientists and engineers obtained from the Census. The proposed comprehensive Doctorate Roster will provide a system for obtaining needed information on newly-awarded doctorates. However, a substantial data gap will exist on new entrants at the bachelor's and master's degree levels. Only a minority of those receiving science baccalaureates directly enter science activities; however, a sizeable proportion of new engineering graduates enter engineering. Many new baccalaureates enter graduate training (though not always in science or engineering fields), some enter other careers (high school teaching, non-technical fields, and even technician jobs), and some (mainly women) do not enter the labor force. At the master's degree level similar patterns exist. Few surveys have been undertaken to obtain information on the activities of new graduates, particularly below the doctorate level.

A basic system exists for obtaining information on new baccalaureates--the American Council on Education's longitudinal survey of college freshmen. Each year since 1966, a cohort of new freshman has been established and a series of follow-up surveys conducted on their undergraduate careers.

In mid 1972, the American Council on Education was awarded a development contract to develop a system of obtaining information on baccalaureates in science and engineering. The surveys of nondoctorate entrants will be based upon the cohorts of freshman sureyed each year by the American Council on Education. The goal of this part of the system include a series of continuing alternate year follow-up surveys of these freshman after graduation to determine trends in the career and educational patterns of new baccalaureates. These surveys will be carried out through the 1970's as a series of overlapping annual collections, including, coverage of a new cohort every two or three years, and on the alternate years coverage of a new cohort every two or three years, and on the alternate years coverage of preceding classes at 2, 3, 4, etc., years after graduation, including after several years coverage of a class of which a substantial proportion had earned a master's degree. Experience gained after several of these surveys could lead to a variation in this pattern.

In FY 1972 and FY 1973, a considerable amount of effort will be expended on the development of these systems. In later years, as the systems mature, it will be necessary to continually carry out systems research and evaluation involving reexamination of the techniques of collection, methodology of analysis, and identification of subtle biases. For example, it is already realized that the method of collecting information on new nondoctorate entrants was chosen because a system existed which could be adapted to these purposes within the resources available. However, some new entrants will be missed, including non-degree and foreign scientists and engineers. It is hoped that the systems research activity will develop means to correct these shortcomings and identify other subtle biases arising out of new circumstances. To these and similar ends, a limited program of research and development on these subjects is proposed. A series of small projects will be designed to be carried out annually beginning in FY 1974.

Thus, the three separate, but interrelated, parts of the Manpower Characteristics system are being established to provide a continuing mechanism for collecting and analyzing information on the science and engineering population of the United States. It has all the features of the Register described above--collection of personal and professional information from individual scientists and engineers, which can be processed on a single time and on a longitudinal basis. Furthermore, the information obtainable from the system will be more truly representative of defined populations, will include essential data elements, and can be provided with less resources than the Register.

1/ Office of Management and Budget, Special Analyses, Budget of the United States, Fiscal Year 1972, p. 89.

THE 1972 POSTCENSAL SURVEY OF PROFESSIONAL SCIENTIFIC AND TECHNICAL MANPOWER

Norman Seltzer, National Science Foundation

The need for meaningful data on a Nation's resources of all types of manpower has become more urgent in recent years as Government (at all levels) has become more involved in the planning, funding, and evaluation of many economic, educational, and scientific policies and programs affecting all sectors of the economy. It is generally taken for granted today that the stock of trained scientific and technical personnel in the U.S. is one of the controlling and essential resources determining scientific and technological activity for economic growth, technological change, and the advancement of knowledge. Furthermore, scientific and technical manpower has, over the past decade, assumed a national and international importance far beyond its numerical size.

In pursuit of more meaningful information about this resource, the NSF decided that the 1970 Census of Population presented a once-in-ten years opportunity (as it did after the 1960 Census) to obtain on a nationwide basis a large array of detailed information on high level manpower.

In the 1970 Census of Population, 20 percent of the enumerated population supplied information on labor force status, occupation, and level of education (i.e., number of years of schooling completed). Utilizing this 20 percent sample, the Foundation planned (with the cooperation of the Census Bureau) to exploit this comprehensive source of information by surveying in detail two sample populations of persons in the civilian labor force: (1) those classified in the Census in scientific, engineering, technical, and related occupations, and (2) those with 4 or more years of higher education (post-H.S.) regardless of current occupation. This project provides the opportunity to obtain, in much greater detail than would be available from the regular Census, data on occupations, functional activities, types of employment, formal education and supplementary training, salary, and mobility, and the relationships among these various characteristics.

Broadly speaking, the Postcensal Project provides three main inputs to assessing our scientific and technical manpower resources:

(1) A nation-wide inventory of the characteristics of U.S. manpower resources, employed and/or trained in scientific, engineering, and technical fields;

(2) Insights into the career patterns of the scientific, engineering and technical manpower resources as related to education and training and other demographic variables for various cohort groups; and

(3) A comprehensive and basic statistical benchmark for evaluating and reconciling data on scientific and technical manpower obtained from a variety of sources.

The postcensal survey also provides the opportunity to obtain two unique sets of information not available from other sources:

(1) Detailed data representing the entire work force of natural and social scientists, engineers, and technicians currently employed for all sectors of the economy at a specific point in time.

(2) General information representing the entire college-educated population (i.e., four or more years of college education) regardless of current occupation at a specific point in time.

A brief review of some of the substantive outputs to be obtained through this project follows:

To begin with, because in the 1970 Census schedule, the amount of information available pertaining to a person's occupation and work activities is quite limited, an attempt is being made to obtain some insight into the extent to which interdisciplinary work in science and technology has resulted in persons classified as engineers working in an area of the physical or life sciences, physicists concerning themselves primarily with some aspect of the medical sciences, or mathematicians calling physics their present field of work.

Another equally important area of job information is the activities or duties that are actually performed; that is, what do people classified in professional and technical occupations "really do" in their jobs. Although we may have some indication that a certain number of engineers may be involved in "research", what the varied job requirements or duties of these personnel are, has not been too well known. For some, this may mean that aside from engaging in research, the job may entail supervising a team of other professional personnel, and writing technical reports; for others, there may be sales duties, and making estimates of markets for new products. For persons in other occupations, there are of course a similar wide range of activities which make up the different types of jobs in which such personnel are engaged.

If our knowledge in the past concerning the current employment and job activities of professional and technical personnel had been limited, this has been even more so about the process over time by which cohorts of such highly trained persons are allocated to various jobs and employers, the career paths which may characterize different groups, and the movement of scientific and technical personnel between various types of employers, occupations, and work activities. Insights into this complex area would be helpful in dealing with an assortment of problems including the one that policy makers always want a definitive answer to--the supply and demand of scientific and technical personnel. To this end, detailed information on employment and work activities has been obtained for the most current job and the two prior jobs held as well as for the entry job into the labor force. It is obvious, of course, that such information cannot encompass complete work histories but it should provide a broad overview of mobility patterns. It will be possible to analyze many factors in relation to changers and non-changers among the various occupational group.

What insights can be provided by such data? The past decade, up to the present time, has been marked by, among other things, changes and shifts in emphasis in vast Federal Government expenditures for research and development, a build-up and cut-back of activities in both government and industry for the space and defense programs, an increase in existing as well as new programs for medical and health research, more emphasis on the development of new products in many technologically-oriented industries, and a considerable expansion and then leveling off of educational facilities to accommodate the influx of new students and provide for ongoing research programs. Against this background, the recent mobility data will provide an evaluation of the movement between employers, jobs, activities, and geographic locations. For example, have more scientists been moving from academic employers to industrial jobs than vice versa? Are a greater proportion of engineers concerned with administrative or supervisory duties than heretofore? Are certain industries attracting a higher proportion of the mobile personnel? Does there appear to be a shifting or upgrading of persons in non-professional jobs (i.e., the technician occupations) to professional occupations? To what extent have these various movements among occupations, employers, functions, etc., resulted in substantial shifts geographically?

By going back to the entry point for job histories, it may be possible to establish typical and variant career histories for specific occupations and occupational groups, for respondents with specific levels and types of training, and for those with certain demographic characteristics.

Several additional items of employment and labor market interest on which data is being obtained relate to reasons for changing jobs, how these persons acquired their different jobs, whether their work was supported or sponsored by Government funds, and levels of salary and income.

In the area of training, a considerable amount of detail was sought on various facets of both formal education and informal types of training. By and large, persons in the occupations covered in this survey have a fairly high level of training, especially when compared to the general population. Requirements for employment in these scientific and technical occupations generally require this background--even more so in the past few decades.

To begin with, since information on training has been obtained as of 1972, it will be possible to update the Census occupational information on number of years of formal training completed. However, the primary interest lies beyond this data, in that we want to determine some of the specifics of higher education obtained in relation to subsequent employment, and data was therefore requested on major fields of study for different types of degrees granted, where appropriate. In addition, because of the general knowledge that a substantial amount of training takes place outside of the formal educational system, inquiry was made about informal types of training received, such as company training programs, military training applicable to civilian occupations, home study correspondence courses, etc.

Some of the more apparent uses of this information includes: a detailed description of the formal education and training of persons in various professional, scientific, and technical occupations; an analysis of current occupation and work activities as well as overall job histories in relation to major fields of study at both undergraduate and graduate levels; the extent to which persons are employed with less than a college in professional occupations, and what types of informal training as well as experience may have contributed to their attaining such positions; and, an analysis of the personal and other background characteristics of the respondents to determine whether any insight can be obtained regarding differences in levels of training and subject matter studied.

Last (but not least), for these occupational groups, information was sought on background and personal characteristics both to supplement data available from Census and other sources and as factors to relate to data obtained in the areas dealing with employment and training.

As indicated previously, a sample of the 4-year college group was surveyed (outside of the target occupations) from whom similar information was obtained in order to provide a general analysis of the Nation's college-educated population. Insights are being sought on the extent to which persons trained in scientific and technical fields were, in 1970 (or beyond), working in occupations seemingly unrelated to this training; the same for persons who started their careers in scientific and technical occupations and were employed elsewhere in 1970, and finally, despite present conditions, what potential might exist for reemployment in scientific and technical fields.

It should also be mentioned here that with all of the information being obtained in this study, there exists the possibility of developing manpower models using multiple regressions and similar techniques which would enable us to relate data on income to other factors such as age, education, industry, geographic location, activities, etc. One additional and important note relating to the problem of occupational information remains to be mentioned. Because data on scientists, engineers, and technicians have been obtained over the years from various sources (i.e., individuals' self-reporting, employer surveys, household surveys, etc.), it has been difficult to reconcile the different counts obtained for persons in such occupations. Using the postcensal data based on the Census nation-wide occupational sample, it is our intention to redefine the populations in the target occupations employing various criteria which will combine such reported items as occupation in different jobs, self-identification in terms of education and experience, level of education attained, fields of study, current national society membership, and valid professional licensure or registration. For example, in the case of engineers, whereas the Census categorized as such only those currently working as engineers, or in the case of the unemployed those last working as engineers, we will use the aforementioned criteria to estimate different engineering populations for specific points in time. This will entail eliminating some persons originally identified as engineers by Census (drop-outs), and transferring into the engineering group some persons originally classified in the other occupational groups or in the broad college-educated group (drop-ins). With newly estimated populations in hand, various inventory and analytical tabulations will be obtained.

In closing, the body of information to be obtained in the Postcensal Survey will be substantial, however, the requirement for timely answers to manpower questions and problems continues to grow. Data collection and analysis of scientific and technical manpower issues will continue to face greater challenges in the years ahead. The growing awareness of the need for better information on these resources are increasing, and it is anticipated that the postcensal survey results will provide a meaningful base for future work.

THE CENSUS ROLE IN SCIENTIFIC AND TECHNICAL MANPOWER PROGRAMS FOR THE 70'S

C. L. Kincannon
Social and Economic Statistics Administration, Bureau of the Census

INTRODUCTION

The extensive data needs of the National Science Foundation and the proximate universality of the Census of Population as a classifying device are the underlying circumstances engendering a supporting role for the Bureau of the Census in the Foundation's statistical programs for scientific and technical manpower in the 1970's. In briefest summary, the Bureau's role is to use the Census of Population as a device to focus on the particular groups in the population in which NSF is interested, then expand the amount of information available by a specialized follow-on survey and finally to extend the data through time by periodic recanvassing of the same population. The task of identifying these particular population groups and expanding the amount of information available for them comprises the Postcensal Manpower Survey (PMS). The task of extending the availability of data through time will be the function of a series of annual or biennial science and engineering manpower surveys.

A single sample selection process served both these tasks by selecting a sufficient number of cases (designated as Sample I) to meet the needs of the Postcensal Manpower Survey and by simultaneously selecting, for specified occupations, an equal number of cases (designated as Sample II) for use in maintaining sufficient panel size for periodic recanvassing. The components of the samples are shown in List A. Both Samples I and II were canvassed in approximately the same span of time, and comparative data on response rates are presented in table 2. However, the focus of this paper is on the procedural and methodological aspects of the Postcensal Manpower Survey, and comment on the work connected with Sample II is limited chiefly to salient differences between the purpose and treatment of the two samples.

POSTCENSAL MANPOWER SURVEY

The Postcensal Manpower Survey is a large-scale survey addressed to persons recorded in the Census of Population as in the experienced civilian labor force and reported in certain occupations or identified as having completed 4 or more years of college. It was designed to obtain detailed information about the occupations and careers of these persons beyond that which is available in the census. Of course, the individual information collected in the survey is held in confidence by the Bureau, and only statistical summaries are made available to the sponsor and the public. The survey is basically similar to the postcensal survey undertaken by the Bureau for NSF in 1962, but it does represent a major effort to profit by and improve upon the experience gained in that earlier survey. The major phases of the Postcensal Manpower Survey are four:

1. Design of the questionnaire and related survey documents;
2. preparation of a mailing list, including sample design and selection, and locating the census questionnaire of each selected person;
3. data collection; and
4. data processing, including the editing, weighting, and tabulation of the information submitted on the questionnaires.

Pretest.--A pretest of the questionnaire content and procedures for this survey was conducted in Dane County, Wisconsin, in the spring of 1970. From the 1968 special census of Dane County, questionnaires were located for persons in target scientific and technical occupations and with 4 or more years of college, and their names and addresses were transcribed for use in mailing the pretest questionnaires. The data collection phase comprised an initial mailing, a reminder postal card, and two followup mailings by first class postage. Selected panels of nonrespondents were treated to experimental followups using certified mail and the telephone. Returned questionnaires underwent clerical coding and editing and were key coded on magnetic tape. The resulting data tapes were subjected to a computer editing program, and item tallies were run for the purpose of evaluating the effectiveness of the questionnaire. The pretest included 10,942 persons of whom 7,549, or 68.8 percent, returned completed questionnaires. Comments on particular experiences in the pretest as they apply to the postcensal survey will be made under the appropriate sections below.

DESIGN OF QUESTIONNAIRES [1]

Survey goals.--The preparation of the questionnaires and other forms to be used in the survey involved a process of design, testing, and review over a period of 2 years. Two major requirements guided the development of these forms. First, the PMS questionnaires should obtain key information about scientists and engineers beyond that which would be available in the census. Second, the National Science Foundation desired to be able to identify persons who might be considered to be engineers or scientists according to definitions other than that prevailing in the census. The census identifies as such only those persons who reported themselves to be currently working as engineers or scientists (or unemployed or in the labor reserve and having last worked as engineers or scientists). Meeting this second requirement, then, involved in addition being able to identify persons who had been trained or otherwise qualified in engineering and science fields. These goals required the addition of many questions about educational experience and career patterns, including specific degrees worked for or acquired, the fields of study for each of these degrees, the sources of financial support for academic work, and, in addition, information about jobs held at periods other than the time of the census, first professional job held, self-identification of one's profession, membership in national societies, and the holding of valid professional licenses. Finally, changes in the general economic situation and particular dislocations in industries employing significant numbers of scientific and technical personnel, which occurred in the course of our planning, emphasized the importance of obtaining updated information on the employment situation for highly qualified manpower in general. This did not require significant modification of the basic design of the questionnaire, for it included questions which permitted the determination of current labor force status, recent employment experience, incidence of part-time employment, layoffs, reduction in wage rates, and pejorative occupational changes; however, these developments did affect our analytical view of certain questions.

Sources of content.--In designing the questionnaires, we sought to take advantage of the experience gained in other surveys. Certain questions, for example, were taken from the 1970 Census of Population and Housing forms. These were mainly of two types: (1) Questions on unchanging characteristics (e.g. date of birth, sex) which were repeated to make certain that we had indeed obtained the questionnaire data from the intended person and (2) questions on characteristics which could legitimately change between the time of the census and the time of the Postcensal Manpower Survey (e.g., educational attainment, employment status, marital status). Naturally, another major source of questions was the 1962 Postcensal Study of Professional and Technical Manpower. Other census surveys which contributed were the 1958 Survey of Professional Manpower, the Consumer Anticipations Survey, and the Subject Response Study for the 1970 Census. Although not an exhaustive list, the following also were sources for content of the PMS questionnaire: The 1968 National Engineers Register and National Register of Scientific and Technical Personnel (NSF), Transferability and Retraining of Defense Engineers (U.S. Arms Control and Disarmament Agency and Stanford Research Institute), the 1967 Professional, Scientific, and Technical Manpower Survey (Canada), and the 1966 Survey of Professional Engineers (United Kingdom). In addition, there were some topics which, although not appearing on any of the surveys which we reviewed, were added in order to contribute to meeting the goals of the survey. These included questions on such subjects as receipt of bonuses or participation in profit-sharing plans, means of acquiring specific jobs, reasons for leaving specific jobs, and the association of changes in residence and changes in jobs.

The process of questionnaire development included consultation with more than 2 dozen interested parties from private organizations and such Government agencies as the National Institutes of Health, the Bureau of Labor Statistics, the Manpower Administration, the Office of Education, and the Office of Management and Budget.

Retrospective inquiries.--At least two major differences between the 1962 survey questionnaire and the 1972 PMS questionnaire should be pointed out. One is the difference in handling retrospective questions about careers, and the other is the method of coding and reporting industry and occupation entries. In 1962, there were batteries of questions dealing with the respondent's current (1962) employment, his employment at the time of the census (April 1960), and the full-time civilian job held upon reaching age 24. Because of the acknowledged difficulties in obtaining accurate information when asking about employment at a given point in the past, it was decided to approach the topic of past employment much on the order of a resume; that is, we chose to ask the respondent to give us information about his current job (or his last job if not currently employed) and the two jobs immediately preceding it. The beginning and ending dates associated with each of the three jobs are also requested and from those it is possible to determine occupation at a specific point in time (e.g., April 1970) as well as to calculate the duration of jobs and the periods between jobs. In addition, the questionnaire sent to professional workers also asks about the occupation and beginning date of the first full-time professional job ever held and the occupation and ending date of the last regular full-time job prior to that. The point of this is to more precisely determine the date and nature of the professional career entry job.

Respondent coding of industry and occupation.--The wording of the questions on industry and occupation was essentially the same for 1962 and 1972, as both are designed to be consistent with Census practice. However, in 1962 the question was open ended; the respondent was allowed to enter whatever response he desired, and that entry was coded clerically using the techniques which served in the Decennial Census. In 1972, the National Science Foundation strongly desired to avoid the clerical coding step and preferred to have the respondent choose from a set of lists enclosed with the questionnaire the code which best described his industry and occupation as well as field of specialization for degrees. This procedure has the advantage of standardizing response and, of course, eliminates a major clerical expense. This procedure was tried in the pretest, and a comparison was made of the code given to the respondent's occupation on the Dane County Special Census questionnaire and the code picked from the list by the respondent for the job he reported on the PMS pretest form as being held in April 1968. About 70 to 80 percent of these comparisons fell within the same occupational group (i.e., engineers, social scientists, etc.), and since there may have been differences in the actual job reported or in the descriptions rendered, this self-coding technique was considered to be acceptable for use in the national survey.

Transmittal letters.--All transmittal letters in the pretest were sent over the signature of Bureau officials, whereas in the national survey, the initial transmittal letter was sent over the signature of a National Science Foundation official with subsequent transmittal letters signed by Bureau officials. One rather important difference in the transmittal letters for the pretest and the national survey should be noted. Bureau administrative regulations introduced after the pretest required the inclusion in the transmittal letters for the national survey of a specific statement that participation in the survey was voluntary. This change doubtless weakened response in the national survey relative to the pretest; however, the pretest transmittal letter clearly stated that the survey was a test confined to Dane County, and the lack of such qualifications in the national survey letter probably favored response. The precise net effect of these changes on the response rate is not known, as we have not yet prepared a separate tally of response rates for Dane County in the national survey.

PREPARATION OF THE MAILING LIST

Sample design and selection.--The sample of approximately 108,000 persons for the Postcensal Manpower Survey (Sample I) was selected from 41 groups of census occupations as shown in List A. These 41 groups were composed of 65 target occupations or strata defined by the Census of Population. Within each stratum, a systematic sample of persons was selected across all 50 States and the District of Columbia. The sampling frame for the survey was the 20-percent census sample records. Each record had been assigned a weight by an involved ratio estimation procedure which is briefly described in the 1970 Census of Population publications. The probability of selection of a person within a stratum was proportional to this 20 percent census weight. The selected cases were confined to the records of persons 16 years old and over and in the experienced civilian labor force. Persons living in group quarters were excluded from the sample. Sample groups 1-31 comprise 54 primary target occupations, and the sample as selected will support detailed tabulations of respondents' characteristics in all but sample group 21. Sample groups 32-40 comprise 9 secondary target occupations with a known propensity for persons who, by training or other qualifications, could be included in the primary target occupations under alternative occupational criteria. The sample size for each of these 9 groups will also permit detailed tabulations of respondents' characteristics in all but sample group 34. The sampling fractions for groups 1-40 do not differ by more than a factor of four, so that it will be possible to reclassify the respondents by alternative occupational criteria (other than the "working as" concept) and tabulate the detailed characteristics of the resulting new groups without increasing the variance for a given sample group by more than 50 percent. Sample group 41 is the residual population in the experienced civilian labor force reporting 4 or more years of college. The size of the sample for this group will permit (1) the tabulation of aggregate statistics for the group and (2) the calculation of the aggregate number of persons who would be in primary target occupations (engineers or scientists only) if no coding errors were made or if alternative occupational criteria were applied.

Sample groups 1-40 include only persons with those occupations actually reported in the census. Persons with imputed entries in those occupations are in the scope of sample group 41.

As stated above, simultaneously with the selection of Sample I for the Postcensal Manpower Survey, another panel of about 58,000 cases, Sample II, was selected for possible future longitudinal surveys. Sample II was selected only from sample groups 1-25 and is equivalent in size to Sample I for those groups. The main purpose in selecting the supplementary sample was to create a reserve to maintain sufficient panel size to permit recanvassing this population annually or biennially to ascertain changes in labor force status and other characteristics. As a byproduct, it is also possible to combine Samples I and II and, for items appearing on the questionnaires for both surveys, produce tabulations with about half the variance as for either sample separately. This feature will be especially useful in tabulating unemployment rates.

The actual sample selection was made in two stages: First, all in-scope records on the Census of Population sample detail file were stripped off and tallied to permit calculation of precise sampling fractions. Second, the sample was selected from the stripped file, and printouts (sample listing sheets) were prepared containing the serial numbers and necessary identifying information for locating the census questionnaire for each selected case.

Obtaining names and addresses.-- The process of searching for the appropriate census questionnaire and transcribing the names and addresses for the PMS sample was a major component of the survey's clerical workload, representing the expenditure of more than 12,000 man-hours over a 6-week period beginning in January 1972. The searching was done in two phases. First, all the enumeration district boxes for the indicated sample cases were isolated and the individual census schedule for each sample case was pulled. Second, the correct sample person was identified by matching characteristics on the sample listing sheet to the census questionnaire. Then the sample person's name and address were transcribed to the sample listing sheet. The serial number of each case was preprinted on the sample listing sheet to obviate the possibility of error in transcribing this number. It was also necessary to use address registers in conjunction with the census questionnaires for all rural areas and for selected urban areas. After the transcription operation was completed and verified, the census questionnaire sample pages pertaining to each sample person were microfilmed for future research use. The serial numbers, names, and addresses on the sample listing sheets were then key coded directly onto magnetic tape; a check-digit feature virtually prevented error in keying the serial number. The resulting data tape comprised the survey master control file and was used to create mailing labels and to check in responses. Of the original sample of 108,000 selected, Census of Population questionnaires were located and names and addresses were transcribed for almost 102,000 (94.2 percent).

DATA COLLECTION

Summary of response.-- The data collection period extended from February 17 to July 17, 1972. In summary, for 12.4 percent of the persons mailed questionnaires, we received no response whatsoever; an additional 4.8 percent of the questionnaires mailed out were returned as not deliverable by the post office; 7.6 percent of the persons to whom questionnaires were mailed refused specifically and categorically to participate in the survey; 0.7 percent were reported as deceased; 0.8 percent returned questionnaires with so little data that it was impossible to process them; and 73.1 percent returned questionnaires with sufficient data for processing. The latter have been termed "completed" questionnaires in this paper. About 0.2 percent of the panel were deleted as being out of scope or by reason of illness so severe that they should be excused from the survey.

Mail canvass.-- The initial mailing of the survey was on February 17, 1972, with followups going out to nonrespondents on March 9 and March 28. All three of these mailings were sent by first class postage. Preliminary results are shown in table 1. As expected, the successive mailings produced dinimishing returns with the proportion returning completed questionnaires declining from 26.1 percent in the initial mailing to 21.4 percent of those in the second mailing and 18.3 percent of those in the third mailing. An additional followup was sent by certified mail on April 24. Proportionately, this mailing was more effective than any of the earlier mailings with 31.2 percent returning completed questionnaires. However, the proportion refusing to complete the questionnaires increased sharply from less than 1 percent of each of the first three mailings to 3.7 percent of the certified. In fact, one-fifth of the total refusals of the survey were received in response to the certified mailing.

The total number of cases returned by the post office as undeliverable (postmaster returns) was approximately 6,000. While successive followups were being sent to nonrespondents, a program was underway to reduce this number. Each piece of mail returned by the post office was remailed to the same address to make sure that its return was not a result of mishandling. If returned a second time for the same reason, the microfilm record of that person's census questionnaire was looked up. These questionnaires were from either the 5 percent or the 15 percent sample in the Census of Population. The 15 percent questionnaire included an item on the name and address of the person's 1970 employer. The PMS package was remailed to the respondent's name in care of that employer. This program reduced the number of postmaster returns by about one-fifth to a total of 4,920.

Telephone followup.--We began telephoning all remaining nonrespondents on May 9. Approximately 32,000 persons were delinquent at that time. All these persons could not be called simultaneously, of course, and 3,100 of them responded before they could be telephoned. In addition, nearly 12,000 could not be contacted either because no telephone number was available or because the number we did obtain was not a good number. The purpose of the telephone call was to ascertain whether or not the respondent had received the questionnaire originally, to obtain the correct address and mail another one if he had not, and to secure his agreement to return a completed form. Of the 17,000 persons actually called, 4,900 refused to complete the questionnaire but agreed to answer an abbreviated list of questions on the telephone; 5,000 refused to cooperate to any extent. However, a good number, approximately 7,000, agreed to complete and return the regular questionnaire. Later, a check was made and those who promised to return a questionnaire but did not (4,900) were called again. These repeat calls yielded an additional 2,000 telephone interviews and almost 500 refusals. In total, of those actually contacted by telephone, 51.6 percent either returned completed questionnaires or gave answers to a list of questions over the telephone. But an additional 32.1 percent refused to participate in the survey; this comprised 71.0 percent of the total refusals for the survey.

Response by occupation and geographic area.--There were some notable differences among the occupation groups in the return rates, as can be seen from the preliminary results shown in table 2. Engineering and science technicians had the lowest rate of completed questionnaires returned (68.8 percent) of all the primary target occupations; this rate was significantly lower than the average for engineers and scientists of 75.0 percent. It was also lower than the rate (72.3 percent) for the secondary target occupations. Life and physical scientists had the highest rate of return of any occupation group at 79.4 percent. The other engineering and science occupations all ranged between 73.0 percent and 74.8 percent, with the exception of computer specialists at 71.9 percent.

Note that response rates for Sample II are higher than for Sample I in every occupation group, although the differences between occupation groups are approximately the same. Since the only significant difference between the techniques used to survey Sample II and sample groups 1-25 in Sample I is the length of the questionnaire, the disparity in response rates indicates the price one pays for increasing the respondent's burden. Furthermore, since the rate of completed questionnaires for Sample I includes cases for which only the short list of questions was answered on the telephone, the actual disparity between the two samples is masked. Presently, more refined data will be available for the investigation of this point.

The lowest rate of return of completed questionnaires for any of the 50 states was 65.8 percent for Nevada, followed by 69.9 percent for Arizona and 70.2 percent for New York. The highest rate of return of completed questionnaires was 81.9 percent for Nebraska followed by 81.1 percent for Delaware and 80.7 percent for West Virginia. As might be expected, the differences among geographic divisions were less, with the highest percentage of completed questionnaires being 75.9 percent for the West North Central Division and the lowest, 71.4 percent, for the Pacific Division. In all, 15 states and the District of Columbia had rates of return of completed questionnaires below the mean of 73.1 percent. Preliminary results by geographic division and State are shown in table 3.

DATA PROCESSING

Precomputer processing.--Mail receipts in the survey were initially classified and sorted by type of receipt: Refusals, correspondence, postmaster returns, completed questionnaires, incomplete questionnaires, and deceased. The postmaster returns, as previously noted, were researched and remailed in care of the employer's address. Deceased, refusals, and good receipts (completed questionnaires) were sent to check-in punch wherein the serial number of the questionnaire was keyed for updating the master control file. Correspondence and incomplete questionnaires were set aside for review by an analyst. These incomplete questionnaires were screened out from other receipts because certain minimum entries were lacking. The standards required an entry which would indicate current educational attainment (either the highest year of schooling completed or the highest degree obtained by the person) and an entry for current labor force status and current or last occupation. When reviewed by an analyst, approximately half of the returns set aside as incomplete were salvaged based on information reported elsewhere on the questionnaire or by a telephone call made directly to the respondent. For refusals, incomplete returns, and deceased, the check-in punch was the terminal operation. Postmaster returns retained that status if no response was received after the mailout to the employer. Completed questionnaires, after check-in, were routed through two coding and verification processes. The first was general coding in which nonstandard entries were converted to proper form, degree level was coded, and conformity to certain standards on such items as wage and salary rates was enforced. Of the 74,000 questionnaires processed, approximately 15,000 were set aside during general coding for professional review, mostly for foreign degrees, wage entry failures, and job sequence problems. The second was geographical coding in which State or foreign country and, depending on the items, county or SMSA were coded for place of birth, location of higher education, location of last three jobs, place of residence before and after each of last three jobs, and current residence.

Computer edit.--After general coding and geographical coding, the information on the 74,000 questionnaires was key coded directly onto magnetic tape. The resulting data tape was passed through a computer editing program which checked for acceptability of codes and ranges of entries, made substitutions for certain missing items, performed certain recodes (for age, highest degree held, highest degree worked on, etc.), checked for consistency between items, and made a more refined edit of wage and salary rate entries. The specifications for the edit program called for the rejection of questionnaires which had passed the clerical screening without the minimum necessary entries or which contained unacceptable income or wage and salary entries. In addition, records which contained 10 or more minor errors, such as out-of-range entries or inconsistent entries, were also rejected, but records with fewer than 10 of the latter sort of errors were allowed to stay in the file with the erroneous items blanked. The rejects from the computer operation were reviewed, appropriate corrections were made, and the records were reentered on the file.

Matched data file.--The edited data file was then matched with (1) the survey master control file (the computer file used to control mailing and followup actions) so that each record would have an associated indication of its response code, and (2) the census sample detail file containing Census of Population data for each selected PMS case. The resulting matched edited data file therefore contains all census sample information for each person originally selected for the PMS sample, even if we may have been unable to find the census questionnaire (and therefore a name and address for mailing). It contains, in addition, the receipt code classification for all cases mailed out in the survey whether or not a completed questionnaire was returned. Finally, for cases in which completed questionnaires were returned in the survey, all reported data that passed the edit also are recorded in the file.

Weighting.--One further step is necessary before the final merged data can be used for tabulation: weighting up to census totals. Actually, each record in the data file will contain four separate weights. The first weight will not make any adjustment for noninterview and will be used for comparing the census characteristics of persons not responding in the postcensal survey to those of respondents. The second weight will make an adjustment for nonresponse and a ratio adjustment to known age, sex, and color census control totals by occupation group. This weight will be used for detailed tabulations by the original census occupation categories and for the tabulations by redefined occupation groups. The third weight will be used for purposes of variance computation. The fourth weight, similar in form to the second weight, will be used only when the Postcensal Manpower Survey sample and Sample II are combined for selected computations.

Tabulations.--The tabulations planned for the Postcensal Manpower Survey fall into three series reflecting relationship of content and priorities in design, programming, review, and publication. The first series of tables is planned to show changes in the employment status of engineers and scientists between 1970 and 1972. It will also provide the information necessary to evaluate the quality of reported data and to estimate the degree of nonresponse bias. Programming is substantially completed on the first part of the tables in this series, and we expect to be reviewing the initial data this fall. The second series of tables is planned to show the basic inventory of scientific, engineering, technical, and other highly qualified manpower for 1970 and 1972. These tables, similar to those in Technical Paper No. 21[2] which were based on the 1962 postcensal study, will include data on fields of study, job activity, citizenship, financial support for college work, industry, class of worker, salary rate, professional identification, and so forth, by age, sex, highest degree held, and occupation. This series will be run by the 1970 Census occupation. It will likely also be run for the redefined occupation groups. The design of the second series of tables is substantially completed, although programming is not yet begun. We hope to publish these tables during the winter of 1972-73. The third series of tables will present data focusing on more analytic relationships than the earlier series of tables: Exposure to science courses in high school and subsequent college work in a scientific field of study; field of study for first (or highest) degree, current occupation, and professional identification; father's occupation and field of study for first degree; salary rate for current job and average duration of last three jobs; and so forth. The design of the third series of tables is underway, although no programming has been initiated.

These tables are planned for publication some time during 1973.

1/ Copies of the questionnaire and reference lists used in the survey are available upon request to the author, Social and Economic Statistics Administration, Bureau of the Census, Washington, D.C. 20233.

2/ U.S. Bureau of the Census, Characteristics of America's Engineers and Scientists: 1960 and 1962, Technical Paper No. 21, U. S. Government Printing Office, Washington, D.C. 1969.

LIST A

1972 POSTCENSAL MANPOWER SURVEY

LIST OF SAMPLE GROUPS

No.	Title	Census Occupation Code	Sample Size
	Computer specialists		
1.	Computer programmers	003	3188
2.	Computer systems analysts	004	3172
3.	Computer specialists, n.e.c	005	684
	Engineers		
4.	Aeronautical and astronautical engineers	006	3173
5.	Chemical engineers	010	2703
6.	Civil engineers	011	3178
7.	Electrical and electronic engineers	012	3673
8.	Industrial engineers	013	3176
9.	Mechanical engineers	014	3177
10.	Metallurgical and materials engineers	015	796
11.	Mining and petroleum engineers	020,021	806
12.	Sales engineers	022	2817
13.	Engineers, n.e.c., and engineering teachers*	023,111	3179
	Mathematical specialists		
14.	Actuaries, mathematicians, statisticians, and mathematics teachers*	034,035,036,112	3179
	Life and physical scientists		
15.	Agricultural scientists and agriculture teachers*	042,102	893
16.	Foresters and conservationists (with 4 or more years of college)	025	823
17.	Atmospheric, space, and marine scientists, geologists, and atmospheric, earth, marine, and space teachers*	043,051,052,103	1725
18.	Biological scientists and biology teachers*	044,104	2589
19.	Chemists and chemistry teachers*	045,105	3183
20.	Physicists, astronomers, and physics teachers*	053,110	1866
21.	Life and physical scientists, n.e.c	054	65
22.	Operations and systems researchers and analysts	055	3184
	Social scientists		
23.	Economists and economics teachers*	091,116	3169
24.	Psychologists and psychology teachers*	093,114	2104
25.	Political scientists, sociologists, urban and regional planners, social scientists, n.e.c., sociology teachers,* and social science teachers,* n.e.c	092,094,095, 096,121,122	1742
	Engineering and science technicians		
26.	Agricultural, biological, and chemical technicians, except health	150,151	3175
27.	Draftsmen	152	3780
28.	Electrical and electronic engineering technicians	153	3177
29.	Industrial and mechanical engineering technicians and numerical control tool programmers	154,155,172	1931
30.	Surveyors	161	3162
31.	Mathematical technicians and engineering and science technicians, n.e.c	156,162	3183
32.	Personnel and labor relations workers	056	3801
33.	Health specialties teachers*	113	1425
34.	Trade, industrial, and technicial teachers*	134	162
35.	Miscellaneous teachers*	135	991
36.	Teachers,* subject not specified	140	3189
37.	Technicians, n.e.c	173	1955
38.	Research workers, not specified	195	3177
39.	School administrators, college (with 4 or more years of college)	235	1506
40.	Managers and administrators, n.e.c. (with 4 or more years of college	245	9489
41.	Persons in the experienced civilian labor force, with 4 or more years of college, and not in any of the above occupations		5758

n.e.c. - not elsewhere classified

* College and university

Table 1.--1972 POSTCENSAL MANPOWER SURVEY: RESPONSE RATES BY TYPE AND COLLECTION PERIOD TO WHICH ATTRIBUTED (PRELIMINARY)

Collection period	Total at risk in the collection period	Completed question-naires	Deceased	Refusal
Total number	101,835	74,483	675	7,689
First mailing (2/17/72-3/14/72)	101,835	26,603	270	47
Second mailing (3/15/72-3/31/72)	74,194	15,906	169	211
Third mailing (4/3/72-4/28/72)	54,309	9,946	120	405
Certified mailing (5/1/72-7/17/72)	42,444	13,234	103	1,564
Telephone calls (5/15/72-7/17/72)	17,042	8,794	13	5,462
Total percent	100.0	73.1	0.7	7.6
First mailing (2/17/72-3/14/72)	100.0	26.1	0.3	(1)
Second mailing (3/15/72-3/31/72)	100.0	21.4	0.2	0.3
Third mailing (4/3/72-4/28/72)	100.0	18.3	0.2	0.7
Certified mailing (5/1/72-7/17/72)	100.0	31.2	0.2	3.7
Telephone calls (5/15/72-7/17/72)	100.0	51.6	0.1	32.1
Total percent	100.0	100.0	100.0	100.0
First mailing (2/17/72-3/14/72)	100.0	35.7	40.0	0.6
Second mailing (3/15/72-3/31/72)	72.9	21.4	25.0	2.7
Third mailing (4/3/72-4/28/72)	53.3	13.4	17.8	5.3
Certified mailing (5/1/72-7/17/72)	41.7	17.8	15.3	20.3
Telephone calls (5/15/72-7/17/72)	16.7	11.8	1.9	71.0

1/ Less than 0.1 percent.

Table 2.--1972 POSTCENSAL MANPOWER SURVEY: RESPONSE RATES BY TYPE AND OCCUPATION GROUP FOR SAMPLES I AND II (PRELIMINARY)

Occupation group	Total mailed out	Completed question-naires	Other responses			
			Total 1/	Deceased	Postmaster returns	Refusal
SAMPLE I						
All occupations						
Number	101,835	74,483	14,741	675	4,920	7,689
Percent	100.0	73.1	14.5	0.7	4.8	7.6
Primary target occupations	100.0	73.5	14.2	0.6	4.7	7.5
Engineers and scientists	100.0	75.0	13.6	0.6	4.4	7.2
Computer specialists	100.0	71.9	14.4	0.2	6.1	6.9
Engineers	100.0	74.8	14.2	0.8	3.6	8.5
Mathematical specialists	100.0	74.2	14.0	0.6	5.0	6.9
Life and physical scientists	100.0	79.4	11.3	0.5	4.9	4.7
Operations and systems researchers and analysts	100.0	73.0	14.9	0.8	3.7	9.2
Social scientists	100.0	73.1	13.2	0.7	5.0	6.2
Engineering and science technicians	100.0	68.8	16.4	0.6	5.8	8.5
Secondary target occupations	100.0	72.3	15.1	0.8	5.1	7.6
Other occupations, with 4 or more years of college	100.0	71.7	14.8	0.6	4.6	7.7
SAMPLE II						
All occupations						
Number	54,281	44,718	5,400	363	2,308	2,086
Percent	100.0	82.4	10.0	0.7	4.3	3.8
Primary target occupations	100.0	82.4	10.0	0.7	4.3	3.8
Engineers and scientists	100.0	82.4	10.0	0.7	4.3	3.8
Computer specialists	100.0	79.9	11.2	0.3	5.8	4.3
Engineers	100.0	82.7	9.8	0.8	3.4	4.4
Mathematical specialists	100.0	82.2	10.7	0.2	5.0	4.2
Life and physical scientists	100.0	84.6	8.7	0.5	4.5	2.4
Operations and systems researchers and analysts	100.0	81.5	10.3	0.6	3.4	4.8
Social scientists	100.0	80.8	10.6	0.7	5.5	3.0
Engineering and science technicians	(X)	(X)	(X)	(X)	(X)	(X)
Secondary target occupations	(X)	(X)	(X)	(X)	(X)	(X)
Other occupations, with 4 or more years of college	(X)	(X)	(X)	(X)	(X)	(X)

Not applicable

1/ Includes questionnaires returned with insufficient data for processing and cases deleted as out of scope.

Table 3.--1972 POSTCENSAL MANPOWER SURVEY: RESPONSE RATES BY TYPE FOR DIVISIONS AND STATES (PRELIMINARY)

Division and state	Total	No response	Completed questionnaire	Other response: Total[1]	Other response: Deceased	Other response: Postmaster return	Other response: Refusal
UNITED STATES							
Number	101,835	12,611	74,483	14,741	675	4,920	7,689
Percent	100.0	12.4	73.1	14.5	0.7	4.8	7.6
NEW ENGLAND:	100.0	11.5	73.3	15.2	0.6	4.2	8.9
Maine	100.0	10.2	75.9	13.9	0.6	3.1	7.7
New Hampshire	100.0	10.2	76.4	13.4	0.8	4.7	6.8
Vermont	100.0	7.9	74.7	17.4	1.3	2.6	11.4
Massachusetts	100.0	12.3	72.6	15.1	0.5	4.7	8.6
Rhode Island	100.0	10.9	72.0	17.1	0.2	3.1	10.7
Connecticut	100.0	11.0	73.8	15.2	0.8	3.8	9.2
MIDDLE ATLANTIC:	100.0	12.5	72.8	14.7	0.8	4.4	8.1
New York	100.0	13.7	70.2	16.1	0.7	5.4	8.4
New Jersey	100.0	12.1	75.3	12.6	0.7	3.3	7.1
Pennsylvania	100.0	10.8	75.2	14.0	0.9	3.8	8.3
EAST NORTH CENTRAL:	100.0	12.0	73.5	14.5	0.7	4.2	8.2
Ohio	100.0	11.8	74.3	13.9	0.9	3.4	8.1
Indiana	100.0	12.0	74.2	13.8	0.7	4.4	7.4
Illinois	100.0	14.0	70.7	15.3	0.4	5.1	8.5
Michigan	100.0	11.4	74.1	14.5	0.7	4.0	8.5
Wisconsin	100.0	8.6	77.5	13.9	0.5	4.1	7.6
WEST NORTH CENTRAL:	100.0	11.1	75.9	13.0	0.7	4.2	6.7
Minnesota	100.0	11.2	74.7	14.1	0.8	4.8	7.2
Iowa	100.0	9.5	79.4	11.1	0.3	4.2	5.4
Missouri	100.0	13.2	72.1	14.7	0.8	4.5	7.5
North Dakota	100.0	10.3	79.5	10.2	1.1	3.2	4.9
South Dakota	100.0	12.8	77.7	9.5	2.1	2.7	4.8
Nebraska	100.0	8.7	81.9	9.4	0.0	2.6	6.1
Kansas	100.0	9.9	77.9	12.2	0.6	3.8	6.3
SOUTH ATLANTIC:	100.0	12.9	73.4	13.7	0.6	4.7	6.8
Delaware	100.0	7.4	81.1	11.5	0.4	3.5	7.0
Maryland	100.0	11.7	75.5	12.8	0.5	3.6	7.4
District of Columbia	100.0	20.7	60.6	18.7	0.7	10.4	5.9
Virginia	100.0	11.3	75.5	13.2	0.4	4.3	7.2
West Virginia	100.0	8.5	80.7	10.8	1.0	2.2	5.7
North Carolina	100.0	12.9	72.5	14.6	0.7	5.3	6.5
South Carolina	100.0	12.6	75.0	12.4	0.9	3.4	6.3
Georgia	100.0	14.3	71.0	14.7	0.6	6.3	6.5
Florida	100.0	15.4	70.3	14.3	0.6	5.2	6.8
EAST SOUTH CENTRAL:	100.0	11.6	75.1	13.3	0.9	4.8	6.4
Kentucky	100.0	11.3	76.9	11.8	0.6	3.7	6.4
Tennessee	100.0	11.5	75.5	13.0	0.5	4.2	6.7
Alabama	100.0	11.1	74.8	14.1	1.2	5.5	6.4
Mississippi	100.0	13.5	72.3	14.2	1.5	6.6	5.5
WEST SOUTH CENTRAL:	100.0	13.1	72.4	14.5	0.6	5.8	6.7
Arkansas	100.0	11.3	77.2	11.5	0.2	5.5	5.1
Louisiana	100.0	12.7	73.5	13.8	0.6	4.9	6.7
Oklahoma	100.0	13.3	71.9	14.8	0.5	5.2	7.3
Texas	100.0	13.3	71.9	14.8	0.6	6.2	6.6
MOUNTAIN:	100.0	11.5	73.6	14.9	0.6	6.2	6.7
Montana	100.0	12.3	77.1	10.6	0.3	3.4	6.2
Idaho	100.0	11.5	75.0	13.5	0.0	6.9	5.6
Wyoming	100.0	7.4	75.7	16.9	0.0	8.9	4.5
Colorado	100.0	9.5	76.6	13.9	0.6	5.9	5.9
New Mexico	100.0	12.9	71.8	15.3	0.6	5.4	7.4
Arizona	100.0	14.2	69.9	15.9	0.4	6.2	8.9
Utah	100.0	9.7	74.1	16.2	0.9	6.9	6.9
Nevada	100.0	17.1	65.8	17.1	1.7	9.4	4.7
PACIFIC:	100.0	13.2	71.4	15.4	0.6	5.7	7.7
Washington	100.0	11.5	73.6	14.9	0.7	5.6	7.3
Oregon	100.0	10.9	75.6	13.5	0.7	5.3	6.0
California	100.0	13.7	70.6	15.7	0.6	5.7	8.0
Alaska	100.0	9.4	74.5	16.1	0.0	12.8	2.7
Hawaii	100.0	14.7	74.5	10.8	0.5	5.4	4.1

1/ Includes questionnaires returned with insufficient data for processing and cases deleted as out of scope.

THE INTERNAL VALIDITY OF THE KUHN-McPARTLAND TST

R. M. Alm, W. F. Carroll, and G. A. Welty, American University

INTRODUCTION

The sociological theory called symbolic interactionism derives from the writings of George Herbert Mead. It is a theory of the development of the self from one's social milieu, and from perceptions of and interaction with others. (See, e.g. Blumer [1] for a succinct account of the Meadian theory.)

It has been charged that Mead, and by implication his followers, eschewed empirical research ([6], pp. 238-239). This is presumably to the detriment of symbolic interactionist theory. While this charge is unquestionably true of Mead, it is not an accurate assessment of present day symbolic interactionists.

In an attempt to operationalize one of the key concepts of symbolic interactionist theory, that elusive concept of the "self", and to further empirical research on the self-concept, Manford Kuhn developed the Twenty Statements Test [4]. Kuhn and his collaborator McPartland supposed that human behavior is organized, the organization supplied by the individual's self-attitudes or self-concept. Following the direction provided by Mead's insistance on language as the basis of socialization and the self, it was straightforward to propose that self-attitudes could be assessed by asking an individual verbally to characterize himself.

The test (hereafter TST) is a sheet of paper, at the top of which are the instructions:

> There are twenty numbered blanks on the page below. Please write twenty answer to the question "Who am I?" in the blanks. Write the answere in the order they occur to you; don't worry about logic or importance.

Up to twenty responses to this generic question provided <u>subjective</u> definitions of the self for Kuhn, which he understood as internalizations of a person's <u>objective</u> social status. Hence the organization of behavior could be analysed in terms of the self-conception, which Kuhn (and Mead) understood as a plan for behavior. Thus the subjective, and perhaps more importantly, the anticipatory aspect of the self, based in part upon <u>ex ante</u> data configurations permitted by symbolization, could be compared with the objective aspect of behavior, known only in terms of <u>ex post</u> data.

The test has the virtues of simplicity and of providing a relatively direct measure of one's self-concept. For such reasons among others, the TST has been utilized in a number of social scientific studies. (cf. [7] for citations)

TWO MAJOR METHODOLOGICAL PROBLEMS

However, the substantive interest in the TST has not been matched by a concern for several rather striking methodological problems that bear on the test as an instrument to assess the self-concept. There has been one major study of the TST which was properly methodological in focus. This was Tucker's examination of Kuhn's theory of the self and the TST as an instrumentality for operationalizing this theory. After characterizing Kuhn's conception of "Self" as a theoretic entity, Tucker notes that it followed for Kuhn that

> self-attitudes are derived within a particular "context of behavior" and are "meaningless" without the explication of that "context." ([12], p. 312)

Kuhn's inference would appear to be wholly compatible with Mead's own thoughts on the self and the self concept. Particularly, Mead notes the relevance and importance of other socially proximate symbol-using actors to the development of the self and one's conception of his self. (cf[5],pp.160-161).

Unfortunately for systematic empirical research, however, such an implication is at variance with one requirement that must obtain for any test. Tucker points out another implication of the use of the TST in operationalizing social psychological theory is that

> it is assumed that the responses from the question, "Who am I?", are applicable to a <u>variety of situations</u>. This assumes that the others who are "present" and "contemporary" are irrelevant to the person's behavior ([12], p. 312, italics his).

These two implications are clearly incompatible. Of particular concern is the relevance to the test subject of the presence of the researcher employing the TST. Either the responses will reflect the idiosyncracies of a "context of behavior" which for our purposes happens to include a "researcher"and is characterized by the label "social scientific research," and the first implication Tucker draws from the self-theory is affirmed and the second denied. Alternatively, the responses will <u>not</u> exhibit variation due to the subject's sensitivity towards a unique social situation. Since in most social contexts there is no researcher present, if the presence of the researcher administering the TST is without effect, the second implication drawn by Tucker will be affirmed, and the first denied.

Another major problem is the incompatibility with theoretical premises of the self theory of the "content analysis procedures" or coding techniques used in categorizing the TST responses. Tucker notes that three basic assumptions of self theory are the slight predictive utility of fixed responses since the self is a function of the social context; that the Twenty Statement Test gives a valid representation of the subjects' self concept; and that the focus of the research is explicitly perspectival. ([12], p. 313).

A methodological problem arises if, when coding the TST responses, a coder other than the subject "imposes the meaning of each of [the responses] from his own perspective" ([12], p.313). Obviously, this is incompatible with the basic premises of self-theory.

If direct acquaintance with the self under study is required for the valid categorization of responses then the basic assumptions that Tucker notes of self theory are affirmed; thus the assumption that a coder other than the subject can impose his own meanings on the responses is denied. Otherwise, the responses can be considered to be relatively fixed, hence invariant to the identity of the coder. In the latter case, the second conditions noted by Tucker will be affirmed, and the premises of self theory denied.

THE PROBLEM OF SITUATIONALITY

When an instrument such as the TST is employed in measuring the self-concept, it is possible that the presence of the researcher may result in different responses from the subjects than the responses that would be occasioned were the researcher absent. The outcome of differential effects of the presence of the researcher requires differential response to the research situation itself, rather than the substantive variations of independent variables, induced by the researcher, within the research situation. One way that differential response to the research situation can be ascertained is by measuring the willingness on the part of the subjects to volunteer for a social scientific study.

It has been documented that there is a volunteer effect, where the subject's sensitivity increases to the "cues" that the experimenter unconsciously provides as to the hypothesis being tested, as motivation to participate as evidenced by volunteering increases (cf.[9]for references and discussion). Professor Orne calls these cues the "demand characteristics" of the experiment, i.e., that which the experimenter (not necessarily consciously) demands of the subject in terms of the substantive variables of the experiment. Orne has been quite explicit: "Within the context of our culture the roles of subject and experimenter are well understood and carry with them well-defined mutual role expectations" ([8]p.777).

Not only is the volunteer subject more sensitive to these cues, but in terms of his perceptions of the cues, he strives (again, not necessarily consciously) to "please" the experimenter by helping realize the prediction, substantiate the hypothesis, etc., by his own behavior. The generic hypothesis we sought to test was that the volunteer effect results in variations in the emerging self-concept measured by the TST.

The subjects were 56 undergraduates enrolled in an introductory sociology course. The volunteer - non-volunteer status of these subjects were ascertained by a procedure similar to that reported by Rosnow and Suls [10]. Immediately subsequent to the administration of the TST, a colleague of the investigator entered the classroom and solicited volunteers for a fictitious experiment in psycholinguistic research to take place several months later. He told the subjects that he would pay each of them $1.00 for a half-hour's participation. Participation in the (fictitious) experiment could be arranged for times mutually convenient to the pretended researcher and the subject. Following Rosnow's rationale, the financial incentive was offered to lend credibility to the experiment, yet was not large enough to provide a motive for participation. Sixteen of the subjects volunteered to participate. At the same time as he took the names and addresses of these volunteer subjects, the pretended researcher collected the anonymously completed TST sheets. Forty of the subjects did not volunteer. Since the subjects were unaware of the hypotheses being tested, they were essentially participating in a single-blind experiment.

There are two major schemes for coding the responses to the TST. The first is the compilation of the subjects locus score, which has been taken as a measure of the extent to which the subject is anchored in the social system ([11], p.50). The second is the categorization of the responses into a set of five analytical categories developed by Kuhn. We have utilized the second scheme in this investigation.

Full definitions of the five categories are given in Schwirian's essay ([11], p.51). We merely list the categories with brief explication. They include (1) Consensual responses, i.e. statuses in social categories and social groups, (2) Ideological beliefs, i.e. religious and philosophical orientations, etc. (3) Aspirations, i.e. future-tensed statements of personal goals and achievements, (4) Preferences, i.e., the respondent's interests and aversions, and (5) Self-evaluation, i.e. evaluative statements assessing one's own mental and physical abilities and demeanor. The 56 TST sheets were coded in terms of these five categories.

Six specific hypotheses were generated. A null hypothesis stated that there would be no significant difference between volunteer and non-volunteer subjects in the mean number of responses for each of the five categories listed above. In addition to these five null hypotheses, one null hypothesis stated that there would be no significant difference between volunteer and non-volunteer subjects in the total number of responses given. It was further supposed that this sixth hypothesis was directional. Volunteer subjects would strive more diligently than non-volunteer subjects to comply with the instructions to "write twenty answers to the question 'Who Am I?'."

The data is summarized in the following table where the mean number of responses per analytical category (and associated hypothesis) as well as the standard deviation of responses per category is given for the two groups. Student's t was utilized to test the differences in means. Data on the significances of the observed differences is given in the right hand column.

Response Category	Non-Volunteers (n=40)		Volunteers (n=16)		Significance of Students's t
	$\bar{X}$	σ	$\bar{X}$	σ	
1. Consensual	6.70	5.12	6.19	4.00	n.s.
2. Beliefs	0.93	1.25	2.25	2.25	p .01
3. Aspirations	0.25	0.62	0.81	1.07	p .05
4. Preferences	1.68	1.68	2.94	2.84	p .05
5. Evaluations	6.15	5.05	5.88	4.62	n.s.
6. [Omissions]	4.40	4.97	1.94	3.70	p .05

The following null hypotheses were sustained: (1) There is no significant difference between the volunteer and non-volunteer subjects, in the mean number of consensual responses given, and (5) there is no significant difference between the volunteer and non-volunteer subjects, in the mean number of evaluative responses given.

The other four null hypotheses were not sustained, and we retained the following research hypotheses: (2) There are significantly more ideological belief responses given by volunteer than by non-volunteer subjects, (3) there are significantly more aspiration responses given by volunteer than by non-volunteer subjects, (4) there are significantly more preference responses given by volunteer than by non-volunteer subjects, and (6) there are significantly fewer omissions given by volunteer than by non-volunteer subjects.

THE PROBLEM OF CODING

Another question arises from Tucker's discussion that bears on the methodological problems of the TST, in considerations of theoretical constructs on which the instrument is based, versus the content analytical (or coding) procedures. ([12], p. 303).

Tucker pointed out and Franklin and Kohout have noted at the time of analysis, the researcher imposes his own meanings on the subjects' statements by coding them according to a set of "a priori" categories ([2], p.82).

The question then arises, "Must the subject score his own test for test validity?" This question would appear to be methodologically germaine to any explicitly perspectival test. Franklin and Kohout compared the locus scores (consensual statements) calculated on the basis of the subject's own coding and those calculated from the researchers' codings. These results were obtained by having the subjects and the researchers code the statements, with the "other" coders who were the researchers scoring the tests independently and later resolving inter-coder discrepancies. Franklin and Kohout's data indicated that the empirical consequences of having the TST responses coded by using the "self" versus the "other" codings "are probably not significantly different." ([2], p. 88)

The present research likewise focuses on the possible discrepancies between "self" and "other" coding. We replicated the Franklin and Kohout test circumstances. However, we have categorized the statements into the set of five analytical categories developed by Kuhn, ([4]) instead of the two broad categories of consensual and subconsensual statements.

The subjects were thirty undergraduates enrolled in an introductory sociology course. After completing the TST in a 12-minute period, the subjects were given coding instructions based on Kuhn and McPartland's definitions of each of the five categories ([4], pp. 40-41). The "other" coders were researchers familiar with both the TST and its analytical procedures and with self-theory. The researchers also coded the test independently, later resolving inter-coder discrepencies as specified by the Franklin and Kohout coding procedure ([2], p. 85). In addition, an average of the other coder analyses was made. This coding procedure was introduced to ascertain if a linear composite of coding outcomes was equivalent to the self coding, thus providing an alternative to the Franklin and Kohout coding procedure. First we will discuss the comparison involving the Franklin and Kohout procedure.

For each of the five coding categories, there was one null hypothesis developed, each one stating there would be no significant difference between the mean number of responses for that category when coded by the subject himself, and when derived from the Franklin and Kohout resolution of coding behavior of several researchers.

Response Category	Self-Coding		Other	
	$\bar{X}$	σ	$\bar{X}$	σ
1. Consensual	5.227	2.861	5.136	3.299
2. Beliefs	1.364	1.965	1.409	1.501
3. Aspirations	.727	1.162	.364	.581
4. Preferences	1.409	1.563	.909	1.444
5. Evaluations	10.273	4.600	10.909	4.524

All the null hypotheses were sustained; the test of significance was Student's t at p=.05.

We will now discuss the comparison of self coding with the linear composite of the coding behavior of the several researchers.

For each of the five coding categories, again there was one null hypothesis developed, each one stating there would be no significant difference between the mean number of responses for that category when coded by the subject himself, and when derived by averaging the coding behavior of the several researchers.

Response Category	Self-Coding		Average	
	$\bar{X}$	σ	$\bar{X}$	σ
1. Consensual	5.227	2.861	5.091	3.366
2. Beliefs	1.364	1.965	1.485	1.352
3. Aspirations	1.727	1.162	.440	.548
4. Preferences	1.409	1.563	1.152	1.093
5. Evaluations	10.273	4.600	10.380	4.372

All the null hypothesis were sustained; again, the test of significance was Student's t at p=.05.

Thus our results to this point corroborate the earlier findings of Franklin and Kohout that the score is invariant under other coding procedures operationally equivalent to subject coding methods, either when inter-coder discrepancies were resolved as specified in the Franklin and Kohout procedure, or when the "self" coders were compared with the average of the "other" coders ([2], p. 82).

However, both Franklin and Kohout procedure, and the procedure of linear composites might significantly and artificially suppress inter-coder variance. Thus, we developed five null hypotheses, each one stating that there would be no significant difference between the mean number of responses as coded by each of the researchers for that category. This data is summarized in the following table.

Response Category	Coder #1		Coder #2		Coder #3		Significance of Fisher's F
	$\bar{x}$	σ	$\bar{x}$	σ	$\bar{x}$	σ	
1. Consensual	4.682	3.414	5.773	3.664	4.818	3.290	n.s.
2. Beliefs	1.727	2,074	.455	.671	2.273	2.142	p .05
3. Aspirations	.409	.590	.682	.839	.227	.528	n.s.
4. Preferences	.636	1.217	2.318	1.810	.500	,964	p. .05
5. Evaluations	11.273	4.682	9.500	4.296	10.364	4.924	n.s.

The test of significance was Fisher's F at p=.05.

We found significant differences between the "other" coders in two of the five categories (the ideological and interest categories) and an almost significant difference in the category of aspirations. This suggests that the resolution of coding disagreements by either the Franklin and Kohout coding procedure or the linear composite procedure arbitrarily suppresses significan variance. This finding was corroborated by a pairwise posterior analysis of the five categories for the self coding and the coding behavior of each of the researchers. Again the test of significance was Student's t at p=.05, and for two of the researchers significant differences were found between their coding behavior and self coding in one category each.

Thus we conclude that in the absence of further research into coder differences one cannot take any particular "other" coder to code the TST since there does appear to be an individual coder effect.

In conclusion, there are several propositions that appear warranted. First, there are significant volunteer effects on the responses to the TST. Thus, the responses to the TST do not appear to be unconditionally applicable to a variety of situations. Second, the Franklin and Kohout coding procedures appear to be operationally equivalent to self coding of the TST. Finally, significant inter-coder variation appears to exist, both between research coders themselves, and between them and the subjects as coders.

REFERENCES

(1) Blumer, Herbert. "Society as Symbolic Interaction" in Rose, A, Human Behavior and Social Processes, Boston; Houghton Mifflin (1962), pp. 181-184.

(2) Franklin, Billy J. and Kohout, Frank J., "Subject-Coded Versus Researcher-Coded TST Protocols:Some Methodological Implications", Sociological Quarterly,XII (Winter,1971), 82-89.

(3) Hickman, C.A. and Kuhn, Manford. Individuals, Groups, and Economic Behavior, New York Dryden Press (1956).

(4) Kuhn, Manford and McPartland, T.S. "An Empirical Investigation of Self-Attitudes", American Sociological Review, Vol. 19 (1954) pp. 68-76.

(5) Mead, George Herbert "A Behavioristic Account of the Significant Symbol", Journal of Philosophy,(1922) Vol. 19, pp. 157-163.

(6) Merton, Robert K. Social Theory and Social Structure, Glencoe: Free Press (1957).

(7) Nass, Gilbert. "Literature on Measurement of the Self-Concept",Journal of Counseling Psychology,Vol. 8 (1961), p. 368.

(8) Orne, Martin. "On the Social Psychology of the Psychological Experiment",American Psychologist,Vol. 17, (1962), pp. 776-783.

(9) Rosenthal, Robert and Rosnow, Ralph. "The Volunteer Subject",in Artifact in Behavioral Research ed. R. Rosenthal and R. Rosnow, New York: Academic Press (1969), pp. 59-118

(10) Rosnow, Ralph and Suls, J. "Reactive Effects of Pretesting in Attitude Research", Journal of Personality and Social Psychology,Vol. 15, (1970), pp. 338-343.

(11) Schwirian, Kent. "Variations in Structure of the Kuhn-McPartland Twenty Statements and Related Response Differences",Sociological Quarterly,Vol. 5 (1964), pp. 47-59.

(12) Tucker, Charles. "Some Methodological Problems in Kuhn's Self Theory" Reprinted in Manis, Jerome G. and Meltzer, Bernard N. Symbolic Interaction: A Reader in Social Psychology. 2nd ed. Boston: Allyn and Bacon, 1972, pp. 303-15.

LARGE SAMPLE AND JACKKNIFE PROCEDURES FOR SMALL SAMPLE ORTHOGONAL LEAST SQUARES INFERENCE

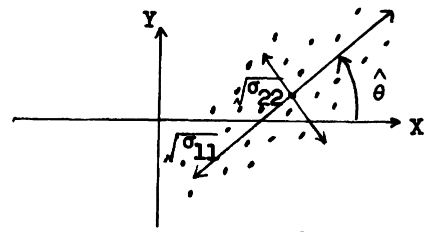
Peter O. Anderson, Ohio State University

Exact distribution theory for orthogonal least-squares estimators is quite difficult (this is a special case of principal components estimation). In general the estimators have asymptotic normal distributions. Using this fact and using a consistent estimator for the variance, approximate confidence intervals can be obtained. An alternative is to use the jackknife technique to obtain confidence intervals. A Monte Carlo study shows that even for small samples the two procedures tend to give similar results.

$$\hat{\theta} = \frac{1}{2} \tan^{-1} \frac{2A_{12}}{A_{11} - A_{22}}$$

$$A_{11} = \sum_{1}^{N} (X_i - \bar{X})^2, \quad A_{12} = \sum_{1}^{N} (X_i - \bar{X})(Y_i - \bar{Y}),$$

$$A_{22} = \sum_{1}^{N} (Y_i - \bar{Y})^2$$



slope = tan $\hat{\theta}$

$\hat{\theta}$ asymptotically $N(\theta, \sigma^2_{\theta,N})$. X,Y bivariate normal implies

$$\sigma^2_{\theta,N} = \frac{1}{N} \cdot \frac{\sigma_{11}\sigma_{22}}{(\sigma_{11} - \sigma_{22})^2} = \frac{1}{N} \cdot \frac{\sigma_{11}/\sigma_{22}}{(\sigma_{11}/\sigma_{22} - 1)^2}$$

Large Sample Confidence Interval: $\hat{\theta} \pm z_{\alpha} \cdot \hat{\sigma}_{\theta,N}$

Jackknife Confidence Interval: $\tilde{\theta} \pm t_{\alpha,K-1} \cdot s_{\theta}/\sqrt{K}$

where $\tilde{\theta} = \frac{1}{K} \sum_{1}^{K} \tilde{\theta}_i, \quad s^2_{\theta} = \frac{1}{K-1} \sum_{1}^{K} (\tilde{\theta}_i - \tilde{\theta})^2$

Numbers of 90% Confidence Intervals for θ which do not contain the true value.

Normal - Normal		$\sigma_{11}/\sigma_{22} = 4$		$\sigma_{11}/\sigma_{22} = 100$	
N	K	Lg.Sample	Jack-knife	Lg.Sample	Jack-knife
30	30	13	11	11	9
30	10	13	11	11	14
10	10	27**	21**	24**	11

Uniform Uniform		$\sigma_{11}/\sigma_{22} = 4$		$\sigma_{11}/\sigma_{22} = 100$	
N	K	Lg.Sample	Jack-knife	Lg.Sample	Jack-knife
30	30	10	10	10	9
30	10	10	13	10	11
10	10	19**	11	16*	8

Uniform Exponential		$\sigma_{11}/\sigma_{22} = 4$		$\sigma_{11}/\sigma_{22} = 100$	
N	K	Lg.Sample	Jack-knife	Lg.Sample	Jack-knife
30	30	13	10	11	9
30	10	13	11	11	9
10	10	19**	10	16*	7

*Significantly different from 10 at 5% level
**Significantly different from 10 at 1% level

Numbers of 95% Confidence Intervals for θ which do not contain the true value.

Normal - Normal		$\sigma_{11}/\sigma_{22} = 4$		$\sigma_{11}/\sigma_{22} = 100$	
N	K	Lg.Sample	Jack-knife	Lg.Sample	Jack-knife
30	30	7	3	5	4
30	10	7	7	5	6
10	10	20**	13**	11**	6

Uniform Uniform		$\sigma_{11}/\sigma_{22} = 4$		$\sigma_{11}/\sigma_{22} = 100$	
N	K	Lg.Sample	Jack-knife	Lg.Sample	Jack-knife
30	30	9	9	9	9
30	10	9	9	9	8
10	10	12**	7	9	4

Uniform Exponential		$\sigma_{11}/\sigma_{22} = 4$		$\sigma_{11}/\sigma_{22} = 100$	
N	K	Lg.Sample	Jack-knife	Lg.Sample	Jack-knife
30	30	9	5	6	5
30	10	9	6	6	6
10	10	13**	4	11**	4

*Significantly different from 5 at 5% level
**Significantly different from 5 at 1% level

VARIANCE COMPONENT ANALYSIS OF CENSUS DATA AND THE USE OF THE HARMONIC MEAN

Ibtihaj S. Arafat, City College of City University of New York

Like every other statistical quantity, averages are used for comparative purposes. Specifically, averages' are used to compare the locations of distinct groups or distributions on the same scale. However, it should be obvious that the different types of averages are not comparable. The ancient Greeks were interested in the phenomenon of the progress of numbers from low to high in regular but different ways. In pursuing this interest, they noticed in particular that the interval between adjacent values may be constant as in the progression 1, 2, 3, 4, 5; that the ratio of adjacent values may be uniform as in the progression 2, 4, 8, 16, 32; that for any three consecutive numbers, A, B, and C, the percentage drop from A to B may be equal to the percentage jump from C to B, as in the progression 1, 1/2, 1/3, 1/4, 1/5. Today, we designate these progressions as arithmetic, geometric, and harmonic respectively (Mueller, et. al., 1970: 146-150). The middle term of each of the series could be calculated by specific formulas. We sum values and divide by N to get the arithmetic mean. Consider the aforementioned geometric series: 2, 4, 8, 16, 32. The middle term is 8, but the arithmetic mean of this progression is 12.4, a value which misrepresents the "central tendency" of the series. To get the middle term, we find the product of all terms and take the Nth root of that product: $\sqrt[5]{(2x4x8x16x32)} = 8$. We refer to a result obtained in this manner as the geometric mean. Thus instead of summing and dividing by N, as in the case of the arithmetic mean, we multiply and find the Nth root.

Consider the harmonic series as given above: 1, 1/2, 1/3, 1/4, 1/5. Neither the arithmetic nor the geometric mean correspond to the middle term of this series. However, the middle term may be obtained by taking the reciprocal of the arithmetic mean of the reciprocals:

Mean of the Reciprocals = $\frac{1+2+3+4+5}{5} = 3$

Reciprocal of Mean of Reciprocals = 1/3.

Although the geometric mean and the harmonic mean occur with some regularity in advanced statistical analysis, they are not likely to be encountered in simple descriptive statistics (Mueller, et.al., 1970: 149). Occasionally, the geometric mean is used to find the size of a population at the midpoint of an interval of time, on the assumption that its growth during that period has been at a constant rate. To illustrate, if the population of a particular community is 5,000 at the time of the first observation and 10,000 at the time of the second observation, the population at the midpoint of the interval would have been approximately 7,000 provided growth was at the same rate throughout the period.

Then what is the harmonic mean, and how could we manipulate it in analyzing census data? According to Marks (1971: 78), "The harmonic mean 'H' is the middle term in a harmonic progression. The harmonic mean of n numbers is 1/nth of the sum of their reciprocals." Spiegel (1961: 49) explained it as follows: "The harmonic mean H of a set of N numbers $X_1, X_2, X_3, \dots X_N$ is the reciprocal of the arithmetic mean of the reciprocals of the numbers:

$$H = \frac{1}{\frac{1}{N}\sum_{j=1}^{N}\frac{1}{X_j}} = \frac{N}{\sum\frac{1}{X}}$$

In practice it may be easier to remember that

$$\frac{1}{H} = \frac{\sum\frac{1}{X}}{N} = \frac{1}{N}\sum\frac{1}{X}\text{ ."}$$

As a computed average, the harmonic mean gives more weight to the smaller values (this is just the opposite of the arithmetic mean), it is capable of algebraic manipulation, and it yields consistent results in most cases (Chou, 1969: 67,82). Chou also mentions that for values which are not all the same and do not have a value of zero, the harmonic mean is always smaller than both the arithmetic mean and the geometric mean. Spiegel (1961: 49) illustrates the relative value of the three means as follows: $H \leq G \leq \bar{X}$ (the equality signs hold only if all the numbers $X_1, X_2, \dots, X_N$ are identical).

In population data, it is not always possible to have an equal number of observations on all treatment combinations. When disproportionality occurs, the analysis of the data becomes difficult. The objective of the researcher in this study is to show how the harmonic mean "H" could be utilized in analyzing census data, which consist of unequal numbers of observations on all subclass combinations. Since the level of measurement of the samples in general includes nominal, ordinal, and ratio measures, the devices used by most of the users of the census data were general statistical descriptive techniques with some ratio and graphic devices. According to Steel and Torrie (1960: 15), the main use of the harmonic mean is in averaging ratios and rates. Steel and Torrie (p.274) also states that in a test of significance for interaction, where unequal subclass numbers are used, the method of weighted squares of means may be used to estimate mean squares for main effects. An alternative procedure is to use the harmonic mean of all subclass n_{ij}'s. They add that calculations are less involved in this case. While the general methods are still applicable, special computing procedures should be used, which depend upon the presence or absence of interaction, although the initial steps are the same. In this case, an analysis of variance could be used with the estimation of variance components rather than F test.

To illustrate the procedure, let us study the different components of the total variation of age distribution by sex, color, and ethnicity. The data are taken from the one-in-a-thousand sample of the 1960 census of the United States of America. (See Table I)

An analysis of variance of the cell means of sex, color, and ethnicity combinations could be performed as illustrated in Table II. The means were considered in the analysis because the number of observations was unequal in the different cells, as shown in Table I. The error degrees of freedom was obtained as the pooled degrees of freedom within cells. The error sum of squares was obtained in the same manner. It was weighted for the inequality of the number of observations in cells by multiplying by the inverse of the harmonic mean of the number of observations within cells

$$\frac{\sum_{i=1}^{K} \frac{1}{N_i}}{K}$$

where K is the number of cells. The effects due to color, ethnicity, and sex are fixed, and hence the effects due to their different interactions are fixed. For this reason θ^2 in the column of the expected mean square (EMS) denotes the sum of the squares of the true effects of the factor or interaction indicated by the subscript of θ^2, divided by the corresponding degrees of freedom. On the other hand, the error was assumed to be random, and its EMS is σ^2_e, designating the error variance. The mean squares were equated with their corresponding EMS's, and the resulting equations were solved to obtain estimates of the error variance and the variation due to different factors and interactions. The variation due to different factors and interactions was obtained in percent of the total variation. The completed analysis is given in Table II.

In interpreting the results, it could be seen clearly that age distribution in this example was mainly influenced by ethnicity, as illustrated about 50 percent of the total variation was due to ethnicity. The next highest variation, 37.98 percent, was due to color. The estimated θ^2_s was a small negative quantity and hence could be considered zero, which implies that there is no significant variation between males and females with respect to the average age. The same could be noted about the estimate of color-ethnicity-sex interaction which was a small negative quantity. And so is the case with the estimates of variation due to color-sex and ethnicity-sex interaction. This indicated that the behavior of color under the two sexes was the same, and the behavior of ethnicity was also the same under the two sexes. However, color-ethnicity interaction showed higher variation, 10.74 of the total percent variation. This implied that the behavior of color was somewhat different under different ethnic groups. These observations could be seen clearly by looking on Table II. From the analysis obtained by the use of the harmonic mean it was easy to forecast the magnitude of variation in age distribution due to sex, color, and ethnicity.

The method of the harmonic mean which was used and which was illustrated above could be applied on any data which uses actual numbers such as age or income with unequal number observations on all subclass combinations.

In conclusion, the harmonic mean deserves more attention than it receives in most of the behavioral statistics textbooks. However, the proper use of the harmonic mean depends upon two main considerations. First of all, it must be remembered that an average refers to some class of units that must be appropriate to the use that the average is to serve. And next, it is specially adapted to a situation where the observations are expressed inversely to what is required in the average (Chou, 1969: 68-69).

References

Chou, Ya-Lun.
1969 Statistical Analysis. New York Holt,Rinehart, and Winston, pp. 67-69.

Marks, Robert.
1971 New Mathematics Dictionary and Handbook. New York: Bantam Books, pp. 78-79.

Mueller, John, et.al.
1970 Statistical Reasoning in Sociology. Boston: Houghton Mifflin Co., pp. 146-150.

Spiegel, Murray.
1961 Schaum's Outline Series: Theory and Problems of Statistics. New York: Schaum Publishing Co., p.49.

Steel, Robert and James Torrie.
1960 Principles and Procedures of Statistics. New York: McGraw-Hill, pp. 15, 252-257, 274-276.

TABLE I

AGE DISTRIBUTION BY SEX, COLOR, AND ETHNICITY

Age Interval in Years	White											
	3rd Generation Sex				2nd Generation Sex				1st Generation Sex			
	M No	%	F No	%	M No	%	F No	%	M No	%	F No	%
Less than 1	1803	1.43	1628	1.29	96	0.40	80	0.34	3	0.03	2	0.02
1- 4	6559	5.18	6452	5.10	357	1.51	322	1.36	35	0.39	42	0.47
5- 9	7841	6.19	7433	5.87	416	1.75	399	1.68	60	0.67	79	0.88
10-14	5913	5.46	6928	5.47	444	1.87	371	1.57	114	1.27	86	0.96
15-19	5382	4.25	5311	4.19	441	1.86	414	1.75	94	1.05	81	0.90
20-24	4156	3.28	4209	3.32	503	2.12	492	2.08	104	1.16	147	1.64
25-29	3758	2.97	3836	3.03	693	2.92	681	2.87	163	1.81	172	1.92
30-34	4022	3.18	4113	3.25	936	3.95	1063	4.48	176	1.96	237	2.64
35-39	4134	3.27	4269	3.37	1239	5.23	1293	5.45	244	2.72	306	3.41
40-44	3606	2.85	3760	2.97	1334	5.63	1349	5.69	210	2.34	207	2.30
45-49	3303	2.61	3389	2.68	1255	5.29	1273	5.37	284	3.16	327	3.64
50-54	2951	2.33	2990	2.36	1078	4.55	1074	4.53	369	4.11	427	4.75
55-59	2357	1.86	2588	2.04	848	3.58	871	3.67	482	5.37	495	5.51
60-64	1881	1.49	2112	1.67	614	2.59	780	3.29	500	5.57	545	6.07
65-69	1569	1.24	1830	1.45	545	2.30	672	2.83	564	6.28	516	5.75
70-74	1146	0.91	1415	1.12	366	1.54	503	2.12	469	5.22	450	5.01
75 or more	1221	0.96	1739	1.37	360	1.52	543	2.29	478	5.32	513	5.71
Total	62602	49.45	64002	50.55	11525	48.62	12180	51.38	4349	48.42	4632	51.58
Total	126604				23705				8981			
Color %	79.48				14.88				5.64			
% in Grand Total	70.61				13.22				5.01			

Age Interval in Years	Non-White											
	3rd Generation Sex				2nd Generation Sex				1st Generation Sex			
	M No	%	F No	%	M No	%	F No	%	M No	%	F No	%
Less than 1	315	1.65	289	1.51	8	1.58	4	0.79	0	0.0	0	0.0
1- 4	1059	5.54	1062	5.55	21	4.16	19	3.76	4	1.00	2	0.50
5- 9	1216	6.36	1183	6.18	17	3.37	14	2.77	6	1.50	4	1.00
10-14	1050	5.49	1075	5.62	25	4.95	26	5.15	5	1.25	7	1.75
15-19	767	4.01	826	4.32	22	4.36	22	4.36	8	2.01	6	1.50
20-24	576	3.01	669	3.50	16	3.17	14	2.77	19	4.76	14	3.51
25-29	547	2.86	596	3.12	24	4.75	20	3.96	20	5.01	24	6.02
30-34	567	2.96	674	3.52	21	4.16	36	7.13	21	5.26	18	4.51
35-39	575	3.01	600	3.14	37	7.33	30	5.94	16	4.01	15	3.76
40-44	525	2.74	592	3.10	22	4.36	27	5.35	13	3.26	9	2.26
45-49	504	2.64	525	2.74	17	3.37	18	3.56	14	3.51	17	4.26
50-54	417	2.18	458	2.39	7	1.39	7	1.39	22	5.51	5	1.25
55-59	342	1.79	401	2.10	7	1.39	8	1.58	31	7.77	15	3.76
60-64	255	1.33	287	1.50	2	0.40	2	0.40	15	3.76	16	4.01
65-69	238	1.24	253	1.32	3	0.59	3	0.59	9	2.26	11	2.76
70-74	139	0.73	184	0.96	3	0.59	1	0.20	17	4.26	3	0.75
75 or more	168	0.88	193	1.01	1	0.20	1	0.20	10	2.51	3	0.75
Total	9260	48.41	9867	51.59	253	50.10	252	49.90	230	57.64	169	42.36
Total	19127				505				399			
Color %	95.49				2.52				1.99			
% in Grand Total	10.67				0.28				0.22			

Grand Total = 179321 = 100.0% Sample Total
White = 159290 = 88.83% of Sample Total
Non-White = 20031 = 11.17% of Sample Total
*Percentages for each generation equal 100.0

TABLE II

ANALYSIS OF VARIANCE OF AGE DISTRIBUTION BY SEX, COLOR, AND ETHNICITY

Source	Degrees of Freedom	Sum of Squares	Mean Square	Expected Mean Square
Total	19	2162.2746		
Color (C)	1	642.8056	642.8056	$\sigma_e^2 + 10\ \theta_C^2$
* Ethnicity (E)	4	1345.8717	336.4679	$\sigma_e^2 + 4\ \theta_E^2$
Sex (S)	1	0.5139	0.5139	$\sigma_e^2 + 10\ \theta_S^2$
Col. x Ethn. (CE)	4	155.4215	38.8554	$\sigma_e^2 + 2\ \theta_{CE}^2$
Col. x Sex (CS)	1	4.2421	4.2421	$\sigma_e^2 + 5\ \theta_{CS}^2$
Ethn. x Sex (ES)	4	10.7637	2.6909	$\sigma_e^2 + 2\ \theta_{ES}^2$
Col. x Ethn. x Sex (CES)	4	2.6561	0.6640	$\sigma_e^2 + \theta_{CES}^2$
+ Error	179,301	473,489.9080	2.6408	σ_e^2

TOTAL VARIATION OF PERCENTAGES

Parameters	Estimates of Parameters	Estimate in percent of total
θ_C^2	64.0165	37.98
θ_E^2	83.4568	49.51
θ_S^2	0.0	0.0
θ_{CE}^2	18.1073	10.74
θ_{CS}^2	0.3203	0.19
θ_{ES}^2	0.0251	0.01
θ_{CES}^2	0.0	0.0
σ_e^2	2.6408	1.57
Total	168.5668	100.00

+ The error was calculated as within cells, the error sum of squares was calculated as within cell sum of squares, divided by the harmonic mean of the number of observations within cells.

σ_e^2 = the variance of the error

θ^2 = sum of the squares of the true effect of the factor or the interaction shown by the subscript, divided by the corresponding degrees of freedom.

* In the statistical analysis, Ethnicity was divided into five categories: (1) Native-born with native-born parents, (2) Native-born with foreign-born father and native-born mother, (3) Native-born with native-born father and foreign born mother, (4) Native-born, with both parents, foreign born, and (5) Foreign-born. (Categories 2, 3, and 4 combined make the 2nd generation American group).

SEASONAL PATTERNS OF BIRTHS IN PUERTO RICO

Ishver S. Bangdiwala, Abelardo Fuertes-de la Haba
University of Puerto Rico

It has been said that man is the only creature who eats when he is not hungry, drinks when he is not thristy and makes love all the time. If the last of these three characteristics referred to about man were true, one would expect that during, say, a period of one year, the results of this behavior of love making in terms of births would be an uniformly distributed probabilistic phenomenon. However, when the data of births are studied over a certain period of time, there are deviations noted. These deviations in case of births, during the calendar year of 12 months, are markedly different from month to month. When this monthly phenomenon is observed from year to year, there is seen to be a certain fixed pattern in monthly distribution. Does this pattern show any seasonal variation? If it does, man, just like other creatures on this earth. may be following the seasonal conception pattern. Knowing, however, the superiority of man's intelectual capacity, one would attribute this seasonal pattern of conception to the traditional activities for customs which are in turn determined according to the season of the year.

To have an illustration of this pattern, the percent distribution of yearly births by month of the United States from year 1950 to 1961 is shown in Table 1, and Graph 1. The months of August (9.04%) and September (8.92%) are the ones having, on an average, the highest percent of the births during this period under study. The percent goes down later during the year and fluctuates at a lower level during the months from January to April and starts rising again. If one studies such a pattern and tries to fit a cyclic equation, it is possible to arrive at logical, meaningful implications explaining the seasonal phenomenon. In this paper, the authors have made an attempt to present the preliminary overall analysis of the births pattern in Puerto Rico over a period of 22 years, from 1950 to 1971.

Annual Births in Puerto Rico

In year 1960, the registered births in Puerto Rico amounted to slighly over 85,000 (Graph 2). The total of births has been decreasing slowly and except for a few upward values, has been around 70,000 in the last four years. The births rate in Puerto Rico has been on the decline, from 38.5 in 1950 to 25.6 in 1971. (Rates per 1,000 population)

Monthly Percentage Distribution

The pattern of monthly births within the year, however, has a peculiar form. In order to be able to compare such pattern from year to year the registered births by months are distributed as percentage of the yearly total. Table 2 and Graph 3, show these percentages. Deviations of the monthly percentage from the expected value 8.33% can be seen to be markedly similar since 1954 to 1969. It may be noted that the proportion of births increase from July until October and starts declining for the next four months, whereafter the births stay fluctuative at the same low level for the next four months. There are such three distinct seasonal periods in the year. Table 3 and Graph 4 show the average trend of these percentages over a period of 22 years. The figures indicate three distinct seasonal periods of the year, namely

July to October : Increase
November to February: Decline
March to June: Fluctuating at a low level

Interpretation of the Pattern

In order to interpret the birth pattern, one would go back to the possible time of conception. The normal time of 9 months and 10 days places the highest conception period in case of the above dates, between October to January, the winter time and the lowest one between May to August, the summer time. The winter time generally keeps people more at home with less outdoor activities. Besides, it includes several long periods of social festivities such as Christmas. In Puerto Rico, Christmas celebrations generally start almost in the beginning of the month of December and lasts until the first half of January next year. The summer time is passed more outdoor due to vacation time for children and also adults. These activities accompanied by less privacy may induce low frequency of love-making opportunities and hence of conception compared with the winter time's privacy and the reaction in the mood after the festive days of Christmas time.

Fitting a Cyclic Curve

The pattern is such a regular one, at least for the data on hand that it was decided to determine the equation for this seasonal phenomenon. A sine-cosine curve was tried to fit to the average percentages. The curve fitted turned out to be as follows:

$Y = 8.33 - 0.6025 \sin (30x)^{\circ} + 0.2439 \cos (30x)^{\circ}$
where Y = percent monthly births
and X = coded number of month (January=1, December =12)

The fit is highly significant (Prob. level 1%), 75.2 percent of the variation explained by the time variable, X.

The maximum value of Y occurs at X= 9.7 months. If one considers the middle of the month as the point of coincidence for the coded number, this maximum value would represent the first week of October as the peak point of birth, thus indicating roughly the Christmas time as the highest incidence of conception.

Some Implications of this pattern:

In a democratic society like the one we are living in, one cannot force to change the human behavior which does not interfere in the freedom of the other fellow man. Hence the society has to accept this and must try to meet with the demand it would create in terms of hospital beds, doctors and nurses for the peak delivery period.

On the other hand, the society which is planning a birth control program, may have to carry out this propaganda campaign during this peak period of conception making people more conscious in this matter and may gain some impact on their control program. Anyway, man is supposed to have his own way of living, behaving and carrying on his personal activities as he wishes. However, he is born in a certain type of society with customs and traditions already prevailing from long time before he was born, and therefore, he, in general, is not as free as he thinks he is. He is influenced directly by this society that has customs and traditions which in turn are influenced by the climate and other natural surroundings. Thus, even if man is said to be a creature who makes love all the time, one would say, from the above data, that he may be following, indirectly, some pattern of seasonality of nature like other creatures.

ACKNOWLEDGEMENTS

The authors wish to thank Mr. Luis Collazo, of Puerto Rico Health Department for supplying the data and Mr. Shrikant I. Bangdiwala, of the University of North Carolina, for his help in computing and preparing graphs.

Table 1
Average Percent Distribution of Births, U.S.A., 1950-61

Month		Average Births (Percent)
January		8.19
February		7.63
March		8.26
April		7.64
May		7.95
June		8.09
July		8.84
August		9.04
September		8.92
October		8.76
November		8.21
December		8.47
	Total	100.00
Annual Average (1950-1961)		4,042,618

Table 3
Average Percent Distribution of Births Observed and Estimated Puerto Rico, 1950-1971

Month	Average Births Observed (Percent)	Estimated Trend Value (Percent)
January	8.38	8.24
February	7.40	7.93
March	8.03	7.73
April	7.86	7.69
May	8.11	7.82
June	7.74	8.09
July	8.20	8.42
August	8.67	8.73
September	9.21	8.94
October	9.22	8.98
November	8.60	8.85
December	8.58	8.58
Total	100.00	100.00
Average Yearly Births (1950-1971)	76,100	76,100

Table 2

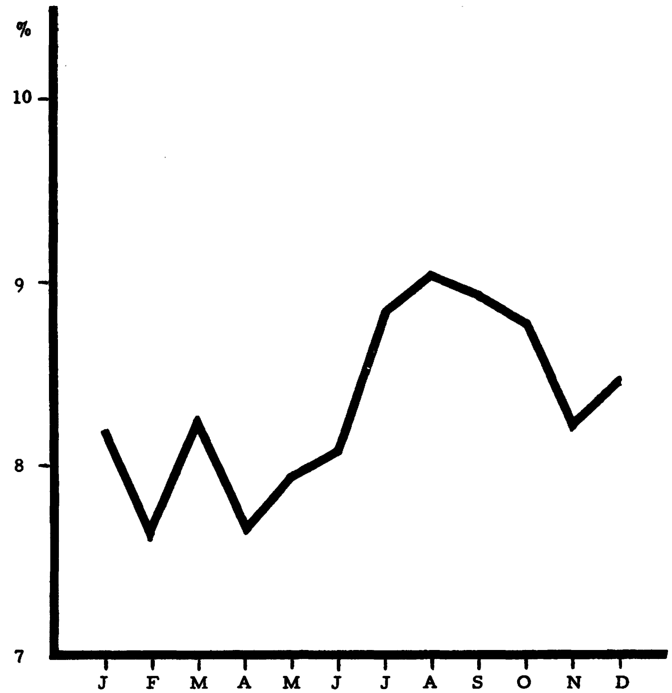
Percentage Distribution of Births by month, Puerto Rico
from 1950 to 1971

	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960	1961
January	8.77	8.61	7.79	8.19	8.00	8.51	8.56	8.68	8.26	8.38	8.54	8.42
February	7.85	7.51	7.09	6.87	7.40	7.29	7.91	7.52	7.40	7.35	7.60	7.38
March	8.52	8.44	8.39	7.71	7.99	7.88	8.17	8.23	8.02	7.64	7.74	7.88
April	8.62	8.48	8.64	8.25	7.61	8.01	7.61	7.77	7.71	7.69	7.76	7.53
May	8.62	8.79	8.99	8.49	8.20	8.29	7.81	7.58	7.78	7.90	8.06	7.98
June	8.00	8.12	8.18	8.33	7.87	8.03	7.70	7.64	7.07	7.55	7.62	7.77
July	8.22	8.47	8.23	8.59	8.29	8.47	8.10	8.19	7.99	7.99	8.26	8.39
August	8.11	8.56	8.69	8.79	8.80	8.75	8.75	8.83	8.82	8.89	8.89	8.61
September	8.55	8.72	8.85	8.96	9.19	8.96	9.21	9.47	9.54	9.49	9.34	9.42
October	8.59	8.51	8.63	9.09	9.36	9.10	9.26	9.18	9.55	9.47	9.22	9.26
November	7.87	7.73	8.22	8.53	8.71	8.28	8.41	8.51	8.97	8.79	8.45	8.71
December	8.28	8.06	8.30	8.20	8.58	8.43	8.51	8.40	8.89	8.86	8.52	8.65
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Total Births	85455	84007	80200	77380	78008	79221	78177	76067	76128	74933	76015	75563

	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971
January	8.02	8.22	8.28	8.32	8.74	8.58	8.31	8.38	8.05	8.80
February	7.20	7.18	7.47	7.36	7.76	7.14	7.45	7.46	6.82	7.74
March	8.01	7.91	8.04	8.12	8.11	8.22	7.91	8.06	7.45	7.98
April	7.79	7.60	7.58	7.60	7.74	8.11	7.49	7.76	7.34	7.84
May	8.05	8.03	7.68	7.87	8.15	8.11	8.28	7.98	7.93	7.76
June	7.72	7.85	7.31	7.63	7.77	7.94	7.69	7.55	7.07	7.76
July	8.26	8.29	7.56	8.29	8.04	8.22	8.37	8.01	7.85	8.12
August	8.89	8.67	8.62	8.71	8.57	8.76	8.38	8.51	8.67	8.47
September	9.17	9.43	9.42	9.47	9.15	9.26	9.37	9.34	9.49	9.13
October	9.37	9.34	9.66	9.51	9.14	9.16	9.35	9.50	9.69	9.25
November	8.88	8.80	9.28	8.61	8.55	8.23	8.57	8.78	9.82	8.65
December	8.64	8.68	9.10	8.51	8.28	8.27	8.83	8.67	9.82	8.50
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Total Births	76677	77382	78837	79586	75735	70735	67989	67577	67438	71117

GRAPH 1

Average Percent Distribution of Births , U.S.A., 1950 - 61



GRAPH 4

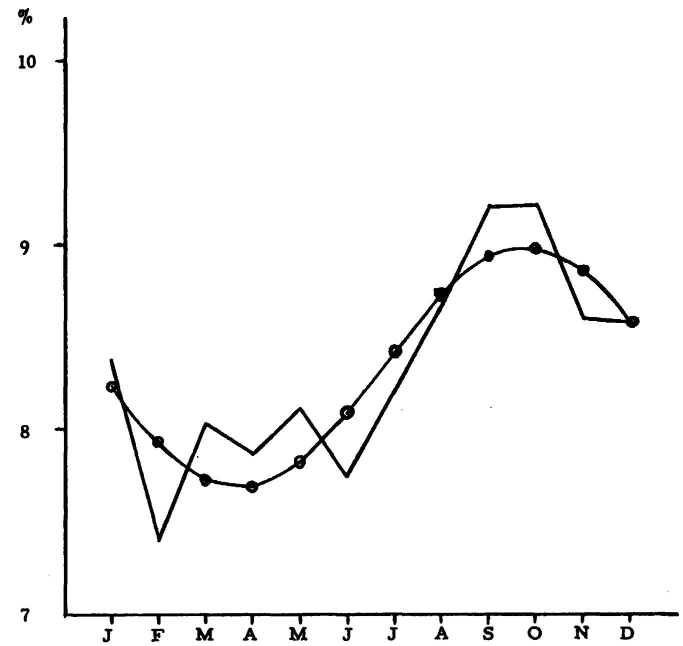
Average Percent of Yearly Births by Months (Data 1950 to 1971), Puerto Rico

Observed %

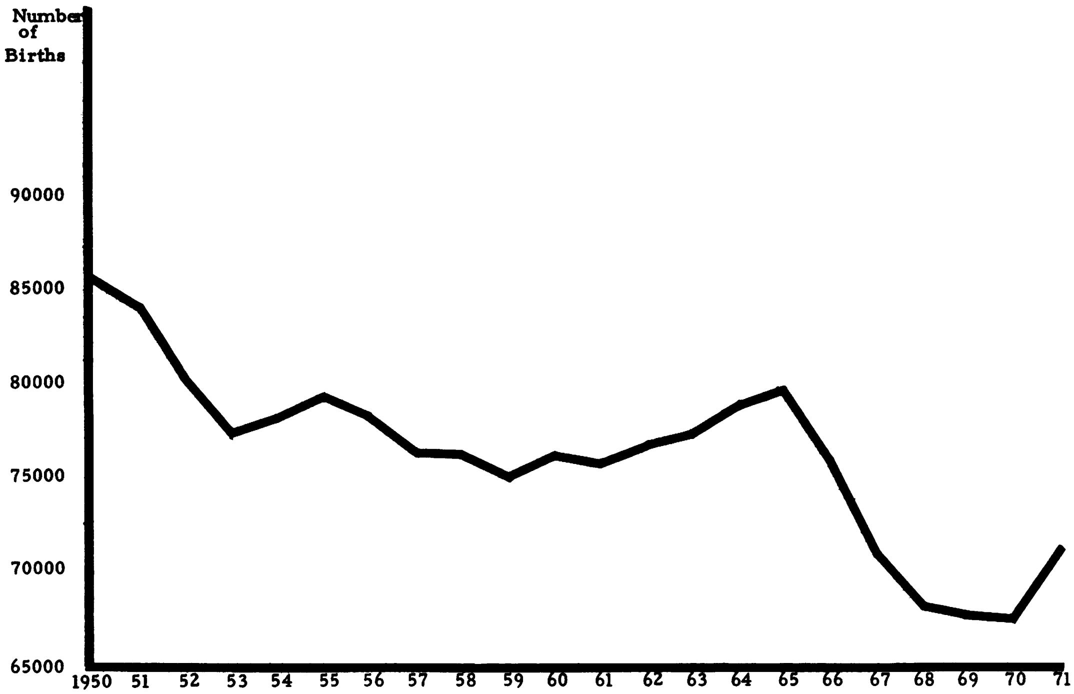
Estimated %

Estimated Percent Birth for the Month = 8.333 - 0.6025 Sine (30 X) ° + 0.2439 Cosine (30 X) °

(X = Number of Month, January=1 December=12)

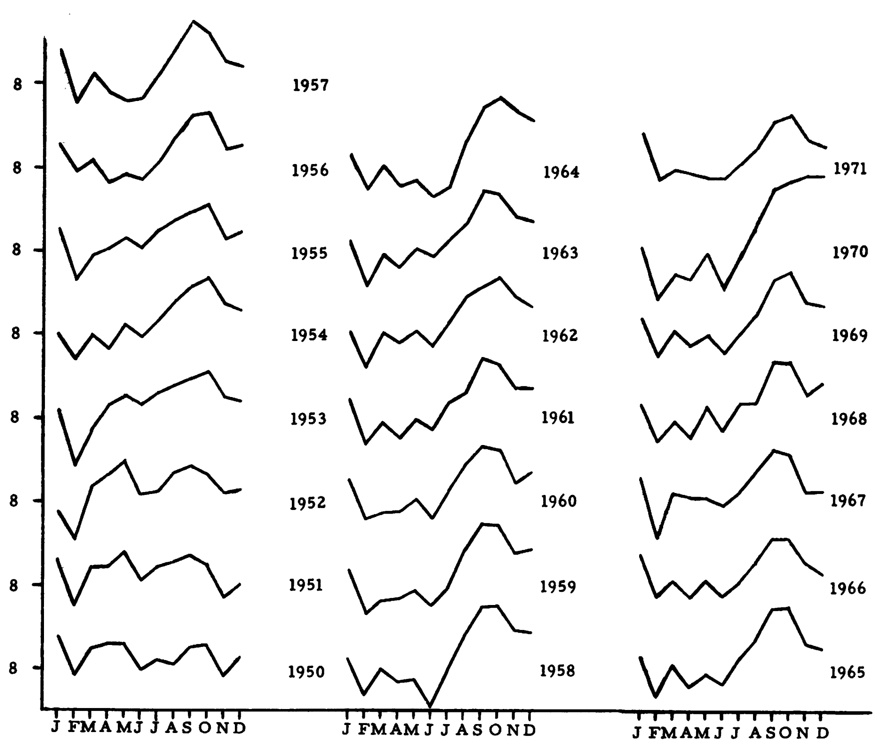


GRAPH 2 Annual Births, Puerto Rico, 1950 to 1971



GRAPH 3

Percent Distribution of Monthly Births, Puerto Rico, 1950 to 1971



THE CLEARINGHOUSE AND LABORATORY FOR CENSUS DATA AND STATISTICS ON FEDERAL STATISTICS USERS

John C. Beresford, National Data Use and Access Laboratories

The report of the President's Commission on Federal Statistics (see /1/) contains some recommendations which are being developed into regular activities of the CLCD. These activities will result in the creation of statistics on users of census statistics. It is somewhat unusual for a private organization to be following recommendations of a Presidential commission. This paper explains why the CLCD is doing so and how it hopes to achieve its results.

The CLCD is an activity of the National Data Use and Access Laboratories (DUALabs), a private non-profit organization concerned with increasing the application of publicly available data to public problems and national needs. The CLCD is an outgrowth of the work of DUALabs and the Center for Research Libraries in supporting the efforts of researchers at universities and elsewhere to acquire and apply the computer tape output of the 1970 Census. Prior to the emergence of the CLCD, over 50 universities and research organizations had joined in DUALabs' Summary Tape Assistance Research and Training (START) Program. These groups had contributed funds of their own, and received additional funds from the Ford Foundation to make possible the community acquisition of the data base from the Bureau of the Census. DUALabs prepared a library of computer programs which greatly simplified the problems of the use of the 1970 Census on computer tape. Also, the tape files were reduced in bulk by 85 to 90 percent, and the cost of acquisition to each participant was similarly reduced.

DUALabs then undertook responsibility for training three persons from each organization in the START Community in the use of the software, problems and errors in the data base, and the general characteristics of the census files which they had to learn to use the files effectively. The basic purpose of this training has been to equip the three persons at each organization to train other users in a similar manner, and to provide support to users at their universities in retrieving, evaluating, and applying census data to their research problems.

The START Program was designed to run for five years, through 1974, at which time we anticipate that strong census use capabilities permanently will be in place at many major research sites around the nation. The START Community, as it is called, uses a common computer technology. An important advantage of the common technology is that researchers can move from one site to another and continue their work with a minimum of technology transfer problems because the same user software, techniques, and information support systems exist at all sites. If access to a particular portion of the data files is not available at a specific site, it can be obtained from the community store at DUALabs at minimum time and cost.

/1/ Federal Statistics Report of the President's Commission, Vol. 1, pages 6,7, GPO, 1971.

An interesting advantage of the community concept is that members can contribute feedback in the form of new data files, software improvements, information about data base errors, and new techniques for using the data in research applications. This feedback has already resulted in the creation of compatibility between important segments of the 1960 census data base and the START software system. It has added 1960 census data to the community storehouse. It has increased the capability of the START software.

It was recognized from the beginning that a great many researchers who were directly connected with the START Program of their university, as well as many more researchers at non-START locations, would require a different form of support to use the 1970 Census. These researchers would be in need of something akin to the services of a research librarian. They would believe census data could help solve their problem but they would need an expert to help them decide if this were true and to advise them how to overcome technical problems to use the data. They would not wish to become census data experts, would not wish to get involved in computer systems, and would not want to be apprised of all errors and all potential applications. But, in the absence of expert assistance, the contents and potentialities of the summary tapes and the Public Use Samples would remain a mystery to these researchers because of the complexity, size, and variety of information resources required to use the materials. The CLCD was established with support from the National Science Foundation for the express purpose of facilitating research applied to national needs which could benefit from the use of the 1970 Census. The general objective of the CLCD is to guide a researcher through the seemingly impenetrable thicket of problems associated with the use of the files and quickly bring him to a point of applying data to a specific research objective. Therefore, the CLCD has a consulting role to orient a user, to guide him, to offer special training as needed, to bring him in contact with those who can provide data, and to help insure that resources exist to accomplish his research objective. (The CLCD is not funded to supply computer or clerical services to researchers.)

To reach as many users as possible, the CLCD has enlisted the aid of groups who have established themselves as competent sources of assistance in the use of the 1970 Census, with proven experience in meeting users application requests in a timely and efficient manner. These groups are called User Contact Sites of the CLCD. They all offer a no-cost orientation service to any researchers. There are now 14 User Contact Sites from Massachusetts to California, and more will be added in the future. (See attachment)

When potential users of the census data base contact the CLCD or any of the User Contact Sites they are informed of consulting and training activities available to them and asked to fill in a Census Use Questionnaire. This questionnaire serves to identify the major interests of the user and to asses the knowledge he possesses of the data base. The replies form the basis for the initial orientation (by letter, telephone, or in person). In many cases this orientation directs the potential users to appropriate printed materials from the Census Bureau, DUALabs, or elsewhere, providing the required data or information about data.

If the user requires further guidance, it is provided by additional personal consultations. Detailed guidance can be arranged if the user wishes to become a "visiting scholar" at the CLCD, meaning that he is assigned desk space and afforded the opportunity for interaction with CLCD staff while he works out the census use and access details of his research project. Group training sessions are also provided and the publications called Census Use Technical Bulletins and Data Access News are used to keep the researcher informed of current developments.

In reviewing the needs of researchers and the requirements of efficiency in CLCD services, it became apparent to the CLCD staff that the greatest possible benefit in applying the results of the census to national needs would be obtained only if a vigorous program were set up to evaluate CLCD services and users. As a result, two additional components of the CLCD work have been constructed. One is an advisory system, and the other component is an information system on users and uses.

Advisory groups have been established representing researchers, State and Local government users, legislators, librarians, computer experts and the User Contact Sites.

The advisory groups have been reviewing the CLCD services to date and have suggested a variety of improvements. A number of these suggestions either parallel or extend some recommendations of the President's Commission on Federal Statistics.

Item: The Commission's report suggests that small area data uses of the census required by law be catalogued. CLCD advisors have suggested that the legally required uses be identified and that an effort be made to establish specific services to enable users quickly and cheaply to acquire the data specified in the laws in a form which matches the reporting requirements of Federal agencies.

Item: The Commission recommends that field studies of the uses of small area data be made. The CLCD advisers have suggested compiling information of this sort from CLCD Census Questionnaires.

Item: The Commission recommends that schemes of cooperative collection and analysis be tried out to determine who uses the census, for what purpose and whether the census provides the information the users want. CLCD advisers have suggested that analyses of census users be made on these points.

Item: The Commission recommends that interchange of knowledge about data for small areas should be encouraged among workers at all levels of government, with special emphasis on the need for analytical specification of data requirements and procedures for using data. CLCD advisers have noted the cooperative possibilities in the exchange of analytic approaches afforded by the CLCD network of information exchange and the Clearinghouse function of referral of users to one another for assistance. They further noted the large number of organizations trained by DUALabs in procedures for using the data and have recommended a great expansion of this effort to reach out to analysts with specific research applications, in addition to those whose emphasis is on computer and data base technology.

Item: The Commission recommends that serious attention should be given to coordinated examination of the demands for data made by the Federal government on State and Local governments. CLCD advisers have noted these demands, have identified some States in which the responses are coordinated, and have suggested the CLCD attempt to educate the users in other States in systems for coordinating their efforts to respond to Federal demands. (This assumes that no coordination will be attempted at the Federal level.)

In response to its advisers suggestions, the CLCD is now planning several new programs. These programs would begin to accomplish the tasks mentioned above which the Commission had seen as responsibilities of the Federal government. Where success in accomplishing these tasks depends on an awareness of users needs, the CLCD may be more effective than the Federal government. (Because the CLCD exists to serve users' needs, it has a far greater desire to respond to these needs than the Federal statistical agencies, whose missions are more complex and less likely to emphasize user requirements.)

The greatest weakness in the CLCD as an instrument for surveying user needs is that it collects information only from those users who come to it seeking help.

The CLCD is creating an information system on users of its services and on the uses they make of the data. The information system will focus on the people and organizations using the 1970 Census. It will have organizational classification codes, and it will classify the characteristics and frequency of census data applications. The kinds of data used, the purposes of the use, and needed data not present in the census will be recorded. The user's capabilities in terms of data bases on site, software and equipment available, and willingness to assist others will be noted. Also noted will be past experience of the user in terms of research activities, special training, data acquired, and software used.

The information will be made available in the following kinds of registers, catalogues, and lists.

Uses made of small area census data required by statute:
- Problems in such use
- Successful solutions to problems

Uses of small area data not required by statute:
- Type of area; type of data
- Purpose of use

Research problems using census (major classifications):
- By researcher and organization

Expert persons available for assistance:
- By location, by area of speciality

Data bases available for use of others:
- By location; software (with information on costs)

Training services available:
- By location; frequency; purpose

Errors in the census:
- By geographic area and subject matter to which errors apply

Special analytic techniques:
- By purpose of analysis, sources of assistance

Although some of this information is now available from the CLCD (for example, an annotated catalogue of sources of assistance in use of the census) it will require extensive additional resources to bring the information system to a fully accomplished project capable of generating reports on demand. In the meantime, the CLCD is collecting enough data on all users it serves to anticipate most of the demands that will be placed on the information system. When the system is developed, it will be possible to create a variety of statistics on users of statistics. A prime purpose of such statistics would be to serve as an information resource for planning future Federal census data collection and processing operations.

Unfortunately, the cases included in the information system will be a biased sample of a universe of data users of unknown size. However, the information collected from the CLCD users will be standard and systematic. Because the users will be making specific uses at the time the information is collected, the resulting statistics could be used in the next decennial census planning process. Such information might be more valuable than the anecdotal information which served as the planning inputs to the last census, and which was distilled from notes taken at public meetings.

<u>CLCD User Contact Sites</u>

CONCORDIA TEACHERS COLLEGE
Victor B. Streufert
7400 August Street
River Forest, Illinois 60305

ILLINOIS STATE UNIVERSITY
Dr. Vernon C. Pohlmann
Department of Sociology-Anthropology
Normal, Illinois 61761

THE KROGER COMPANY
Daniel I. Neidorf
1014 Vine Street
Cincinnati, Ohio 45201

METROPOLITAN PLANNING COMMISSION OF NASHVILLE-DAVIDSON COUNTY
Gerald F. Nicely
P.O. Box 39
Nashville, Tennessee 37202

THE PENNSYLVANIA STATE UNIVERSITY
Dr. Jay L. Politzer
Computation Center - Computer Bldg.
University Park, Pennsylvania 16802

PRINCETON-RUTGERS CENSUS DATA PROJECT
Judith S. Rowe
Princeton University Computer Center
87 Prospect Avenue
Princeton, New Jersey 08540

TEXAS A&M UNIVERSITY LIBRARY
Joe Davis
College Station, Texas 77843

UNIVERSITY OF CALIFORNIA, LA
Mary Ryan
University Research Library
Los Angeles, California 90024

UNIVERSITY OF FLORIDA
Ray Jones
University Libraries, Library West
Gainesville, Florida 32601

UNIVERSITY OF HOUSTON
Dr. Ernest J. Henley
Urban Systems Laboratory
Cullen College of Engineering
Houston, Texas 77004

UNIVERSITY OF MISSISSIPPI
Max W. Williams
Institute of Urban Research
University, Mississippi 38677

UNIVERSITY OF MISSOURI
Dr. Rex Campbell
Demographic Data Service
104 Sociology Building
Columbia, Missouri 65201

UNIVERSITY OF NORTH CAROLINA
Dr. Richard C. Rockwell
Social Science Data Library
Institute for Research in Social Science
Chapel Hill, North Carolina 27514

UNITEL Census Program
Walter I. Nissen
Massachusetts Institute of Technology
Room 145-310
Cambridge, Massachusetts 02139

RACIAL DISCRIMINATION AND INCOME DIFFERENTIALS AS RELATED TO EDUCATION AND AGE, 1939 TO 1967

Bruce Bosworth, St. John's University

Introduction

This paper describes an attempt to investigate an interesting social and economic hypothesis. Originally discussed in the book, The Economics of Discrimination(1), it can be stated as follows: that discrimination against nonwhites, as shown by differences in their incomes when compared to that of whites, systematically increases with the same education and age.(2) This statement implies that the gap in incomes betweeen races widens as members of each race become more educated and grow older.

This investigation compares mean income data for white and nonwhite males within and between four time periods to determine if discrimination did in fact systematically increase with the same education and age.

The Data Used

The empirical investigation uses data from the 1940, 1950, and 1960 censuses. The 1967 income data was obtained from the Current Population Reports, and was derived by manipulation on a computer so that the data could be of the same format as the data of the other years. Table 1 shows the complete set of data used.(3)

Derivation of 1967 Mean Income Data

The mean incomes for 1967 presented in Table 1 are the result of a computer calculation performed on data in the Current Population Reports, Series P-20, No. 182, April 28, 1969, "Educational Attainment: 1968," Table 7, pp. 26-28.(4) The original form of the data was that of 40 cells for each age group and race- five income classes, by eight educational classes- the cell values were per cent by years of school completed. The total population in each income interval was also given. Conversion of this data to mean incomes for each education group was achieved by assuming that midpoints of the income class intervals were the means of the classes, and the mean of the open-end class of \$15,000 was \$25,000; then the population and per cent distributions were used as weights, so that for each education class the mean income was a weighted average calculated as follows:

$$\text{Mean Income} = \frac{\Sigma(\text{Income X Population X Per Cent})}{\Sigma(\text{Population X Per Cent})}$$

Measures of Discrimination

The mean income data in Table 1 are analyzed by using two measures of discrimination. The first measure is the "Market Discrimination Coefficient" or MDC.(5) This measure is defined as

$$\text{MDC} = \frac{Y(W)}{Y(N)} - \frac{Y_o(W)}{Y_o(N)} \qquad (1)$$

where Y(N) and Y(W) represent the actual incomes of N and W, $Y_o(N)$ and $Y_o(W)$ represent the incomes of N and W without discrimination. If it is assumed that W and N are perfect substitutes with perfect compitition in the market place, then $Y_o(W) = Y_o(N)$ and (1) reduces to

$$\text{MDC} = \frac{Y(W)}{Y(N)} - 1 = \frac{Y(W) - Y(N)}{Y(N)} \qquad (2)$$

In the form given by (2), the MDC represents the percentage difference between the incomes of W(white males) and N(nonwhite males) with respect to the income of N(nonwhite males). The MDC has a range from zero or no discrimination, to values above zero representing higher levels of discrimination.

Because the MDC goes from zero upward, without a fixed upper limit, it was modified so that comparisons could be more easily made from within and between time periods being investigated. This modification is called the "Discrimination Measure" or DM.(6) This measure of discrimination is defined as

$$\text{DM} = \frac{Y(W) - Y(N)}{Y(W)} = 1 - \frac{Y(N)}{Y(W)}$$

where Y(N) and Y(W) are as before. The range of the DM is from zero or no discrimination, to one or complete discrimination. The DM represents the percentage difference between the incomes of W (white males) and N(nonwhite males) with respect to the incomes of W(white males).

Analysis of Mean Income Data

The analysis of the mean income data for the four time periods is based on the two measures of discrimination described above.

Discrimination Measures(DM). Table 2 presents the DM's based on the detailed levels of education that are given in Table 1. The following should be noted: (a) Males- 25 years old and over- The DM's in each year increase as education increases from less than 8 years of elementary education, where it is lowest,

TABLE 1

MEAN INCOME FOR MALES 25 YEARS OF AGE AND OVER, BY EDUCATIONAL ATTAINMENT, COLOR, AND AGE FOR THE UNITED STATES, 1939, 1949, AND 1967

Education and Age	White Income				Nonwhite Income			
	1939	1949	1959	1967	1939	1949	1959	1967
25 years and over								
Total-All	1419	3376	6112	8902	609	1664	3260	4961
Elementary:	1125	2540	4422*	6094	566	1466	2730*	4088
Less than 8 years	NA	2234	3983	5395	NA	1386	2562	3862
8 years	NA	2875	4837	6655	NA	1909	3318	4763
High School:	1538	3586	5938*	8299	772	2133	3732*	5336
1 to 3 years	1413	3306	5555	7476	728	2030	3522	4931
4 years	1685	3840	6250	8718	859	2298	4021	5746
College:	2297	5399	9012*	12128	1064	2739	4958*	6772
1 to 3 years	1964	4501	7554	10208	948	2413	4355	6488
4 or more	2646	6250	10238	13491	1208	3177	5671	7222
25 to 44 years								
Total-All	1360	3431	6000	8781	600	1790	3367	5126
Elementary:	1058	2618	5322*	5833	555	1569	2707*	4073
Less than 8 years	NA	2316	3814	5246	NA	1492	2517	3782
8 years	NA	2890	4656	6305	NA	1915	3249	4758
High School:	1436	3425	5739*	8006	745	2145	3714*	5311
1 to 3 years	1336	3203	5351	7247	711	2037	3491	4854
4 years	1552	3607	6008	8338	816	2317	4009	5724
College:	2127	4836	8062*	11365	1016	2662	4839*	6738
1 to 3 years	1850	4159	6938	9612	903	2413	4309	6484
4 or more	2411	5512	8948	12500	1163	3032	5473	7220
45 to 64 years								
Total-All	1562	3715	6289	9050	630	1680	3102	4729
Elementary:	1235	2843	4534*	6255	590	1535	2749*	4098
Less than 8 years	NA	2537	4090	5459	NA	1458	2597	3912
8 years	NA	3185	4957	6857	NA	2027	3408	4756
High School:	1955	4176	6323*	8705	896	2222	3793*	5422
1 to 3 years	1718	3712	5860	7745	814	2137	3615	5112
4 years	2245	4702	6837	9298	1039	2365	4080	5827
College:	2921	6871	11155*	13456	1212	3094	5341*	6858
1 to 3 years	2361	5555	8755	11095	1099	2572	4495	6516
4 or more	3552	8054	13454	15429	1335	3688	6277	7265

NA: not available. *Values derived by aggregations.

TABLE 2

DISCRIMINATION MEASURES FOR MEAN INCOME OF WHITE AND NONWHITE MALES, BY EDUCATIONAL ATTAINMENT, AND AGE, FOR 1939, 1949, 1959 AND 1967

Educational Attainment and Age	1939	1949	1959	1967
25 years old & over				
Elementary:				
Less tahn 8 years	NA	.380	.357	.280
8 years	NA	.336	.314	.284
High School:				
1 to 3 years	.485	.386	.366	.340
4 years	.490	.402	.357	.341
College:				
1 to 3 years	.517	.464	.423	.364
4 or more	.543	.492	.446	.465
25 to 44 years				
Elementary:				
Less than 8 years	NA	.356	.340	.279
8 years	NA	.337	.304	.245
High School:				
1 to 3 years	.468	.364	.348	.330
4 years	.474	.358	.332	.314
College:				
1 to 3 years	.512	.420	.379	.325
4 or more	.518	.450	.388	.422
45 to 64 years				
Elementary:				
Less than 8 years	NA	.425	.365	.283
8 years	NA	.364	.312	.306
High School:				
1 to 3 years	.526	.424	.383	.340
4 years	.537	.498	.403	.373
College:				
1 to 3 years	.535	.537	.487	.413
4 or more	.624	.542	.533	.529

Source: Based on data from TABLE 1. NA: not available.

TABLE 3

DISCRIMINATION MEASURES AND MARKET DISCRIMINATION COEFFICIENTS, BY BROAD EDUCATION CLASSIFICATIONS, AND AGE, FOR 1939, 1949, 1959 AND 1967

Educationa Attainment and Age	1939	1949	1959	1967
	Discrimination Measures			
25 years old & over				
All	.571	.507	.467	.443
Elementary	.497	.423	.383	.329
High School	.498	.405	.371	.357
College	.537	.493	.450	.442
25-44 years old				
All	.559	.478	.439	.416
Elementary	.475	.401	.491	.302
High School	.481	.374	.353	.337
College	.522	.450	.400	.407
45-64 years old				
All	.597	.548	.507	.478
Elementary	.523	.460	.394	.345
High School	.542	.468	.400	.376
College	.585	.550	.521	.490
	Market Discrimination Coefficients			
25 years old & over				
All	1.330	1.029	.875	.794
Elementary	.988	.733	.620	.491
High School	.992	.681	.591	.555
College	1.159	.971	.818	.791
25-44 years old				
All	1.267	.917	.782	.713
Elementary	.906	.669	.966	.432
High School	.928	.597	.545	.507
College	1.093	.817	.666	.687
45-64 years old				
All	1.479	1.211	1.027	.914
Elementary	1.093	.852	.649	.526
High School	1.182	.879	.668	.608
College	1.410	1.221	1.089	.962

Source: Based on data from TABLE 1.

to 4 or more years of college, where the DM's are highest. But, between years, within each level of educational attainment the DM's decline except in one case. The DM value in 1967 for 4 or more years of college is slightly larger than the comparable DM in 1959, but below the one in 1949. This higher value in 1967 is due to the incompleteness of the original source data.(7)
(b) Males- 25 to 44 years old- For this age group the DM's follow the patterns shown by the 25 years old and over group. That is, an increase in DM's within each year as education increases, but declines from year to year, with the highest DM's in 1939 and the lowest ones in 1967.
(c) Males- 45 to 64 years old- For this age group the DM's follow the same patterns mentioned above. Increases in the DM's within each year as education increases, and declines from year to year in each education level.

Comparison of the DM's of the two broad age groups, 25 to 44 years old, and 45 to 64 years old, shows the younger group having lower DM's than the older group. This difference suggests an increase in discrimination against nonwhite males as their age increases.

Market Discrimination Coefficients (MDC). MDC's were calculated for the same set of data as used to calculate the above DM's. The results noted for the DM's apply for the MDC's. In general, these MDC's suggest a rise in discrimination against nonwhite males within each year as education increases, with declining discrimination over time in each education level.(8)

Additional Data Analyses

The mean income data in Table 1 were analyzed by detailed education classifications as shown in Table 2. An additional analysis by major educational classifications was also carried out. That is, for those males who have completed elementary education, or high school education, or college education.

For these broad groups DM's and MDC's were calculated. The results are given in Table 3. Analysis and interpretation of these two measures gives support to findings described earlier in this paper. From Table 3 it can be seen that within each year discrimination against nonwhite males with the same education and age as white males increases. Discrimination is lowest for those with only an elementary education, and is highest for those with a college education in each year.

Figure 1 illustrates these increases within each year by the plotted DM values for both age groups. Over time, from 1939 to 1967, the plotted DM's for the mean incomes show a decline suggesting a decrease in discrimination against nonwhite males in each education group.

Figure 2 further illustrates this decline in discrimination over time by showing the plotted DM values for each age group by their level of education.

A further suggestion of the decline in discrimination against nonwhite males over time is shown by Figure 3. The plots in this graph are based on the assumption that those males who were 25-44 years old in 1939, will be 45-64 years old in 1959. The DM values in Table 3 for these two groups are plotted in Figure 3 by level of education. Each level of education exhibits a decline over the 1939-1959 period. Also, as seen in the previous two figures, the greatest discrimination is found at the college level of education, with the lowest discrimination faced by those nonwhite males having only an elementary level of education.

Summary and Conclusions

The analyses based on DM's and MDC's calculated for mean income data of males in 1939, 1949, 1959, and 1967 showed that in these years both measures of discrimination increased within each time period as both education and age increased. But over time there were declines in these measures. These results suggest that <u>within</u> each year discrimination against nonwhite males increases with their education and age. But they also suggest that <u>over time</u> discrimination against nonwhite males decreases with their education and age.

When the same group of men are studied in 1939 and 1959, the findings suggest that although discrimination increases with increased education, it declines with increased age. Thus, nonwhite males when compared to white males with the same education and age are faced with less discrimination as both groups get older.

With respect to the hypothesis stated earlier; if discrimination against nonwhites with regard to their education and age was systematically increasing, it was expected that the two measures of discrimination would show either stable patterns or increasing patterns over time. All of the above analyses do not show measures of discrimination that are increasing over time that would suggest systematic increases in discrimination with regard to education and age. Rather, the declines over time of the measures of discrimination suggest that discrimination against nonwhite males with the same

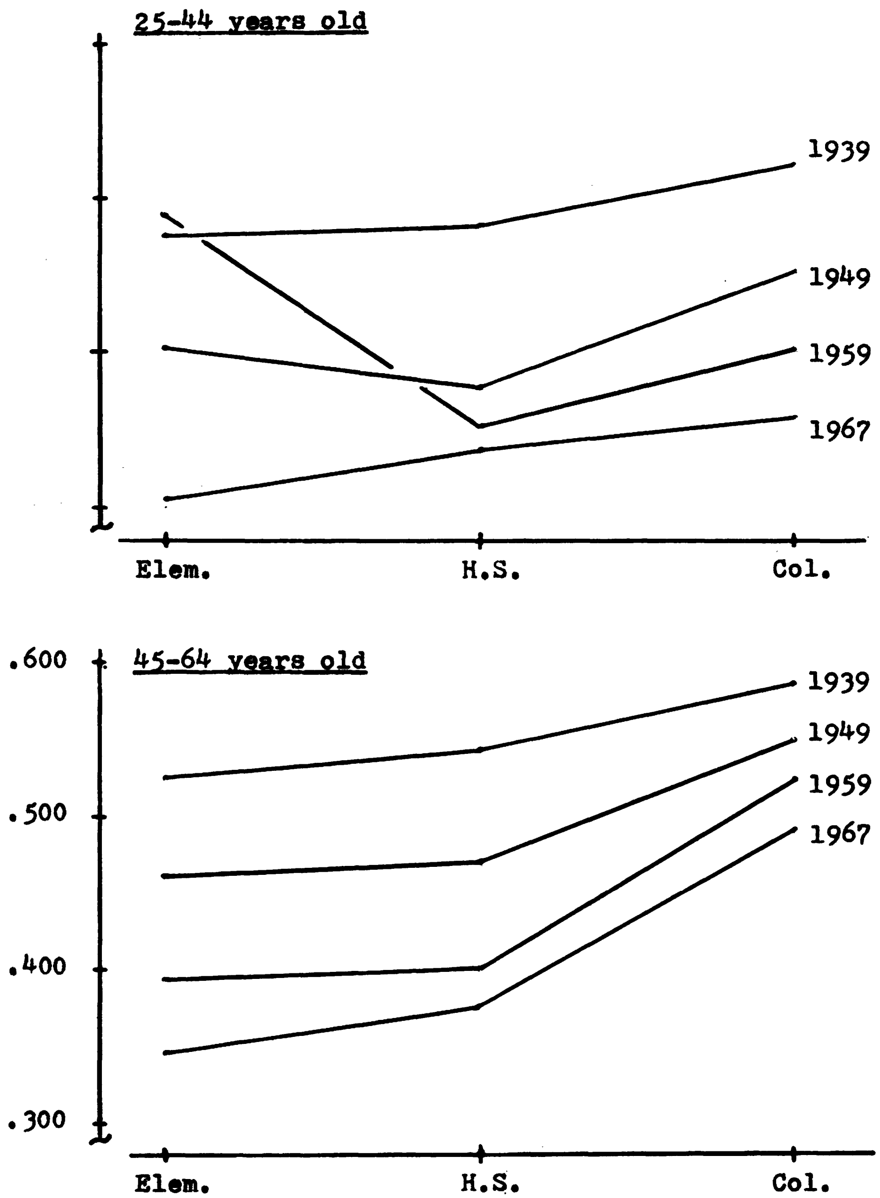


Figure 1. Plotted DM Values as Education Increases, By Age Group, For 1939, 1949, 1959 and 1967
(Source: TABLE 3)

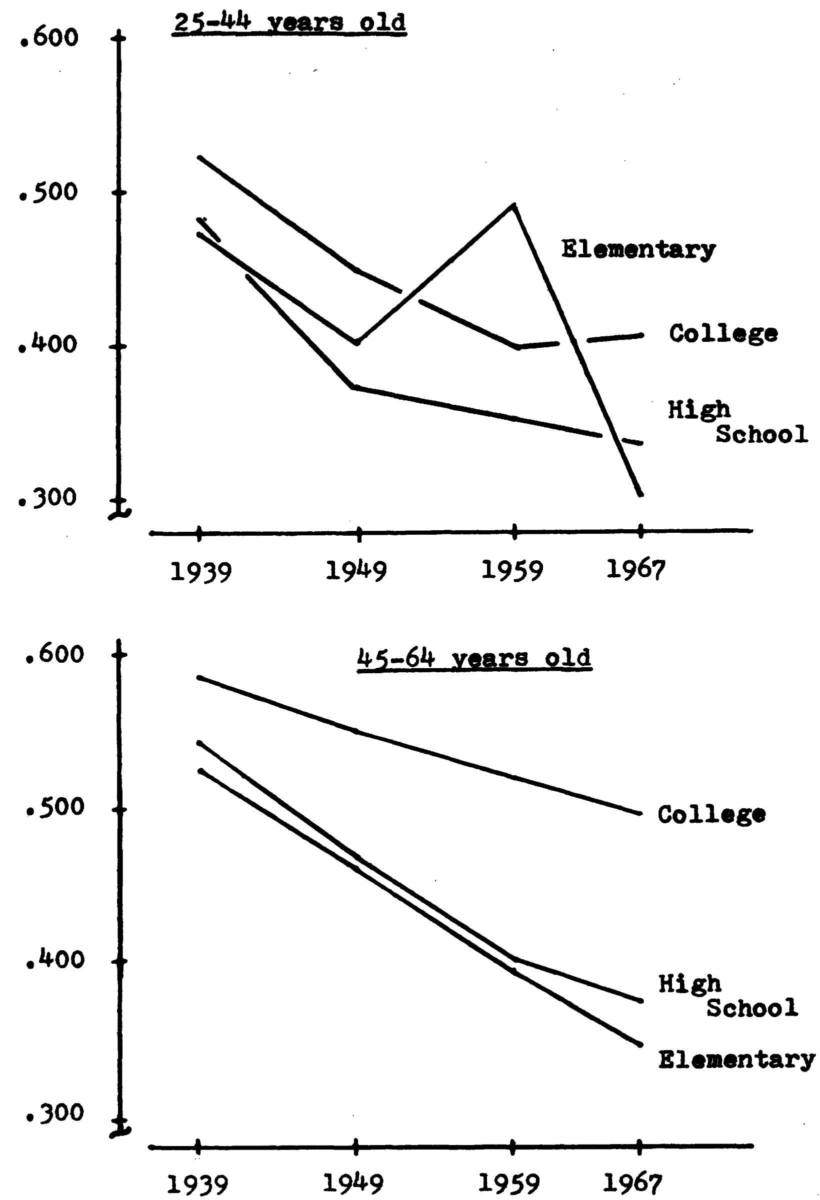


Figure 2. Plotted DM Values Over Time- 1939, 1949, 1959 and 1967, By Age Group, For Each Level of Education
(Source: TABLE 3)

education and age as white males may be systematically decreasing.

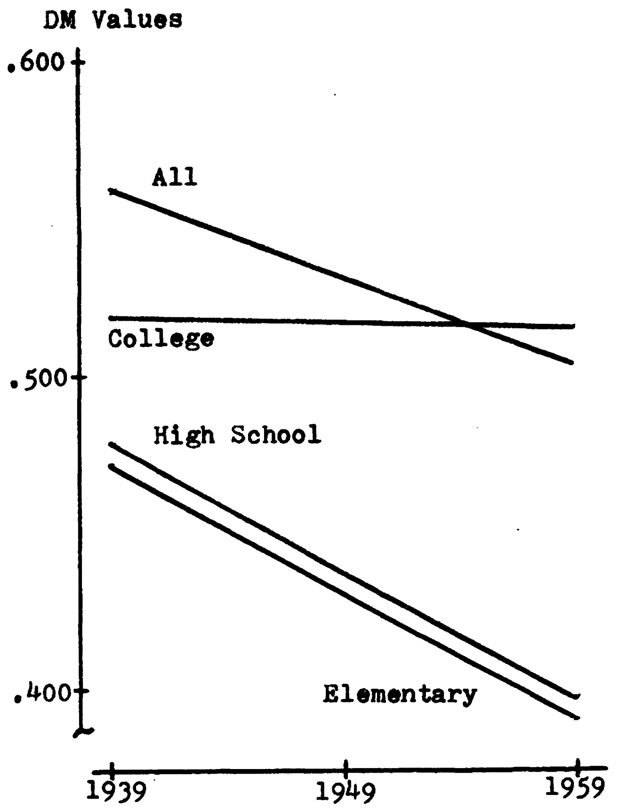


Figure 3. Plotted DM Values: 25-44 Years Old in 1939, and 45-64 Years Old in 1959, For Each Level of Education (Source: TABLE 3)

Footnotes

(1) Becker, Gary S., The Economics of Discrimination, (The University of Chicago Press, Chicago, 1959), pp. 129-131.
(2) Bosworth, Bruce,"A Quantitative Analysis of White-Nonwhite Income Differentials as Related to Education and Age," (Unpublished doctoral dissertation, New York University, Graduate School of Business Administration, 1971), p.4.
(3) Table 1 sources: the 1939 and 1949 data appear as "Table 3. - Mean Income for Males 25 Years of Age and Over by Educational Attainment, Color, and Age, United States, 1939, 1949, and 1956," in Herman P. Miller, "Income and Education: Does Education Pay Off?," an article in Selma Mushkin (ed.), Economics of Higher Education, (U. S. Department HEW, Washington, D. C., 1962), p. 137. The original source indicated for this data is: 1939- the Sixteenth Decennial Census of Population, "Education: Educational Attainemnt by Economic Characteristics and Marital Status"; 1949- the U. S. Census of Population: 1950, Series P-E, No. 5B, "Education." The 1959 data source is "Table 1. Occupation and Earnings of Males 25 to 64 Years Old in the Experienced Civilian Labor Force with Earnings in 1959, by Years of School Completed, Age, and Color, for the United States: 1960," pp. 2-3 in "Occupation by Earnings and Education," PC(2)-7B, U. S. Census of Population: 1960. For the 1967 data source see text.
(4) For a detailed discussion of the computer calculations performed see Bosworth, Op. Cit., p. 46.
(5) Becker, Op. Cit., p. 14.
(6) Bosworth, Bruce, "White-Nonwhite Income Differentials as a Measure of Discrimination," 1970 Proceedings of the American Statistical Association, Social Statistics Section, p. 251.
(7) For Negroes within the age groups 25 to 44 years old and 45 to 64 years, the 1967 source data was incomplete due to the sample base being less than 75,000. Thus, no percent by years of school completed was given for higher income values. To overcome the downward bias that would result because of the lack of these distributions, I assumed that the distribution given for the Negro males 25 to 64 years old applied. An indication of this downward bias and its adjustment is illustrated by the follwoing data actually calculated:

Mean Incomes of Negro Males with 1 to 3 Years of College

Age	Before Adj'm't	After Adj'm't
25-64	$6488	$6488
25-44	5950	6484
45-64	5690	6516

(8) These MDC values can be found in Table 45, page 190, of the work cited in footnote (2).

A NEW APPROACH FOR THE MANAGEMENT OF HEALTH

Martin K. Chen

National Center for Health Services Research and Development

A fairly exhaustive review of the available literature reveals that not much progress has been made in the measurement of health since Stouman and Falk (1) developed quantitative indices in three areas: vitality and health, environment, and public health activity. In recent years, the work of Chiang (2) and that of Fanshel and Bush (3) represent the more notable efforts in this area using, respectively, stochastic and deterministic models. The indicators developed by these authors, however, present conceptual and methodological problems that tend to limit their utility in their present state.

Chiang's index is based on the probability distribution of the population at risk with respect to what he assumes to be three independent random variables--the frequency of illness, the duration of illness in number of days, and the time lost due to death in a given year, again in number of days. Conceptually, the duration of illness is purported to be a measure of the severity of illness, but in reality it is more appropriately a measure of the chronicity or acuteness of illness. Furthermore, his assumption of independence of the three random variables is hardly tenable in view of the fact that it is a common phenomenon for many individuals to be sick before their death and that within a given year the duration of illness cannot be independent of the time lost due to death.

The index of Fanshel and Bush, termed the Health Status Index (HSI), is operationally defined as the mean value of 11 weighted functional states of health of the population at a given time. The weights for the different functional states vary from 0 for the lowest state, death, to unity for the highest state, well-being. The other intermediate weights are assigned by a technique similar to Thurstone's paired comparison scaling technique (4), based on the proportions of times a panel of judges rate one attribute (in this case functional state) over the other, pair by pair.

While the paired comparison technique is a rigorous way of estimating personal preferences--indeed, Mosteller (5) has shown that under certain conditions the estimates are identical with the least squares estimates, Fanshel and Bush's technique permits of modifications of the estimated values according to expert prognoses and expected benefits from health program intervention given limited resources. In so doing the authors make their HSI contingent on prognoses and judgmentally-determined optimum resourse allocations, with the result that the HSI is no longer an invariant measure of health status per se.

Philosophically and conceptually, the authors of the two indices implicitly, if not explicitly, define health in absolute terms, when in reality health is more appropriately regarded as a relative concept. Dubos (6), who has written extensively on health from the philosophical, biological and sociological points of view (7,8), states:

"Health ... cannot be absolute and permanent values, however careful the social and medical planning."

The World Health Organization (9), after a lengthy discussion of the concept of health in which it frankly admits that its definition of health as "a state of complete physical, mental and social wellbeing" does not lend itself to objective measurement, comes to the conclusion that health may best be expressed as "a degree of conformity to accepted standards of given criteria in terms of basic conditions of age, sex, community and region, within normal limits of variation."

Health as a relative concept, then, is the basis for the formulations of the two health indicators that are to follow. Two key factors are clearly implied by this concept and both are required for the validity of the formulation. One key factor is that an "ideal norm", however derived, is available in each health dimension. These norms would be the "accepted standards of given criteria" as stated by WHO. The second key factor is that members of a given group or population must vary around, not necessarily symmetrically, the "ideal norm" along each health dimension. This factor is related to a measure of variation as conceived of by WHO.

Before we proceed to derive the indicators, we would like to emphasize that the "ideal norm" of a health dimension, such as age and height adjusted bodily weight, may or may not be the mean value of a group of people. For instance, it is recognized that in our affluent society where food is plentiful, a large number of middle-aged males and females have an overweight problem. In such a case, the mean weight of an age and sex group obviously cannot be the "ideal norm", in the form of a point or in the form of a range. Another norm, less than the mean, must be sought.

Since health is multi-dimensional (See, for example, Sullivan (10)), it is essential that the dimensions, which may be quite disparate, be measured on the same scale and with the same degree of variability. This is, of course, impossible because the units of measurement for different health dimensions, as for example, visual acuity and blood pressure, cannot be the same. The only feasible solution, then, is to transform the raw measures, taken as deviations from the "ideal norms," into standard scores.

Notice that the deviations from the "ideal norm" divided by the "standard deviation" are not standard scores unless the norm happens to coincide with the mean. The sum of the deviations from an origin other than the mean is, of course, not zero, and the "standard deviation" is not a true standard deviation in that the squared deviations summed are not taken from the mean. Thus the transformed scores do not possess the properties of the standard scores. In fact, it is algebraically proved (See appendix) that the variance of these transformed scores is always less than unity.

The real problem posed by using the transformed scores is that neither the means nor the standard deviations of the transformed scores on different health dimensions are comparable, and simple aggregation of the transformed scores would not ensure equal weights to the health dimensions, because the weights would vary directly in proportion to the magnitudes of the means of the health dimensions. The solution to this problem is to treat the deviation scores from the "ideal norm" as raw scores, and where positive and negative deviations connote different degrees of seriousness from the health point of view, differentially weight the positive and negative deviations. Then the positive and negative signs of the scores can be discarded and the absolute values used in transforming them into z scores.

DERIVATION OF -H INDEX

Symbolically, if we let $\pm X_{ij}$ be individual i's deviation score from the "ideal norm" on dimension j, the corrected deviation score after weighting is $X'_{ij} = V_{+,-} \left|\pm X_{ij}\right|$, where $V_{+,-}$ are the differential weights for the positive and negative deviation scores, and $\left|\pm X_{ij}\right|$ is the absolute value of individual i's deviation score on dimension j. $V_{+,-}$ can be any number or zero for deviation scores that are zero, because any number times zero is zero. Then treating the corrected deviation scores as raw scores, we perform the following statistical operations to derive an individual's z score on a single health dimension and his health index, which is the simple aggregation of the weighted health scores in standard form. We shall designate the index as -H to indicate that as an index it has a negative relation with the health level of the individual.

We write $x'_{ij} = X'_{ij} - \bar{X}'_j$, where $\bar{X}'_j = \sum_1^n X'_{ij}/n$, and n is the size of the group. We further write $S_j = (\sum_1^n x'^2_{ij}/n)^{\frac{1}{2}}$. Then individual i's z score on dimension j is $z_{ij} = X'_{ij}/S_j$. To eliminate negative signs, we add a constant 5 to all the z scores to transform them into Z scores. Finally, we write: $-H = \sum_j^m w_j Z_{ij}$, where w_j are the weights, however derived, for the component health dimensions. $w_j = 1$ if no particular weights are assigned to the dimensions.

It can be seen from the equation for -H that if an individual's score fell on the "ideal norms" on all the health dimensions, his deviation scores would all be zero, his corrected deviation scores would also be zero, and his z scores would be negative and the highest among all the negative z scores in his group. After transformation, however, his Z scores would be the lowest among all the z scores. The magnitudes of the other Z scores are a function of the magnitudes of the corrected deviation scores. The larger the corrected deviation scores, the larger the Z scores. This phenomenon fully justifies the rationale of -H as an index of "negative health" in relative terms.

Further examination of -H shows, assuming some degree of normality of the distributions of the Z scores, that the lowest -H value that indicates optimum health in the aggregated dimensions would be approximately $1.5 \sum_j^m w_j$, the produce of 1.5 times the sum of the weights, because practically all the Z scores would fall within 3.5 standard deviations of the means, which is 5 (the constant added to the z scores to derive Z scores). Similarly, the highest -H value would be $8.5 \sum_j^m w_j$. If no differential weights were assigned, then the range of -H would be approximately 1.5m to 8.5m.

If a measure of variability of -H is desired, its variance can be computed as follows (See, for instance, Ghiselli, 1964): $S^2{-H} = \sum_j^m w_j^2 S_j^2 + 2\sum_j^m w_i w_j S_i S_j r_{ij}$, where $i \neq j$, r_{ij} are the correlations between all possible pairs of health dimensions in the composite and the summation is taken over all m dimensions. However, since all the standard deviations of the Z scores on the m dimensions are unity (if this is not clear, it is recalled that the standard deviation of z scores is unity and that adding a constant to the scores does not affect the standard deviation), we can simplify the formula by writing:

$$S^2_{-H} = \sum_j^m w_j^2 + 2\sum_j^m w_i w_j r_{ij}.$$

If it could be safely assumed that the health dimensions are uncorrelated (which would probably be true of hearing acuity and visual acuity), then we could further simplify by writing: $S^2_{-H} = \sum_j^m w_j^2$, since the last term would vanish.

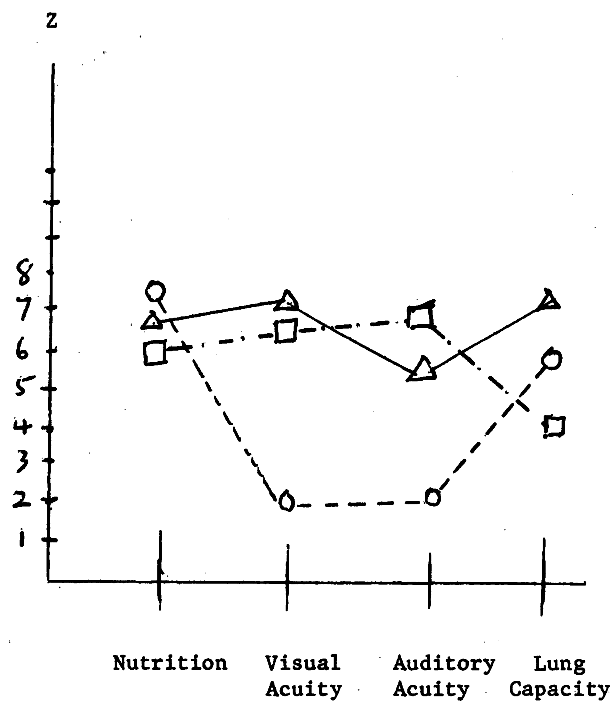
The variance of -H would be useful in computing the standard error of measurement that would provide some information as to the accuracy of -H as an individual index for a particular group. Using the analogy of a single test, the standard error of measurement for -H, which is a composite measure, would approximately be: $S_{e(-H)} \cong S_{-H}(1 - \bar{r}_{xx})^{\frac{1}{2}}$, where $\bar{r}_{xx}$ is the mean of the reliabilities of the component dimensional measures.

Note that once the Z scores for each health dimension is derived, the decision whether or not to aggregate the Z scores into a single index for each individual must hinge upon the purpose of the index. If the purpose is to have a single index as a measure of the outcome of a health program, then the single index should consist of those health dimensions on which the program is designed to have an effect. For example, if a health program had as its goals the elevation of nutrition standards and the improvement of visual acuity of the citizens of a community, the single index would be composed of two dimensions: nutrition level and visual acuity.

On the other hand, the purpose may simply be to compare a group of individuals on several health dimensions that are deemed important to the performance of a particular task or job. In that case, a profile of the Z scores of these individuals may give a better picture than a single aggregated index. A hypothetical profile of Z scores of three individuals on four health dimensions is given in Figure 1.

FIGURE 1

Hypothetical Profile of Three Individuals on Four Health Dimensions



This profile shows that Individual A is fairly high in nutrition, pretty low in visual acuity and auditory acuity, and above average in lung capacity. Individual B is below average in nutrition, average in visual acuity, slightly above average in auditory acuity, and average in lung capacity. Individual C is average in nutrition, slightly above average in visual acuity, slightly below average in auditory acuity, and above average in lung capacity.

It might be argued that in combining the health dimensions perhaps a multiplicative (non-linear) model would be more appropriate than an additive (linear) model that has been suggested. Our response to this argument would be that in our present state of knowledge there is no theoretical or empiric basis for choosing one model over the other. At any rate, if a multiplicative model were indeed found to be more appropriate, we could always perform a logarithmic transformation of the Z scores and still apply the additive model.

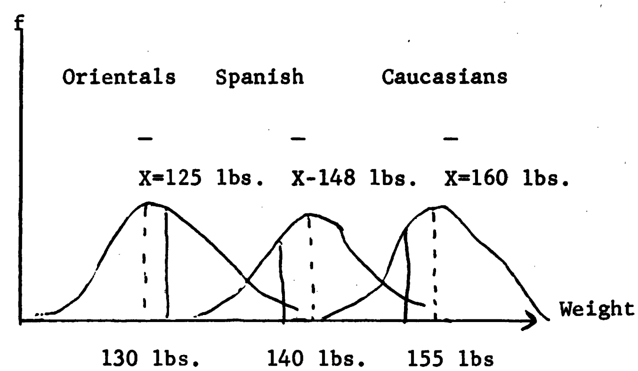
One thing that may not be obvious is whether or not aggregation of the Z scores as a single index makes the index invariant over occasions or communities. Assuming that identical instruments are used in the measurement of the health dimensions, and further assuming that the distributions of the health dimensions are normal or similar in shape across communities, then the single index does posses the property of invariance. This statement is based on the fact that when the scores which are measured on equal intervals, are derived from deviations from comparable norms, the scores take on the property of a ratio scale, with what is psychometrically known as a relative zero point (11,12).

DERIVATION OF -T INDEX

Although the rationale for developing -H for individuals is also applicable to a community or nation, the heterogeneity of the population in a community or nation does not permit a straight forward aggregation of the individual -H values into a single index. The situation may be illustrated in Figure 2.

FIGURE 2

Theoretical Representation of the Heterogeneity of the Male Adult Population in the United States with Respect to a Health Dimension: Weight



In this figure, the numbers along the abscissa are the "ideal normative weights" for the three subpopulations. The mean weights of the three subpopulations are given atop the three curves. In this theoretical example, more people in the Oriental subpopulation are underweight than overweight, more Spanish Americans are overweight than underweight, and more Caucasians are overweight than underweight. This figure is, of course, an oversimplification, because in reality the "ideal" weights within each subpopulation must be contingent upon bone structure or height. Nevertheless, it does serve to indicate the complexity of the problem of generating a single index of health for a population.

Theoretically, however, it should be feasible to derive a single index for a community or a nation by using the same rationale that is used in developing an individual index. Assuming that m dimensions of health are used and that there are k subpopulations in a community, each subpopulation being of size n_l, we could obtain the deviation scores from the "ideal norms" of the subpopulations in each dimension, correct the deviation scores by differential weighting of positive and negative deviations, transform the corrected scores into z scores and further transform the z scores into Z scores by adding the constant 5 to get rid of the negative signs.

Before we proceed to derive the index, we would like to emphasize that while the "ideal norms" for a health dimension may differ from community to community or from nation to nation, the transformation of the corrected deviation scores into z scores is performed by treating the corrected deviation scores from all the communities or nations on one dimension as one variable. As long as the individuals from each subgroup within a community or nation are properly identified by some coding scheme, a mean Z score on a given dimension is obtainable for each of the subgroups. The magnitudes of the mean scores are a function of the magnitudes of the corrected deviation scores from their "ideal norms." For purposes of simplification, we assume there are k number of subgroups within each community or nation.

We write $\bar{Z}_{..l} = \sum_j^m \sum_i^{n_l} w_j Z_{ijl}/n_l$, the mean of subgroup$_k$l for all weighted m dimensions of health, and $\sum_l n_l = N$, the total population of a community or nation. Then the index for a community or nation, which we shall designate as -T, would be:

$$-T = \sum_l^k n_l \bar{Z}_{..l}/N = \sum_l^k \left(n_l \sum_l^k Z_{..l}/n_l\right)/N = \sum_l^k Z_{..l}/N.$$

It is seen that -T is actually a weighted mean of all health dimensions for all subgroups, the weights being the sizes of the subgroups. The variance of -T is:

$$S^2_{-T} = (1/N)^2 \left(\sum_j^m w_j^2 + 2\sum_j^m w_i w_j r_{ij}\right), \text{ where } i \neq j,$$

and the summation is taken over all m dimensions in the aggregate. Where the dimensions are statistically independent, the second term drops out and we have: $S^2_{-T} = (1/N)^2 \left(\sum_j^m w_j^2\right)$, for the general case. If no special weights are assigned to the dimensions, then we have:

$S^2_{-T} = (1/N)^2 (m + 2\sum_j^m r_{ij})$. For the independent case we have simply: $S^2_{-T} = (1/N)^2 m$.

The standard error of measurement for -T would be: $S_{e(-T)} = S_{-T}(1-\bar{r}_{xx})^{\frac{1}{2}}$, where $\bar{r}_{xx}$ is the mean of the reliabilities of the measured health dimensions.

The assumption that there are k subgroups within a community in no way implies that k is a constant for all the communities to be compared with the -T index. In fact, Community A may have three subgroups and Community B five subgroups. The number of subgroups within a community is determined by the number of "ideal norms" that are required on a health dimension. For example, if the "ideal norms" for weight were the same for blacks and whites in the United States, then there would be no reason to consider blacks and whites as two subgroups on that particular dimension. However, the number of health dimensions used must be the same at all times for -T to be invariant over communities.

PROBLEMS OF ESTABLISHING NORMS

Since in most cases the mean of a measured health dimension would not be the "ideal norm," some means of establishing the norm had to be found. One approach would be to convene a large panel of medical experts and assign to it the task of recommending the "ideal norms" for age-sex specific groups on various health dimensions. This panel could also be charged with the responsibility of determining the weights of the various dimensions for the purpose of aggregating them into a single index when such an index is required. In the absence of objective data for decision making,a consensus of the opinions of these medical experts based on the largest body of available medical evidence could constitute a reasonable approximation to objective criteria.

Where there is any theoretical or empirical basis for believing that the health dimensions are related to mortality and/or morbidity, one can easily determine the points or ranges of these measured dimensions within which mortality and morbidity are at the lowest, given that the appropriate data are either available or obtainable. For purposes of aggregating the health dimensions into a single index, one can apply regression or discriminant analysis to determine the optimum weights for these dimensions that are most predictive of mortality and morbidity.

SOME CLOSING REMARKS

It has been pointed out that the decision to use an aggregated index and the kinds of health dimensions in the index must depend on the purpose which the index is to serve. In terms of mortality and morbidity data, the index composed of such data cannot appropriately serve as an outcome measure unless a health program is aimed at overall reduction of mortality and morbidity from all causes. Also, in using the index for cross-community comparisons, it may be necessary to differentiate mortality and morbidity by cause, so that the index comprises only those dimensions that are applicable to all the communities compared. For example, it would not be fair to compare New York City with a quiet rural town in Wyoming in terms of an index that is composed of mortality and morbidity data that include automobile deaths and injuries, for the simple reason that there may not be any automobile in the Wyoming town.

A look at the formula for standard error of measurement for either of the two indices reveals that it is a function of two computed statistics: S_{-H} (or S_{-T}) and $\bar{r}_{xx}$. If $\bar{r}_{xx}$ is unity; that is, if all measured dimensions have perfect reliability, then $S_{e(-H)}$ or $S_{e(-T)}$ reduces to zero, indicating no measurement error, regardless of the magnitude of S_{-H} or S_{-T}. On the other hand, if $\bar{r}_{xx}$ is zero, then $S_{e(-H)} = S_{-H}$ and $S_{e(T)} = S_{-T}$, the standard deviation of the -H or -T values. The lesson to be learned here is that one can always increase the precision of the index by increasing the reliabilities of the various measures of the component health dimensions, by reducing the variability of the group or subpopulation through judicious selection of "ideal norms," or by doing both.

A couple of caveates are now in order. First, the -H or -T index is meant to be a general index, hopefully useful for research and administrative purposes. It is not meant to be a diagnostic tool for ascertaining the degrees of threat of life from a variety of potential risk factors. For example, the fact that Individual A has a lower -H value than Individual B does not necessarily mean that Individual A has a longer age corrected life expectation than that of Individual B, for the simple reason that Individual B may have undiagnosed cancer of the lung and this information is not incorporated into the -H index. Such information would be useful for the development of indices that are ill-oriented rather than health-oriented, as the -H and -T indices are.

Another caveat is that, like most indices that have been designed, the -H or -T index does not possess inherent validity. In fact, with the exception of a few indicators in the mental health area, most health indicators that are in existence do not provide information on validity and reliability. If these indicators are to win acceptance in research quarters, they must be accompanied by such information, in the same way that standardized aptitude and achievement tests must be accompanied by such information.

APPENDIX

Algebraic Proof that the Variance of Transformed Scores Not Based on First and Second Moments of the Distribution of a Random Variable Is Less than Unity.

Let X be a random variable, $\bar{X}$ the means of this variable, $\bar{X}'$ the "ideal norm" that differs from $\bar{X}$ by d.

We write:

$x = X - \bar{X}$

$x' = X - \bar{X}'$

$d = \bar{X} - \bar{X}'$

Then $x' - x = (X - \bar{X}') - (X - \bar{X}) = \bar{X} - \bar{X}' = d$

$$x' = x + d$$

$$\Sigma x' = \Sigma x + nd = nd \qquad (1)$$

$$x'^2 = (x + d)^2 = x^2 + 2dx + d^2$$

$$\Sigma x'^2 = \Sigma x^2 + 2d\Sigma x + nd^2 = \Sigma x^2 + nd^2 \qquad (2)$$

$$S_{x'}^2 = \Sigma x'^2/n = \Sigma x^2/n + d^2 = S_x^2 + d^2 \qquad (3)$$

$$z_{x'} = x'/S_{x'}$$

$$\Sigma z_{x'} = \frac{1}{S_{x'}} \Sigma x' \qquad (4)$$

Substituting (1) into (4), we ave

$$\Sigma z_{x'} = \frac{1}{S_{x'}} (nd) \qquad (5)$$

$$\Sigma z_{x'}^2 = \frac{1}{S_{x'}^2} \Sigma x'^2 \qquad (6)$$

Substituting (2) into (6), we have

$$\Sigma z_{x'}^2 = \frac{1}{S_{x'}^2} (\Sigma x^2 + nd^2) \qquad (7)$$

$$S_{z_{x'}}^2 = \left(\frac{1}{n} \Sigma z_{x'}^2 - \frac{(\Sigma z_{x'})^2}{n}\right) \qquad (8)$$

Substituting (3), (5) and (7) into (8), we have

$$S_{z_{x'}}^2 = \frac{1}{n}\left(\frac{\Sigma x^2 + nd^2}{S_x^2 + d^2} - \frac{n^2 d^2}{n(S_x^2 + d^2)}\right)$$

$$= \frac{1}{n}\left(\frac{\Sigma x^2 + nd^2 - nd^2}{S_x^2 + d^2}\right)$$

$$= \frac{1}{n}\left(\frac{\Sigma x^2}{S_x^2 + d^2}\right)$$

If $d^2 = 0$; that is, if $\bar{X}' = \bar{X}$, then

$$S_{z_{x'}}^2 = \frac{1}{n}\frac{\Sigma x^2}{S_x^2} = \frac{S_x^2}{S_x^2} = S_{z_x}^2 = 1$$

Since for any given sample Σx^2, n, S_x^2 are all constants, $S_{z_{x'}}^2 < S_{z_{x'}}^2 = 1$ whenever $d^2 > 0$.

REFERENCES

(1) Stouman, K. and Falk, I. S., "Health indices-a study of objective indices of health in relation to environment and sanitation," Bulletin of Health of the League of Nations, 8(1939)63.

(2) Chiang, C. L., An Index of Health: Mathematical Models. PHS Publication No. 1000 - Series 2, No. 5. Washington, D.C.: U.S. Government Printing Office, 1965.

(3) Fanshell, S. and Bush, J. W., "A health index and its application to health research outcomes," Operations Research, 18(1970) 1021-1066.

(4) Thurstone, L. L., "A law of comparative judgment," Psychological Review, 34(1927) 273-286.

(5) Mosteller, F., "Remarks on the method of paired comparisons: I. The least squares solution assuming equal standard deviations and equal correlations," Psychometrika, 16(1951)3-9.

(6) Dubos, R. J., Mirage of Health, Utopias, Progress and Biological Change. New York: Harpers, 1959.

(7) Dubos, R. J. and Dubos, J., Bacterial and Mycotic Infections of Man. Philadelphia: Lippincott, 1948.

(8) Dubos, R. J. and Dubos, J., The White Plague: Tuberculosis, Man and Society, Boston: Little, Brown, 1948.

(9) World Health Organization, Measurement of Levels of Health. WHO Technical Report Series No. 137. Geneva: World Health Organization, 1957.

(10) Sullivan, D. F., Conceptual Problems in Developing an Index of Health. PHS Publication No. 1000, Series 2, No. 17, Washington, D.C.: U.S. Government Printing Office, 1966.

(11) Stevens, S. S., "On the theory of scales of measurement," Science, 103(1946)677-680.

(12) Ghiselli, E. E., Theory of Psychological Measurement. New York: McGraw-Hill, 1946.

SELECTED SOURCES OF PUBLIC TRANSFER MONEY INCOME FROM THE CURRENT POPULATION SURVEY

by

John F. Coder
Social and Economic Statistics Administration, Bureau of the Census

Introduction

In recent years, there has been rising interest in the distribution of income by income class intervals for families and unrelated individuals in the United States. Much of this interest has been directed toward the distribution and effect of public transfer payments. The purpose of this paper is to describe a source of more detailed information on four types of public transfer payments and to explain the methods that can be used to extract this information from the March Current Population Survey data file. The four types of public transfers to be discussed in this paper include (1) unemployment compensation, (2) workmen's compensation, (3) government employee pensions, and (4) veterans' payments.

Since 1947, the Bureau of the Census has collected annual money income information in the March supplement to the Current Population Survey (CPS). As the need for more accurate and more detailed income data has grown, the income supplement program to the March CPS has been revised several times to improve and expand the quality and quantity of income data collected by the Bureau of the Census. Shown in Figure 1 is the March 1971 income questionnaire. The use of the detailed "yes-no" response circles as shown in Questions 48b, 48d, and 48e was adopted beginning with the March 1969 CPS for the purpose of reducing the chance that an interviewer or respondent would overlook a particular type of income that the respondent received but would otherwise fail to report. Although, these "yes-no" response circles were first introduced to aid in the information collection process, recent inquiries have been made as to how the "yes-no" response circles could be used to obtain more income information for detailed money income sources.

Data Collection and Availability

In March of each year, the eight income questions, as shown in Figure 1, are asked of each person 14 years old and over in each CPS sample household. (Sample size for the March 1971 CPS was about 50,000 households.) The interview procedure requires the interviewer to ask each question, word for word, as printed on the questionnaire, for each eligible person in the sample household. If the person did not receive a specific type of income, the appropriate "none" or "no" circle is marked. If the person reports that he (she) did receive a specific type of income, the "yes" circle is marked and the dollar amount is recorded. For questions containing more than one set of "yes-no" response circles, (Questions 48b, 48d, and 48e), each "yes" or "no" response for each income type is recorded in sequence after which the amounts of income received are summed and recorded. This procedure allows only one amount to be recorded for each question, even though the respondent reported receiving income from two or more of the types covered in that question.

Beginning with the March 1969 CPS, the response to each set of "yes-no" circles has been carried on the tape file for each person 14 years old or over. Data for income years 1968 through 1972 can be tabulated using the "yes-no" response indicators on the March CPS file. This paper, however, is limited to a discussion of 1970 income data.

Definitions of Public Transfer Income Types in Question 48d

The Bureau of the Census definition of income is restricted to only the periodic receipt of money income.1/ Therefore, any lump sum payments, regardless of their sources are conceptually excluded. In order to insure that all procedures and concepts in collecting income data are fully understood, the CPS interviewers are retrained in March of each year prior to the March CPS operation.

Following is a description of the four types of public transfer income covered in Question 48d of the March CPS supplement questionnaire. Although the majority of the income included in these types is paid by governmental agencies, some non-governmental payments such as union strike benefits are also included.

Unemployment Compensation.--Amounts to be included as unemployment compensation are money received from governmental unemployment insurance agencies or private companies and strike benefits received from union funds. Also included are money received for transportation and/or subsistence by persons participating in the Area Redevelopment Act and the Manpower Development Training Act.

Workmen's Compensation.--Amounts to be included as workmen's compensation are money received periodically for injuries incurred at work. This money must have been paid under workmen's compensation laws required by all states or from private insurance companies where the insurance was paid by the employer, not by the employee.

Government Employee Pensions.--Amounts to be included as government employee pensions are money received from retirement pensions paid by Federal, State, county, or other government agencies to former employees, including pensions paid to retired members of the Armed Forces and their survivors.

Veterans' Payments.--Amounts to be included as veterans' payments are money payments made periodically by the Veterans Administration to disabled members of the Armed Forces or to the survivors of deceased veterans. Also included are subsistence allowances paid to veterans for education and on-the-job training.

Excluding Unusable Sample Cases

As a result of the design of the income question (i.e., grouping of the income types within one overall question using "yes-no" response circles) and method of data processing, there are a number of cases which must be excluded from the overall sample before data for individual "yes-no" responses can be tabulated.

First, the indicators of a "yes-no" response on each person's computer data record for each income question having "yes-no" response circles are "unedited." "Unedited" means that these indicators have not been checked and edited to be consistent with the reported dollar amount. Due to very tight quality control, inconsistent cases occur infrequently. However, since the amount reported for an inconsistent "yes-no" response cannot be attributed to any single "yes" response, these cases should be excluded from any analysis of income using "yes-no" response circles.

Also, nonrespondents need to be excluded from the sample. Since the CPS income processing procedures allocate2/ income to nonrespondents by each overall question and not by individual "yes-no" response circle, allocated cases must be excluded. For the March 1971 CPS, the nonresponse rate for Question 48d was about 5.6 percent. This means that 5.6 percent of all persons did not report the amount of income they received from this source.

Finally, because only a single dollar amount is carried on each person's computer record for each income question regardless of the number of "yes" responses for income types within a single question, amounts for persons with more than one "yes" response cannot be assigned to any specific income type, e.g., if "yes" responses were made to both government employee pensions and veterans' payments there is no good method of separating the amount reported for each source. Therefore, in order to examine the individual income types in Question 48d, persons reporting the receipt of more than one type of income must also be excluded. About 4.3 percent of the "fully reported" persons with income (all "yes-no" responses filled and work experience unallocated) reported receiving more than one of the income types covered in Question 48d. After eliminating persons with the above problems, the remaining universe in 1970 represents about 84 percent of all persons 14 years old and over with income from these sources.

Analysis of Data from the March 1971 CPS

Shown below in table 1 is the distribution of "yes" responses in Question 48d.

Table 1.--Question 48d, Public Transfer Income by Type of "Yes" Response for "Fully Reported" Persons Reporting a Dollar Amount: March 1971 CPS

Type of "YES" Response	Number (thousands)	Percent
Total	9,749	100.0
Only One "yes" response marked	9,334	95.7
Unemployment Compensation	3,657	37.5
Workmen's Compensation	785	8.1
Government Employee Pensions	1,335	13.7
Veterans' Payments	3,557	36.5
All "yes" response marked	-	-
Other combinations1/	415	4.3

- Represents zero

1/ The other combinations group includes combinations of any two or three "yes" responses. There are no persons receiving all four types.

The most important finding in table 1 is the large proportion of persons reporting only one "yes" response. For March 1971 CPS, about 95.7 percent reported receiving only one of the four types covered in Question 48d. This means that, as previously mentioned, only about 4.3 percent of all "fully reported" persons reported receiving more than one of the income types covered in Question 48d. If the proportion of persons receiving more than one type of income for a question designed such as Question 48d becomes too large, information is lost because the income for these persons cannot be attributed to one specific income type. This would suggest a need for either rearranging the types of income into different groupings or completely separating the types of income into individual questions.

Of the total "fully reported" persons with income in Question 48d, 37.5 percent reported receiving unemployment compensation only, 8.1 percent reported receiving workmen's compensation only, 13.7 percent reported receiving government employee pensions only, and 36.5 percent reported receiving veterans' payments only. As would be expected, no person reported receiving all four sources (all "yes" responses).

Shown in table 2 are the income distributions for the four types of public transfer income for "fully reported" persons with a single "yes" response for the March 1971 CPS (1970 income). The 1970 median incomes for these public transfers were: 1) Unemployment compensation, $449, 2) workmen's compensation, $445, 3) government employee pensions, $2,626, and 4) veterans' payments, $857.

Shown in table 3 are data on the recipiency rate and median income for each of the public transfer income type by selected persons' characteristics. These data indicate that in 1970 about 4.9 percent of all male family heads with income received some income from unemployment compensation, for which the median income amount received was $451. Also, about 5.0 percent of all male family heads with income received some income from veterans' payments, for which the median income amount received was $907.

It should be noted that, because the data presented here are estimates based on a sample, they are subject not only to sampling variability but errors of response and nonreporting. In most cases the income questions are based on the respondent's memory rather than on records. This memory factor in data derived from field surveys of income probably produces underestimates, especially of irregularly received types of income such as unemployment and workmen's compensations.

Statistical Comparisons Between Household Survey and Administrative Record Data

At present, work is being undertaken to compare statistically, public transfer income data collected from the March CPS to data obtained from administrative records of Federal and State governments which administer these programs. As these analyses are completed, they will be made available in future publications.

1/ March CPS money income represents income prior to any deductions such as income tax, Social Security, health insurance, union dues, etc.

2/ Nonrespondents are allocated an income amount for each unanswered question by assigning the income reported for that question by a respondent with similar social and economic characteristics.

Table 2.--QUESTION 48d, PUBLIC TRANSFER INCOME IN 1970 - ALL PERSONS 14 YEARS OLD AND OVER REPORTING ONLY ONE "YES" RESPONSE: MARCH 1971 CPS
(Persons 14 years old and over as of March 1971)

Type of Income	Number fully reported with income (thousands)	Percent Distribution														Median income (dollars)	Mean income (dollars)
		Total	Under $100	$100 to $299	$300 to $499	$500 to $699	$700 to $999	$1,000 to $1,499	$1,500 to $1,999	$2,000 to $2,499	$2,500 to $2,999	$3,000 to $4,999	$5,000 to $6,999	$7,000 to $9,999	$10,000 and over		
Unemployment Compensation	3,657	100.0	11.7	24.9	18.1	12.7	12.2	11.8	6.0	1.8	0.3	0.4	0.2	-	(Z)	449	620
Workmen's Compensation	785	100.0	12.0	27.9	13.7	9.0	7.8	10.2	6.7	4.2	3.1	3.9	0.6	0.8	-	445	877
Government Employee Pensions	1,335	100.0	0.4	1.5	3.8	3.7	7.3	10.4	10.5	10.0	9.5	25.0	10.9	4.5	2.4	2,626	3,166
Veterans' Payments	3,557	100.0	3.3	8.4	14.8	12.6	20.6	16.8	6.7	6.1	3.3	4.0	2.5	0.4	0.3	857	1,230

\- Represents zero Z = Less than 0.05 percent

Table 3.--PERSONS WITH PUBLIC TRANSFER INCOME IN 1970 AS A PERCENT OF ALL PERSONS WITH INCOME, AND MEDIAN PUBLIC TRANSFER INCOME, BY TYPE OF PUBLIC TRANSFER INCOME, BY SELECTED PERSON'S CHARACTERISTICS

Selected Characteristics 1/	Unemployment Compensation		Workmen's Compensation		Government Employee Pensions		Veteran's Payments	
	Persons with this source as a percent of all persons with income 2/	Median income (dollars)	Persons with this source as a percent of all persons with income 2/	Median income (dollars)	Persons with this source as a percent of all persons with income 2/	Median income (dollars)	Persons with this source as a percent of all persons with income 2/	Median income (dollars)
Family Relationship								
Male family head	4.9	451	1.3	441	1.8	3,254	5.0	907
Female family head	2.9	432	0.7	(B)	1.5	(B)	5.1	939
Other family member	5.5	414	0.7	357	0.9	1,838	2.6	640
Unrelated individual	2.8	571	0.6	633	2.7	2,150	5.6	852
Race and Sex								
White male	4.4	466	1.0	436	1.5	3,232	4.8	881
Negro and other races, male	4.4	425	1.2	(B)	0.8	(B)	3.5	1,269
White female	2.5	437	0.4	425	1.2	1,945	2.0	754
Negro and other races, female	2.5	352	(Z)	(B)	(Z)	(B)	2.0	860
Age								
14 to 24 years	2.7	375	0.4	374	(Z)	(B)	1.4	552
25 to 64 years	4.4	455	1.0	453	0.9	2,493	3.2	1,164
65 years and over	0.7	768	(Z)	(B)	4.6	2,237	4.3	891

B = Base less than 75,000 Z = Less than 0.05 percent

1/ Characteristics as of March 1971.

2/ In order to obtain the recipiency rates, the number of persons who reported only one type of income in Question 48d was raked to the total number of persons receiving income in this question assuming those persons excluded from the sample received these types of income in the same proportion as did the fully reported persons who reported receiving only one type of income.

Figure 1.--The March 1971 CPS Income Supplement Questionnaire

34. In 1970 how many weeks did ... work either full time or part time not counting work around the house? *(Include paid vacation and paid sick leave)*

(Mark weeks)

If 1-49 *(Skip to 38)*
If 50-52 *(Skip to 40)*

None ○ *(Ask 35)*

35. Even though ...did not work in 1970, did he spend any time trying to find a job?

Yes ○ *(Ask items 36-37)* No ○ *(Skip to item 37)*

36. How many different weeks was... looking for work or on layoff from a job?

1-4 wks. ○ 27-39 wks. ○
5-14 wks. ○ 40 wks. or more ○
15-26 wks. ○
(Ask item 37)

37. What was the main reason... did not work in 1970? Was he...

Ill or disabled and unable to work ○
Taking care of home or family ... ○
Going to school .... ○
Could not find work .... ○
In Armed Forces............. ○
Retired .... ○

38. You said...worked about *(entry in item 34)* **weeks in 1970. How many of the remaining** *(52 minus entry in item 34)* **weeks was... looking for work or on layoff from a job?**

(Mark weeks and ask 39)

None ○ *(Skip to 41A)*

39. Were the ____ weeks ... was looking for work (or on layoff) all in one stretch?

Yes – 1 stretch ○
No – 2 stretches ○
No – 3+stretches ○
(Skip to item 41A)

40. Did... lose any full weeks of work in 1970 because he was on layoff from a job or lost a job?

Yes ○
No ○
(Skip to 42)

41A. INTERVIEWER CHECK ITEM
Are all weeks accounted for in items 34 and 38?

Yes ○ *(Skip to 42)* No ○ *(Ask 41B)*

41B. What was... doing most of the remaining weeks in 1970? Was he ...

Ill or disabled and unable to work ...... ○
Taking care of home or family ○
Going to school .... ○
In Armed Forces .... ○
Retired .... ○
Other *(Specify)*........... ○

(Ask item 42)

42. When... was working in 1970, did he usually work full time or part time?

Full-time ○ Part-time ○

43. What was...'s longest job in 1970? *(Compare with entry in item 23)*

Same as item 23.... ○ *(Skip to item 44)*
Different from item 23 ○ *(or item 23 blank)* *(Specify below and go to item 44)*

43A. NAME OF EMPLOYER

43B. INDUSTRY

43C. OCCUPATION

43D. CLASS OF WORKER

PRIVATE ....P ○
GOV'T...........G ○
SELF-EMP.........O ○
INC. Yes ○ No ○
WITHOUT PAY... WP ○

(For persons with "P" or "G" in 43D, or "Same" in 43 and "P" or "G" in 23D)

44. Does... belong to a labor union?

Yes ○ *(Specify)* No ○

FOR OFFICE USE ONLY

INDUSTRY

Ø ○ ○ A ○
1 ○ ○ B ○
2 ○ ○ C ○
3 ○ ○ D ○
4 ○ ○ E ○
5 ○ ○ F ○
6 ○ ○ G ○
7 ○ ○ ○ H ○
8 ○ ○ ○ J ○
9 ○ ○ ○ K ○
L ○
M ○

OCCUPATION

Ø ○ ○ ○ N ○
1 ○ ○ ○ P ○
2 ○ ○ ○ Q ○
3 ○ ○ ○ R ○
4 ○ ○ ○ S ○
5 ○ ○ ○ T ○
6 ○ ○ ○ U ○
7 ○ ○ V ○
8 ○ ○ W ○
9 ○ ○ X ○
Y ○
Z ○

INCOME IN THE YEAR 1970 *(Write actual amount in the space provided in items 45-48c and mark the appropriate dotted numbers.)*

LAST YEAR (1970) How much did... receive:

45. In wages or salary before any deductions?

CPS-581

$

None

46. In net income from his own business or professional practice or partnership?

CPS-581 MARK NET INCOME

Gross income $
minus
Business $
equals
Net income $

None
Lost Money

47. In net income from his own farm?

CPS-581 MARK NET INCOME

Gross income $
minus
Business expense $
equals
Net income $

None
Lost Money

48. During 1970, did... receive any money from:

48a. –Social Security or Railroad Retirement checks from the U.S. Government?

Yes ○ No ○

How much altogether? $

CPS-581

48b. –Estates, trusts or dividends?

Yes ○ No ○

–Interest on savings accounts or bonds?

Yes ○ No ○

–Net rental income?

Yes ○ No ○

How much altogether? $

CPS-581

Lost money

48c. –Welfare payments or other public assistance *(aid to families with dependent children, old age assistance, or aid to the blind or totally disabled)?*

Yes ○ No ○

How much altogether? $

CPS-581

48d. –Unemployment compensation?

Yes ○ No ○

–Workmen's compensation?

Yes ○ No ○

–Government employee pensions?

Yes ○ No ○

–Veteran's payments?

Yes ○ No ○

How much altogether? $

CPS-581

48e. –Private pensions or annuities?

Yes ○ No ○

–Alimony?

Yes ○ No ○

–Regular contributions from persons not living in this household?

Yes ○ No ○

–Anything else?

Yes ○ No ○

How much altogether? $

CPS-581

TWO SAMPLE PROBLEMS
FOR A DICHOTOMOUS VARIABLE WITH MISSING DATA

Janet Dixon Elashoff, Stanford University
amd Robert M. Elashoff, University of California, San Francisco

Introduction. Incomplete or missing data is a major problem in many fields. Data may be incomplete due to subject nonresponse or refusal to cooperate, transcription errors, random loss, or to a variety of other reasons. As a consequence, statistical techniques to deal with incomplete data are necessary. One common approach is simply to delete and ignore the incomplete cases. To select an appropriate technique, however, something must be known about the kind of observations which are missing and the variables influencing the loss of certain observations.

Two sample problems with binary data are considered in this study. These problems are formulated in section 2 along with specific probability models to describe the missing observations. In later sections, we describe and explain the difficulties encountered in parameter estimation for the sampled populations, discuss estimation and tests for differences and ratios of the parameters, and give recommendations for data analysis.

Problem formulation, Models, Notation. Many studies to compare the effectiveness of different treatments have nonrespondents. For example, suppose patients with a certain disease are assigned either an active drug ($x = 2$) or a placebo ($x = 1$) in a double blind study. The placebo has the same side effects as the active drug, but presumably it does not have the same curative or palliative effect as the active drug. A follow-up study is made and each patient is scored as improved ($y = 1$) or unimproved ($y = 0$) on a response variable y. Lack of improvement may cause some patients to drop out of the study or refuse to cooperate further. Improvement also may give patients a reason to drop out or a chance to leave the area. In either case the y measurements are unknown. Clearly, under these circumstances, a missing y may be influenced by whether or not the patient is improved but not directly by the drug the patient received. The goal of the clinical study is to compare the probability of improvement for the active drug (p_2) with the probability of improvement for the placebo (p_1).

A statistical model for such problems is defined in this way. A random sample of N_i, $i = 1, 2$, individuals is given treatment $x = i$, and n_i of these individuals are observed on the binary response variable y. No measurement error exists. Denote by p_i the probability that $y = 1$ when $x = i$. The value of x is known for each of the $(N_1 + N_2)$ individuals. Define

$q(x,y) = P(\text{an individual's } y \text{ is recorded} | x,y)$.

We can distinguish four particular specifications of $q(x,y)$:

(1) $q(x,y) = q$,
the missing data occur at random;

(2) $q(x,y) = q_y$,
the probability of missing data depends on an individual's y value but not his x value;

(3) $q(x,y) = q_x$,
the probability that an individual's score is recorded depends only on the population he is from; or

(4) $q(x,y) = q_{xy}$,
the general case where the probability of missing data depends on both x and y.

The second specification, case 2, is more appropriate for problems like the drug example described above. We investigate the question of how much it matters whether we base our statistical analysis on case 1 or case 2.

Our goals are to estimate the population quantities p_1 and p_2, to estimate the comparative population measures

1. $D = p_1 - p_2$,
2. $R = p_1/p_2$,
3. $OR = \dfrac{p_1/(1-p_1)}{p_2/(1-p_2)}$

and to propose significance tests and confidence intervals for these three population measures.

Under cases 1 and 3, standard estimators and inference procedures for p_1, p_2, D, R, and OR may be used conditional on the observed sample sizes, n_1 and n_2. Under case 4, there is insufficient information to estimate p_1 and p_2 or the functions D, R, or OR since only 4 of the 6 parameters can be estimated from the data. Our study investigates the maximum likelihood estimators for the parameters p_i, D, R, and OR under case 2, examines asymptotic and small sample results for conditional and unconditional means, variances, and mean squared errors of the estimators, and compares their behavior to that of case 1 estimators.

Estimation of the p_i. Under case 1, when $q(x,y) = q$, and the missing observations occur at random, the $(N_i - n_i)$ missing observations can be ignored and the remaining observations regarded as random samples of size n_i. Let r_i be the number of individuals for whom $y = 1$ out of the n_i actually observed in population i. Then the ml estimators of p_1, p_2, and q are

(5) $\hat{p}_{1i} = r_i/n_i$

(6) $\hat{q} = (n_1 + n_2)/(N_1 + N_2)$.

Standard distribution theory applies to these estimators.

When $q(x,y) = q_y$ and case 2 holds, the ml estimators are derived by the following argument. Let

$\alpha_1 = q_1p_1$, $\alpha_2 = q_1p_2$;

$\beta_1 = q_0(1-p_1)$, $\beta_2 = q_0(1-p_2)$.

Then, the likelihood of the two samples becomes

(8) $L = C \prod_{i=1}^{2} \alpha_i^{r_i} \beta_i^{(n_i-r_i)} [1-\alpha_i-\beta_i]^{(N_i-n_i)}$

After differentiating L with respect to the α_i and β_i and setting the resulting equations to zero, we find

(9) $\hat{\alpha}_i = r_i/N_i,\ \hat{\beta}_i = (n_i-r_i)/N_i,\ i = 1, 2.$

Since the parameter set $(\alpha_1, \alpha_2, \beta_1, \beta_2)$ is a one-to-one transformation of (p_1, p_2, q_1, q_0) when $p_1 \neq p_2$ and neither q_1 nor q_0 equal to zero, the ml estimators of p_i, q_0, q_1 can be obtained from equations (7) and (9) as

$$\hat{p}_{2i} = (r_i/N_i)[N_2(n_1-r_1)-N_1(n_2-r_2)](n_1r_2-r_1n_2)$$

$$(10)\quad \hat{q}_0 = (n_1r_2-n_2r_1)/(N_1r_2-N_2r_1)$$

$$\hat{q}_1 = (n_1r_2-n_2r_1)/(N_2(n_1-r_1)-N_1(n_2-r_2)).$$

We note that these estimators of p_1, p_2, q_1, q_0 break down when $p_1 = p_2$. Clearly, when $p_1=p_2=p$, we have only a single sample from which it is impossible to obtain even a consistent estimator of p under case 2. Mathematically, when $p_1=p_2=p$, then $\alpha_1 = \alpha_2 = \alpha$ and $\beta_1 = \beta_2 = \beta$ and although ml estimators of α and β, or q_1p and $q_0(1-p)$, exist the parameter set (α,β) is not a one-to-one transformation of (p, q_0, q_1) and ml estimators for p, q_1, q_0 do not exist.

In practice, additional difficulties arise with the use of this ml estimation method even when $p_1 \neq p_2$. Samples in which $(\hat{\alpha}_1/\hat{\alpha}_2) < 1$ and $(\hat{\beta}_1/\hat{\beta}_2) < 1$ or in which both ratios are greater than one lead to some of the $(\hat{p}_1, \hat{p}_2, \hat{q}_0, \hat{q}_1)$ being negative. This situation not infrequently occurs when p_1 is close to p_2 or the n_i are small. Such a difficulty implies that the ml estimators are not very precisely determined. If we constrained the $\hat{\alpha}_i$, $\hat{\beta}_i$ so that the preceding inequalities would not occur, we would essentially be setting $p_1 = p_2$ in which case p cannot be estimated.

Under case 2 where $q(x,y) = q_y$ both $\hat{p}_{1i}$ and $\hat{p}_{2i}$ have asymptotic normal distributions for $p_1 \neq p_2$. The mean of $\hat{p}_{1i}$ is θ_i which equals $p_iq_1/(p_iq_1 + (1-p_i)q_0)$ and the asymptotic mean of $\hat{p}_{2i}$ is p_i. The asymptotic conditional and unconditional variances of $\hat{p}_{1i}$ and $\hat{p}_{2i}$ are given in Elashoff and Elashoff (1971). The asymptotic variance formulas for $\hat{p}_{2i}$ contain the term $(p_1-p_2)^2$ in the denominator indicating that for $|p_1-p_2|$ small the variance of $\hat{p}_{2i}$ will be large.

For tests of the null hypothesis $H_0 : p_1 = p_2$ under cases 1 or 2 $\theta_1 = \theta_2$ if and only if $p_1 = p_2$ (for neither $q_0 = 0$ or $q_1 = 0$), and thus a test of $p_1 = p_2$ may be carried out by standard methods such as the Fisher-Irwin test conditional on n_1, n_2, and $r_1 + r_2$.

<u>Estimation of D, R, and OR</u>. The population measures D, R, and OR, or log OR are frequently used quantities for comparing p_1 and p_2. In this section we discuss their estimation when $q(x,y)=q_y$. Although inferences about D, R, or OR will usually be based on conditional variance formulas, a detailed numerical study of asymptotic formulas and small sample behavior conditional on possible n_1 and n_2 pairs is unwieldy and consequently discussions will focus on unconditional results. To evaluate the usefulness of asymptotic formulas for comparisons and to examine the behavior of the estimators in small samples, exact unconditional means and variances were calculated. For example, the exact mean of $\hat{D}_1$ is

$$(11)\quad \sum_{n_2=1}^{N_2} \sum_{n_1=1}^{N_1} \sum_{r_2} \sum_{r_1} \hat{D}_1 \frac{p(r_1)p(r_2)p(n_1)p(n_2)}{1 - P\{n_1=0 \text{ or } n_2=0\}}$$

where

$$p(r_i) = \binom{n_i}{r_i} \theta_i^{r_i}(1-\theta_i)^{(n_i-r_i)} \text{ and } p(n_i) =$$

$$\binom{N_i}{n_i} [p_iq_1+(1-p_i)q_0]^{n_i}[1-p_iq_1-(1-p_i)q_0]^{(N_i-n_i)}$$

Note that results were obtained conditional on $n_1 \neq 0$, and $n_2 \neq 0$ and that for $n_1r_2 = n_2r_1$ we defined $\hat{D}_2 = 0$. Calculations were made for $N_1=N_2=$ 20, 50, for p_1, p_2 = .10, .25, .50, .75, .90 and q_1,q_0 = .50, .75, .90, 1.0. Summary results may be found in Elashoff and Elashoff (1971).

Estimators for D are obtained by substitution of $\hat{p}_{1i}$ or $\hat{p}_{2i}$ in $D = (p_1-p_2)$ and are given by

$$(12)\quad \hat{D}_1 = \left(\frac{r_1}{n_1} - \frac{r_2}{n_2}\right)$$

for case 1, and

$$(13)\quad \hat{D}_2 = \left(\frac{r_1}{N_1} - \frac{r_2}{N_2}\right)\frac{N_2(n_1 - r_1) - N_1(n_2 - r_2)}{(n_1r_2 - n_2r_1)}$$

for case 2, when $p_1 \neq p_2$.

Both $\hat{D}_1$ and $\hat{D}_2$ have asymptotic normal distributions under case 2. Asymptotic conditional and unconditional means and variances are given in Elashoff and Elashoff (1971).

The estimator $\hat{D}_2$ does not exist when $p_1 = p_2$; the presence of the term $(\theta_2-\theta_1)^4$ in the denominator of the conditional variance of $\hat{D}_2$ demonstrates that $\hat{D}_2$ will have a large variance when p_1 is near p_2. $\hat{D}_2$ is a consistent estimator of D while $\hat{D}_1$ is not consistent unless $q_1 = q_0$ or $p_1 = p_2$; the bias in $\hat{D}_1$ is $(\theta_1-\theta_2) - (p_1-p_2)$ independent of N.

Examination of unconditional asymptotic and small sample results indicate that neither $\hat{D}_1$ nor $\hat{D}_2$ provides a good estimate of D in general. Unless $q_1 = q_0$ or $p_1 = p_2$, $\hat{D}_1$ may have considerable bias, and unless $N|p_1-p_2|$ is large $\hat{D}_2$ has a relatively large variance.

In small samples, $\hat{D}_2$ is biased; both the bias and the variance of $\hat{D}_2$ decrease as N increases, so for sufficiently large N, $mse(\hat{D}_2) < mse(\hat{D}_1)$ unless $p_1 = p_2$ or $q_1 = q_0$. We note however that for $N_1 = N_2 = 50$, $mse(\hat{D}_2) < mse(\hat{D}_1)$ only for $|p_1-p_2| > .4$ and $|q_1-q_0|$ large.

Let us now consider the estimation of R. Maximum likelihood estimators for $R = p_1/p_2$ are

(14) $\hat{R}_1 = \dfrac{r_1 n_2}{r_2 n_1}$

and

(15) $\hat{R}_2 = \dfrac{r_1 N_2}{r_2 N_1}$

for cases 1 and 2, respectively. Under case 2, R_1 and R_2 have asymptotic normal distributions with means and conditional and unconditional variances as given in Elashoff and Elashoff (1971).

The estimator $\hat{R}_1$ is consistent if and only if $q_1 = q_0$ or $p_1 = p_2$, otherwise the bias is

(16) bias $(\hat{R}_1) = (q_1-q_0)(p_2-p_1)\dfrac{\theta_1}{p_2 p_1}$;

$\hat{R}_2$ is consistent. Note that the ratio of the asymptotic conditional variances, var $\hat{R}_1$/var $\hat{R}_2$, equals

$$\left(\frac{\tau_2}{\tau_1}\right) = \left(\frac{n_2 N_1}{N_2 n_1}\right)^2$$

which should approach

$$\left(\frac{p_2 q_1 + (1-p_2)q_0}{p_1 q_1 + (1-p_1)q_0}\right)^2$$

for large N. Thus var $\hat{R}_1$/var $\hat{R}_2 < 1$ for $(q_1 - q_0)(p_2 - p_1) < 0$ or $(q_1 - q_0)(1 - R) < 0$. The ratio of asymptotic unconditional variances varies with $(q_1 - q_0)(p_2 - p_1)$ in a similar way but is generally smaller than the ratio of conditional variances.

Asymptotic unconditional formulas for the mean squared errors of $\hat{R}_1$ and $\hat{R}_2$ were compared for N = 200 for the parameter sets defined earlier. Except for cases where $q_1 = q_0$ or $p_1 = p_2$, when $\hat{R}_1$ is unbiased, $mse(\hat{R}_1)/mse(\hat{R}_2)$ was generally greater than .84 and frequently greater than 1.0, which suggests that use of $\hat{R}_2$ will prove generally satisfactory in large samples.

Both $\hat{R}_1$ and $\hat{R}_2$ are biased in small samples. The bias in $\hat{R}_2$ is independent of p_1 and decreases slowly with increasing N and increasing q_1 (it is almost unaffected by q_0). The range of percentage bias in $\hat{R}_1$ is similar to that of $\hat{R}_2$ when $q_1 = q_0$ but generally larger when $q_1 \neq q_0$. For investigators interested in using $\hat{R}_2$, the estimator can be corrected for bias using standard methods.

Comparisons of exact mean squared errors for $\hat{R}_1$ and $\hat{R}_2$ when N = 20 and N = 50 demonstrate that asymptotic formulas provide good indications of the size of $mse(\hat{R}_1)/mse(\hat{R}_2)$ in small samples. The ratios of exact to asymptotic unconditional variances are quite similar for $\hat{R}_1$ and $\hat{R}_2$. The exact variances are generally larger than the asymptotic variance formulas for both $\hat{R}_1$ and $\hat{R}_2$ except for $p_1 = p_2$ and N = 20; for N = 50, the ratios vary from 1.0 to 3.7, being close to 1.0 for R < 1.0 and larger for R > 1.0.

On the whole then, $\hat{R}_2$ should provide a reasonable estimator of R for N not too small.

The estimators of OR for case 1 and case 2 both reduce to

(17) $\hat{OR} = \dfrac{r_1(n_2 - r_2)}{r_2(n_1 - r_1)}$.

This estimator is asymptotically unbiased under both cases. The asymptotic conditional and unconditional variances of $\sqrt{N_1 + N_2}$ OR under case 2 are given in Elashoff and Elashoff (1971).

The independence of the form of the estimator from q(x,y) suggests that the use of $\hat{OR}$ will be robust to q(x,y). Although asymptotically unbiased, $\hat{OR}$ may have a substantial bias for $N_1 = N_2 = 20$. Generally the bias is of the order of 20% to 50% of OR, although it does not contribute appreciably to the mean square error. The behavior of $\hat{OR}$ in small samples does not seem to depend particularly on $|p_1 - p_2|$ or $|q_1 - q_0|$. For $N_1 = N_2 = 20$, the exact variance may be from 2 to 5 times larger than the asymptotic variance for the parameter sets investigated.

To estimate OR, $\hat{OR}$ (or a modification to reduce bias) can be used for either case 1 or 2. Uniformly most-accurate confidence intervals can be constructed for OR using the noncentral distribution of r_1, r_2 conditional on $(r_1 + r_2)$, N_1, N_2 (see Lehmann, 1959). This noncentral distribution is the same for both cases.

Some authors prefer log OR to OR. Of course the estimator of log OR has the same property of invariance under cases 1, 2, and 3 as does the estimator of OR. Haldane (1955), Anscombe (1956) and Gart and Zwiefel (1967) have recommended the substitution of $\hat{p}_i + \frac{1}{2n_i}$ for $\hat{p}_i$ and $(1-\hat{p}_i) + \frac{1}{2n_i}$ for $(1-\hat{p}_i)$ to reduce the bias of the estimator of the logit. This would result in the estimator

(18) $\log \hat{OR} = \log \dfrac{(r_1 + 1/2)(n_2 - r_2 + 1/2)}{(r_2 + 1/2)(n_1 - r_1 + 1/2)}$

Uniformly most accurate confidence intervals for log OR could be constructed in the same way as for OR.

Conclusions. We have studied two-sample problems with dichotomous data in which the probability that an individual score will be missing depends on the value of that score.

Estimation of the p_i and q_1, q_0 and estimation of $D = p_1 - p_2$ break down unless $N|p_1-p_2|$ is large. The case 2 estimator of R does not perform especially well for $N|p_1 - p_2|$ small. This suggests preceding attempts to estimate D or R by a test of $H_0 : p_1 = p_2$, since conditional upon n_1, n_2, and $r_1 + r_2$ such a test may be carried out by standard methods even when case 2 holds. Alternatively, since the estimator of OR is the same under cases 1 and 2, estimation of OR or log OR rather than of D or R should be considered when more than a small fraction of the data is missing.

Acknowledgements. The authors would like to thank B. L. Welch for his helpful comments.

REFERENCES

Anscombe, F. J. (1956). "On estimating binomial response relations." Biometrika, 43, 461-464.

Elashoff, J. D. and Elashoff, R. M. (1971). "Missing data problems for two samples on a dichotomous variable." Research and Development Memorandum No. 73, Stanford Center for Research and Development in Teaching, Stanford University, Stanford, Calif.

Gart, J. J. and Zweifel, J. R. (1967). "On the bias of various estimators of the logit and its variance with application to quantal bioassay." Biometrika, 54, 181-187.

Haldane, J. B. S. (1955). "The estimation and significance of the logarithm of a ratio of frequencies." Annals of Human Genetics, 20, 309-311.

Lehmann, E. L. (1959). Testing statistical hypotheses. New York: John Wiley and Sons.

A THEORETICAL MODEL FOR PREDICTING INTERNAL MIGRATION IN THE UNITED STATES

M. E. El Attar, Mississippi State University
J. D. Tarver, The University of Georgia

Interest in internal migration in the United States can be traced to the state-of-birth data first collected in the population census of 1850 (Lee and Lee, 1960). As a sociological phenomenon, however, internal migration in the United States has been seriously studied only since the second quarter of the twentieth century (Thomas, 1938). Previous studies of internal migration may be characterized as follows:

1. In most cases, the relation between selected variables and migration was a segmental one, in the sense that it was impossible to infer the relation between migration and the major factors in a changing economy.
2. The geographic units of analysis varied from study to study, making it difficult to generalize about migration from different areal units.

The primary aim of the present paper is two-fold: First, to develop a theoretical model via the formulation of a systematic conceptual scheme that embraces those variables to which internal migration is most highly related; and second, to apply this theoretical model to areal units in which commuting is reduced to a minimum, economic heterogeneity is increased to a maximum, and labor markets tend to be considerably smaller than their corresponding areas. These areas were delineated in 1969 by the Office of Business Economics and termed OBE Economic Areas (OBEA's) of the United States (Office of Business Economics, 1967).

Construction of the Theoretical Model

According to the present theoretical model, the process of migration is conceptualized to be a result of one or any combination of the following three major factors: First, net migration produced by the continuous development and change in automation and technology in the United States; second, net migration generated by attractive or repellent characteristics; third, net migration produced by compulsory agencies. In other words, the model will use as a basic framework the concepts of mechanization, automation, and technological change; attraction and repellency; and compulsion. For a definition of these concepts, see Jaffe and Froomkin (1968); Bates (1969); Dunlop (1962); Rezler (1969); Lee (1966); Tarver and others (1967); Blanco (1962); Gossman and others (1967); Zipf (1946); Stouffer (1940, 1960); Isard (1960); and Ravenstein (1885, 1889).

Concerning the interrelationship of these concepts, one may assume that migration is stimulated by those occupations and industries most affected by technological developments and changes. This assumption has been substantiated by several studies in which professional, technical, and kindred workers showed the highest geographic mobility, especially for long distance migration, and the farmers for short distance (Lively and Taeuber, 1939; Tarver, 1964b; Beshers and Nishiura, 1961; Miller, 1966, 1967; Ladinsky, 1967 a and b).

In this respect, one can assume that the higher the technological developments and changes in certain occupations or industries relative to others, the greater is the geographic mobility of persons engaged in them. Consequently, areas with differentiated occupational and industrial structures are expected to yield different selective migration patterns. Phrased in this general statement, migration in the United States is viewed as the result of two polar types of decisions by migrants--voluntary impellent decisions and involuntary impellent decisions.

The voluntary impellent decisions of the workers to migrate stem from their response to the attractive incentives they expect from other labor market areas to maximize their earnings, or the desire to improve their chances of finding a job. For the purpose of this study, then, it seems justifiable to represent the attraction or repellency of an area by those factors which are believed to carry great weight among the other factors that impel the person to migrate. Figure 1 shows four of the relevant factors. These are change in civilian employment, estimated underemployment, change in unemployment, and change in real median income of families and unrelated individuals. The dashed line originating from unemployment implies that the relationship between unemployment and migration is not simultaneous, in the sense that there exists a time lag between the unemployment of a person and his migration. A study of labor mobility in Great Britain based on unemployment data for 1923-36 showed that the time lag between unemployment and migration ranged between a half year and a year and a half (Makower, <u>et al</u>., 1939).

In the involuntary-impellent decision to migrate, the individual encounters factors which compel him in the sense that he is ejected by them rather than being rejected by them. For the purpose of this study, only one factor was selected to represent this type of migration; namely, movements among armed forces personnel.

Based on the above conceptual framework as schematically outlined, the process of migration is conceived in this study as an epiphenomenal behavior produced by socio-cultural compulsions whose influences on individuals vary according to the social division of labor in which they are involved and the incentives to move, which vary in time and space.

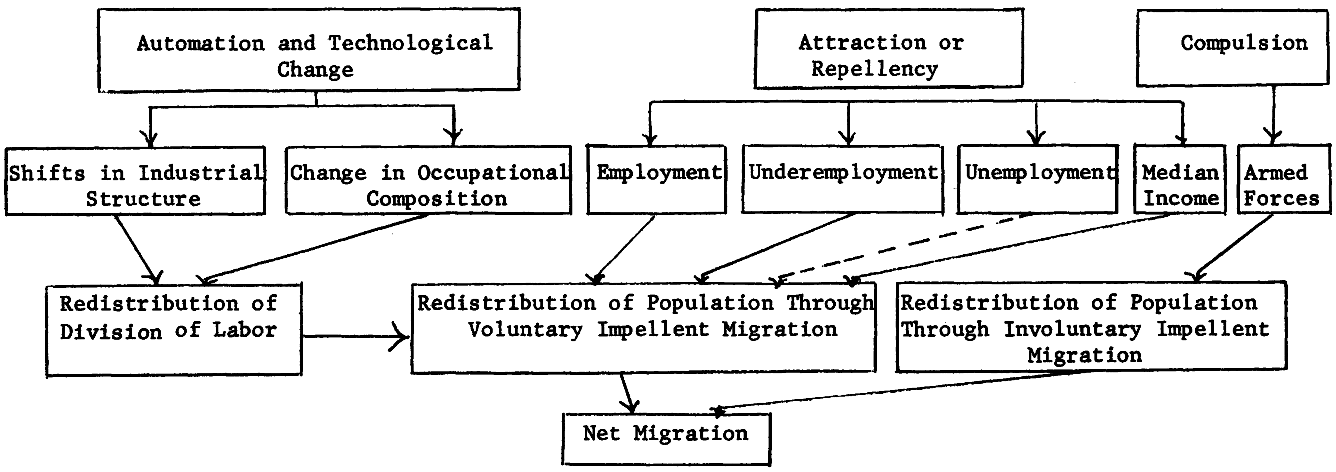
The Hypotheses

In view of the above conceptualization, one can state the hypotheses of this study as follows:

<u>Hypothesis 1</u>: Net migration is functionally related to shifts within occupational and industrial structures. Implicit in this hypothesis is the assumption that the changes within these structures contribute differentially to migration.

<u>Hypothesis 2</u>: There is a relationship between changes in civilian employment and net migration.

Figure 1. Schema of the Theoretical Model for Predicting Internal Migration in the United States



a. If the change in number of employed civilians is used, the relationship is direct.
b. If the change in number of unemployed persons is used, the relationship is inverse.

Hypothesis 3: There is a direct relationship between change in real median income of families and unrelated individuals and net migration.

Hypothesis 4: There is an inverse relationship between underemployment and net migration.

Hypothesis 5: There is a direct relationship between change in armed forces personnel and net migration.

Universe of Analysis and the Data

For purpose of illustration, an aggregation of eight OBEA's are selected. These nodal economic areas constitute the whole state of Georgia and portions of the contiguous states of Alabama, Florida, South Carolina, and Tennessee. The observational unit in this study is the county. There are 201 counties in the selected eight areas. For application on the national level, however, a whole area may be taken as a single observation (Trott, 1971). The data were compiled from Bowles and Tarver (1965); U. S. Bureau of the Census (1952, 1961, 1962); Ashby (1965); Kampe and Lindamood (1969). The underemployment estimates are based on "man-years economically unutilized labor"; and the income purchasing power was adjusted by using the reciprocals of the U. S. consumer price indices for 1949 and 1959, respectively.

The Variables and the Statistical Model

The variables among which the various hypothesized relationships are to be measured and analyzed are symbolized as follows:

Dependent Variable:

m_i = estimate of the net migration of persons 25 to 64 years of age, 1950-1960, of the ith county; i = 1,..., 201.

Independent Variables:

0_{ij} = occupational change of the jth occupational group in the ith county, 1950-1960; j = 1,...,11.

I_{ik} = industrial change of the kth industrial group in the ith county, 1950-1960; k = 1,...,14.

E_i = change in total civilian employment of the ith county, 1950-1960.

U_i = change in total unemployment of the ith county, 1950-1960.

R_i = real change in median income for families and unrelated individuals in the ith county, 1949-1959.

D_i = underemployment estimate in the ith county, 1960.

A_i = change in the number of armed forces in the ith county, 1950-1960.

The relation between net migration and the independent variables is assumed to be a functional linear relationship. The regression model specifying the relation between migration and occupational change may be stated as follows:

$$m_i = B_o + B_j\, 0_{ij} + e_i$$

where m_i and 0_{ij} are defined above.

The stepwise regression method is a suitable statistical technique for this research, since it shows those independent variables which account for the most variation in the dependent variable by order of entry of each independent variable. The computations were performed by means of the BMD computer programs (Dixon, 1968).

Analysis of Findings

The findings of the study and the tests of the hypotheses are provided in three major parts as follows:

Occupational and Industrial Changes

Here one is concerned with testing the first hypothesis. Thus, in the case of occupational change, the null hypothesis to be tested is that

$$H_o\colon B_1 = B_2 = \ldots = B_{11},\ \text{and}\ B_j = 0 (j=1,\ldots 11).$$

This is equivalent to saying that there is no relationship between the net migration of persons 25-64 years of age, and changes in the occupational structure of employed civilians during 1950-1960. The same conceptualization is applicable to industrial changes by analogy.

Occupational change.--Table 1 presents the

TABLE 1. CUMULATIVE R^2, REGRESSION COEFFICIENTS AND THEIR STANDARD ERRORS, AND PARTIAL AND SIMPLE CORRELATIONS BETWEEN NET MIGRATION OF PERSONS 25-64 YEARS OLD AND OCCUPATIONAL CHANGES IN EIGHT SELECTED OBEA's OF THE UNITED STATES, 1950-1960

Occupation and Intercept	Coefficients of				
	Cumulative R^2	Regression and Intercept	Standard Error	Correlation Partial	Correlation Simple
Sales	0.778	6.152	.549	.632	.882
Craftsmen	0.878	4.565	.477	.571	.872
Farmers	0.924	2.128	.228	.562	.199
Professional	0.935	-3.657	.410	-.544	.698
Not Reported	0.954	1.051	.286	.258	.398
Managers	0.956	1.327	.803	.120*	.847
Farm Laborers	0.957	0.666	.365	.131*	.210
Serv. Workers	0.957	0.432	.464	.068*	.551
Laborers	0.957	0.553	.581	.069*	.212
Clerical	0.957	0.290	.476	.044*	.843
Operatives	0.957	0.114	.267	.031*	.722
Intercept		-472.415			

*Not significant at .05 level.

cumulative R^2, the coefficients of regression, the standard errors, and the coefficients of partial and simple correlations, for the relation between the net migration of persons 25-64 years of age and occupational changes in eight selected OBEA's of the United States, 1950-1960.

Differentiation in the effect of occupational changes on migration is reflected in the patterns of the arrangements of the occupational categories and in the magnitude and signs of the regression coefficients. The direction of the relationship is given by the signs of the partial correlation coefficients. In all, the first hypothesis is confirmed. For instance, the "sales" category was the first in importance since it accounted for 78 percent of the variation in net migration. Other occupational categories of major importance were craftsmen, farmers, and professional workers. Change in the employment of these three occupational categories accounted for 16 percent of the variation of net migration. The four occupational categories taken together accounted for 94 percent of the variation of net migration. This level of R^2 is high, which indicates the success of the theoretical model, especially when we state that this high value of R^2 is not affected by either "saturation" or "multicollinearity," as detected from the covariance matrix and standard errors of estimated regression coefficients.

All regression coefficients are significantly different from zero at the five percent level. Values and signs of the regression coefficients show differentiated relations between shifts in occupational structure and migration. The regression and partial correlation coefficients of the occupational category of professional, technical, and kindred workers are negative. There are two possible explanations for this phenomenon.

According to the first explanation, the reversed relation of the professional category from a direct relationship in the zero order correlation to inverse relationship in the partial correlation (and consequently a negative regression coefficient), implies that the professional category disturbs rather than enhances the prediction of migration (Table 1).

The second explanation which, in fact, clarifies the first, is that professional workers have a much higher propensity to migrate than other workers (Tarver, 1964; Saben, 1964; Miller, 1967; U. S. Department of Labor, 1965). Professional, technical, and kindred workers are very heterogeneous group, not only in terms of detailed occupational categories, but also with regard to age, sex, color, education, marital status, and family size. Moreover, the heterogeneity of this group means that the substitution among its members is inelastic.

Industrial change.--The null hypothesis to be tested is as follows:

H_0: $B_1 = B_2 = \ldots = B_{14}$, and $B_k = 0$ $(k=1,\ldots,14)$.

The multiple regression model used in estimating the regression coefficients and testing the hypothesis is as follows:

$$m_i = B_0 + B_k I_{ik} + e_i.$$

Table 2 presents the cumulative R^2, the regression coefficients, the standard errors of the estimates, and the coefficients of partial and simple correlations, for the relation between net migration of persons 25-64 years of age and industrial changes in the eight selected OBEA's of the United States, 1950-1960. As was the case with occupational change, the industrial categories were arranged according to the importance of the variable in explaining the variation in net migration.

The relative importance of the industrial categories and the magnitude and signs of the regression coefficients support our hypothesis. As an example, the trade category was the first variable of importance since it accounted for 85 percent of the variation in net migration. Other industrial categories of major importance were agriculture, construction, professional services,

TABLE 2. CUMULATIVE R^2, REGRESSION COEFFICIENTS AND THEIR STANDARD ERRORS, AND PARTIAL AND SIMPLE CORRELATIONS BETWEEN NET MIGRATION OF PERSONS 25-64 YEARS OLD AND INDUSTRIAL CHANGES IN EIGHT SELECTED OBEA's OF THE UNITED STATES, 1950-1960

Industries and Intercept	Coefficients of: Cumulative R^2	Regression and Intercept	Standard Errors	Correlation: Partial	Correlation: Simple
Trade	.845	1.904	.390	.338	.919
Agriculture	.876	1.590	.115	.713	.263
Construction	.916	2.495	.447	.378	.826
Professional Services	.942	-2.123	.195	-.623	.591
Manufacturing	.947	.944	.135	.456	.754
Public Administration	.952	1.100	.212	.355	.520
Finance, etc.	.957	3.642	.676	.367	.770
Services	.959	1.560	.532	.210	.655
Transportation, etc.	.960	1.587	.447	.237	.675
Business Services	.962	-2.774	1.353	-.149	.667
Entertainment Services	.963	4.108	2.890	.103*	.735
Armed Forces	.963	.103	.108	.070*	.418
Mining	.963	.827	1.117	.054*	.059*
Forestry	.963	-.299	1.055	-.021*	-.015*
Intercept		-574.189			

*Not significant at .05 level.

and manufacturing. Employment changes in these four industrial categories accounted for 10 percent of the variation of net migration. Taken as a unit, industrial changes of the five categories accounted for 95 percent of the variation of net migration, a one percent improvement over the occupational categories. The partial correlation coefficient of professional services, not only reversed the sign of the total correlation coefficient from positive to negative but also uncovered the real strength of the inverse relationship. This negative direction is also apparent in the sign of regression coefficient of this industrial category. This negative relationship is consistent with the relation obtained from the occupational category of professional, technical, and kindred workers. It is also consistent with the percentage change of employment in this industrial category (62.3 percent) and with the inference of migration selectivity.

Other Selected Variables and Migration

In this section one is concerned with a set of five variables which influence the decision of an individual to migrate. These five variables are (1) change in total civilian employment,1950-1960; (2) change in total unemployment, 1950-1960; (3) estimated underemployment, 1960; (4) change in real median income of families and unrelated individuals, 1949-1959; and (5) change in armed forces personnel. Computation of results was achieved by the following multiple regression model.

$m_i = B_0 + a_iE_i - b_iD_i + c_iA_i - f_iU_i + g_iR_i + e_i$, where a, b, c, f, and g are regression coefficients. Table 3 gives the basic results on the relationship between the five selected variables and migration.

Employment and migration.--The basic goal here is in testing the hypothesis that there is a direct relationship between changes in the number of employed civilians and net migration. This hypothesis is confirmed, since the regression coefficient and simple and partial correlation coefficients between net migration and changes in total employment are positive and significant at the five percent level. Moreover, the change in total civilian employment emerged as the first variable of importance by explaining 70 percent of the variance in migration.

Underemployment and migration.--The hypothesis to be tested here states that migration is inversely related to underemployment. This hypothesis emerged to be true as shown by the partial correlation and regression coefficients. Although the total correlation is positive, the computation of the partial correlation reversed the sign and provided a higher association. The interpretation of this is that, other things being equal, the lower the number of man-years of economically utilized labor in these areas, the greater is the out-migration from them. With regard to the total variation explained in migration, underemployment was the second variable in importance where it accounted for 13 percent of variation in migration. These results lead to the conclusion that underemployment is useful in predicting net migration.

Armed forces and migration.--A direct relationship between changes in the number of armed forces personnel and net migration was hypothesized. This direct relationship is proved to be true, as provided by the zero order correlation, the partial correlation and regression coefficient in Table 3. All coefficients are significant at the five percent level. Moreover, changes in armed forces personnel seem to be a useful variable in predicting migration, for it accounted for more than one percent of the variation in mi-

TABLE 3. CUMULATIVE R^2, REGRESSION COEFFICIENTS, STANDARD ERRORS OF ESTIMATES, AND PARTIAL AND SIMPLE CORRELATIONS BETWEEN NET MIGRATION OF PERSONS 25-64 YEARS OLD AND SELECTED VARIABLES IN EIGHT SELECTED OBEA's OF THE UNITED STATES, 1950-1960

Selected Variables and Intercept	Cumulative R^2	Coefficients of: Regression and Intercept	Standard Errors	Correlation: Partial	Correlation: Simple
Employment	.696	.976	.047	.831	.834
Underemployment	.828	-1.269	.102	-.666	.196
Armed Forces	.841	.698	.159	.300	.418
Unemployment	.843	-1.620	.908	-.127*	.606
Median Income	.844	.369	.404	.065*	.371
Intercept		411.015			

*Not significant at .05 level.

gration with a small magnitude of the standard error.

Unemployment and migration.--The hypothesis to be tested states an inverse relationship between unemployment and net migration. This hypothesis is confirmed by the partial correlation coefficient and coefficient of regression. However, the partial association is not significant at the five percent level. This conclusion is consistent with other findings (Tarver, 1964a; Lowry, 1966). The association might be improved if the lagged relationship between unemployment change and migration were considered, for there is a time lag between unemployment as a cause and migration as an effect (Makower, et al., 1939). Our conclusion is that the change in unemployment as presented in this study is not a useful variable in the prediction of migration.

Real median income and migration.--The last hypothesis to be tested states a direct relationship between net migration and changes in real median income of families and unrelated individuals. Direction of the relationship is confirmed (Table 3). However, the relationship between the two variables became insignificant after the effects of the other variables (employment, unemployment, underemployment, and armed forces) had been removed, as given by the partial correlation coefficient. From this conclusion we can infer that changes in median income may not be very useful in explaining migration.

Summary and Conclusion

In this paper, a theoretical model which embraced those variables to which internal migration was believed to be most related was developed, and relationships were examined between net migration and the stated variables in the theoretical model. The analysis supports all hypotheses. The regression model which was fitted accounts for more than 95 percent of the variability in the 1950-1960 net migration of the eight selected OBEA's in terms of occupational and industrial changes, and 84 percent in terms of selected variables.

In conclusion, the authors believe that the theoretical model developed in this study proved useful in predicting net migration, and that its application to all the OBEA's of the United States will contribute to the refinement of our present knowledge of internal migration.

REFERENCES

Ashby, Lowell D. 1965. Growth Patterns in Employment by County 1940-1950 and 1950-1960, Southeast, Washington, D.C.: U. S. Government Printing Office.

Bates, Frederick L. 1969. "The Impact of Automation on the Organization of Society," pp. 107-148 in Ellis L. Scott and Roger W. Bolz (eds.), Automation and Society, Athens, Georgia: The Center for the Study of Automation and Society.

Beshers, James M. and Eleanor N. Nishiura. 1961. "A Theory of Internal Migration Differentials," Social Forces, 39 (March):214-218.

Blanco, Cicely. 1962. The Determinants of Regional Factors Mobility, Den Haag, Rotterdam: by the author.

Bowles, Gladys K. and James D. Tarver. 1965. Net Migration of the Population, 1950-1960 by Age, Sex, and Color, Vol. 1, Part 3, South Atlantic States, Washington, D.C.: U. S. Government Printing Office.

Dixon, W. J., ed. 1968. BMD Biomedical Computer Programs, Berkeley and Los Angeles: University of California Press.

Dunlop, John T. 1962. Automation and Technological Change, Englewood Cliffs: Prentice-Hall.

El Attar, Mohamed. 1970. "Migration and Occupational and Industrial Changes in Georgia, 1950-1960," Athens: The University of Georgia, unpublished Ph.D. dissertation.

Gossman, Charles S. et al. 1967. Migration of College and University Students, State of Washington, Seattle: Washington State Census Board.

Isard, Walter, et al. 1960. Methods of Regional Analysis: An Introduction to Regional Science, Massachusetts: M.I.T. Press.

Jaffe, A. J. and Joseph Froomkin. 1968. Technology and Jobs: Automation in Perspective, New York: Frederick A. Praeger.

Kampe, Ronald E. and William A. Lindamood. 1969. Underemployment Estimates by County, United States, 1960, Washington, D.C.: U. S. Department of Agriculture.

Ladinsky, Jack. 1967a. "Source of Geographic Mobility Among Professional Workers: A Multivariate Analysis," Demography, 4(1):293-309.

Ladinsky, Jack. 1967b. "Occupational Determinants of Geographic Mobility Among Professional Workers," American Sociological Review, 22(April): 253-264.

Lee, Everett S. and Ann S. Lee. 1960. "Internal Migration Statistics," Journal of the American Statistical Association, 55(December):664-697.

Lee, Everett S. 1966. "A Theory of Migration," Demography, 3(1):47-57.

Lively, C. E. and Conrad Taeuber. 1939. Rural Migration in the United States, Washington, D.C.: U. S. Government Printing Office.

Lowry, Ira S. 1966. Migration and Metropolitan Growth: Two Analytical Models, San Francisco: Chandler Publishing Co.

Miller, Ann R. 1966. "Migration Differentials in Labor Force Participation: United States, 1960," Demography, 3(1):58-67.

Miller, Ann R. 1967. "The Migration of Employed Persons to and from Metropolitan Areas of the United States," Journal of the American Statistical Association, 62(December):1418-1432.

Makower, H., J. Marschak and H. W. Robinson. 1939. "Studies in Mobility of Labor: A Tentative Statistical Measure," Oxford Economic Papers, 2(May):70-97.

Office of Business Economics. 1967. "OBE Economic Areas of the United States," Washington, D.C. (Mimeographed).

Petersen, William. 1969. Population, New York: The MacMillan Company.

Ravenstein, E. G. 1885. "The Laws of Migration," Journal of the Royal Statistical Society, 48(June):167-277.

Ravenstein, E. G. 1889. "The Laws of Migration," Journal of the Rural Statistical Society, 52(June):241-301.

Rezler, Julius. 1969. Automation and Industrial Labor, New York: Random House.

Saben, Samuel. 1964. "Geographic Mobility and Employment Status, March 1962-March 1963," Monthly Labor Review, 87(August):873-881.

Stouffer, Samuel A. 1940. "Intervening Opportunities: A Theory Relating Mobility and Distance," American Sociological Review, 5(December):845-867.

Stouffer, Samuel A. 1960. "Intervening Opportunities: A Theory of Competing Migrants," Journal of Regional Science, 2:1-26.

Tarver, James D. 1964a. "Metropolitan Area Inter-County Migration Rates: A Test of Labor Market Theory," Industrial and Labor Relation Review, 18(January):213-223.

Tarver, James D. 1964b. "Occupational Migration Differentials," Social Forces, 43(December):234-241.

Tarver, James D., William R. Gurley, and Patrick M. Skees. 1967. "Vector Representation of Migration Streams Among Selected State Economic Areas During 1955 to 1960," Demography, 4(1):1-18.

Thomas, Dorothy S. 1938. Research Memorandum on Migration Differentials, New York: Social Science Research Council.

Trott, Charles E. 1971. "An Analysis of Out-Migration," Proceedings of the American Statistical Association, Business and Economic Statistics Section (1971), 192-199.

U. S. Bureau of the Census. 1952. U. S. Census of Population: 1950. Characteristics of the Population, states of interest, Washington, D.C.: U. S. Government Printing Office.

U. S. Bureau of the Census. 1961. U. S. Census of Population: 1960. General Social and Economic Characteristics, states of interest, Washington, D.C.: U. S. Government Printing Office.

U. S. Bureau of the Census. 1962. Statistical Abstract of the U. S. 1962, Washington, D.C.: U. S. Government Printing Office.

U. S. Department of Labor. 1965. Manpower Report of the President, 1964, Washington, D.C.: U. S. Government Printing Office, 143-158.

Webster, Noah. 1968. Webster's Unabridged Third New Dictionary of the English Language, Springfield, Mass.: G. & C. Merriam Company.

Zipf, George K. 1946. "The P_1P_2/D Hypothesis: On Intercity Movements of Persons," American Sociological Review, 11(December):677-686.

THE DISCOURAGED WORKERS' EFFECT AS MEASURED THROUGH THE CURRENT POPULATION SURVEY

Paul O. Flaim, Bureau of Labor Statistics

Until a couple of decades ago the many millions of working-age persons outside the labor force were of limited concern to the labor economists. The notion that many of these persons might have wanted work but were not seeking it because of a belief that their search would be fruitless was not widely entertained. It gradually became evident, however, that millions of persons shift in and out of the labor force each year, not only because of personal reasons, but also in apparent response to changing labor market conditions. Recognizing this, the President's Committee to Appraise Employment and Unemployment Statistics (more familiarly known as the Gordon Committee) concluded in 1962 that "the relatively simple dichotomy between those in and out of the labor force...(no longer provides)...a satisfactory measure of the labor supply." What was especially needed to construct a better measure, added the Committee, were data on the so-called "discouraged workers" or "hidden unemployed"--those persons who want work but who are not looking for a job because of a belief that their search would be in vain. (In so doing, it should be added, the Committee also recommended that those persons not be included in the unemployment count.)

As the Gordon Committee was issuing its recommendations, some labor economists were already grappling with the problem of estimating the extent to which the discouragement over job prospects was affecting the growth of the labor force. Notable among the early pioneers in this field of research were Alfred Tella 1/and Thomas Dernburg and Kenneth Strand 2/. Thereafter, other economists, using a variety of econometric techniques, also undertook similar research. Essentially, they all attempted to measure the elasticity of labor force participation rates in response to the intensity of the demand for labor as reflected by the unemployment rate, the wage rate, and other variables. Optimal participation rates, those consistent with conditions of "full employment," were then applied to the population to obtain a "full employment labor force." To the extent that the actual labor force, as measured through the Current Population Survey, failed to match this theoretical labor force, the gap would be ascribed to the discouraged workers' phenomenon or hidden unemployment.

While the various estimates of hidden unemployment emanating from these econometric exercises were being discussed, 3/ the Bureau of Labor Statistics was taking steps to measure the extent of hidden unemployment through the Current Population Survey (CPS). A special set of questions designed to elicit detailed information on the reasons for nonparticipation in the labor force was tested experimentally in the 1964-1966 period and was finally incorporated into the regular CPS questionnaire in January 1967. The data derived through this set of special questions have already shed new important light on the status and attitudes of persons outside the labor force. Although the availability of these relatively new data has evidently not been widely known, they have been published quarterly by the Bureau of Labor Statistics since 1969 in a special set of tables added to Employment and Earnings.

The early analyses of these new data were, by necessity, limited to cross-sectional examinations, done in snap shot fashion. 4/ No time-series analysis could have been undertaken until a sizeable number of observations had been accumulated. This paper is a step in that direction. Its purpose is not only to describe how these data are collected and what they show, but also to determine the extent to which their behavior has been influenced by the changes in the unemployment situation. Two variables will be of particular interest, (1) the number of so-called "discouraged workers" and (2) the number of workers leaving the labor force because of "slack work."

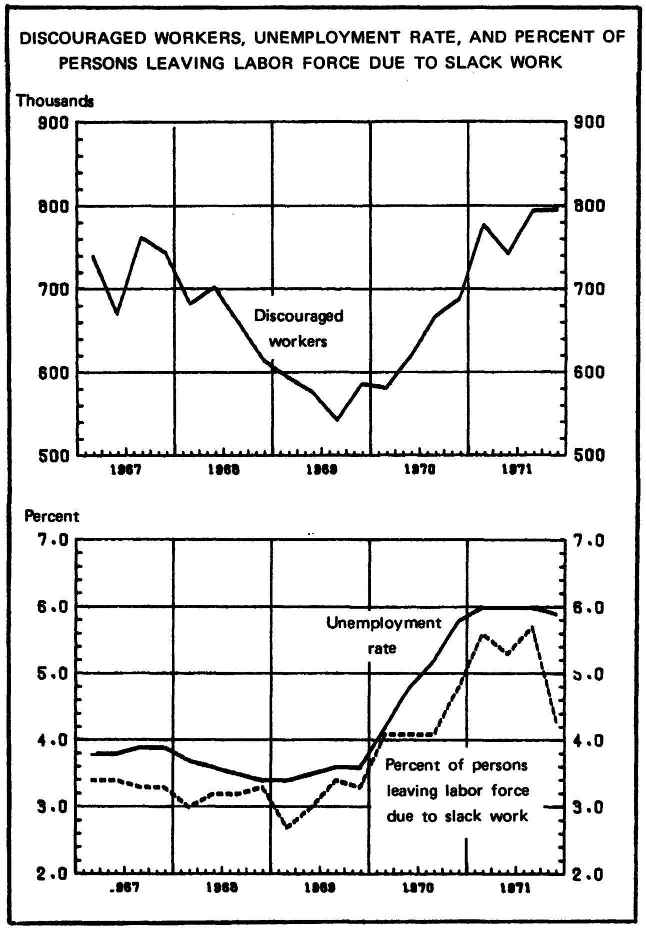
An overview of the 1967-1971 trends

The early analyses of the new not-in-the-labor force data showed that less than one-tenth of the more than 50 million nonparticipants professed any desire to be holding a job. Even among these, only about 700,000 or less than one-fifth, were classified as "discouraged workers." The others who claimed they wanted a job turned out to be either in school, in poor physical conditions, or to be prevented from seeking work by a burden of home responsibilities (table 1). Furthermore, the ranks of the discouraged workers were found to contain relatively few men of prime working age--less than 200,000. The great majority consisted, instead, of teenagers, housewives and elderly persons. These findings seemed to fly in the face of the contentions that there were virtually millions of discouraged workers and that they included large numbers of men.

It should be noted, however, that the data being analyzed in the late 1960's had been collected in a period of unusually low unemployment, when the jobless rate was below 4 percent. Any conclusion drawn on the basis of data for such a period could not be applied rationally to the situation of the early 1960's, when unemployment had been much higher. It thus remained to be seen how the discouraged workers' series would behave when the unemployment rate rose to a much higher level.

Unfortunately, we did not have to wait long for such a setting. As we all know, unemployment rose substantially during 1970. This rise was also accompanied by at least a temporary slackening in labor force participation among some groups. The question is the extent to which this slackening in participation may have stemmed from a rise in discouragement over job prospects.

As can be discerned from the chart, there is,



indeed, some positive relationship between the unemployment rate and the number of discouraged workers. Both series trended downward, though in differing degrees, during the 1967-69 period; both rose substantially during 1970; and both can be said to have shown little distinct movement during 1971. Despite the positive relationship between unemployment and discouragement, the coefficient of correlation between these two variables, derived on the basis of seasonally adjusted monthly data for the 1967-1971 period, 5/ was a rather low 0.53. Nor was the coefficient raised when the relationship between the two series was tested on the basis of data disaggregated in terms of major age-sex-color components. By lagging the discouraged workers' variable, respectively, by 3 and 6 months behind the unemployment rate, the coefficient of correlation was raised somewhat--to 0.61 in both cases--but was still far from indicating a very close relationship between the two variables. (Table 2)

"Cyclical" vs. "structural" discouragement

A close examination of the disaggregated data on discouraged workers for the 1967-1971 period revealed a significant change in composition in terms of the specific reason cited by these persons for their belief that they could not obtain a job. Specifically, there was an increase in the proportion whose discouragement appears to have been directly related to the conditions of the job market. Conversely, there was a decline in both the number and proportion of persons attributing their discouragement to rather personal situations or deficiencies.

This change in composition is clearly shown in table 3, where the discouraged are grouped into two broad categories. The first category includes the workers reported as believing that there were no jobs in their line of work or area and those who had tried unsuccessfully to find a job and had then given up the search. The second category includes those thinking they wouldn't get a job due to their very young or very advanced age, those who saw their lack of education or training as the major obstacle, and those who cited other personal handicaps (i.e. language difficulties).

It would appear, given the different nature of the reasons for discouragement, that the first category of discouraged workers should be much more cyclically sensitive than the second. Discouragement among the second category appears to be more of a "structural" nature and thus not necessarily related to the tightness, or "looseness," of the job market. A glance at table 3 would tend to confirm this hypothesis. As shown, all of the 200,000 increase in the number of discouraged workers between 1969 and 1971 took place among those blaming their situation on job-market weaknesses.

Regression analysis also confirmed this hypothesis. Whereas, as noted above, the regression of the total number of discouraged workers against the overall unemployment rate yielded a coefficient of correlation of only 0.53, a regression of the number of persons discouraged because of job market reasons against the jobless rate yielded a much higher correlation coefficient--0.79. (Table 2) On the other hand, a regression against the unemployment rate of the number of persons whose discouragement hinged largely on personal factors yielded a negative relationship with a correlation coefficient of -0.47. There is no ready explanation for this negative relationship, but some possibilities may be raised. For example, the passage of legislation designed to reduce job discrimination because of age may have reduced the number of elderly workers who think that they could not get a job due to their advanced age. It may also be hypothesized that when unemployment rises some workers who had previously been attributing their discouragement to personal reasons may find it more respectable to point to the deteriorating job market.

Whatever the reasons for the behavior of the data for the second category of discouraged workers, it is clear that if we limit our comparison to the first category, that is, those who attribute their discouragement to the conditions of the job market, we find that their number did increase and decrease in line with the underlying movement of the unemployment rate during the 1967-1971 period. 6/

Who are the discouraged workers?

As was the case during the first years of availability of data on discouraged workers, the proportion of men of prime working age among them continues to be relatively small. Of the 775,000 persons classified as discouraged workers in 1971, only about 75,000, or one-tenth, were men 25 to 59 years of age. (See table 4) Evidently a man of prime working age who really wants a job is not easily discouraged from his search but perseveres even after repeated failures.

Negroes are even more over-represented among the discouraged workers than they are among the unemployed. Although they make up only one-tenth of all the persons of working age both inside and outside the labor force, they accounted for about one-fifth of the unemployed and for nearly one-third of the discouraged workers in 1971.

In terms of previous work history, about two-fifths of the discouraged workers had been out of the job market less than one year when interviewed. Only 14 percent had never worked before. There are, of course, large differences among age-sex-color groups in this respect. (See table 4)

Evidently most discouraged workers regard their status as only temporary. Although they do not deem it worthwhile to look for a job at the time of the interview, they are apparently more hopeful in terms of their future prospects. As also shown on table 4, nearly 80 percent were reported as planning to actively seek work within the next 12 months. It would thus be erroneous to assume that most discouraged workers have permanently given up on the job market.

Problems of measuring discouragement

Determining the extent of discouragement over job prospects is, admittedly, a very difficult task, inasmuch as it involves the measurement of what are essentially subjective phenomena, namely one's desire for work and one's perception of the conditions of the job market. These facts are by no means easy to get at even through a large survey such as the CPS. To begin with, the housewife is typically the only person interviewed in each CPS household. While she may be quite informed about the activity of the other household members in terms of working or seeking work, she is not the best judge of their attitudes in terms of "wanting" work or "intending" to seek a job. Secondly, not all persons who become discouraged over their failure to find a job may disclose their true reasons for leaving the labor force, even if interviewed individually. Even among those who may truly want work, some could prefer, for example, to attribute their nonparticipation status to ill health or other "socially accepted reasons" rather than to admit, in effect, that they regard themselves as failures in the job market. On the other hand, some other persons whose desire for work is, at best, of questionable intensity may deem it proper to report themselves as wanting a job and then explain their failure to look for one in terms of unavailability.

To get at the discouraged workers, the CPS interviewer asks first whether the persons not in the labor force "want a regular job now, either full or part time." If the answer is yes, or even a tentative yes, there is a follow-up question as to the reasons why they are not looking for work. In order to be classified as discouraged, a person's principal reasons for not looking for work must fall in one of the following 5 categories:

1. Believes no work available in line of work or area.
2. Had tried but couldn't find any work.
3. Lacks necessary schooling, training, skills, or experience.
4. Employers think too young or too old.
5. Other personal handicap in finding a job.

It may be argued that the requirement that a person must first be reported as wanting a job in order to be questioned about possible discouragement, yields a rather restrictive definition of hidden unemployment. What about those persons, one might ask, who, upon losing their job, may decide to return to school and who would then not want a job "now"? Should they not also be regarded as discouraged workers? In answer to this it must be noted that if the discouraged workers' data are to be useful as a measure of underutilization of manpower for policy purposes, they should hardly include persons who say they do not want a job--and whose activity may actually prevent them from taking a job.

It is also important to note that, even before inquiring about the nonparticipants' current desire for jobs, the interviewer asks when they last worked and why they left their last job. As we shall see below, the data thus obtained on the reasons why persons not in the labor force left their last job may be as valuable in terms of understanding the dynamics of the labor force under changing economic conditions as are the data on discouraged workers discussed above. 7/

Unexpected discontinuity

One of the most interesting, though unexpected, findings from the 5 years of experience in obtaining statistics on labor force nonparticipants is that it apparently makes quite a bit of difference whether the questions about their current desire for work and future job-seeking plans are asked in the first month in which they are visited by the CPS interviewer or in subsequent months.

Since a person's reasons for nonparticipation in the labor force are not likely to change from one month to another, this information is asked in only one of the four consecutive monthly interviews conducted in households falling in the CPS sample. From 1967 through 1969, the questions were asked in the month in which a given household first entered the CPS sample and then again one year later when the same household reentered the sample for the second and final 4-month stint after an 8-month hiatus. In January 1970, the questions were switched from the first and fifth month-in-sample to the fourth and eighth. In effect, instead of being asked when a household enters or re-enters the sample, they are now being asked only when a household leaves the sample. 8/

This switch turned out to have a noticeable impact on the data for the persons not in the labor force. Following the switch, proportionately fewer persons, particularly among the housewives, were reported as either wanting a job at present or as planning to look for work in the near future. Possibly, having become increasingly more at ease with the interviewer with each passing month a respondent is less likely to exaggerate his (or her) attachment to the labor force in the fourth monthly interview than in the first one. This discontinuity in the data is a good illustration of the difficulties which arise in the measurement of what are essentially attitudes on the part of workers or potential workers. In this case, however, it appears that the data obtained since January 1970 are more realistic than those obtained in the 1967-1969 period.

Examining flows out of labor force

Some appreciation of the extent to which cyclical changes in the employment situation may be affecting the dynamics of the labor force may also be obtained by simply examining the changes in the gross flows out of the labor force and in the reasons for these outflows.

At any given time, there are about 10 million persons outside the labor force who have left their last job within the previous 12 months. Through the special set of questions asked since 1967, it has been possible to group these persons according to their reasons for leaving their job, regardless of whether or not they want a job when interviewed. This information is presented in table 5. Of particular interest in terms of cyclical behavior are the data on the number of persons whose jobs have been terminated, either temporarily or permanently, because of economic reasons. Of the four categories under the "economic" heading, "slack work" appears to be the most cyclically sensitive. As also shown on the chart, the changes in this variable --being expressed as a percent of the total leaving the job market during the previous 12 months--are, indeed, closely related to the changes in the unemployment rate. In fact, the coefficient of correlation between the over-all unemployment rate and the number of persons reporting they had left the labor force after having lost their jobs due to "slack work" was 0.83 on the basis of monthly data for the 1967-71 period. The substitution of data on unemployment due to job loss for the over-all measurements of unemployment yielded coefficients of roughly similar magnitude. (See table 6).

Summary and conclusion

After 5 years of experience in the collection of data on discouraged workers through the Current Population Survey, it appears that the survey is, indeed, a very viable vehicle for such a purpose. Although everyone might not agree with the definition of "discouragement" used for the purposes of the survey, the data gathered so far have shed important light both on the discouraged workers' phenomenon and other aspects of labor force dynamics.

Although the accumulated data are not yet sufficient to enable us to establish with any certainty the relationship between given variables, it can also now be said at least tentatively that changes in the number of discouraged workers are postively related to changes in the unemployment rate. The same can also be said for changes in the number of workers leaving the labor force because of slack work. To the extent that this is true, it would appear that we should take into account these variables, as well as the data on unemployment and underemployment, when assessing the waste of manpower which accompanies an economic recession.

FOOTNOTES

1/ A. Tella, "The Relation of Labor Force to Employment," Industrial and Labor Relations Review, XVII (April 1964), pp. 454-469.

2/ T. Dernburg and K. Strand, "Cyclical Variation in Labor Force Participation," Review of Economics and Statistics, XLVI (November 1964), p. 378.

3/ For an analysis of these early estimates of "hidden unemployment," see Jacob Mincer, "Labor Force Participation and Unemployment: A Review of Recent Evidence" in R.A. Gordon and M.S. Gordon, Editors, Prosperity and Unemployment (New York, Wiley, 1966), p. 73.

4/See Robert L. Stein, "Reasons for Nonparticipation in Labor Force," Monthly Labor Review, July 1967, p. 22, and Paul O. Flaim, "Persons not in the Labor Force", Monthly Labor Review, July 1969, p. 3.

5/ Although the not-in-the-labor-force data are published only quarterly, they are tabulated monthly. They have also been seasonally adjusted experimentally, although not yet published in this form.

6/It would be useful at this point to compare the CPS data on discouraged workers with econometrically derived estimates. However, such comparisons are currently being prepared for publication by Joseph L. Gastwirth (currently at George Washington University).

7/The persons not in the labor force are also queried about their plans to look for work. This information is published quarterly by the BLS, but its predictive value has yet to be evaluated.

8/The switch was an attempt to see if these questions had increased the rotation group bias in the unemployment figures. This refers to the higher incidence of joblessness in households entering or reentering the CPS sample than in households which have been in the sample for two or more consecutive months. This "first-month bias" became larger around 1967, and it was hypothesized that this was related to the introduction of the not-in-the-labor-force questions. The reduction in the reported incidence of unemployment for the first and fifth month-in-sample groups and the concomitant rise for the fourth and eighth following the January 1970 switch of the not-in-the-labor-force questions seems to have confirmed this hypothesis.

Table 1. Persons not in labor force by desire for jobs and reasons for nonparticipation, 1967-71
(In thousands)

Reasons for nonparticipation	Annual averages			Quarterly averages, seasonally adjusted							
				1970				1971			
	1967	1968	1969	1st	2nd	3rd	4th	1st	2nd	3rd	4th
Total, not in labor force	52,484	53,289	53,596	53,703	54,120	54,583	54,714	55,033	55,828	55,945	55,872
Want job now	4,698	4,478	4,459	3,811	3,670	4,108	4,015	4,428	4,443	4,508	4,363
In school	1,104	1,115	1,126	1,028	1,043	1,226	1,130	1,215	1,272	1,364	1,192
Ill health, disability	768	656	627	503	411	520	500	580	638	486	541
Home responsibilities	1,325	1,263	1,257	917	876	933	982	1,048	979	994	1,075
Think cannot get a job	732	667	574	582	621	668	689	778	743	795	796
All other reasons	769	777	875	781	719	761	714	807	811	869	759
Do not want job now	47,786	48,809	49,137	49,892	50,450	50,475	50,699	50,605	51,385	51,437	51,509
In school	5,641	5,892	5,958	6,056	6,122	5,580	6,161	6,141	6,634	6,660	6,378
Ill health, disability	3,741	3,684	3,826	3,749	4,037	3,980	3,828	4,047	4,090	4,036	4,119
Home responsibilities	31,239	31,667	31,384	31,910	32,259	32,246	32,227	31,977	32,351	32,269	32,221
Retirement, old age	5,313	5,540	5,795	5,791	5,766	6,217	5,961	6,004	5,788	6,108	6,352
All other reasons	1,853	2,027	2,174	2,386	2,266	2,452	2,522	2,436	2,522	2,364	2,439

Table 2. Regressions of selected categories of discouraged workers against various measurements of unemployment

Variables		Regression results					
Independent (X)	Dependent (Y)	Regression equation	r	r^2	s	T ratio	Durbin-Watson
Unemployment rate, overall	Discouraged, total	Y = 472 + 47.0X (43.7) (10.0)	0.53	0.28	77.8	4.70	0.88
" " men 20+	" men 20+	Y = 145.6 + 6.8X (11.8) (4.1)	0.21	0.05	29.8	1.66	1.51
" " women 20+	" women 20+	Y = 257.4 + 31.7X (44.1) (9.9)	0.39	0.15	62.2	3.18	1.31
" " teens 16-19	" teens 16-19	Y = -14.4 + 9.4X (31.9) (2.3)	0.47	0.22	35.8	4.10	1.90
" " overall	" total lagged 3 months	Y = 432.7 + 57.5X (42.2) (9.9)	0.61	0.36	72.7	5.80	1.02
" " overall	" total lagged 6 months	Y = 413.2 + 63.8X (44.9) (10.8)	0.61	0.37	72.3	5.90	1.08
Persons unemployed 15 weeks and over	" total	Y = 568.1 + 142.3X (20.2) (26.4)	0.58	0.34	74.5	5.41	0.98
Average duration of unemployment	" total	Y = 285.6 + 43.1X (65.6) (7.2)	0.62	0.38	71.9	5.97	1.18
Unemployment rate, overall	" job market reasons	Y = 77.2 + 76.0X (33.8 (7.7)	0.79	0.62	60.3	9.81	1.54
" " "	" personal reasons	Y = 387.8 - 27.1X (31.2) (7.1)	-0.45	0.20	55.5	-3.80	0.67
" " "	" job market reasons lagged 3 months	Y = 58.6 + 82.4X (33.6) (7.9)	0.81	0.65	58.0	10.41	1.73
" " "	" job market reasons lagged 6 months	Y = 44.1 + 88.1X (37.7) (9.1)	0.79	0.62	60.7	9.70	1.61

Table 3. Composition of discouraged workers by reason for believing they cannot find a job, annual averages, 1967-71

(Numbers in thousands)

Reason	1967	1968	1969	1970	1971
Total (in thousands)	732	667	574	638	774
Job-market factors	383	371	311	437	537
Had looked but could not find job	168	161	161	244	300
Thinks no job available	215	210	150	193	237
Personal factors	349	297	263	201	236
Employers think too young or too old	216	171	139	105	112
Lacks education, skills, training	84	74	78	60	85
Other personal handicap	49	52	46	36	39
Percent distribution	100.0	100.0	100.0	100.0	100.0
Job-market factors	52.3	55.5	54.2	68.5	69.5
Had looked but could not find job	23.0	24.1	28.0	38.2	38.8
Thinks no job available	29.4	31.4	26.1	30.3	30.7
Personal factors	47.7	44.5	45.8	31.5	30.5
Employers think too young or old	29.5	25.6	24.2	16.5	14.5
Lacks education, skills, training	11.5	11.1	13.6	9.4	11.0
Other personal handicap	6.7	7.8	8.0	5.6	5.0

Table 4. Discouraged workers by time elapsed since last job, intentions to seek work in future, and sex, age, and color, 1971 annual averages

Sex, age, and color	Total discouraged (In thousands)	Percent distribution by time elapsed since last job					Percent who intend to seek work within 12 months
		Total	Less than 1 year	1 to 5 years	More than 5 years	Never worked	
Total, 16 years and over	774	100.0	40.5	26.7	18.5	14.3	78.7
Male, 16 years and over	238	100.0	51.5	26.8	8.8	12.6	82.4
16-19 years	59	100.0	49.2	11.9	--	39.0	89.8
20-24 years	34	100.0	65.6	18.8	--	15.6	96.9
25-59 years	73	100.0	58.7	29.3	9.3	2.7	84.9
60 years and over	73	100.0	41.9	39.2	18.9	-	66.2
Female, 16 years and over	536	100.0	35.4	26.7	22.6	15.3	77.4
16-19 years	80	100.0	40.5	8.9	1.3	49.4	86.3
20-24 years	74	100.0	45.9	28.4	5.4	20.3	83.8
25-59 years	308	100.0	34.3	27.9	29.9	8.1	79.5
60 years and over	74	100.0	24.7	39.7	32.9	2.7	54.1
White 1/	589	100.0	--	--	--	--	--
Negro and other races 1/	185	100.0	--	--	--	--	--

1/ Breakdown of discouraged workers in terms of time elapsed since last job and future job-seeking intentions is not available separately for whites and Negroes.

Table 5. Persons not in labor force who stopped working during previous 12 months, by reason for leaving last job, 1967-71
(In thousands, not seasonally adjusted)

Number leaving labor force by reason	Annual averages			Quarterly averages							
				1970				1971			
	1967	1968	1969	1st	2nd	3rd	4th	1st	2nd	3rd	4th
Total: Left job previous 12 months	9,327	9,752	10,175	10,944	9,761	9,514	10,302	11,091	9,869	9,468	9,965
Percent distribution by reason	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
School, home responsibilities	49.2	50.3	50.5	50.7	47.0	46.7	52.5	47.5	45.1	46.0	52.0
Ill health, disability	9.5	9.2	9.6	7.8	8.9	10.9	8.4	8.5	8.2	9.3	8.9
Retirement, old age	5.3	6.0	6.1	5.8	6.5	7.4	7.0	7.1	7.3	7.6	7.6
Economic reasons	17.1	17.8	16.6	18.8	17.5	19.1	16.7	21.1	19.4	20.6	16.9
End of seasonal job	9.2	9.1	8.5	9.1	7.8	8.6	7.1	9.6	8.0	7.8	8.5
Slack work	3.3	3.1	3.1	4.4	4.2	4.1	4.3	6.0	5.4	5.7	3.9
End of temporary job	4.6	5.6	5.1	5.3	5.6	6.4	5.4	5.5	6.1	7.1	4.5
All other reasons	18.9	16.7	17.2	16.9	20.1	15.9	15.4	15.8	20.0	16.5	14.7

Table 6. Regression of selected categories of workers leaving labor force against various measurements of unemployment

Variables		Regression results					
Independent (X)	Dependent (Y)	Regression equation	r	r^2	s	T ratio	Durbin-Watson
Unemployment rate........................	**Total leaving due to economic reasons**	**Y = 1271.4 + 113.5X (68.7) (15.7)**	**0.69**	**0.47**	**122.3**	**7.21**	**1.17**
Unemployment rate........................	**Left due to slack work**	**Y = 2.1 + 86.3X (33.0) (7.6)**	**0.83**	**0.69**	**58.8**	**11.4**	**1.92**
Number of unemployed who lost last job...	**Left due to slack work**	**Y = 140.7 + 0.15X (21.8) (1.46)**	**0.82**	**0.67**	**60.6**	**10.9**	**1.88**
Job-losers rate..........................	**Left due to slack work as percent of total leaving labor force**	**Y = 1.4 + 1.3X (0.2) (0.1)**	**0.82**	**0.67**	**0.6**	**11.0**	**2.04**

FAMILY SIZE AND ECONOMIC WELFARE IN A DEVELOPING ECONOMY

Deborah S. Freedman*, The University of Michigan

Discussions of the relationship between population growth and economic development usually assume, either explicitly or implicitly, that couples with fewer children to support will be better off financially. Improvement in economic welfare is itself one of the goals of development. In addition, the increased availability of income could stimulate new patterns of economic behavior, some of which may be favorable to the development effort. While these problems are recognized as important, little empirical work at the micro-level has been done on them for developing countries. As a small contribution to such analysis, this paper examines how income and other measures of economic behavior are related to family size in Taiwan.

Presumably, couples with fewer rather than more children, assuming equal family incomes, should have more economic options, e.g., for increases in per capita consumption or for more savings. Of course, couples with smaller families could choose to enjoy increased leisure, working only enough to maintain their customary levels. Such a pattern seems unlikely in Taiwan; most Taiwanese couples are interested in working hard and improving their living standards.

The present paper will consider whether, associated with varying numbers of children, there are differences in:

(1) income, initially looking at total family income, but more importantly, examining per capita or per consumption unit income

(2) use of available income for modern consumption--durables and recreation--and for savings

(3) economic attitudes

The data come from a 1969 island-wide survey of 2300 husbands with wives in the child-bearing years (18-24). A previous survey in 1966 had collected extensive demographic data from the wife; in this survey a substantial number of economic questions were included. The present analysis is limited to the 1323 couples who have been married at least ten years, since more recently married couples are less likely to have completed their families and may not have had their children long enough to feel the full effect on the family budget.

The analysis of the relationship of actual income levels to family size is further restricted to couples who live in nuclear units--husband, wife and unmarried children; two-thirds of the couples married 10 years or longer fall in this category. It is more difficult to estimate the effective income of couples who share joint living arrangements since (1) an assumption must be made about how total family income is shared--what proportion is available for this particular couple and their children, and (2) the data for joint families are less precise with regard to age and sex composition. A few nuclear families included unmarried children who earned incomes which they did not pool with the family; these cases were excluded from the analysis.

Small families are better off, even in terms of total income. While the two intermediate income categories are not ordered, the smallest families do have the highest incomes and the largest families have the lowest incomes, (See Table 1, Col. 1). This ordering reflects the higher education of husbands with small families; once education is controlled by Multiple Classification Analysis (MCA) [1], there is no consistent relationship between income and family size (Col. 2). Income per capita probably is a better measure of well being and, with this measure, the larger the family, the less resources they have available for each family member. This relationship is monotonic and persists, only slightly lessened, when an adjustment is made for husband's education.

An even more relevant measure of family welfare may be income per consumption unit, which recognizes basic differences in consumption needs between adults and children. Weights were assigned to the children in each family, based on their age and the relative caloric needs of different age groups.1/ The income of the smallest families, on a per consumption unit basis, is more than double the income of the largest families, with three and four child families intermediate. Having three instead of four children is only slightly more advantageous in terms of income per consumption unit; the initially lower total income of three-child families offsets their lesser burden of support. Adjusting for husband's education does not appreciably alter the relationship (See Table 1, Columns 5 and 6).

Thus, couples with fewer children in Taiwan are better off than others, both in terms of per capita income and income per consumption unit. Over the full range of family size, the differences in well-being associated with parity are considerable, but there is no appreciable advantage in having three instead of four children. The more favorable per capita income position of small families should make possible additional options with regard to expenditures and saving and could affect economic attitudes; these in turn could have an impact on the development effort.

We look next at economic behavior--are there differences with regard to consumption and saving patterns that are associated with family size? Our analysis is restricted to several kinds of modern consumption, particularly consumer durables and recreation expenditures.2/ Of course, it is possible that the additional income per consumption unit might be expended to increase traditional expenditures for food and clothing. Such expenditures, insofar as they improve health standards, also may be beneficial to development - as an investment in human capital. However, it seems more likely that additional income would be expended on modern goods and services; such items, being new and highly valued, have high marginal utility. Expenditures on modern consumption can have a direct impact on development. In the first place, the use of modern durables can exert a modernizing influence; for example, a radio or television set brings listeners in contact with a broader world while a motorcycle increases mobility. Modern recreation expands the horizons of participants by providing enjoyable contacts outside of the immediate family group. On the production side, the demand for modern durables requires the development of technical skills and the creation of new marketing channels. Saving differentials associated with family size can provide resources for investment.

Small families do rank higher than larger families on consumption of modern durables and use of modern recreational facilities. They also are more likely to have saved[3/] (see unadjusted deviations in Table 2). These relationships are as expected; couples with fewer children to support devote more of their income to new forms of consumption or to savings. The relationships shown are for nuclear families, since only in these families does the interviewed respondent, together with his wife and children, unequivocally comprise both the income and the consumption unit. However, in unpublished tabulations these relationships also were found to hold for all couples married ten years or more, including those in joint or stem families.

The higher consumption and savings patterns of low parity families reflect, to some extent, their higher education and somewhat higher family incomes, since education and income are highly correlated with consumption and saving. However, after an adjustment is made for the effect of family income and education, using MCA, meaningful relationships remain between family size and both modern consumption and saving, although the magnitude of the relationships is diminished. The per capita income advantage which smaller families enjoy does affect their saving and modern consumption (Cols. 2, 4 and 6 of Table 2).

Although family size influences modern consumption and saving, its effect is modest, particularly when compared to the effect of total family income. This is not surprising since the variations in purchasing power stemming from family size differences are much less than those attributable to differences in family income. Seventy-five percent of the families married ten years had either 3, 4, or 5 children; 87 percent were in the 3-6 child range. Supporting six children instead of three (the largest variation possible for most couples) would, at most, cut consumption possibilities in half. The range in total income, on the other hand, is far greater.

Husband's education, after adjusting for the effects of income, has a stronger net relationship than does family size to consumption of modern durables and to recreation expenditures. Education serves to bring individuals in contact with the modern sector and influences their tastes. These differences in tastes could be expected to influence purchases of new kinds of goods and services more than the variations in purchasing power associated with family size differences. Both education and family size show a similar, fairly sizeable, relationship to saving behavior.

Our rather arbitrary assignment of consumption weights to Taiwanese children may be an improper generalization of Western experience and may overstate their marginal cost to the family. The marginal cost of an additional child is probably less in Taiwan than in Western countries; more food is home produced, supplemental child care is probably performed by relatives and lesser expenditures are made on lessons and recreation. Perhaps the marginal cost is small and the situation is as expressed in a Chinese proverb--"an additional baby requires only an extra pair of chopsticks." If so, the effective income advantage of small families would be lessened and this might account for the moderate sized differences in consumption and saving attributable to family size.

Even though the differences in modern consumption and saving associated with family size are not large, they are not inconsequential. For example, the adjusted consumption and savings measures for couples with three children are 12 to 30 percent higher than those for couples with five or more children; couples with two or less children rank even higher. In addition, the patterns of these relationships are consistent and monotonic.

We next look at economic attitudes; are there attitudinal differences associated with family size? The attitudinal measures used are: consumption aspirations[4/], perceptions of income change (both past and future), college expectations for sons, and two measures of expectations of support and assistance from children. Of course, cross-sectional studies can not establish lines of causality. Achieved family size may be one result of a couple's prior assessment of the net economic value of children or their desire for a higher standard of living. At the same time, an individual's attitudes reflect his experiences and achievements and it may well be that success in limiting family size, with its attendant improvement in family welfare, would affect a couple's expectations and aspirations. Economic attitudes can have an impact on development. For example, studies of consumer behavior have shown a positive relationship between the level of optimism and aspirations [3]; aspirations, in turn should stimulate greater work effort. Contrariwise, couples who expect help from their children might feel less pressed to earn and save during their productive years.

We have restricted this part of the analysis, which deals with attitudes, to couples who have at least three children. There is reason to believe that for couples with fewer children, family size is less likely to reflect their basic attitudinal patterns. Most couples in Taiwan want at least three children, with the result that smaller families usually reflect fecundity impairments or late marriages, rather than choice. For example, 65 percent of the couples with less than three children had fewer than their ideal number of children; this compares to less than 20 percent for all nuclear families. Many of these couples do have physical impairments which limit child-bearing; 39 percent cited specific reasons for subfecundity and the actual level of subfecundity, including cases with no specific cause, is undoubtedly much higher. On the other hand, these families were included in the analysis of actual consumption patterns, inasmuch as consumption possibilities depend on achieved family size, irregardless of whether the couple wanted more than this specific number.

One pair of attitudinal indices measures the husband's and wife's perception of the usefulness of children--that is, the potential benefits to be derived from them. For the wife, this combines her expectations of living with her married sons and of being supported by them in her old age. For the husband, the measure includes, not only these two indications of traditional reliance on children but several other specific benefits.5/ Our data (Cols. 5-8 of Table 3) show that there is a positive relationship between family size and either the husband's or wife's expectations of the usefulness of children. The relationship is modest but regular, and it persists after adjusting for income and education. It seems reasonable that couples who expect to reap considerable benefits from their children might be motivated to have a large family. Of course, since this is a cross section study, we do not know whether they had many children because they foresaw their usefulness, or whether this reflects only an unplanned for reality--that couples with many children can more realistically expect help.

We also have some measures of aspirations for improvement in living standards - for modern consumption and for better education for children. Such aspirations are one measure of the opportunity cost of children, since they represent possible alternative uses of income as a result of limiting family size. Columns 1-4 of Table 3 show that aspirations, both for modern consumption and for education, are negatively related to family size. Couples with moderate sized families--3 or 4 children--express higher educational aspirations than do couples with 5 or more children. However, the difference is not large (a net difference of about 8 percentage points) and there is no appreciable difference in educational aspirations associated with having 3 instead of 4 children. There also is evidence that the educational aspirations of fathers with few children may be more reality oriented than is true for husbands with larger families. Only 40 percent of all fathers who expressed college plans had some idea of the costs involved; the smaller the family, the more likely it was that the father had some concrete notion about the costs of his educational ambitions. Family size is similarly, but less strongly, related to aspirations for modern consumption; large families do have lower consumption aspirations, but the magnitude of the differential is very modest. A high evaluation of the opportunity costs of children could motivate couples to have fewer children. An alternate possibility is that couples who had small families for other reasons might, as a result, feel they could aspire to a higher standard of living.

A third set of indices measures perceptions about the family's current financial situation (See Table 3 cols. 9-12). Couples with small families are more likely than others to feel their financial situation has improved in the last five years. Despite an increase in real income of 25 percent during the past 5 years, most families in Taiwan did not recognize any improvement in their current financial situation. The optimism of small families could reflect either their relatively better per capita income situation or the confidence gained by success in limiting family size. On the other hand, the fewer children a couple has, the less likely the husband is to feel his future income situation will improve. This may accurately reflect the situation of Taiwan parents for whom children are still the most likely source of old age support. Their present financial situation may be strained because of the expense of raising children, but these same children may be their source of support in the future. There is some evidence for this, in that thirty percent of the husbands who expected to be better off 5 years hence, cited help from children as the basis for this expectation.

In sum we have found that there are income differentials associated with family size in Taiwan. Family size is inversely related to economic welfare; couples with few children are better off than those with more children to support, in that they have more income available per family member, either on a per capita or per consumption unit basis. Family size is also related to differences in economic behavior and attitudes in ways which seem favorable to development, but the magnitude of these relationships is modest. Couples with small families are more likely to save and somewhat more likely than others to enjoy modern consumption. As expected, differentials in family size are associated with attitudinal differences; couples who expect financial support from children are likely to have large families, but attitudes measuring the opportunity cost of children--consumption and educational aspirations--show only weak associations to family size. It may well be that the costs and benefits of children in Taiwan are such that they diminish the welfare differential associated with having fewer children.

Footnotes

* The author is an Assistant Professor of Economics at the University of Michigan. The data for this study were collected with the assistance of a grant from the Population Council. The field work for the survey was conducted by the Taiwan Provincial Institute of Family Planning under the direction of Dr. Tom Sun.

1) Income per consumption unit was obtained by dividing family income by the weighted sum of family members, weights being assigned on the basis of the U.S. Department of Agriculture need standards. [(For example, adults are assigned a weight of one, while children under four are weighted .3.) Only the age of the youngest and oldest child in each family were readily available, but they provided a rough estimate of the age composition of the children and an appropriate average weight was applied. Thus, income per consumption unit for each family = total family income ÷ (number of children x average weight) + 2 (for the parents)].

2) Modern durables include the following: electric fan, sewing machine, bicycle, motorcycle, rice cooker, clock, radio, record player, television, air conditioner, refrigerator, gas burner, camera and washing machine. Modern recreation expenditures include travel, meals in restaurants and movie attendance.

3) The measure of savings was the respondent's statement as to whether or not he had accumulated savings since marriage.

4) Consumption aspirations included interest in obtaining more durables, improving housing and enjoying more modern recreation. For details on index construction, see D. Freedman [2]. Perceptions of income change included statements about whether they were better off relative to five years hence. One measure of expectation of support from children was the proportion of wives with traditional attitudes about support from and living with children in old age. The father's measure was a more elaborate index based on many statements made by the father about possible assistance from his children. For a more detailed description of the father's index see E. Mueller [4].

5) The index measure included statements about help from children around the house or in farm or business and whether he expected to share in the children's earnings, even before retirement.

References

[1] Andrews, F., Morgan, J., and Sonquist, J., Multiple Classification Analysis Ann Arbor, Survey Research Center.

[2] Freedman, Deborah S., "Consumption Aspirations as Economic Incentives in a Developing Economy" Human Behavior in Economic Affairs (Essays for George Katona) North Holland Press, 1972.

[3] Katona G., Strumpel B., and Zahn E., Aspirations and Affluence, Comparative Studies in the United States and Western Europe, McGraw-Hill, 1971.

[4] Mueller, Eva, "Economic Motives for Family Limitation." To appear in Population Studies.

Table 1

Income of Nuclear Families by Size of Family
(Women married 10 years or more)

Size of Family	Number of Families	Mean Total Income (NT $000) Unadjusted	Adjusted[a]	Income per Capita (NT $000) Unadjusted	Adjusted[a]	Income per Consumption Unit (NT $000) Unadjusted	Adjusted[a]
		(1)	(2)	(3)	(4)	(5)	(6)
0-2 children	46	34.3	30.4	8.6[b]	7.6	9.78	8.99
3 children	164	28.5	26.0	5.7	5.2	6.96	6.45
4 children	243	31.6	30.9	5.3	5.1	6.63	6.49
5 or more children	351	26.1	28.2	3.4[b]	3.7	4.42	4.87
Total	805	28.7					

[a]Adjusted for husband's education

[b]The average size of families with 0-2 children is almost four persons while that for families with five or more children is 7.6 persons.

Table 2

Various Economic Behavior Measures for Nuclear Families, Married 10 years or More, By Size of Family, Education of the Husband, and Family Income

	Number of Couples	Ownership of Modern Objects GM = 5.7 Unadjusted	Adjusted[1]	Consumption of Services GM = 2.0 Unadjusted	Adjusted[1]	Percentage Saving GM = 44% Unadjusted	Adjusted[1]
		(1)	(2)	(3)	(4)	(5)	(6)
Size of Family							
0-2	46	6.0	5.6	2.3	2.1	50	48
3	164	5.7	5.4	2.0	1.8	45	41
4	243	5.3	5.2	1.6	1.5	42	40
5+	352	4.5	4.8	1.3	1.4	29	31
Family Income[2]							
Under $12,000 per year	170	2.9	3.3	.6	.9	9	14
12,000-23,999	257	3.8	4.1	1.0	1.2	24	26
24,000-35,999	185	6.0	5.9	2.0	1.9	49	48
36,000-47,999	89	7.1	6.7	2.0	1.7	61	56
48,000-71,999	62	8.1	7.3	3.2	2.7	73	65
72,000 and over	42	9.0	8.2	4.0	3.4	79	71
Husband's Education							
Less than Primary graduate	193	3.4	4.4	.6	1.0	17	30
Primary Graduate	399	4.6	4.8	1.3	1.4	33	35
Junior High	96	6.6	5.8	2.6	2.3	59	49
High School graduate or more	117	8.2	6.6	3.2	2.5	68	48
Total	805						

[1]Adjusted for income, education, and family size

[2]40 NT$ = 1 $ U.S.

Table 3

Selected Economic Attitudes for Nuclear Families, Married at Least Ten Years, by Family Size

Size of Family	Number of Cases	Consumption Aspirations GM = 2.0 Unadjusted	Adjusted[1]	Percentage With Educational Expectations GM = 63% Unadjusted	Adjusted[1]	Expectations of help from children: Father's Measure GM = 6.0 Unadjusted	Adjusted[1]	Mother's Measure GM = 50% Unadjusted	Adjusted[1]
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
3	164	2.2	2.2	70	66	5.5	5.7	42	45
4	243	2.1	2.0	68	67	5.8	5.9	49	50
5+	351	1.9	1.9	55	58	6.3	6.2	56	53
Total	758								

Size of Family	Number of Cases	Percentage Better off 5 years ago GM = 21% Unadjusted	Adjusted[1]	Percentage Better off 5 years from now GM = 21% Unadjusted	Adjusted[1]
		(9)	(10)	(11)	(12)
3	164	29	28	18	17
4	243	19	20	21	20
5+	351	17	18	22	24
Total	758				

[1]Adjusted for family income and husband's education

ESTIMATORS FOR SAMPLES SELECTED FROM TWO OVERLAPPING FRAMES

Wayne A. Fuller, Iowa State University
Leon F. Burmeister, University of Iowa

1. INTRODUCTION

The use of two frames or lists to select a sample of a population was, perhaps, first used by the United States Bureau of the Census in the "Sample Survey of Retail Stores" conducted in 1949 and described by Bershad in Hansen, Hurwitz and Madow [1, p. 516]. Hartley [7] considered the design and estimation for such surveys. Lund [10], Cochran [3] and Williams [12] have also studied the problem and offered alternative estimation procedures.

We assume two frames, A and B, containing N_A and N_B elements respectively are available. We adopt the notation of Hartley [7] and denote by N_{ab} the number of elements included in both frame A and frame B. N_a is the number of elements occurring only on frame A and N_b is the number of elements occurring only on frame B. Thus

$$N_A = N_a + N_{ab},$$
$$N_B = N_b + N_{ab}$$

and the total number of elements in the population, N, is given by

$$N = N_a + N_b + N_{ab} = N_a + N_B = N_b + N_A.$$

We call the elements contained only on frame A domain a, the elements only on frame B domain b and those elements on both frames A and B domain ab, or the <u>overlap domain</u>. We assume that simple random samples of sizes n_A and n_B are drawn from frames A and B respectively. The number of elements sampled from frame A and contained in domain a is denoted by n_a. The number of elements sampled from frame A and contained in domain ab is denoted by n'_{ab}. The number of sampled elements in domains ab and b drawn from frame B are denoted by n''_{ab} and n_b respectively. Thus

$$n_A = n_a + n'_{ab}, \text{ and}$$
$$n_B = n''_{ab} + n_b. \tag{1.1}$$

We consider the estimation of the number of elements in the overlap domain, N_{ab}, and the estimation of the total of a characteristic, Y. We develop estimators which to order (n) are superior to those appearing in the literature.

2. ESTIMATION OF N_{ab}

2.1 Estimation of N_{ab} when Duplicated Items are Ignored

We first consider the problem of estimating the number of elements in the overlap domain, N_{ab}. Hartley [7] proposed the estimator

$$\hat{N}_{ab,H} = \frac{pn'_{ab}N_A}{n_A} + \frac{qn''_{ab}N_B}{n_B}, \tag{2.1}$$

where p and q are constants with the property

$$p + q = 1.$$

This estimator makes no use of elements in the sample that were selected both from frame A and from frame B. The variance of Hartley's estimator is given by

$$\mathrm{Var}(\hat{N}_{ab,H}) = p^2 f_A^{-2}\,\mathrm{Var}(n'_{ab}) + q^2 f_B^{-2}\,\mathrm{Var}(n''_{ab}). \tag{2.2}$$

Since n'_{ab} and n''_{ab} are hypergeometric random variables, it follows that the p which minimizes the variance is given by

$$p'_{oH} = \frac{n_A N_b g_B}{n_A N_b g_B + n_B N_a g_A}, \tag{2.3}$$

where

$$g_A = \frac{N_A - n_A}{N_A - 1}$$

and

$$g_B = \frac{N_B - n_B}{N_B - 1}.$$

For this value of p the variance of the estimated overlap domain is given by

$$\mathrm{Var}_o(\hat{N}_{ab,H}) = \frac{N_{ab}N_a N_b g_A g_B}{n_A N_b g_B + n_B N_a g_A}. \tag{2.4}$$

To obtain an estimator of N_{ab} that is a function of the sample data only we set

$$\hat{N}_b = N_B - \hat{N}_{ab},$$
$$\hat{N}_a = N_A - \hat{N}_{ab}$$

and substitute expression (2.3) for p into (2.1). Expression (2.1) then reduces to a quadratic in $\hat{N}_{ab}$ and we define a new estimator by this quadratic,

$$[n_A g_B + n_B g_A]\hat{N}^2_{ab,s} - [n_A N_B g_B + n_B N_A g_A + n'_{ab}N_A g_B + n''_{ab}N_B g_A]\hat{N}_{ab,s} + [n'_{ab}g_B + n''_{ab}g_A]N_A N_B = 0. \tag{2.5}$$

It can be shown that the roots of (2.5) are always real and that the largest root is always greater than or equal to the minimum of N_A and N_B. The smallest root is always contained in the interval zero to minimum (N_A, N_B), inclusive. Hence we take the left (smallest) root of (2.5) as the estimator of N_{ab}. Note that the Hartley estimator (2.1) with fixed p, unlike the estimator defined by (2.5), does not always fall in the range of feasible values for N_{ab}.

The large sample properties of $\hat{N}_{ab,s}$ are given in the following theorem.

Theorem 1: The sequence of estimators, $\hat{N}_{ab,s}$, defined by (2.5) satisfies

$$\operatorname{Var}(\hat{N}_{ab,s}) = \frac{N_{ab}N_aN_bg_Bg_A}{n_AN_bg_B + n_BN_ag_A} + O(1), \tag{2.6}$$

and

$$E(\hat{N}_{ab,s}) = N_{ab} + O(\tfrac{1}{n}).$$

Thus $\hat{N}_{ab,s}$ has the same limiting variance as the linear estimator (2.1) employing optimal weights. Further the bias of $\hat{N}_{ab,s}$ is of low order compared to the variance of $\hat{N}_{ab,s}$. Based on these results we recommend $\hat{N}_{ab,s}$ as an estimator for N_{ab} when it is judged impractical to identify elements entering the sample from both frames.

For replacement sampling the estimator $\hat{N}_{ab,s}$ reduces to the maximum likelihood estimator presented by Williams [12].

2.2 Estimation of N_{ab} when duplicated elements are identified

Assume that the n'_{ab} sample elements from frame A are compared with the n''_{ab} elements from frame B and it is determined that n_d of these elements are common. The probability that an element in domain ab will be selected in the sample from frame A and from frame B is $(n_A/N_A)(n_B/N_B)$. Hence $n_df_A^{-1}f_B^{-1}$ is an unbiased estimator of N_{ab}. We consider the estimator of N_{ab} constructed as the linear combination of three unbiased estimators:

$$\hat{N}_{ab,1} = pf_A^{-1}n'_{ab} + rf_B^{-1}n''_{ab} + (1 - p - r)f_A^{-1}f_B^{-1}n_d. \tag{2.9}$$

The variance of $\hat{N}_{ab,1}$ is given by

$$\operatorname{Var}\{\hat{N}_{ab,1}\} = p^2C_A + r^2C_B + (1 - p - r)^2 [C_A + C_B + C_{AB}] + 2p(1 - r - p)C_A + 2r(1 - r - p)C_B, \tag{2.10}$$

where

$$C_A = f_A^{-2}\operatorname{Var}\{n'_{ab}\} = g_An_A^{-1}N_aN_{ab},$$

$$C_B = f_B^{-2}\operatorname{Var}\{n''_{ab}\} = g_Bn_B^{-1}N_bN_{ab}$$

and

$$C_{AB} = N_{ab}f_A^{-1}f_B^{-1}g_Ag_B(1 - \frac{1}{N_A} - \frac{1}{N_B} + \frac{N_{ab}}{N_AN_B}) = N_{ab}f_A^{-1}f_B^{-1}g_Ag_B + O(1).$$

Minimizing the variance with respect to p and r we obtain

$$p_o = \frac{C_AC_B + C_{AB}C_B}{C_AC_B + C_AC_{AB} + C_{AB}C_B}$$

and

$$r_o = \frac{C_AC_B + C_{AB}C_A}{C_AC_B + C_AC_{AB} + C_{AB}C_B} . \tag{2.11}$$

Substituting the optimum values of p and r into equation (2.10), we obtain

$$\operatorname{Var}_o\{\hat{N}_{ab,1}\} = \frac{N_{ab}N_aN_bg_Ag_B}{f_Af_BN_aN_b + n_BN_ag_A + n_AN_bg_B} . \tag{2.12}$$

Comparing the variance expression (2.12) with that of $\hat{N}_{ab,H}$, (2.4), we see that identifying the duplicated items reduces the variance by a fraction depending upon the sampling rates and the proportion of the population in the overlap domain. That is,

$$\operatorname{Var}\{\hat{N}_{ab,1}\} = [f_B(1 - \alpha)g_A + f_A(1 - \beta)g_B] \operatorname{Var}\{\hat{N}_{ab,H}\}[f_Af_B(1 - \alpha)(1 - \beta) + f_B(1 - \alpha)g_A + f_A(1 - \beta)g_B]^{-1} . \tag{2.13}$$

We now consider the maximum likelihood estimator of N_{ab}. The probability of obtaining a sample with the given number of elements selected from the overlap domain is given by

$$L\{n'_{ab},n''_{ab},n_d;N_{ab}\} = L(n'_{ab};N_{ab})L(n''_{ab},n_d|n'_{ab};N_{ab})$$

$$= \frac{\binom{N_a}{n_a}\binom{N_{ab}}{n'_{ab}}}{\binom{N_A}{n_A}} \; \frac{\binom{N_b}{n_b}\binom{N_{ab} - n'_{ab}}{n''_{ab} - n_d}\binom{n'_{ab}}{n_d}}{\binom{N_B}{n_B}} .$$

Setting the ratio of $L(n'_{ab},n''_{ab},n_d;N_{ab})$ to $L(n'_{ab},n''_{ab},n_d;N_{ab} - 1)$ equal to one, we find that the value of N_{ab} maximizing the likelihood is given as the solution to the quadratic equation[1]

$$(n_a + n_{ab} + n_b)\hat{N}^2_{ab,m} - [n_aN_B + n_{ab}(N_A + N_B) + n_bN_A - n_an_b]\hat{N}_{ab,m} + n_{ab}N_AN_B = 0, \tag{2.14}$$

where

$$n_{ab} = n'_{ab} + n''_{ab} - n_d.$$

It is possible to show that $N_{ab,m}$ is always given by the left root of (2.14).

Theorem 2: The sequence of maximum likelihood estimators defined by (2.14) satisfies

$$Var\{\hat{N}_{ab,m}\} = \frac{N_{ab}N_aN_bg_Ag_B}{f_Af_BN_aN_b + n_BN_ag_A + n_AN_bg_B} + 0(1)$$

and

$$E\{\hat{N}_{ab,m} - N_{ab}\} = (f_Af_BN_{ab}^2N_aN_bg_Ag_B)/$$

$$\{(f_Af_BN_aN_b + n_BN_ag_A + n_AN_bg_B)$$

$$[-2N_{ab}(n_A - n_B + f_Af_BN_{ab}) + (n_A + f_BN_{ab}$$

$$- f_Af_BN_{ab})N_B + (n_B + f_AN_{ab} - f_Af_BN_{ab})N_A$$

$$- f_Af_BN_aN_b]\} + 0(\frac{1}{n}).$$

3. ESTIMATORS OF THE POPULATION TOTAL

3.1 Duplicated items not identified

We now consider methods of estimating the population total, Y. It is well known that, for example

$$E\{\bar{y}_a | n_a, n'_{ab}, n''_{ab}, n_b\} = \bar{Y}_a.$$

Therefore given an estimator of N_{ab} we consider the estimator

$$\hat{Y}_s = (N_A - \hat{N}_{ab,s})\bar{y}_a + \hat{N}_{ab,s}\bar{y}_{ab,s}$$

$$+ (N_B - \hat{N}_{ab,s})\bar{y}_b, \qquad (3.1)$$

where

$$\bar{y}_{ab,s} = w\bar{y}'_{ab} + (1 - w)\bar{y}''_{ab},$$

$$\bar{y}'_{ab} = \frac{1}{n'_{ab}} \sum_{i=1}^{n'_{ab}} y_i$$ = mean of elements in domain ab selected from frame A,

$$\bar{y}''_{ab} = \frac{1}{n''_{ab}} \sum_{j=1}^{n''_{ab}} y_j$$ = mean of elements in domain ab selected from frame B

and

$$w = \frac{n'_{ab}(1 - f_B)}{n'_{ab}(1 - f_B) + n''_{ab}(1 - f_A)} .$$

Clearly w has been chosen to minimize the variance of $\bar{y}_{ab,s}$. If the finite correction term can be ignored the estimator of $\bar{Y}_{ab}$ reduces to the mean of all $n'_{ab} + n''_{ab}$ elements selected from domain ab.

Note that, given $\hat{N}_{ab,s}$, the estimator $\hat{Y}_s$ is linear in the observations y_i. Since the weights are not a function of the characteristic they apply equally well for all y-characteristics.

Theorem 3: The sequence of estimators defined by (3.1) satisfies

$$Var\{\hat{Y}_S\} = N_a(f_A^{-1} - 1)S_a^2 + [(1 - f_B)f_A +$$

$$+ (1 - f_A)f_B]^{-1}(1 - f_A)(1 - f_B)N_{ab}S_{ab}^2$$

$$+ N_b(f_B^{-1} - 1)S_b^2$$

$$+ (\bar{Y}_{ab} - \bar{Y}_a - \bar{Y}_b)^2 \frac{N_{ab}N_aN_bg_Ag_B}{n_AN_bg_B + n_BN_ag_A} + 0(1) \qquad (3.2)$$

and

$$E\{\hat{Y}_S\} = Y + 0(\frac{1}{n}) ,$$

where

$$S_a^2 = \frac{1}{N_a - 1} \sum_{i=1}^{N_a} (y_{ai} - \bar{Y}_a)^2,$$

$$S_{ab}^2 = \frac{1}{N_{ab} - 1} \sum_{i=1}^{N_{ab}} (y_{ab,i} - \bar{Y}_b)^2$$

and

$$S_b^2 = \frac{1}{N_b - 1} \sum_{i=1}^{N_b} (y_{bi} - \bar{Y}_b)^2.$$

Hartley [7] originally suggested the estimator

$$\hat{Y}_H = \frac{N_A}{n_A}(n_a\bar{y}_a + pn'_{ab}\bar{y}'_{ab})$$

$$+ \frac{N_B}{n_B}(qn''_{ab}\bar{y}''_{ab} + n_b\bar{y}_b),$$

and Lund [10] suggested the modification

$$\hat{Y}_L = f_A^{-1}n_a\bar{y}_a + (pf_A^{-1}n'_{ab} + qf_B^{-1}n''_{ab})\bar{y}_{ab,L}$$

$$+ f_B^{-1}n_b\bar{y}_b, \qquad (3.3)$$

where

$$\bar{y}_{ab,L} = \frac{n'_{ab}\bar{y}'_{ab} + n''_{ab}\bar{y}''_{ab}}{n'_{ab} + n''_{ab}} .$$

The variance of Lund's estimator

$$Var(\hat{Y}_L) = N_a(f_A^{-1} - 1)S_a^2 + (f_A + f_B)^{-2}[f_A(1 - f_A)$$

$$+ f_B(1 - f_B)]N_{ab}S_{ab}^2 + N_b(f_B^{-1} - 1)S_b^2$$

$$+ g_A n_A^{-1} N_{ab} N_a [\bar{Y}_a - p\bar{Y}_{ab}]^2$$

$$+ g_B n_B^{-1} N_{ab} N_b [\bar{Y}_b - q\bar{Y}_{ab}]^2$$

is never greater than the variance of Hartley's estimator.

Lund gave the optimum value of p as

$$p_{oL} = \frac{f_A^{-1}(1-\alpha)\bar{Y}_a + f_B^{-1}(1-\beta)(\bar{Y}_{ab} - \bar{Y}_b)}{[f_A^{-1}(1-\alpha) + f_B^{-1}(1-\beta)]\bar{Y}_{ab}}$$

and suggested the estimator

$$\hat{p}_{oL} = \frac{n_B f_A^{-1} n_a \bar{y}_a + n_A f_B^{-1} n_b (\bar{y}_{ab} - \bar{y}_b)}{[n_B f_A^{-1} n_a + n_A f_B^{-1} n_b]\bar{y}_{ab}} .$$

Substitution of $\hat{p}_{oL}$ into $\hat{Y}_L$ (3.3) gives an estimator of the same form as $\hat{Y}_s$ of equation (3.1).

The bias in Lund's estimator is $O(1)$ while that of $\hat{Y}_s$ is $O(1/n)$. Further the estimator of $\bar{Y}_{ab}$ employed in $\hat{Y}_s$ is more efficient than that in $\hat{Y}_L$ for nonreplacement sampling. Also the linear form of (3.1) is clearly a computational advantage in large scale surveys.

3.2 Duplicate Items Identified

If the items entering the sample from both frame A and frame B are identified, the estimator $\hat{N}_{ab,m}$ is superior to $\hat{N}_{ab,s}$. Also the mean of the distinct units is superior to $\bar{y}_{ab,s}$ as an estimator of the mean of the overlap domain. Therefore the estimator

$$\hat{Y}_m = (N_A - \hat{N}_{ab,m})\bar{y}_a + \hat{N}_{ab,m}\bar{y}_{ab} + (N_B - \hat{N}_{ab,m})\bar{y}_b,$$

where

$$\bar{y}_{ab} = \frac{1}{n_{ab}} \sum_{i=1}^{n_{ab}} y_i = \text{mean of distinct units}$$

and

$$n_{ab} = n'_{ab} + n''_{ab} - n_d,$$

clearly has smaller variance than $\hat{Y}_s$. We have

$$\mathrm{Var}(\hat{Y}_m) = N_a f_A^{-1}(1-f_A)S_a^2 + (1 - f_A - f_B + f_A f_B)$$

$$(f_A + f_B - f_A f_B)^{-1} N_{ab} S_{ab}^2 + N_b f_B^{-1}(1-f_B)S_b^2$$

$$+ \frac{(\bar{Y}_{ab} - \bar{Y}_a - \bar{Y}_b)^2 N_{ab} N_a N_b (1-f_A)(1-f_B)}{[f_A f_B N_a N_b + n_B N_a(1-f_A) + n_A N_b(1-f_B)]}$$

$$+ O(1).$$

4. ESTIMATION FOR GENERAL SAMPLING PROCEDURES

We now relax the assumption of simple random sampling in each frame. We assume that sampling is such that unbiased estimators of totals whose error is $O_p(n^{1/2})$ and estimators of the variances whose error is $O_p(n^{1/2})$ of these estimators are available. Let

- $\hat{Y}_a$ be an unbiased estimator of the total for domain a constructed from the sample of frame A,
- $\hat{Y}_b$ be an unbiased estimator of the total for domain b constructed from the sample of frame B,
- $\hat{Y}'_{ab}$ be an unbiased estimator of the total of domain ab constructed from the sample of frame A, and
- $\hat{Y}''_{ab}$ be an unbiased estimator of the total of domain ab constructed from the sample of frame B.

We assume that the observational unit is the same for the sample selected from frame A or for that selected from frame B[2] and let

- $\hat{N}'_{ab}$ be an unbiased estimator of the number of observational units in domain ab constructed from the sample of frame A,
- $\hat{N}''_{ab}$ be an unbiased estimator of the number of observational units in domain ab constructed from the sample of frame B,
- $\hat{N}_a$ be an unbiased estimator of the number of observational units in domain a constructed from the sample of frame A, and
- $\hat{N}_b$ be an unbiased estimator of the number of observational units in domain b constructed from the sample of frame B.

In the previous sections an estimator of N_{ab} was constructed from the two unbiased estimators $\hat{N}'_{ab}$ and $\hat{N}''_{ab}$. Similarly an estimator of $\bar{Y}_{ab}$ was constructed from the two unbiased estimators $\bar{y}'_{ab}$ and $\bar{y}''_{ab}$. Since the expected value of $\bar{y}'_{ab}$ and $\bar{y}''_{ab}$ given $\hat{N}'_{ab}$, $\hat{N}'_{ab} \neq 0$, and $\hat{N}''_{ab}$, $\hat{N}''_{ab} \neq 0$, is $\bar{Y}_{ab}$, it was possible to treat the estimation of N_{ab} and $\bar{Y}_{ab}$ separately.

If the sampling is other than simple random, $\hat{N}'_{ab}$ and $\hat{N}''_{ab}$ may be correlated with $\bar{y}'_{ab} = \hat{Y}'_{ab}/\hat{N}'_{ab}$ and $\bar{y}''_{ab} = \hat{Y}''_{ab}/\hat{N}''_{ab}$. Therefore, to be efficient, it is necessary to use all auxiliary information in the estimation of, for example, Y. This suggests the estimator

$$\tilde{Y}_r = \hat{Y}_a + \hat{Y}_B + \beta_1(\hat{N}'_{ab} - \hat{N}''_{ab}) + \beta_2(\hat{Y}'_{ab} - \hat{Y}''_{ab}), \qquad (4.1)$$

where

$$\hat{Y}_B = \hat{Y}_b + \hat{Y}''_{ab}.$$

The optimal $\hat{\beta}' = (\hat{\beta}_1, \hat{\beta}_2)$ is given by

$$\begin{bmatrix}\hat{\beta}_1\\ \hat{\beta}_2\end{bmatrix} = \begin{bmatrix}\hat{\mathrm{Var}}(\hat{N}'_{ab} - \hat{N}''_{ab}) & \hat{\mathrm{Cov}}(\hat{N}'_{ab} - \hat{N}''_{ab}, \hat{Y}'_{ab} - \hat{Y}''_{ab})\\ \text{sym.} & \hat{\mathrm{Var}}(\hat{Y}'_{ab} - \hat{Y}''_{ab})\end{bmatrix}^{-1}$$

$$\begin{bmatrix}-\hat{\mathrm{Cov}}(\hat{Y}_a, \hat{N}'_{ab}) + \hat{\mathrm{Cov}}(\hat{Y}_B, \hat{N}''_{ab})\\ -\hat{\mathrm{Cov}}(\hat{Y}_a, \hat{Y}'_{ab}) + \hat{\mathrm{Cov}}(\hat{Y}_B, \hat{Y}''_{ab})\end{bmatrix} . \qquad (4.2)$$

A consistent estimator of the variance of the estimator is

$$\hat{\mathrm{Var}}(\tilde{Y}_r) = \hat{\mathrm{Var}}(\hat{Y}_a) + \hat{\mathrm{Var}}(\hat{Y}_B) +$$

$$+ \hat{\beta}_1[\hat{\mathrm{Cov}}(\hat{Y}_a, \hat{N}'_{ab}) - \hat{\mathrm{Cov}}(\hat{Y}_B, \hat{N}''_{ab})]$$

$$+ \hat{\beta}_2[\hat{\mathrm{Cov}}(\hat{Y}_a, \hat{Y}'_{ab}) - \hat{\mathrm{Cov}}(\hat{Y}_B, \hat{Y}''_{ab})]. \qquad (4.3)$$

The estimator (4.1) is clearly recognizable as a multiple regression estimator. Hartley proposed the estimator

$$\hat{Y}_H = \hat{Y}_a + \hat{Y}''_{ab} + \hat{Y}_b + p(\hat{Y}'_{ab} - \hat{Y}''_{ab}). \qquad (4.4)$$

Hartley's estimator is inefficient relative to $\tilde{Y}_r$ if the partial correlation between $\hat{Y}_a + \hat{Y}_B$ and $\hat{N}'_{ab} - \hat{Y}''_{ab}$ is not zero.

It is also clear that if other y-characteristics are observed in the survey it may be possible to further decrease the variance of the estimator (4.1) by including other unbiased estimators of zero in the regression estimator.

The estimator (4.1) is not linear in y since $\hat{\beta}_2(\hat{Y}'_{ab} - \hat{Y}''_{ab})$ is not a linear function of y. However it is 'nearly linear' in the sense that given the differences $\hat{N}'_{ab} - \hat{N}''_{ab}$ and $\hat{Y}'_{ab} - \hat{Y}''_{ab}$ the estimator is linear in y. Thus in a large scale survey one might be willing to choose a few y values to serve as auxilliary variables for all computations. This would simplify the computations and guarantee numerical consistency between tables.

Acknowledgment

This research was partially supported by the Bureau of the Census through the Joint Statistical Agreement J.S.A. 72-4.

FOOTNOTES

1. In order to simplify our presentation we ignore the fact that the estimate of N_{ab} should be an integer.

2. This assumption can be modified so that conclusions are applicable for cases wherein the observational units are different.

REFERENCES

[1] Bershad, Max A., "A Sample Survey of Retail Stores," *Sample Survey Methods and Theory*, Vol. 1, New York: John Wiley and Sons, Inc, 1953, 516-558.

[2] Bryant, Edward C. and King, Donald W., "Estimation from Populations Identified by Overlapping Sample Frames," Unpublished paper presented at the American Statistical Association meeting, Palo Alto, California, 1960.

[3] Cochran, Robert S., "Theory and Application of Multiple Frame Surveys," Ph.D. dissertation, Iowa State University Library, Ames, 1965.

[4] Cochran, Robert S., "The Estimation of Domain Sizes When Sampling Frames are Interlocking," Proceedings of the Social Science Section of the American Statistical Association meeting, Washington, D. C., 1967.

[5] Cochran, William G., *Sampling Techniques*, 2nd ed., New York: John Wiley and Sons, Inc., 1963.

[6] Fuller, W. A., "Properties of Some Estimators for the Errors in Variables Model. Submitted for Publication, 1972.

[7] Hartley, H. O., "Multiple Frame Surveys," Proceedings of the Social Science Section of the American Statistical Association meeting, Minneapolis, Minnesota, 1962.

[8] Horvitz, D. G. and Thompson, D. J., "A Generalization of Sampling Without Replacement from a Finite Universe," *Journal of the American Statistical Association*, 47, 1952, 663-685.

[9] King, Donald W., "Variance Estimation in Populations Identified by Multiple Sampling Frames," M.S. Thesis, University of Wyoming Library, Laramie, 1960.

[10] Lund, Richard E., "Estimators in Multiple Frame Surveys," Proceedings of the Social Science Section of the American Statistical Association meeting, Pittsburg, Pennsylvania, 1968.

[11] Steinberg, J., "A Multiple Frame Survey for Rare Population Elements," Proceedings of the Social Science Section of the American Statistical Association meeting, Philadelphia, Pennsylvania, 1965.

[12] Williams, Robert E., "Estimation of Overlapping Strata Boundaries," M.S. Thesis, University of Wyoming Library, Laramie, 1957.

SENSITIVITY OF GENERATION TABLES OF WORKING LIFE FOR MEN TO DIFFERENT PROJECTIONS OF LABOR FORCE PARTICIPATION RATES AND MORTALITY RATES

Howard N. Fullerton, Bureau of Labor Statistics

Tables of working life currently have three principal uses. First, measures developed in the tables are used to estimate the entrants to and withdrawal from the labor force. Second, as an hypothetical work force, they may be used as models whereby the effects of various social and demographic factors may be studied. Finally, they are used in court cases to aid in the determination of the potential loss of earnings incurred in cases of work-related deaths or injuries.

Generation tables of working life for men were first developed by this writer earlier this year. (3) Their potential usefulness had been suggested in a Canadian study, and our own comparison of such tables with conventional period tables has confirmed their usefulness. (2) In particular, generation tables have been found to be more realistic than period tables for measuring labor force flows and for certain kinds of through-time analyses and model construction. The superiority of the generation table was seen in each of the four measures most commonly utilized: the expectation of work-life remaining, the expected length of retirement, the rate of retirement, and the rate of accession to the labor force.

As generation tables of working life come to be used more frequently, there is a need to see how sensitive they are to changes in rates of mortality and labor force participation. Because the relations of the events affecting labor force entry and withdrawal to age and mortality have not been described in theoretical terms, it was not feasible to carry out a theoretical analysis as has been done by Keyfitz for stable populations. (4) Instead, an empirical analysis has been attempted. There are difficulties in generalizing from an empirical study. It is also difficult to know if the postulated changes in the mortality rates are of the same order of magnitude as the postulated changes in the labor force rates. However, the rates postulated are considered to be within the range of probable occurrences. Although this paper concentrates on the effects of these factors on generation tables, most of the conclusions also apply to period tables.

This study was made for two work-cohorts of men, those entering the labor force in 1960 and in 1970. For each cohort, two assumptions concerning labor force participation rates in the future were made. The first was that the set of rates observed for men in 1970 would continue unchanged. This has been termed the "constant" labor force participation rate (LFPR) assumption. A second assumption was that the age-specific labor force participation rates would continue the gradual declines which have been observed, but that the change would be less for each successive period, ending with constant rates after 2020. Three sets of mortality rates were used. The first two were the low and high projections of mortality to 2000 made by the Social Security Administration (2) and the third was an assumption that mortality would continue at the 1970 level after 1970. The Social Security Administration (SSA) low projection has the highest expectation of life, and the 1970 mortality assumption has the lowest, while the SSA high projection falls in between. The latter assumption is the one used by the Bureau of the Census in its recent projections.

Thus, for each cohort, there are six tables of working life (or sets). In tables 1 through 4, these measures are arranged in the same order, first the three sets for the constant labor force assumption, then the three sets for the declining labor force assumption. The order of the three sets within the particular labor force assumption is in order of lowest to highest mortality. For convenience in reference, we shall refer to a particular set within a cohort by its column number, i.e., 1960 (5) would be the table formed by the 1960 cohort with the declining labor force assumption and SSA high mortality, and 1960 (6) would be the table formed with the declining labor force assumption and 1970 mortality.

The first measure to be investigated for sensitivity is the expectation of work-life remaining (table 1). For each of the cohorts, and under both projections of labor force participation, the higher the mortality, the lower the expectation of work-life. Also for each mortality assumption, the declining labor force participation has a lower expectation of work-life. Three points can be explored in this connection. First, the ew_x is less sensitive than the expectation of life remaining, e_x, to changes in mortality projections. For example, for the 1970 cohort, assuming constant labor force rates, the expectations of life at age 16 for each group were (in order of increasing mortality) 56.8, 55.0, and 53.3, a range of 3.6 years. The expectations of work-life for the same age were 46.8, 46.1, and 45.2 years, a range of 1.6 years. For the other labor force assumption for the 1970 cohort, and for the 1960 cohort, the same observation applies.

Second, within the 1960 cohort, the expectation of work-life is almost the same for both columns (3) and (4) until age 55. This gives some indication of the difference between the two projections of labor force participation. The highest work-life expectancy in the declining labor force rate group is at the same level with the lowest in the constant labor force rates group. The correspondence is not so strong in the 1970 cohort, but the work-life remaining for 1970 (4) for this cohort is only slightly greater than for 1970 (3).

The effect of changes in the labor force participation rates on the expectation of work-life remaining may be obtained by comparing the measure for a given mortality assumption under each of the labor force projections for a cohort, i.e. by comparing columns (1) and (4). Until age 65, there is a difference of about a year in ew_x. For the 1960 cohort, the difference between columns (1) and (4) is 1.1 years until age 45, and for columns (2) and (5), one year. (The 1970 mortality rates are not as smooth, but the differences are either 1.0 or 0.9 years.) After age 45, the differences are slightly lower, 0.9 or 0.8 years. At age 65, the effect of being well beyond the end of the period of projection (2000 and 2020 for mortality and labor force) override the differences between the cohort and the low expectancy of work-life and result in lower differences, either 0.6 or 0.5 years. However, this is a greater percent of the remaining work-life. At age 70, each of the cohorts and labor force groups has the same expectation of work-life for the same mortality assumption: 6.9 years for SSA low, 6.6 years for SSA high, and 6.3 years for 1970 mortality. This is because work-life for most men ends before age 70, and it also reflects the effects noted for age 65.

The impact of the two different labor force assumptions depends on the cohort. For the 1960 cohort, there is a difference of about a year between the expectation of work-life for the highest and the lowest mortality assumptions until age 45, when the work-life remaining is less than half that for age 16. For the 1970 cohort, the difference between the work-life remaining for the high and the lowest mortality assumptions is about one and a half years and it declines slowly through age 60. Because both mortality and labor force participation are the same for all groups for ages 16-25 for the 1960 cohort, the pattern for that cohort differs from the 1970 cohort for the younger ages. These observations lead to the conclusion that when it is necessary to change either the mortality or labor force participation assumptions when projecting the remainder of a cohort's work experience, the work-life expectancy will be changed by a constant for those younger ages for which the data were not projected. The changes in work-life expectancy that have been observed in this exercise are of the same order as the 1950 to 1960 changes in period tables.

The next measure to be discussed is the expectation of "retirement" (table 2). In the context of this paper, "retirement" refers to a worker's final withdrawal from the labor force, not to participation in any formal program providing annuity benefits. Looking down any column, one observes that there is little change in the figures until age 65. The range in most cases is 0.3 years. The two Social Security Administration projections give a maximum expectation of retirement at age 35, with a relative maximum at age 55. These mortality projections have a relative minimum in the rates at the ages 25-30, and this combined with the maximum labor force participation rate for the following age group, causes the peak expectancy of retirement at age 35. The 1970 mortality rates are even lower at ages 25-30, changing the pattern enough that there is only a relative maximum at age 55. The appearance of maximums at age 55 presumably results from the effects of soon-to-occur labor force withdrawals and greater increases in mortality.

Turning to the impact of the labor force participation rates upon the expectation of retirement, we note that the differences between the same mortality assumptions under the two labor force participation rates projected for each cohort are almost constant until the retirement ages, with only a drop of 0.2 years from age 16 to 60. This is different from what was observed with work-life expectancy. Further, the difference is a smaller percent at the older ages, the opposite of the behavior of the expectation of work-life.

A change in the mortality assumption has a greater effect on the expected duration of retirement than on the expected length of working life. The difference between the expectation of retirement under the highest and the lowest mortality is about two years. It is here that another dissimilarity between the two labor force rates can be observed. For both cohorts, the 1970 labor force rates provide a smaller difference than the declining labor force rates, but the 1960 cohort has slightly smaller difference between the highest and lowest mortality patterns within a particular labor force projection.

Continuing the examination of the effects of different labor force and mortality patterns on mortality, we turn to the retirement rates (table 3). With the possible exception of ages 65-70, mortality assumptions that vary as much as those in the present study do not significantly change the retirement rates. This is an interesting discovery, since retirement at ages before 55 probably reflects forced withdrawal due to illness or injury more than voluntary action. One would therefore have expected to find some sensitivity to mortality.

The retirement rates, however, are more sensitive to varying labor force projections. For ages 40-45 through 65-70, the declining labor force participation rates result in retirement rates that are 22 to 28 percent higher than the constant labor force rates for the same mortality rate in the same cohort. The difference in the assumed participation rates in this case could result in a sizeable over-(or under-) statement of the labor force, if used to estimate retirements or replacement needs.

The final measure under consideration, the accession rate, is part of the actual work experience of the 1960 cohort, and does not vary. Data for the 1970 cohort are present in table 4. As in the case of the retirement rates, the accession rates are not sensitive to mortality. Since the accession rate is discounted for mortality, this is the behavior expected. The relation between the two labor force assumptions results in the sensitivity being different for different ages. The maximum labor force participation rate under the declining labor force projection is almost the same as under the constant labor force, but the route to the maximum takes a different path. For the declining rates, there are fewer accessions at ages 16-20, and more during ages 20-25. There are fewer accessions under the declining pattern for ages 25-30. As a result, there are slightly less (2 percent) accessions for the younger ages, and more accession at ages 20-25 (7 percent). However, the rates are not of the same order for successive age groups. If the accession rates for ages 16-20 and 20-25 are combined, there is little difference, (less than one percent), although the declining labor force always has the lower rate:

Assumed mortality rates	Assumed labor force participation rate	
	Constant	Declining
SSA 1	568.1	563.3
SSA h	567.5	562.7
1970	564.0	559.3

Although the expectation of work-life and of retirement are of interest as indicators of future work patterns, within the Bureau of Labor Statistics generation working life tables are primarily used as sources for estimates of accession and separation rates (including death and retirement). These measures are used to estimate gross flows into and out of the work force and are defined so that they may be applied to the entire population (accession rates) or labor force (separation rates). The retirement rate is used to estimate occupational replacement needs (5). The age distribution of persons employed in a given occupation is multiplied by the corresponding separation rates from the working life table. Until period working life tables were constructed recently from projections of the labor force and mortality, the projected separations were based on the most recent working life table, which might refer to a period almost a decade in the past. For projecting to 1985, the technique contemplated is to use one set of rates to go from the age distribution for an occupation reported in the 1970 Census directly to the separations in 1985 (with an adjustment for the changes expected in the age distribution of the labor force). For this purpose, a generation rate rather than some combination of the 1975 and 1980 period retirement rates would be more realistic.

Summary

The purpose of this paper has been to investigate the impact of various projections of mortality and the labor force on selected measures of labor force activity derived from generation tables of working life for men. The results are summarized in table 5.

Table 1. Expectation of work-life remaining

Cohort and mortality	1960 Cohort						1970 Cohort					
	Constant lfpr			Declining lfpr			Constant lfpr			Declining lfpr		
Age	SSAL (1)	SSAH (2)	1970 (3)	SSAL (4)	SSAH (5)	1970 (6)	SSAL (1)	SSAH (2)	1970 (3)	SSAL (4)	SSAL (5)	1970 (6)
16----------	46.5	46.2	45.4	45.4	45.2	44.4	46.8	46.1	45.2	45.7	45.1	44.3
20----------	42.8	42.3	41.6	41.7	41.3	40.7	43.1	42.4	41.5	42.0	41.4	40.6
25----------	38.1	37.6	37.0	37.0	36.6	36.0	38.4	37.7	37.0	36.3	36.7	36.0
30----------	33.5	32.9	32.3	32.4	31.9	31.4	33.7	33.0	32.3	32.6	32.0	31.4
35----------	28.8	28.2	27.7	27.7	27.3	26.8	28.9	28.3	27.7	27.8	27.3	26.7
40----------	24.3	23.8	23.3	23.3	22.8	22.4	24.4	23.8	23.3	23.4	22.8	22.4
45----------	19.9	19.4	19.0	19.0	18.5	18.1	20.0	19.5	19.0	19.0	18.6	18.1
50----------	15.7	15.3	14.9	14.8	14.4	14.1	15.8	15.3	14.9	14.8	14.4	14.1
55----------	11.7	11.4	11.0	10.8	10.5	10.2	11.7	11.3	11.0	10.8	10.5	10.1
60----------	8.2	8.0	7.6	7.3	7.2	6.8	8.2	7.9	7.6	7.3	7.0	6.8
65----------	5.8	5.6	5.4	5.1	5.0	4.8	5.8	5.6	5.4	5.1	4.9	4.8
70----------	6.9	6.6	6.3	6.9	6.6	6.3	6.9	6.6	6.3	6.9	6.6	6.3

Table 2. Expectation of retirement

Cohort and mortality	1960 Cohort						1970 Cohort					
	Constant lfpr			Declining lfpr			Constant lfpr			Declining lfpr		
Age	SSAL	SSAH	1970	SSAL	SSAH	1970	SSAL	SSAH	1970	SSAL	SSAH	1970
	(1)	(2)	(3)	(4)	(5)	(6)	(1)	(2)	(3)	(4)	(5)	(6)
16----------	9.9	8.6	7.9	11.0	9.6	8.9	10.0	8.9	8.0	11.1	9.9	8.9
20----------	9.9	8.8	8.0	11.0	9.8	8.9	10.0	8.9	8.0	11.1	9.9	8.9
25----------	10.1	8.9	8.1	11.2	9.9	9.1	10.1	9.1	8.1	11.2	10.1	9.1
30----------	10.1	9.1	8.2	11.2	10.1	9.1	10.2	9.1	8.2	11.3	10.1	9.1
35----------	10.2	9.2	8.3	11.3	10.1	9.2	10.3	9.2	8.3	11.4	10.2	9.3
40----------	10.1	9.0	8.2	11.1	10.0	9.1	10.2	9.1	8.2	11.2	10.1	9.1
45----------	10.1	9.1	8.2	11.0	10.0	9.1	10.2	9.0	8.2	11.2	9.9	9.1
50----------	10.1	9.0	8.2	11.0	9.9	9.0	10.1	9.1	8.2	11.1	10.0	9.0
55----------	10.2	9.1	8.4	11.1	10.0	9.2	10.2	9.2	8.4	11.1	10.0	9.3
60----------	10.0	9.0	8.4	10.9	9.8	9.2	10.0	9.1	8.4	10.9	10.0	9.2
65----------	8.9	8.2	7.7	9.6	8.8	8.3	8.4	8.2	7.7	9.6	8.9	8.3
70----------	4.8	4.4	4.2	4.8	4.4	4.2	4.8	4.4	4.2	4.8	4.4	4.2

Table 3. Retirement rates 1/, per thousand

Cohort and mortality	1960 Cohort						1970 Cohort					
	Constant lfpr			Declining lfpr			Constant lfpr			Declining lfpr		
Age	SSAL	SSAH	1970	SSAL	SSAH	1970	SSAL	SSAH	1970	SSAL	SSAH	1970
	(1)	(2)	(3)	(4)	(5)	(6)	(1)	(2)	(3)	(4)	(5)	(6)
30-35-------	0.1	0.1	0.1	0.5	0.5	0.4	0.0	0.0	0.1	0.0	0.1	0.1
35-40-------	1.4	1.4	1.3	1.7	1.6	1.7	1.4	1.3	1.3	1.9	1.9	1.8
40-45-------	2.0	2.1	2.1	2.8	2.9	2.9	2.0	2.0	2.1	2.8	2.8	2.9
45-50-------	3.6	3.3	3.4	4.4	4.2	4.3	3.6	3.6	3.4	4.4	4.4	4.3
50-55-------	4.9	6.4	5.3	6.1	7.5	6.4	5.0	5.0	5.3	6.1	6.1	6.3
55-60-------	17.8	18.4	17.2	21.9	22.5	21.3	17.6	17.6	17.2	21.8	21.7	21.3
60-65-------	57.4	53.7	58.8	73.4	69.7	75.0	57.7	58.0	58.8	74.0	74.5	75.4
65-70-------	160.5	159.9	158.9	196.2	195.1	194.5	160.5	159.6	158.8	197.0	196.0	195.2

1/ Retirement refers to final withdrawal from the labor force.

Table 4. Accession rate, per thousand, 1970 cohort

Labor force:	Constant lfpr			Declining lfpr		
Mortality: Age group	SSAL	SSAH	1970	SSAL	SSAH	1970
	(1)	(2)	(3)	(4)	(5)	(6)
16-20------------------------	486.0	485.8	484.3	474.6	474.4	473.0
20-25------------------------	82.1	81.7	79.7	88.7	88.3	86.3
25-30------------------------	14.0	13.4	11.7	12.2	11.6	9.9

lfpr: labor force participation rate
Mortality: SSAL: Social Security Administration low projection
SSAH: Social Security Administration high projection
1970: Rates remain at the 1970 level after 1970.

Table 5. Sensitivity of various measures of labor force activity

	Measures from table of working life			
	$\mathring{e}w_x$ 1/	$\mathring{e}r_x$ 2/	$5S^r_x$ 3/	5^A_x 4/
A. Sensitivity to mortality rate				
16-30----------------------------------	M	--	--	L
30-55----------------------------------	M	L	L	--
55 and over----------------------------	L	L	L	--
B. Sensitivity to labor force participation rate				
16-30----------------------------------	L	--	--	L
30-55----------------------------------	L	M	M	--
55 and over----------------------------	L	M	M	--

L: low sensitivity
M: moderate sensitivity

1/ $\mathring{e}w_x$ = The expected length of working life remaining at age x.
2/ $\mathring{e}r_x$ = The expected duration of retirement of retirees at age x.
3/ $5S^r_x$ = The rate of separations from the labor force for reasons other than death at age x to x+5, per 1,000.
4/ 5^A_x = The rate of accession to the labor force at age x to x+5, per 1,000.

References

(1) Francisco Bayo, "United States Population Projections for OASDHI Cost Estimates," Actuarial Study No. 62, U.S. Department of Health, Education, and Welfare, Social Security Administration, December 1966.

(2) Frank T. Denton and Sylvia Ostry, "Working-Life Tables for Canadian Males," 1961 Census Monograph Programme, Dominion Bureau of Statistics, Ottawa, Canada, 1969.

(3) Howard N. Fullerton, "A New Type of Working Life Table for Men," Monthly Labor Review, July 1972, pp. 20-27.

(4) Nathan Keyfitz, "Linkages of Intrinsic to Age-Specific Rates," Journal of the American Statistical Association, June 1971, pp. 275-281.

(5) "Estimating Replacement Needs," Tomorrow's Manpower Needs, BLS Bulletin No. 1606, U.S. Department of Labor, Bureau of Labor Statistics, February 1969, pp. 47-58.

SOME DYNAMIC INDICATORS OF NEIGHBORHOOD CHANGE*

Stephen Gale, Northwestern University

Eric G. Moore, Northwestern University

Several recent studies of social and residential mobility have argued that the descriptive properties of discrete space stochastic matrix models provide extremely useful constructs on which to base explanatory inferences concerning both the distribution and transformation of local population characteristics. Furthermore, when these models are time homogeneous (even if only in the short run), it can be shown that they may also be used as a basis for projections of future distributions. Unfortunately, owing to limitations in the availability of data and very weak conceptualizations of most city and regional planning situations, there have been few attempts to employ these models in development of such indicators for use in social planning and decision-making. This has been especially true in those cases when time homogeneity is not demonstrably present and sensitive indicators of change are required. In this paper, we will outline an approach using stochastic matrix models to generate indicators of social and residential changes in small areas. More specifically, we show that by using the types of detailed longitudinal data files currently being developed in several cities in the United States and Europe, selected parameters can be used to present summary descriptions of associated short-range shifts in social structure and neighborhood composition as well as a formal framework for developing indices of structural change.

1. *Social Indicators*

As they are usually conceived of, social indicators are sets of measures which summarize certain (functionally specified) observable properties of a complex system and which may be compared with identical measures for the same properties of other (perhaps ideal) systems. Where the relevant properties of the system are known to be homogeneous in their behavior, there is clearly no need for other than descriptive indicators. Where the system is heterogeneous, alternative measures of observable characteristics must serve to indicate the state of the system and its pattern of change or deviance.[1]

The problem in developing social indicators is to find those measures which accurately and intercomparably describe the states of the social system, their patterns of change, and those instrumental factors which can potentially be used to regulate these changes. This problem, itself, can be regarded as made up of four related subproblems: (a) issues concerned with the relationship between the types of indicators and their uses (i.e., functional questions); (b) the formulation of conceptual models; (c) data availability; and (d) the difficulties encountered in devising uniformly meaningful (and interpretable) measures or scales. The first and last of these problems (a and d) require much deeper philosophical and mathematical analysis than can be provided here and thus neither of these issues will be treated in this paper. We will concern ourselves primarily with the development of a particular functional class of dynamic indicators (i.e., for local area population changes) and, more specifically, with the kinds of conceptual models and data required for their development.

1.1 A Conceptual Model

The proposed indicators of local population change require a conceptual model combining both neighborhood status and change components; this model should include not only measures of current conditions but also parameters which reflect the size, direction, and relationships among the variables which affect shifts in these conditions. Of the several formal models which have been employed in studying local area mobility, perhaps the most significant development has been the use of discrete space stochastic process models based on individual-level (disaggregated) data. Originally, these models seemed appealing mainly because it appeared that the characteristics of the residential location and relocation process bore striking similarities to the processes of social mobility and age-specific population change, both of which have been examined using stochastic models (e.g., Bartholomew (1967), White (1970)). More recently, additional epistemological and methodological justification has been provided for the use of this class of models through an examination of the criteria for explanation and explanatory evidence (Fisher (1960), Gale (1971), Gale (1973)). These models will thus be employed as the formal basis of our subsequent development of social indicators.

Since many of the analytical consequences of using discrete space stochastic matrices as models of residential location behavior are predicated on a close correspondence between the desired theory and the definitions and classification schemes employed, it is particularly important that these aspects of the model be fully comprehensible. In this regard, it is useful to have a notation which both illuminates the definitional and classificatory issues and provides an explicit means for comparing results obtained from several alternatives. Such a notation has been described previously in terms of a general finite-dimensional cross-classification table K^N where $N=\{n_1,\ldots,n_N\}$ is the set of properties or conditional predicates and $K^{n_i}=\{k_1^{n_i},\ldots,k_{K_i}^{n_i}\}$ is the set of classifications associated with each of the properties, n_i $(i=1,\ldots,N)$ (Gale (1972)).[2]

Using this framework, the specific conceptual model of neighborhood change is defined as a K^6 time-dependent contingency table which, in order to satisfy both planning needs and the desire to be consistent with existing theory, includes information on geographic area, the attributes of households, and dwelling-unit type. The model is thus defined by $N=\{n_1,n_2,n_3,n_4,n_5,n_6\}$ and $K^{n_i}=\{k_1^{n_i},\ldots,k_{K_i}^{n_i}\}$ $(i=1,\ldots,6)$. n_1 is geographic area at time t, n_2 is geographic area at t+1, n_3 is type of dwelling unit at t, n_4 is type of dwelling unit at t+1, n_5 is household-type at t, and n_6 is household-type at t+1; K^{n_1} and K^{n_2} represent the classification of $K_1=K_2=n$ mutually exclusive and exhaustive subareas, K^{n_3} and K^{n_4} represent the classification of the $K_3=K_4=s$ mutually exclusive and exhaustive dwelling-unit types; and K^{n_5} and K^{n_6} represent the classification of $K_5=K_6=m$ mutually exclusive and exhaustive household-type subgroups. If we now define the allocation of household types to dwelling-unit types as the *occupancy pattern* at time t (e.g., the set of n x s x m submatrices, A(t), corresponding to the dimensions n_1, n_3, and n_5), we may then designate the focus of our attention as being on both the nature of the occupancy patterns, A(t), and the matrices which transform A(t) to A(t+1).

1.2 Data

The choice of a generalized cross-classification model as the conceptual framework for the development of dynamic social indicators does more than specify the form of the model with which we will work; it also strongly influences the kinds of data required to estimate the parameters on which the indices are based. For example, by a judicious selection of class boundaries so as to coincide with the units on which federal and state agencies collect information, the n x s x m matrix representing the overall occupancy pattern can be obtained directly from published records. However, where alternative classifications and definitions are desired and where the matrix which transforms one state description to another is of interest, more finely disaggregated (i.e., individual level) data are generally required.[3]

For the present purposes, the combination of model-specific and functional requirements strongly implies that data be collected on the level of individual units. For geographic areas this requires a specific spatial referencing system (e.g., coordinate designations); for dwelling-unit types, detailed information on the size and condition of each unit; and for household-type, detailed information on characteristics such as family structure, age, race, and economic condition. Moreover, since the conceptual model with which we will be working involves time dependencies, the data collected must also include information which allows inferences on changes in status at intervals which are appropriate to the problem in hand.

Clearly, the type of data we require is not generally available in the United States (although it is available in some European countries). However, under funding from federal, state, and local sources, one city has recently developed a set of data on which to base the types of dynamic indicators in which we are interested; these data are currently available for both 1971 and 1972. Locational information is given by the address of the dwelling as well as a household identification number. By using address matching procedures it is thus possible to extract information concerning dwelling-specific occupancy changes as well as removals (e.g., demolitions) and additions (e.g., new construction).[4]

2. *A Representational Model of Neighborhood Change and Some Classes of Dynamic Indicators*

We now consider a specific form of the conceptual model described in section 1.1. In particular, we examine the form of the 6-way matrix which transforms A(t) into A(t+1). To this simple model we also wish to add "birth" (e.g., new construction) and "death" (e.g., demolition) components. Moreover, since the size of this 6-way matrix would render direct analysis impracticable in most cases and since our main interest is on the shifts in occupancy patterns for given sub-areas, we will also modify the model such that we consider only area-specific occupancy patterns--i.e., the transformation from ${}_iA(t)$ to ${}_iA(t+1)$ (i=1, ...,n).[5] More specifically, we define the area-specific occupancy matrices as the set ${}_iA(t)=\{{}_ia_{jk}(t)\}$ with marginals ${}_ia_H(t)=\{{}_ia_{.k}(t)\}$ (the vector of household composition of area i at t) and ${}_ia_D(t)=\{{}_ia_{j.}(t)\}$ (the vector of dwelling unit composition in area i at t).

Planning interest is usually focused on the behavior of either ${}_ia_H(t)$ or ${}_ia_D(t)$. For present purposes we will focus on ${}_ia_H(t)$. The basic argument to be pursued is twofold:

i) that the transformation from ${}_ia_H(t)$ to ${}_ia_H(t+1)$ is dependent on the specific occupancy pattern ${}_iA(t)$ and therefore we need to develop measures of the *differential* contribution of different dwelling-unit types to changes in the household composition; and

ii) that the transformation from ${}_ia_H(t)$ to ${}_ia_H(t+1)$ possesses three components and therefore we need to develop measures of the differential contribution of each of these sources to changes in household composition. The set of computed measures for each of these decompositions would then provide the basic set of descriptive indicators of the process of change which can be used as input to planning and policy decisions. (Note that each set of such measures is definition and classification specific.)

2.1 A Representational Model of Neighborhood Change

Three classes of events may be regarded as leading to changes in ${}_iA(t)$:

i) decrements to the stock due to demolition, abandonment, dwelling unit combination and division; ii) additions to the stock due to new construction, dwelling unit combination and division; iii) changes in the household and dwelling unit characteristics of dwelling units whose physical identity remains the same during the interval (t+1-t). This class of events is further split into two components. We first define an *occupancy transfer* as the replacement of one household by another during the interval (t+1-t) and then define the two events as a) changes in household and dwelling characteristics for those dwelling units experiencing an occupancy transfer; and b) changes in household and dwelling characteristics for dwelling units *not* experiencing an occupancy transfer.

The following notation will be used to describe the above conditions:

${}_iD(t)=\{{}_id_{jk}(t):j=1,\ldots,s,k=1,\ldots,m\}$, the matrix of decrements to the occupancy classes jk

${}_iV(t)={}_iA(t)-{}_iD(t)=\{{}_iv_{jk}(t):j,\ldots,s;k=1,\ldots,m\}$

${}_iM(t)=\{{}_im_{jk}(t):j=1,\ldots,s;k=1,\ldots,m\}$, the occupancy pattern at time t of dwelling units experiencing occupancy transfers in the interval (t+1-t)

${}_iM^*(t)=\{{}_im^*_{jk}(t)\}$, the occupancy pattern at time t+1 of those dwelling units experiencing occupancy transfers in the interval (t+1-t)

${}_iS(t)=\{{}_is_{jk}(t)\}$, the occupancy pattern at time t of those dwelling units *not* experiencing occupancy transfers in the interval (t+1-t)

${}_{i}S^{*}(t)=\{{}_{i}s^{*}_{jk}(t)\}$, the occupancy pattern at time t+1 of those dwelling units *not* experiencing occupancy transfers in the interval t+1-t)

${}_{i}b(t)=\{{}_{i}b_{j}(t)\}$, the vector of additions by dwelling type during (t+1-t)

${}_{i}B^{*}(t)=\{{}_{i}b^{*}_{jk}(t)\}$, the occupancy pattern of additions at the end of the interval (t+1-t)

Relations between these patterns are illustrated in Figure 1.

For the present we will assume that the characteristics of ${}_{i}D(t)$ and ${}_{i}B^{*}(t)$ are known and we will therefore focus our attention on the transitions from ${}_{i}M(t)$ to ${}_{i}M^{*}(t)$ and from ${}_{i}S(t)$ to ${}_{i}S^{*}(t)$.

2.1.1 ${}_{i}M(t) \rightarrow {}_{i}M^{*}(t)$

Given the availability of disaggregated data on occupancy patterns at t and t+1 which includes household name and address identification (as discussed in section 1.2), we are able to represent this transfer system as a 4-way contingency table ${}_{i}F(t)$ whose individual elements ${}_{i}f_{jkh\ell}(t)$ represent the number of dwelling units in occupancy class jk at time t (the occupancy of a type of dwelling unit by a type k household which, in experiencing an occupancy transfer, changed to occupancy class $h\ell$ at time t+1. In many practical situations we may make a useful modification by assuming that dwelling characteristics remain stable in the short run. We then define the modified 3-way table ${}_{i}F^{+}(t)$ with elements ${}_{i}f^{+}{}_{jh\ell}(t)$. Expression (4) gives a model of this class of occupancy changes:

$${}_{i}m^{*}_{j}(t)={}_{i}m_{j}(t)\,{}_{ij}P(t) \qquad (4)$$

where ${}_{i}m_{j}(t)$ is the j^{th} row of ${}_{i}M(t)$, ${}_{i}m^{*}_{j}(t)$ is the j^{th} row of ${}_{i}M^{*}(t)$, and ${}_{ij}P(t)$ is the household transfer matrix for the j^{th} dwelling type in area i; its elements are given by

$${}_{ij}P_{h\ell}(t)=\frac{{}_{i}f^{+}_{jh\ell}(t)}{{}_{i}f^{+}_{jh\cdot}(t)}\,.$$

2.1.2 ${}_{i}S(t) \rightarrow {}_{i}S^{*}(t)$

We may also define corresponding representations for those dwelling units *not* experiencing occupancy transfers. The notations corresponding to ${}_{i}F(t)$ and ${}_{i}F^{+}(t)$ and ${}_{i}G(t)$ and ${}_{i}G^{+}(t)$, and corresponding to expression (4) we have

$${}_{i}s^{*}_{j}(t)={}_{i}s_{j}(t)\,{}_{ij}Q(t) \qquad (5)$$

An important rationale for formulating the process in this way is that it is now possible to test a wide variety of hypotheses concerning the structure and components of changes in the system as well as for interactions between household and dwelling classifications. Furthermore, in the short run it is also possible to examine the degree to which ${}_{ij}P(t)$ and ${}_{ij}Q(t)$ are spatially and temporally homogeneous, thus casting greater light on small area forecasting problems. However, in the present context we are concerned mainly with developing sets of measures from this basic structure which will yield more sensitive statements regarding change than those currently available; it is to this problem that we now turn our attention.

2.2 Some Classes of Dynamic Social Indicators

The detailed representational model of neighborhood change outlined above (section 2.1) can be used to generate a wide range of measures relating both to the persistence of particular occupancy classes and to the rates of change in population composition attributable to different sources. For illustrative purposes we will consider two such classes: (i) those relating to gross mobility associated with given occupancy classes; and (ii) component rates of change for occupancy classes.

2.2.1 Gross Mobility by Occupancy Class

The activity rate associated with a given occupancy class provides an indication of the degree of involvement of that class in neighborhood mobility and, indirectly, of the potential for regulation of change in that occupancy class. To obtain such an index from the model of occupancy patterns we define a measure akin to the usual gross movement rate which is used in migration studies (e.g., Gittus (1961)). Let the total activity rate of an occupancy class jk be given by ${}_{i}W_{jk}$, where

$${}_{i}W_{jk}=\frac{({}_{i}f_{jk\cdot\cdot}+{}_{i}f_{\cdot\cdot jk})+({}_{i}g_{jk\cdot\cdot}+{}_{i}g_{\cdot\cdot jk}-2\,{}_{i}g_{jkjk})}{0.5({}_{i}f_{jk\cdot\cdot}+{}_{i}f_{\cdot\cdot jk}+{}_{i}g_{jk\cdot\cdot}+{}_{i}g_{\cdot\cdot jk})}.$$

The overall activity rate of the dwelling type j (${}_{i}W_{j\cdot}$) can also be obtained as a weighted average of the ${}_{i}W_{jk}$.

A measure which is complementary to the total activity rate is the degree of persistence of an occupancy class over time. Although this could be computed by taking the proportion of $v_{jk}(t)$ which is still in occupancy class jk at t+1, such a measure would be misleading in that it is composed of two distinct elements:

i) persistence deriving from the propensity to move; and
ii) persistence deriving from the probability of an out-migrant household being replaced by another household with the same characteristics.

However, if we define the persistence due to the first element as

$${}_{ijk}U_{S}=\frac{{}_{i}g_{jkjk}}{{}_{i}s_{jk}}$$

and, for the second element as

$${}_{ijk}U_{M}=\frac{{}_{i}f_{jkjk}}{{}_{i}m_{jk}}$$

we can obtain the following identity for the overall persistence of occupancy class jk

$${}_{ijk}U_{T}={}_{i}\theta_{jk}\cdot{}_{ijk}U_{M}+(1-{}_{i}\theta_{jk})\cdot{}_{ijk}U_{S} \qquad (6)$$

where ${}_{i}\theta_{jk}$ is the proportion of dwelling units in v_{jk} which experience an occupancy transfer in (t+1-t). The advantage of this type of index is that we can now link the concepts of stability and change with the mobility rate of different subgroups. As has been shown elsewhere (Moore (1972)) there is no *necessary* relation between mobility and change, and one function of developing these classes of indicators is to spell out the range of relations in specific urban relocation situations.

2.2.2 Component Rates of Change

An important set of measures from the standpoint of the local area planner are the rates of change in each occupancy class attributable to the classes of event defined above (section 2.1). At an early stage of inquiry, knowledge of such rates provides perhaps the strongest basis on which to evaluate the consequences of different planning and policy decisions affecting a particular local area. For each of these classes of events we may thus define the following rates of change for a given occupancy class, jk:

i) due to decrements $= \frac{{}_i d_{jk}}{{}_i a_{jk}} = {}_i\gamma_{jk}$

ii) due to additions $= \frac{{}_i b^*_{jk}}{{}_i a_{jk}} = {}_i\delta_{jk}$

iii.a) for dwellings experiencing occupancy transfers

$$= \frac{{}_i m^*_{jk} - {}_i m_{jk}}{{}_i m_{jk}} = {}_i\alpha_{jk}$$

iii.b) for dwellings not experiencing occupancy transfers

$$= \frac{{}_i s^*_{jk} - {}_i s_{jk}}{{}_i s_{jk}} = {}_i\beta_{jk}$$

and the composite measure for events (iii.a) and (iii.b)

$${}_i\lambda_{jk} = {}_i\theta_{jk} \cdot {}_i\alpha_{jk} + (1 - {}_i\theta_{jk}) \cdot {}_i\beta_{jk} \qquad (7)$$

Equation (7) has particular import for understanding the processes underlying neighborhood change. It is quite possible that ${}_i\alpha_{jk}$ and ${}_i\beta_{jk}$ have different signs and change in composition will thus depend critically on the mobility rate ${}_i\theta_{jk}$. For example, in some cases (e.g., inner city apartment areas) it is only a high mobility rate which can maintain the population composition in stable form and it is crucial for a planner to be able to distinguish between this and stability which results from very low overall mobility rates.

2.3 Indicators: A Prospectus

The set of indicators outlined above are for highly disaggregate sub-groups. One of the tasks for subsequent analyses is to identify sub-groups of areas, household, and housing types over which the indicators are homogeneous. We expect to be able to use the definitional properties expressed by the notation of the K^N model in combination with existing theory relating to residential mobility, filtering of house values, condition aging and residential location to provide bases for formulating hypotheses regarding the specification of definitional and classificational procedures which lead to homogeneity. At present, this theory is not sufficiently strong to enable us to forego the use of the kind of disaggregate analysis we have outlined. However, by employing such an analysis in a wide variety of cases, it is hoped that a basis will be provided for modifying and refining existing theory in such a way as to make detailed monitoring necessary in only a few representative cases.

Footnotes

*The support of the Urban Systems Engineering Center, Northwestern University, is gratefully acknowledged.

1. It is possible to treat this problem formally by using time inhomogeneous Markov processes or semi-Markov processes (e.g., Ginsberg (1972), Gilbert (1972), rather than by developing indices but, given our weak understanding of neighborhood change, the assumptions contained in these models are still unnecessarily restrictive.
2. Note that this is especially crucial when sensitivity tests are to be employed to assess the stability of the computed parameters and indices under alternative theoretical conceptualizations.
3. The transformation matrix can be uniquely computed from marginal distributions only under the condition in which there are no higher-order (i.e., third or fourth order) interactions (Goodman (1970)). Since most of the literature on residential relocation indicates that this is not the case, and since there is no *a priori* reason to believe that it should be true for any particular sub-population, we will not treat this aspect in any further detail.
4. Note that this implicitly restricts us to considering those changes that take place at intervals of at least one year. However, for most types of neighborhood changes this appears to be quite adequate.
5. Note that other kinds of mobility models could also be defined using this same framework--e.g., residential mobility, condition aging, occupational mobility, etc.

REFERENCES

Bartholomew, D.J. (1967), *Stochastic Models for Social Processes*, (John Wiley: New York).

Butler, E.W. et al. (1969), *Moving Behavior and Residential Choice: A National Survey*, National Cooperative Highway Research Program, Report No. 81, Highway Research Board, Washington, D.C.

Duncan, O.D., R. Cuzzort and B. Duncan (1961), *Statistical Geography*, (Free Press: Glencoe).

Fisher, F.M. (1960), "On the Analysis of History and the Interdependence of the Social Sciences," *Philosophy of Science, 27*, 147-158.

Gale, S. (1971), "Evolutionary Laws in the Social Sciences." Paper presented at the Fourth International Congress on Logic, Methodology and Philosophy of Science, (Bucharest, Rumania).

____(1972), "Inexactness, Fuzzy Sets, and the Foundations of Behavioral Geography," *Geographical Analysis, 4*, in press.

____(1973), "Explanation Theory and Models of Migration," *Economic Geography*, forthcoming.

Gilbert, G. (1972), "Two Markov Models of Neighborhood Housing Turnover," *Environment and Planning, 4*, 133-146.

Ginsberg, R.B. (1972), "Semi-Markov Processes and Mobility," *Journal of Mathematical Sociology, 1*, 233-62.

Gittus, E. (1961), "Migration in Lancashire and Cheshire," *The Town Planning Review, 32, 2,* 141-152.
Goodman, L.A. (1970), "The Multivariate Analysis of Qualitative Data: Interactions Among Multiple Classifications," *Journal of the American Statistical Association, 65,* 226-256.
Moore, E.G. (1972), *Residential Mobility in the City,* Association of American Geographers, Commission on College Geography, Resource Paper, No. 13, (Washington, D.C.)
Moore, E.G. and S. Gale (1972), "Comments on Models of Occupancy Shifts and Neighborhood Change." Paper presented at the Mathematical Social Sciences Board Working Seminar on Models of Residential Location and Re-Location in Metropolitan Areas, Northwestern University.
White, H.C. (1970), *Chains of Opportunity,* (Cambridge: Harvard University Press).

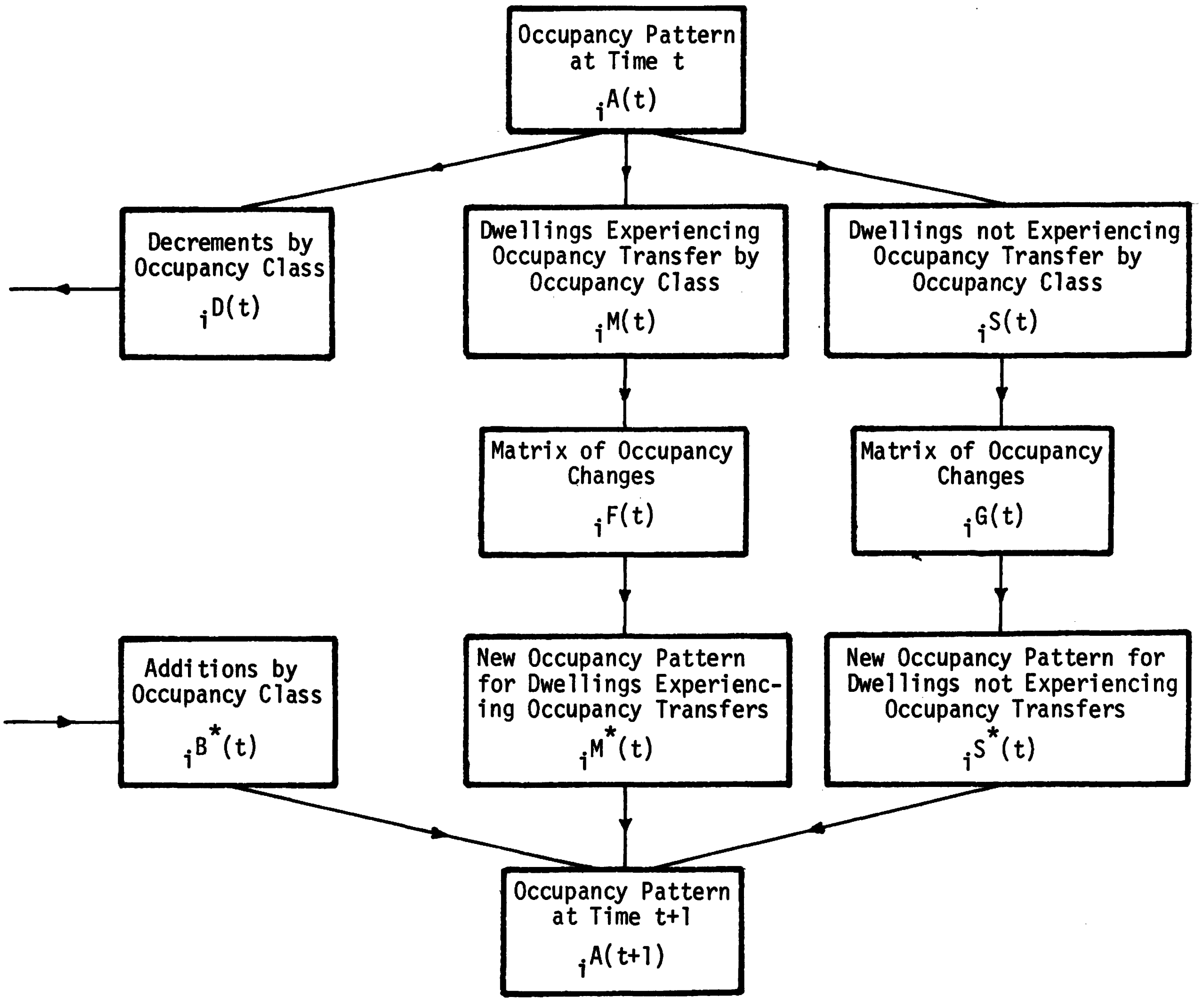


FIGURE 1: SOURCES OF CHANGE IN THE OCCUPANCY PATTERN IN THE i^{th} SUB-AREA

THE INFLUENCE OF LABOR FORCE PARTICIPATION ON FERTILITY IN THAILAND*

Sidney Goldstein, Brown University

Introduction

Increasingly, discussion of the appropriateness and effectiveness of methods to limit population growth have focussed on those measures which go beyond family planning to emphasize the need for adoption of policies "expressly related to family roles and opportunities for legitimate alternative satisfactions and activities" (1). Proponents of such measures argue that they offer greater promise of reduction in family size because they directly assault the motivational framework of reproduction. Among the means suggested for limiting reproduction within marriage as well as postponing marriage is modification of the complementarity of the roles of men and women (3). Of particular interest in this regard is the nature of the relation between female labor force participation and fertility.

In contrast to the generally well established negative relationship between female employment and fertility in the industrialized world (See, for example, 16, 17), research focussing on less developed countries points to no such uniform pattern (See, for example, 5, 10, 11, 13, 17). There is increasing evidence to suggest, however, that the greater the incompatibility between the role of mother and worker, the greater is the differential fertility behavior of workers and women not in the labor force (12, 19). Also, it may not be labor force participation _per se_, but rather a series of other variables associated with labor force participation, such as later age at marriage and longer periods of schooling, which affect fertility levels (20).

If the degree of incompatibility between the roles of mother and worker affects the extent of difference in fertility level between working and non-working women, one would expect such fertility differentials among rural women to be less than those of urban women; participation in farm work as well as the greater availability of parental surrogates would lead to less role conflict among working women in rural places in contrast to the situation in urban places where the likelihood of work in the modern sectors of the economy, the greater separation of work and residence, and the absence of parental surrogates are all probably greater (6).

The analysis of fertility differentials in Thailand provides some opportunity to test the relationship between fertility and labor force participation of women in a developing country. Like many such countries, Thailand has had a high growth rate, over three percent, due largely to continuation of fertility levels exceeding 40 per 1,000 (2, 15, 18), while mortality declined sharply (14). The country remains highly rural in character: three-fourths of its population live in households engaged in agriculture, and over four-fifths of the economically active population 11 years of age and older work on farms. Yet, urbanization in Thailand is assuming increasing importance (9). The 1970 level of urbanization (14.7 percent) was still low, but the urban growth rate is high, averaging about 5 percent a year. Particularly noteworthy has been the very rapid population increase in Greater Bangkok, which grew from only 0.8 million persons in 1947 to 2.9 million by 1970. Containing over half of all of Thailand's urban population, Greater Bangkok accounted for almost two-thirds of all urban population growth in the country. Yet, urban development has also begun to permeate all regions of Thailand and has become an important factor in the complex process of national, social, and economic development. The existence of differential patterns of urban and rural female labor force participation suggests the desirability of assessing the extent to which such differentials are, in turn, associated with differential patterns of fertility.

The Data

Just recently special tabulations based on a 1 percent sample tape of the 1960 Thai census became available; they provide one of the first opportunities for a comprehensive assessment of fertility differentials in Thailand. The data, based on a question on children ever born to ever married women, permit controls for age of mother, urban-rural residence, and agricultural household status,[1] and make possible the evaluation of the relation between labor force participation and fertility in Thailand. Furthermore, a comparative analysis for both the urban and the rural segments of the population provides the opportunity to explore further, although indirectly, the relation between fertility and the compatibility of the roles of mother and worker. (For discussion of the limitations of the data from the 1 percent tape for fertility analysis, see 8.)

Urban-Rural Differentials in Fertility, Labor Force Participation, and Occupation

As in many countries of the world, urban-rural residence in Thailand is related to fertility level. For the kingdom as a whole, the average number of children ever born per 1,000 ever married women was reported as 4,261 in 1960. But the fertility level varied from a high of 4,461 for the rural, agricultural population to a low of 3,375 for those living in Bangkok. Bangkok's level, about 25 percent below that of the most rural part of Thailand, corroborates the significant impact of the urban residence on fertility levels. But it also emphasizes that the still predominantly rural character of Thailand and the high rural fertility account for the high fertility level characterizing the country as a whole.

Thailand's rural character also accounts for the exceptionally high proportion of its adult female population in the labor force (4). In 1960, 78 percent of all adult females and over 85 percent of those between ages 20 and 50 were labor force participants (Table 1). The high proportion of women in the labor force results largely from the inclusion of unpaid family workers, particularly those engaged in agriculture. For Thailand as a whole, 82 percent of all females in the labor force were unpaid family workers. Compared to 87 percent of the rural, agricultural women aged 11 and over who were in the labor force, only one-third of those in Bangkok were reported as working. Moreover, 97 percent of all working women living in rural, agricultural households worked on farms, compared to only 2.4 percent of those in metropolitan Bangkok. As would be expected, the female urban labor force is characterized by considerable variation in occupational composition with one-third employed as sales workers, one quarter as skilled laborers, and almost one-fifth as service workers (Table 2). These differential patterns of labor force participation by urban-rural residence emphasize the desirability of ascertaining the pattern of fertility differentials characterizing the labor force participants and non-participants as well as the impact which urban-rural residence and type of occupation have on fertility.

Labor Force Participation and Fertility

For the kingdom as a whole, the fertility level of women who were in the labor force in 1960, as shown in Table 3, was 15 percent higher than that of women classified as housewives, and 10 percent higher when standardized. Thus, for Thailand as a whole, participation of women in the labor force as measured by "current" activity patterns is not associated with lower fertility. To the contrary, these particular data suggest that factors associated with labor force participation may be conducive to higher average fertility. This pattern of differentials runs counter to that noted for the developed world, as well as for many developing countries.

But the age specific data suggest that the relation between labor force participation and fertility may operate differentially for the various age segments of the population, favoring lower fertility among labor force participants compared to housewives under age 30, and higher fertility for the labor force participants aged 30 and over, with the degree of difference between the economically active women and the housewives tending to increase with increasing age. This pattern suggests that among those in the early and peak ages of childbearing, participation in the labor force is more incompatible with fertility than it is among older women, especially those who have completed their childbearing.

The higher fertility levels of older employed women compared to the housewives may reflect several situations: 1) Higher cumulative fertility may force women to work in order to meet the greater consumption needs of the larger household size. 2) The availability in the household of older children who are able to care for the younger children may permit a higher degree of labor force participation by these mothers (19). 3) For women aged 45 and over, the end of childbearing may facilitate labor force participation.

The higher fertility characterizing the economically active women in the kingdom as a whole does not extend uniformly to all residence categories, but shows a clear-cut pattern in relation to urban-rural status. For both the labor force participants and the housewives in the rural, agricultural category, fertility exceeds that of comparable groups in all other urban-rural categories, but the differential between labor force participants and housewives within the rural, agricultural category is very small for both the unstandardized and standardized levels.

Yet within the rural population there are some age differentials following the same pattern as noted earlier for the kingdom as a whole. Because the rural, agricultural category of employed women includes such a large proportion of unpaid family workers, participation in the labor force and continued maintenance of the role of mother are not strongly incompatible. This may well account for the relatively low level of fertility differentials characterizing the economically active and the housewife group.

For all other rural-urban categories, the age standardized fertility levels of the economically active women were below those of the housewives, but the differential was noteworthy only for Bangkok where the fertility of housewives was over 10 percent below that of economically active women. Moreover, with the exception of two age groups (35-39 and 50 and over) whose levels were almost identical, every age group in Bangkok was characterized by lower fertility for economically active women than for housewives.

That the fertility level of housewives in Bangkok is higher than that of labor force participants in all ages under 35 and in several age groups above 35 indicates that in the urban environment higher fertility does not in itself serve as a stimulus for increased participation of women in the labor force in order to increase family income. The close similarity for the 50 and over age category in the fertility levels of the economically active women and the housewives may well be due to the inclusion of a wide range of older age groups in this category, including many older women who because their children are no longer living at home are more easily able to participate in the labor force. In the more homogeneous 40-49 age group the fertility level of housewives was considerably above that of economically active women. This suggests that high fertility in the urban environment may require women to spend more time at home with their children. In part this may reflect the absence of parental surrogates in the urban environment. In part it may be indicative of a greater ability in urban places to rely upon the income of husbands and other members of the family. It may also reflect less opportunities in the metropolitan area for women to work due to the excessive rate of population growth.

Most likely, however, the pattern of fertility differentials in Bangkok is associated with the different nature of the work experience of economically active women and the fact that employment more often involves separation of work and family roles (Cf., for example, patterns noted for Latin American cities and Puerto Rico, 5, 19).

The fact that the direction of differentials between economically active women and housewives at the two extremes of the urban-rural category are opposite suggests that labor force participation operates quite differently in a highly urban and a highly rural environment, having little, if any, overall effect on fertility in the latter and considerable effect in the former. This is further suggested by the occupational differentials which exist (Table 4). For the kingdom as a whole, women engaged in farming have by far the highest fertility, 4,503 per 1,000 ever married women, standardized for age; and women farmers in the rural, agricultural category have the highest levels of all residence-occupation groups. Overall, the data on occupational differentials reenforce the conclusion that women engaged in agricultural activities tend to have a higher number of children while those not in agriculture have lower levels of fertility, although the pattern of differentials was not altogether consistent for the specific non-agricultural categories.

Overall, this comparison of the fertility levels of employed women and housewives as well as women in specific occupations in Bangkok with those of women in corresponding categories in other urban and rural groups shows that, with minor exceptions, fertility in Bangkok is considerably lower. This lends support to the conclusion that the impact of urban and rural residence on fertility operates independent of the labor force status of the female population. Other factors, such as differentials in literacy or education associated with urban living or the selection of migrants to urban places, result in urban women having lower fertility.

As part of a larger study (7), the relation between education, as measured by both literacy and number of school years completed, and fertility has also been assessed. This showed that both literacy and educational achievement affect fertility. The number of children ever born is lower for literate women in both rural and urban residence categories. The influence of literacy is greater, however, in the highly urban environment of Bangkok, where sharp differentials characterize the total group as well as all age segments. In part, this may reflect the higher level of educational achievement of the literate urban population. This interpretation is supported by the fertility differentials by educational level. Regardless of residence category, education is inversely related to fertility level. For older women, the sharper differential obtains between those with no schooling and those with primary schooling; for younger women achievement of more than a primary education seems to have the greatest effect on fertility levels. Overall, education plays a key role as an instrument of fertility reduction. Moreover, the analysis suggests that the lower fertility characterizing the metropolitan center probably reflects the combined effects of higher educational levels and higher rates of female labor force participation.

In sum, the joint analysis of the impact of labor force participation and education on fertility lends support to the view that adoption of policies expressly related to family roles and opportunities for legitimate alternative satisfactions and activities holds promise of reduction in family size. To the extent that significant labor force, education, and urban-rural differentials exist, even in a country with high overall levels of fertility, policies directed at fostering high rates of educational enrollment of women, greater participation in the non-agricultural labor force, and greater exposure to the urban way of life should be considered as part of any program designed to achieve reductions in fertility.

FOOTNOTES

*This research was initiated while the author served, under sponsorship of the Population Council, as Demographic Advisor to the Institute of Population Studies at Chulalongkorn University, Bangkok. The full cooperation of the Institute's staff and of the National Statistical Office in making unpublished data available is gratefully acknowledged. The research upon which this manuscript is based was performed pursuant to Contract No. NIH-70-2190 with the National Institutes of Health, Department of Health, Education and Welfare.

1. Using both residence in 1960 and household-type data as a joint means of approximating an urban rural "continuum", five categories of urban-rural status have been established: at the most urban level are those living in the metropolis of Bangkok in non-agricultural households. Three intermediary categories consist of (a) all other urban residents living in non-agricultural households; (b) all urban residents in agricultural households; (c) all rural residents in non-agricultural households. At the most rural extreme are all rural, agricultural households. The relative positions of categories (b) and (c) are somewhat arbitrary; had household type rather than residence been used as primary classification variable, the order would have been reversed.

REFERENCES

(1) Blake, Judith, "Demographic Science and the Redirection of Population Policy," Journal of Chronic Diseases, 18 (1965), 1200.

(2) Caldwell, J. C., "The Demographic Structure," in T. H. Silcock, Thailand: Social and Economic Studies in Development, Canberra: Australian National University Press, 1967, 27-64.

(3) Davis, Kingsley, "Population Policy: Will Current Programs Succeed?" Science, 158 (November, 1967), 730-739.

(4) Debavalya, Nibhon, "Labor Force Composition of Thailand," Journal of Social Sciences, 8 (January, 1971), 143-157.

(5) Gendell, Murray, "The Influence of Family Building Activity on Woman's Rate of Economic Activity," World Population Conference, 1965, Vol. IV, New York: United Nations, 1967, 285.

(6) Gendell, Murray, Maria Nydia Maraviglia, and Philip C. Kreitner, "Fertility and Economic Activity of Women in Guatemala City, 1964," Demography, 7 (August, 1970), 273-286.

(7) Goldstein, Sidney, "The Influence of Labor Force Participation and Education on Fertility in Thailand," Population Studies,

26 (November, 1972), forthcoming.
(8) Goldstein, Sidney, "Religious Fertility Differentials in Thailand 1960," Population Studies, 24 (November, 1970), 327-328.
(9) ______________, "Urbanization in Thailand, 1947-1960," Demography, 8 (May, 1971), 205-223.
(10) Heer, David, and Elsa S. Turner, "Areal Differences in Latin American Fertility," Population Studies, 18 (March, 1965), 279-292.
(11) Jaffe, A. J., People, Jobs, and Economic Development, Glencoe, Ill.: The Free Press, 1959.
(12) Jaffe, A. J., and K. Azumi, "The Birth Rate and Cottage Industries in Underdeveloped Countries," Economic Development and Cultural Change, 9 (October, 1960), 52-63.
(13) Miro, Carmen A., and F. Rath, "Preliminary Findings of Comparative Fertility Surveys in Three Latin American Countries," Milbank Memorial Fund Quarterly, 42, No. 2, Part 2 (April, 1965), 36-62.
(14) National Statistical Office, The Survey of Population Change, 1964-1966, Bangkok: National Statistical Office, 1969, 9.
(15) Perkin, Gordon W., et al., Country Profiles: Thailand, New York: The Population Council, May, 1969.
(16) Ridley, Jeanne Clare, "Women's Changing Status," in Fogarty International Center Proceedings No. 3. The Family in Transition: A Round Table Conference, 1969, Bethesda: National Institutes of Health, 1971.
(17) Stycos, J. Mayone, and Robert H. Weller, "Female Working Roles and Fertility," Demography, 4 (1967), 210-217.
(18) Thomlinson, Ralph, Thailand's Population, Bangkok: June, 1971.
(19) Weller, Robert H., "The Employment of Wives, Role Incompatibility, and Fertility," Milbank Memorial Fund Quarterly, 46, (October, 1968), 507-526.
(20) Weller, Robert H., and James M. Sakoda, "A Longitudinal Study of Female Employment and Fertility in Puerto Rico: 1940-1960." Paper presented at annual meeting of the Eastern Sociological Society, New York April, 1971.

Table 1. PERCENT OF FEMALES IN THE LABOR FORCE, BY AGE AND URBAN-RURAL RESIDENCE, 1960

Age	Bangkok	Other Urban, Non-Agric.	Urban, Agric.	Rural, Non-Agric.	Rural, Agric.	Total Kingdom
11 - 12	8.9	5.0	15.0	21.9	42.6	36.4
13 - 19	28.8	33.3	48.3	62.5	88.7	80.2
20 - 24	42.1	51.2	70.3	63.1	97.0	87.5
25 - 29	40.8	46.2	75.2	61.7	96.6	85.8
30 - 34	41.4	50.3	80.2	61.7	96.4	86.6
35 - 39	48.3	49.8	63.1	64.4	96.7	87.1
40 - 44	43.5	61.7	71.2	63.8	97.0	88.3
45 - 49	33.8	64.2	74.7	67.5	96.2	87.9
50 and over	24.9	38.1	53.5	43.3	70.2	62.8
Total	34.2	42.7	60.0	56.8	86.8	77.9
Total standardized for age	34.0	42.6	60.3	56.8	86.8	77.9

Table 2. OCCUPATIONAL DISTRIBUTION OF FEMALES BY URBAN-RURAL RESIDENCE, 1960 (Standardized for Age)

Occupation	Bangkok	Other Urban, Non-Agric.	Urban, Agric.	Rural, Non-Agric.	Rural, Agric.	Total Kingdom
Professional and administrative workers	6.7	7.1	5.4	2.0	0.2	0.9
Clerical workers	5.3	1.2	1.6	0.2	*	0.2
Sales workers	34.8	53.3	14.6	29.6	1.2	6.0
Farmers and miners	2.4	7.0	65.1	42.5	97.2	87.4
Service and transport workers	18.2	12.9	5.0	5.4	0.2	1.5
Craft workers	27.8	15.7	6.8	17.9	1.0	3.6
Unclassified and new workers	4.8	2.7	1.3	2.3	0.1	0.4
Total percent	100.0	100.0	100.0	100.0	100.0	100.0
Total number	1,363	1,550	685	6,180	54,843	64,621

*Less than 0.1 percent.

Table 3. NUMBER OF CHILDREN EVER BORN PER 1,000 EVER MARRIED WOMEN BY AGE, LABOR FORCE STATUS, AND URBAN-RURAL RESIDENCE, 1960

Age	In the Labor Force	Housewife	In the Labor Force	Housewife
	Bangkok		Rural, Non-Agricultural	
13 - 19	*	964	464	569
20 - 24	1,488	1,641	1,417	1,662
25 - 29	2,371	2,605	2,611	2,768
30 - 34	3,127	3,657	3,731	4,012
35 - 39	4,347	4,331	4,929	5,059
40 - 44	4,602	5,140	5,411	5,854
45 - 49	4,093	4,406	5,415	5,172
50 and over	3,688	3,658	5,027	5,373
Total	3,374	3,407	3,794	3,733
Total standardized for age	3,256	3,675	3,957	4,179
	Other Urban, Non-Agricultural		Rural, Agricultural	
13 - 19	450	667	478	667
20 - 24	1,430	1,583	1,347	1,345
25 - 29	2,761	2,873	2,675	2,824
30 - 34	3,820	3,778	4,140	4,420
35 - 39	4,752	4,597	5,415	5,296
40 - 44	4,830	5,736	6,199	5,608
45 - 49	5,246	4,828	6,360	5,176
50 and over	4,507	4,531	6,148	5,922
Total	3,956	3,575	4,336	4,204
Total standardized for age	3,775	3,853	4,508	4,349
	Urban, Agricultural		Total Kingdom	
13 - 19	571	*	478	688
20 - 24	1,500	1,900	1,359	1,598
25 - 29	2,358	2,741	2,661	2,760
30 - 34	3,592	4,625	4,076	3,987
35 - 39	4,784	4,560	5,315	4,862
40 - 44	5,544	5,667	6,044	5,616
45 - 49	5,789	5,400	6,187	4,895
50 and over	5,441	4,964	5,948	5,058
Total	4,183	4,176	4,253	3,742
Total standardized for age	4,038	4,121	4,406	4,022

*Less than 10 women.

Table 4. NUMBER OF CHILDREN EVER BORN PER 1,000 EVER MARRIED WOMEN BY OCCUPATION, BY URBAN-RURAL RESIDENCE (STANDARDIZED FOR AGE) AND FOR TOTAL KINGDOM BY AGE, 1960

	Farmers and Miners	Craft Workers	Professional and Administrative Workers	Service and Transport Workers	Sales Workers
Bangkok	2,659	2,834	2,416	2,881	3,720
Other Urban, Non-Agricultural	3,978	3,072	3,081	3,823	3,988
Urban, Agricultural	4,125	2,231	2,182	3,276	4,131
Rural, Non-Agricultural	4,055	3,922	3,423	4,114	3,919
Rural, Agricultural	4,527	3,598	3,939	3,735	3,780
		Total Kingdom			
13 - 19	477	542	*	*	511
20 - 24	1,365	1,163	974	1,338	1,496
25 - 29	2,683	2,419	2,188	2,375	2,704
30 - 34	4,129	3,535	3,591	3,705	3,714
35 - 39	5,404	4,767	3,671	4,243	5,073
40 - 44	6,196	4,830	4,087	4,672	5,211
45 - 49	6,356	4,831	4,727	4,983	5,298
50 and over	6,126	4,559	5,162	4,549	4,698
Total	4,315	3,460	3,322	3,580	4,143
Total standardized for age	4,503	3,627	3,508	3,592	3,887

*Less than 10 women.

THE EFFECTS OF HIGH SCHOOL CURRICULUM ON AGE-EARNINGS PROFILES

John T. Grasso and John R. Shea*
Center for Human Resource Research, The Ohio State University

1. INTRODUCTION

Much of the support for vocational education evident in recent years stems from its alleged economic advantages for the individual (10). Within a multivariate framework this paper examines the relationship between high school curriculum and one important indicator of labor market success, hourly rate of pay subsequent to graduation. This evaluation, of course, is partial. A complete assessment of "curriculum effects" would take account of other performance criteria such as the dropout rate, the probability of attending college, job satisfaction, and the like.

Two important empirical issues are addressed in this paper. One concerns the shape of age-earnings profiles associated with different curricula. Some studies have shown an initial wage or earnings advantage for vocational graduates vis-a-vis graduates from either the general or academic track (2, 3, 7, 8), while other studies have not. At least two studies reveal a narrowing of an initial vocational wage (or earnings) advantage over time (3, 7, 8), while another well-known study suggests a widening gap (1). A second, related issue is how to treat post-secondary training. The shape of estimated age-earnings profiles can be distorted by research strategies which "control for" differences in post-secondary training by excluding high school graduates who have had such training. As this paper will demonstrate, such an approach ignores important interactions between curriculum and subsequent training. Taussig suggests that "vo ed" graduates may be at a disadvantage in regard to access to subsequent training programs (14). The work of Schriver and Bowlby indicates that post-secondary training in Tennessee Area Vocational-Technical Schools is complementary with subsequent training that paid off (12). The "widening gap" that Carroll and Ihnen found was based on a comparison of graduates of a post-secondary, two-year technical institute with high school graduates who had no additional training (1). The work of Horowitz and Herrnstadt indicates that for at least one important craft occupation (tool and die making), a vocational high school experience combined with apprenticeship leads to better performance ratings than any other path (6).

In the next section of this paper we identify the data base used in the analysis and specify the models that were tested. Results are presented and discussed in Section 3. This is followed in Section 4 by a short summary and conclusions.

2. DATA, MODELS, AND VARIABLES

This paper utilizes data derived from personal interviews with a national probability sample of young men who were part of the civilian, noninstitutional population 14 to 24 years of age when first interviewed as part of the National Longitudinal Surveys in October 1966. Supplementary information from school records, including scores on tests of mental ability, was obtained by the Census Bureau from a mailed questionnaire in 1968.[1] As long as each respondent remained in the civilian, noninstitutional population, an attempt was made to interview him every October through 1971. In October 1969--the most recent survey for which information is available--1,192 of the original 5,225 respondents were not interviewed, an attrition rate of 22 percent for the whites and 27 percent for the blacks.[2] By design, Negro and other races were overrepresented in the original sample by a three-to-one ratio relative to whites. This was done to permit a reasonably confident analysis of the black experience. For this reason, and because we have used unweighted sample cases in this paper, results are presented separately for whites and blacks. (Non-Caucasians, other than blacks, have been excluded from the analysis.) This paper, then, uses information from personal interviews conducted in the fall of 1966, 1967, 1968, and 1969, and from the school survey. For each year, our attention is restricted to respondents who (1) had completed 12 (but no additional) years of "regular" schooling by that time, (2) were not enrolled in school, and (3) were employed for wages or salaries.

We examine hourly rates of pay within the context of multiple linear regression analysis, using a method described by Gujarati to test for differences in relationships among several groups (4). We employ a general wage model with the following dummy variables: REGION (SOUTH = 1, otherwise = 0), SMSA (lives in SMSA = 1, otherwise = 0), HLTH (respondent reports his health limits the kind or amount of his work = 1, otherwise = 0), and MILEXP (has had military experience = 1, otherwise = 0). SES is an index of socioeconomic status of the respondent's family of orientation, consisting of the average of at least three of five normalized and equally weighted components: father's education, mother's education, oldest older sibling's education, Duncan index of father's occupation when the respondent was age 14, and a measure of the presence of reading material in the home at that time. Respondents for whom three of the five components were not available have been excluded from the analysis. IQ is a normalized measure of mental ability obtained in the survey of the high school.[3] TRNG consists of the actual number of months the respondent participated in post-school non-degree programs of various kinds; examples are apprenticeship and military programs. YRS is our proxy for length of potential work experience and equals the respondent's age less 17.

Testing the effects of curriculum within this framework requires the construction of several more variables. We have focused on "interaction effects" involving high school curriculum. A set of K dummy terms that identify our groups was constructed (VOC = 1 for graduates of the vocational curriculum; COLL = 1 for college preparatory; and GEN = 1 for general). One of these, GENERAL, was omitted from the equation and serves as the reference group. The set of (K-1) dummies was multiplied by each of the P variables selected for testing.[4] The (K-1) time P new variables were added to the model, and those which were found to be statistically significant (e.g., $X_i * D_j$) can be interpreted to imply that the relationship between the dependent variable Y and the independent variable X_i is different for group j than for the reference group. (Variables for which no new terms were added are assumed to operate similarly for all groups.)

We have tested for three kinds of potential differences in rate of pay: in starting wage rates, in the pattern of wage rate changes over time, and in the effects of post-secondary training. Our P variables are the intercept (i.e., age 17 when YRS = 0), the YRS variable, and TRNG, respectively. Thus we add six new variables, which are, respectively, VOC, COLL, YRS*VOC, YRS*COLL, TRNG*VOC, and TRNG*COLL. This gives us the following General Cross-Sectional Model:

$$\begin{aligned} WAGE_t = {} & b_1\ REGION_t + b_2\ SMSA_t + b_3\ HLTH_t + b_4\ SES \\ & + b_5\ IQ + b_6\ MILEXP_t + (b_7\ YRS_t + b_7'\ YRS_t{*}VOC \\ & + b_7''\ YRS_t{*}COLL) + (b_8\ TRNG_t + b_8'\ TRNG_t{*}VOC \\ & + b_8''\ TRNG_t{*}COLL) + (b_0 + b_0'\ VOC + b_0''\ COLL) + e, \\ & \text{where } t = 1969. \end{aligned}$$

The results generated sufficient interest to prompt the construction of a variation of the general cross-sectional model. In this modified version, the variable $TRNG_t$, defined as total cumulative months of training received, was replaced by the seven type-specific components from which it was constructed: Business College or Technical Institute (B), Company (C), Apprenticeship (A), Correspondence School (CS), Military (M), non-degree Regular School (R), and Other (O); and $TRNG_t$*VOC and $TRNG_t$*COLL were replaced with their respective counterparts. We have called the result our Training-Specific Cross-Sectional Model.

While cross-sectional estimates illuminate the probable relationships of our variables with rate of pay, these estimates are not completely satisfactory. We can also estimate the influence of high school curriculum on initial wages and on changes in wages over time by examining, first, the wage rates of 17 and 18 year-olds separately and, second, the change in wage rates between 1966 and 1969. These tests permit a validation of the findings in the General Cross-Sectional Model. Our Initial Wage Model, restricted to 17 and 18 year-olds, includes $REGION_t$, $SMSA_t$, $HLTH_t$, SES, IQ, the curriculum dummies and an error term. The Wage Change Model, applicable to respondents out-of-school all four years, takes the following form:

$$\begin{aligned} \Delta\ WAGE_{1966-1969} = {} & b_1\ SES + b_2\ IQ + b_3\ TRNG_{1966} \\ & + b_4\ TRNG_{1966-1969} + b_5\ YRS_{1966} + (b_0 \\ & + b_6\ VOC + b_7\ COLL) + e. \end{aligned}$$

3. RESULTS

Table 1 presents the results based on the General Cross-Sectional Model. Among the variables for which we did not test for differences by curriculum, REGION is significant for both races and SMSA, IQ, and MILEXP are also significant for whites. Coefficients of HLTH and SES, although not statistically significant, display the expected sign.

Among the variables that we tested for differences by curriculum, few significant differentials were found. As indicated by the VOC and COLL variables, there is no gap at the intercept. As for post-secondary training, TRNG is significant for all blacks, while for whites it is significant for vocational graduates only. In fact, training beyond high school is associated with more than two cents per hour for every month of training received. Finally, regarding the YRS variable, each year beyond school is associated with an increase of 17 cents per hour for the reference group of whites, and although it is not significantly different for the other groups, the negative coefficient for vocationals is noteworthy. Among the blacks, the age-wage profile is much flatter and clearly no different by curriculum.[5]

The Training-Specific Cross-Sectional Model provides results almost identical to the General Cross-Sectional Model on the nontraining variables; therefore, only the regression coefficients and t-values for the training variables are presented in Table 2. For the whites, apprenticeship training yields handsome returns for all respondents and perhaps more so for vocational and college preparatory graduates than general. Business college or technical institute (B) and company programs (C) appear to pay off for vocationals, and correspondence school (CS) yields a return for those from the academic track. On the other hand, training in the military (M) adds nothing to the wages of vocationals and general graduates and is negative for the college group. For the blacks, however, this latter relationship is reversed; $TRNG_tM$ is positive and significant for all graduates and even higher for vocationals and academics than for graduates of the general curriculum.

These findings support the results of the General Cross-Sectional Model. White vocational graduates who followed their graduation with additional training in business schools or technical institutes, in company programs, in apprenticeships, etc., experienced increases in rate of pay. Indeed, <u>every</u> training coefficient for white vocational graduates is positive, while in the absence of training, the vocational group's experience was less bright, certainly no better than the general curriculum group, and appears to have worsened over time. For the blacks, the absence of significant relationships may be due to (1) a lower rate of participation in training and shorter average length of training (except within the military category); (2) small sample sizes; (3) our inability within the context of the larger model to identify significant variables for explaining black wage rates; or (4) the lack of market alternatives equivalent to those of whites.

In an effort to test directly for a differential intercept (wage rate in October subsequent to graduation), our Initial Wage Model was fit to the data provided by 17 and 18 year-olds for each year from 1966 through 1969. While the number of sample cases is small (approximately 100 whites and 30 blacks each year), there is no evidence that having graduated from the VOC or COLL curricula had any influence, ceteris paribus, on the rate of pay reported by young high school graduates in the four years. None of the coefficients for the VOC or COLL dummies is statistically significant, and the signs are not consistent across the years.[6]

One of the values of longitudinal data can be to illuminate the determinants of change: in this case, change in hourly rate of pay. In our Wage Change Model, $\Delta WAGE_{1966-1969}$ is the dependent variable. Although results are not shown here,[7] the coefficient of YRS_{1966}, the amount of work experience as of the base year, is negative (reflecting the flattening out of earnings profiles with advancing age). SES, IQ, $TRNG_{1966}$, and $TRNG_{1966-1969}$ are positive, although not all are statistically significant. No differential by curriculum in wage rate change over time is supported.

Since training is an important correlate of rate of pay and interacts with curriculum, a few words should be said about the nature and magnitude of training received. Young white men reported more training than their black counterparts; three-fifths of the former but less than half of the latter had some training since leaving high school. Among those with training, the average duration was much longer for whites than blacks: about 15 months for the former and 9 months for the latter. There are, however, very few differences between curriculum groups in the pattern of training received. Somewhat to our surprise, for example, approximately one-twelfth of each of the three groups in our sample participated in apprenticeship programs, albeit the average duration was longer for VOC than for other graduates. However, since our sample excludes those with one or more years of college, two or three times as many vocational graduates entered apprenticeships as college preparatory graduates.[8]

4. SUMMARY AND CONCLUSIONS

In our judgment, the data do not fit a compelling theoretical framework, be it human capital theory or learning theory. Nor do our findings provide unambiguous answers for the formulation of educational policy. Before interpreting our principal findings, several caveats are in order. First, our evaluation of high school vocational programs has been limited to a single criterion: hourly rate of pay. We have ignored possible differences by curriculum in attitudes toward work, dropout prevention, and so forth. Second, observed wage rates may be a poor reflection of differences in productivity. Third, we have relied on respondents' self-reporting of high school program; and the vocational track includes such diverse areas as metalworking, woodworking, electrical, mechanical, and other fields. Finally, we have ignored potential differences in students that existed prior to choice of curriculum--differences in motivation, commitment to work, and the like.

In view of our failure to find a statistically significant difference in starting wages for any curriculum group, for either whites or blacks, it appears that the vocational high school curriculum per se does not provide skills which lead to immediate market advantages. With respect to change in hourly wage rates over time, the vocational and academic groups are again not significantly different from the general curriculum group. If anything, among whites, the vocational graduates advanced less rapidly over time.

The lower age-wage profiles of white vocational graduates may prompt a conclusion that vocational programs provide less-than-adequate attention to developing general learning abilities and to providing a solid base from which subsequent work experience can build. However, we cannot believe that vocational programs per se impart uniformly better or worse general skills for work, in light of the fact that vocationals appear to gain most from post-secondary training, even though they gain least from experience.

Given the pattern of post-secondary training, high school vocational programs appear to be essentially pre-vocational and must be followed with further, more specific training after graduation for maximum benefit. Although the results indicate that nonvocational graduates may have benefited from selective training experiences, the vocationals who received training beyond high school appear to have gained the most. Although black men participated in training programs less frequently than white (and for shorter durations even when they did), post-school training for blacks--including that taken in the military--was a profitable course of action regardless of curriculum.

We suggest that the key to interpretation of our findings (e.g., that vocationals profit most from training but least from experience) may lie in the structure of jobs and labor markets toward which at least some vocational students are prepared. In some cases, vocational programs may fulfill an important social function by imparting skills and knowledge which would not be provided elsewhere in their absence, particularly in industries composed of many small firms that cannot undertake training efficiently. It may be an advantage for society to provide some vocational programs in schools, so that workers in these industries earn starting wages which are no less than for other workers. Other vocational programs may sometimes permit entry to superior training programs. Yet, students from other curricula enter apprenticeships, and both vocational and nonvocational graduates reap benefits. There also may be a relatively limited number of high-wage, blue-collar jobs, which some graduates with vocational skills obtain, while the remainder choose jobs from poorer alternatives. Finally, union dominance of entry to some sectors of the economy, market, and technological factors--some or all may promote a market structure consistent with the empirical findings presented here. Needless to say, the results of our work thus far have been sufficiently intriguing that we expect to continue our inquiry into the linkages between curriculum in the schools and later labor market success.

FOOTNOTES

*This paper is based on data from The National Longitudinal Surveys, a project sponsored by the Manpower Administration of the U.S. Department of Labor under the authority of the Manpower Development and Training Act. The work is carried out in collaboration with the U.S. Bureau of the Census. Since researchers are encouraged to use their own judgments freely, this paper does not necessarily represent the official position or policy of the Department of Labor. The authors would like to thank Andrew I. Kohen, Herbert S. Parnes, and several other colleagues for helpful comments on an earlier draft of the paper. Responsibility for interpretations and conclusions, of course, rests with the authors.

[1]Results from various kinds of tests--intelligence, achievement, scholastic aptitude, etc.--were standardized and pooled. While somewhat imprecise, we use the terms "mental ability" and "IQ" in referring to these measures, which are described in (9).

[2]Approximately three-fifths of the losses were due to induction into military service. Efforts are made to reestablish contact with these young men as they return to civilian life.

[3]Both measures, SES and IQ, are more fully described in (9), Appendixes A and B. Respondents for whom an IQ score was not available have been excluded from the analysis.

[4]If the intercept term is one of the P variables, the new terms are simply the dummies themselves.

[5]In order to examine the implications of restricting wage comparisons to graduates with no post-secondary training whatsoever, the General Cross-Sectional Model was rerun excluding any respondent with training or military experience. The results are similar to those discussed here. Repeated cross-sectional tests of the basic model using data from the 1966, 1967, and 1968 surveys reveal about the same pattern as the more complete and more recent (1969) data. Results are available from the authors.

[6]The detailed results may be obtained by writing to the authors.

[7]Once again, the results may be obtained from the authors.

[8]If the base were expanded to include all high school graduates, whether or not they attended college, a somewhat different picture would emerge. Data not shown here suggest that about two-thirds of the college preparatory graduates went on to college, 25 to 35 percent of the generals, and 10 to 20 percent of the vocationals. Thus, on a relative basis, two or three times as many vocational graduates entered apprenticeship as college preparatory graduates.

REFERENCES

(1) Carroll, Adger B. and Ihnen, Loren A. Costs and Returns for Investments in Technical Schooling by a Group of North Carolina High School Graduates. Economics Research Report No. 5. Raleigh: North Carolina State University, 1967.

(2) Corazzini, Arthur J. "The Decision to Invest in Vocational Education: An Analysis of Costs and Benefits." Vocational Education, a supplement to the Journal of Human Resources 3(1968):88-120.

(3) Eninger, Max U. The Process and Product of T&I (Trade and Industry) High School Level Vocational Education in the United States: The Process Variables. Pittsburgh: The Educational System Research Institute, 1968.

(4) Gujarati, Damodar. "The Use of Dummy Variables in Testing for Equality between Sets of Coefficients in Linear Regressions: A Generalization." The American Statistician 24(December 1970):18-21.

(5) Hall, Robert E. "Why is the Unemployment Rate So High at Full Employment?" Brookings Papers on Economic Activity 3(1970):393-94.

(6) Horowitz, Morris A. and Herrnstadt, Irwin L. A Study of the Training of Tool and Die Makers. Boston: Northeastern University, 1969.

(7) Kaufman, Jacob J. et al. The Preparation of Youth for Effective Occupational Utilization: The Role of the Secondary Schools in the Preparation of Youth for Employment. University Park, Pennsylvania: Institute for Research on Human Resources, 1967.

(8) Kaufman, Jacob J., and Lewis, Morgan V. The Potential of Vocational Education: Observations and Discussions Based on a Study of Three Selected Cities in Pennsylvania. University Park, Pennsylvania: Institute for Research on Human Resources, 1968.

(9) Kohen, Andrew I. "Determinants of Early Labor Market Success Among Young Men: Race, Ability, Quantity and Quality of Schooling." Ph.D. dissertation, The Ohio State University, 1972.

(10) Marland, Sidney P., Jr. "Educating for the Real World." Mimeographed. U.S. Office of Education, Department of Health, Education, and Welfare, May 26, 1971.

(11) Schaefer, Carl J., and Kaufman, Jacob J. Occupational Education for Massachusetts. Boston: Massachusetts Advisory Council on Education, 1968.

(12) Schriver, William A. and Bowlby, Roger L. The Effects of Vocational Training on Labor Force Experience. Memphis, Tennessee: Memphis State University, Center for Manpower Studies, 1971.

(13) Somers, Gerald G. et al. The Effectiveness of Vocational and Technical Programs: A National Follow-up Survey. Madison, Wisconsin: The University of Wisconsin-Madison, Center for Studies in Vocational and Technical Education, 1971.

(14) Taussig, Michael K. "An Economic Analysis of Vocational Education in the New York City High Schools." Vocational Education, a supplement to the Journal of Human Resources 3(1968):59-87.

Table 1 Hourly Rate of Pay (Cents per Hour), 1969: Means, Standard Deviations, and Estimated Regression Coefficients for General Cross-Sectional Model

Variables and statistics	WHITES				BLACKS			
	Mean	S.D.	b	t value	Mean	S.D.	b	t value
$REGION_t$ (1=South)	.26	.44	-36.4 ***	-3.53	.54	.50	-75.9 ***	-5.30
$SMSA_t$ (1=Yes)	.69	.46	35.4 ***	3.66	.80	.40	8.13	.43
$HLTH_t$ (1=Limits)	.09	.28	-22.5	-1.50	.08	.28	-33.7	-1.31
SES (Index; $\bar{X}$=10; S.D.=1.7)	10.6	1.8	1.03	.41	9.0	1.8	4.34	1.00
IQ ($\bar{X}$=100; S.D.=16)	99.6	11.7	1.46***	3.74	85.8	14.4	.72	1.39
$MILEXP_t$ (1=Yes)	.27	.44	-23.0 **	-2.18	.18	.39	- 8.56	- .39
YRS_t (from age 17)	5.0	3.3	15.7 ***	8.25	4.4	3.1	3.89	1.27
YRS_t*VOC[a]	-	-	- 5.92	-1.56	-	-	- 1.67	- .29
YRS_t*COLL[a]	-	-	- 1.14	- .31	-	-	1.18	.17
$TRNG_t$ (months)	9.2	15.1	.48	1.11	3.6	7.6	2.28**	2.03
$TRNG_t$*VOC[a]	-	-	2.32***	3.14	-	-	3.46	1.23
$TRNG_t$*COLL[a]	-	-	.91	1.17	-	-	- 3.46	- .61
Constant term (at age 17)	-	-	73.5	1.54	-	-	173.2***	3.21
VOC[a]	.21	.41	2.64	.13	.19	.40	-15.0	- .52
COLL[a]	.19	.39	-22.2	-1.13	.12	.32	24.8	.62
$WAGE_t$	316	116	-	-	261	90	-	-
# of observations	532				130			
$\bar{R}^2$	.29				.29			
F	16.69***				4.75***			

a The omitted category is GEN, the group of graduates from the general curriculum.
** Significant at .05 level.
*** Significant at .01 level.

Table 2 Hourly Rate of Pay (Cents per Hour), 1969: Estimated Regression Coefficients for Training Variables in Training-Specific Cross-Sectional Model

Variables and statistics	WHITES		BLACKS	
	b	t value	b	t value
- -	-	-	- -	-
TRNG$_t$B	- 1.55*	-1.92	- 1.85	- .41
TRNG$_t$B*VOC	3.39**	2.09	.64	.04
TRNG$_t$B*COLL	.42	.21	2.14	.27
TRNG$_t$C	- 1.85	-1.42	2.97	.97
TRNG$_t$C*VOC	6.40***	2.81	6.01	.59
TRNG$_t$C*COLL	3.00	1.25	- 9.61	- .88
TRNG$_t$A	1.95**	2.31	3.44	1.04
TRNG$_t$A*VOC	1.95	1.51	.30	.06
TRNG$_t$A*COLL	2.27	1.57	(No respondents)	
TRNG$_t$CS	- .71	- .40	(No respondents)	
TRNG$_t$CS*VOC	2.86	.63	(No respondents)	
TRNG$_t$CS*COLL	10.12***	3.18	(No respondents)	
TRNG$_t$M	.87	.64	2.83**	2.02
TRNG$_t$M*VOC	- .30	- .18	10.36	1.14
TRNG$_t$M*COLL	- 3.61*	-1.81	34.86	1.49
TRNG$_t$R	5.81	1.41	10.71	.80
TRNG$_t$R*VOC	- 5.14	-1.00	(No respondents)	
TRNG$_t$R*COLL	- 5.22	- .76	- 5.37	- .25
TRNG$_t$O	5.81***	2.61	- 1.14	- .27
TRNG$_t$O*VOC	- 1.78	- .43	15.49*	1.67
TRNG$_t$O*COLL	-11.77*	-1.89	62.98**	1.96
# of observations	532		130	
$\bar{R}^2$	.35		.24	
F	9.84***		2.30***	

* Significant at .10 level.
** Significant at .05 level.
*** Significant at .01 level.

JOINT PROBABILITIES OF SELECTION OF UNITS IN SYSTEMATIC SAMPLES[1]

G.B. Gray, Statistics Canada

Introduction

Many schemes for sampling n units out of N with probability proportional to size (pps) have been devised. Examples of pps methods of selection are given by I.P. Fellegi [1], H.O. Hartley and J.N.K. Rao[3], Horvitz and Thompson [4] and many others. In each case, a Horvitz-Thompson estimate of the population total from the sample may be obtained. Upon calculation of joint probabilities of pairs of units in the sample, variance estimates may be calculated by either the Horvitz-Thompson [4] or Yates-Grundy [6] formulas.

The simplest of all sampling schemes is systematic pps sampling as described in [2] and is particularly useful for rotating samples in recurring surveys. The units may be in a fixed pre-determined order or in random order prior to selection but in either case, unbiased estimates of the total may be obtained. In the case of rotating samples, the units may be randomly ordered prior to the first survey and the order may be maintained for the first and subsequent surveys with a partial or complete rotation of units.

In the case of a fixed pre-determined order of listing, most of the joint probabilities of selection are zero so that unbiased estimates of the variance are not possible. However, if the order of listing has been undertaken in such a manner to ensure a negative serial correlation for sampled units, then an over-estimate of the variance may be obtained by treating the units as though they had been sampled with pps with replacement.

In the case of randomly ordered units, nonzero joint probabilities nearly always exist for every pair of units and these may be easily calculable by the algorithm described below for small populations[2]. H.O. Hartley and J.N.K. Rao [2] derived an asymptotic formula in 1962, which holds approximately true only for pairs of units selected from large populations. Alternatively, the algorithm could be applied to a large sample of possible arrangements of units and the joint probability estimated by averaging over the large sample of units.

Systematic Sample Procedure, Notation, and Assumptions

p_i = relative size of i^{th} unit where $i = 1,2, \ldots, N$ (relative number of persons, relative Census sales, for example). In a systematic pps selection procedure, probability of selection of unit no. $i = np_i$ where $n \geq 2$.

It will be assumed here that $np_i \leq 1$. In some universes or strata, the relative size of a unit, say i, may be so large that, by the procedure of systematic sampling, np_i would lie in the range (1,2) so that the unit could be selected once with probability $1/np_i$ or twice with probability $(np_i - 1)/np_i$. In most cases, it would be preferable from cost/variance benefit and operational point of view simply to include the unit with certainty and adjust the relative sizes of the remaining units. In practice, the procedures could be adapted to include units in a certainty stratum if $np_i > a$ for some arbitrary a less than one such as 0.5 for example. In small populations, however, it is difficult to achieve such a condition while maintaining a fixed sample size and ensuring a sampling stratum containing at least two selected units. If n = 1, then no joint probabilities exist and no variance estimate is possible.

Considering a particular order of listing, for the moment, say, in serial number order, we have the units 1,2,3, ..., N together with their relative sizes, p_i, probabilities of selection, np_i and accumulated probabilities $t_i = \sum_{j=1}^{i} np_j$, in Table 1:

Table 1: Set-up of Units for Systematic pps Selection of n Units

Unit No.	Relative Size	Probability of Selection	Accumulated Probabilities
1	p_1	np_1	$t_1=np_1 \quad =0+d_1$
2	p_2	np_2	$t_2=t_1+np_2=I_2+d_2$
3	p_3	np_3	$t_3=t_2+np_3=I_3+d_3$
4	p_4	np_4	$t_4=t_3+np_4=I_4+d_4$
.	.	.	.
.	.	.	.
.	.	.	.
i	p_i	np_i	$t_i=t_{i-1}+np_i=I_i+d_i$
.	.	.	.
.	.	.	.
.	.	.	.
N	p_N	np_N	$t_N=n \quad =I_N+0$

[1] A similar study was undertaken by W.S. Connor [1] in 1966.

[2] Zero joint probabilities however do occur in an example stated by J.N.K. Rao in his thesis [5] and in W.S. Connor's article [1] even when 5 units are randomly ordered and 2 units are selected. (Relative sizes .1, .1, .25, .275, .275).

In the above table, t_i is partitioned into an integer I_i and a decimal d_i component. As will be seen later, this device will facilitate the computation of joint probabilities. Thus $I_2=0$ or 1 in the above table according as $np_1 + np_2 < 1$ or ≥ 1.

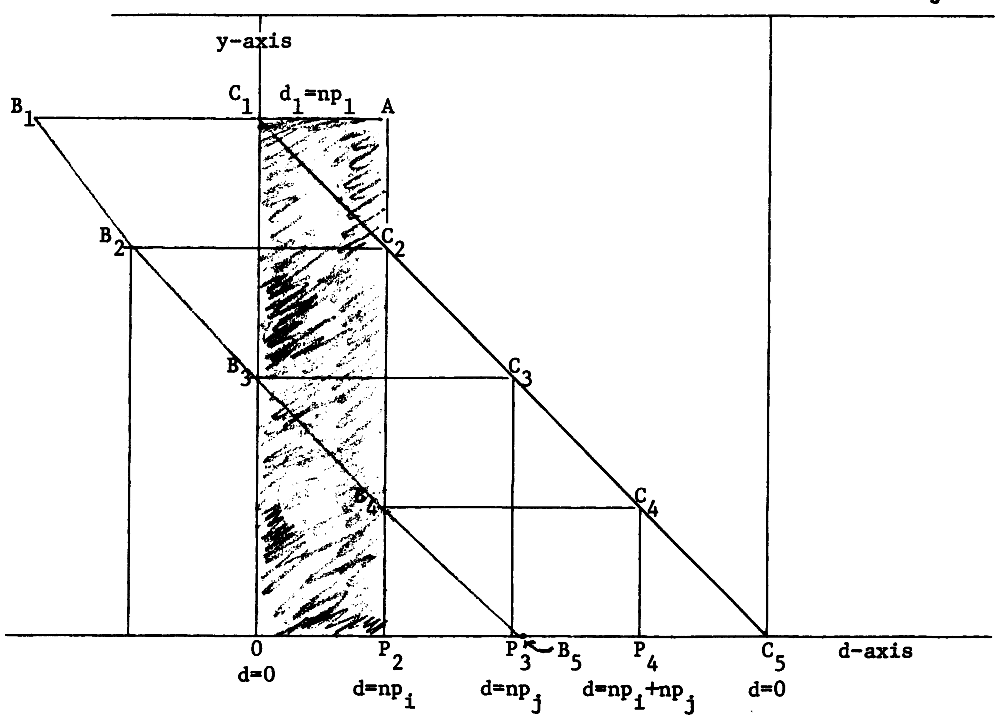
The procedure of systematic sampling is then to select a random number r in the range [0,1) and select n units $i_1, i_2, \ldots, i_n$ such that $t_{i_1} \leq r < t_{i_1+1}$; $t_{i_2} \leq r+1 < t_{i_2+1}$; $t_{i_3} \leq r+2 < t_{i_3+1}, \ldots t_{i_m} \leq r+m-1 < t_{i_m+1}, \ldots,$ and finally $t_{i_n} \leq r+n-1 < t_{i_n+1}$. If $n = N$, t_{i_n+1} may be taken as any arbitrary number greater than n, say $t_{i_N} + np_i$.

In practice, the probabilities of selection and their accumulations are usually replaced by actual sizes $S_1, S_2, \ldots S_N$ and $T_i = \sum_{j=1}^{i} S_j$ and the systematic pps selection is undertaken by selecting a random integer $R\varepsilon[1,T_N/n]$ and selecting units $i_1, i_2, \ldots i_m, \ldots i_n$ so that $T_{i_m} \leq R + (m-1)\, T_N/n < T_{i_m+1}$.
The resultant sample is exactly the same as before with $r = R/(T_N/n)$ and $p_i = S_i/T_N$.

Now if X_i is a characteristic value for the i^{th} unit of some item whose total $X = \sum_{i=1}^{N} X_i$ is to be estimated, an estimate of the total is given by:

$$\hat{X} = \sum_{m=1}^{n} X_{i_m}/np_{i_m}, \tag{1}$$

FIG. I Diagram showing joint probability of units i and j when $np_i + np_j \leq 1$



Explanation of Above Table

C_1 A of length $d_1 = np_i$ is projected upon the d-axis, represented by OP_2 or any parallel line in the shaded area. The abscissa of $AP_2 = d_1$. The distances B_i C_i between the two diagonal lines B_1 B_2 ... B_5 and C_1 C_2 ... C_5 is of length np_j and the abscissae of all points on C_1 C_2 ... C_5 represent all possible values of d_{k+2} while those of B_1 B_2 ... B_5 represent those of d_{k+1}. If B and C represent the positions on B_1 B_5 and C_1 C_5 for a given set of k units between units i and j, then the portion of BC covered in the shaded area represents the joint probability of selection of units i and j.

Here, we are assuming $np_i \leq np_j$ and $np_i + np_j \leq 1$.

and its true and estimated variance are given by:

$$V(\hat{X}) = \sum_{i=1}^{N} \sum_{j>i} (np_i \cdot np_j - \pi_{ij}) \left(\frac{X_i}{np_i} - \frac{X_j}{np_j}\right)^2 \tag{2}$$

and

$$\hat{V}(\hat{X}) = \sum_{m=1}^{n} \sum_{q>m} \left(\frac{np_{i_m}\, np_{i_q}}{\pi_{i_m i_q}} - 1\right) \left(\frac{X_{i_m}}{np_{i_m}} - \frac{X_{i_q}}{np_{i_q}}\right)^2 \tag{3}$$

respectively, applying Yates-Grundy formulas [6].

Here, π_{ij} is the joint inclusion probability of units i and j in the sample.

(3) is an unbiased estimate of (2) only if the joint probabilities of all $\binom{N}{2}$ pairs of units in the population are non-zero. Otherwise, by direct substitution of the individual and joint probabilities and the statistics of the selected units in (3), (2) will be under-estimated by the contribution by the double summation $\sum_{i=1} \sum_{j>i}$ of (2), in which $\pi_{ij} = 0$.

(2) may be written also as:

$$V(\hat{X}) = \sum_{i=1}^{N} np_i \left(\frac{X_i}{np_i} - \frac{X}{n}\right)^2 + \sum_{i=1}^{N} \sum_{j\neq i} \pi_{ij} \left(\frac{X_i}{np_i} - \frac{X}{n}\right) \left(\frac{X_j}{np_j} - \frac{X}{n}\right) \tag{4}$$

and be defining

$$\sigma^2 = \sum_{i=1}^{N} p_i \left(\frac{X_i}{Np_i} - \frac{X}{N}\right)^2 \tag{5}$$

and

$$r_{FP} = \frac{1}{n(n-1)\sigma^2} \sum_{i\neq j}^{N} \pi_{ij} \left(\frac{X_i}{Np_i} - \frac{X}{N}\right) \left(\frac{X_j}{Np_j} - \frac{X}{N}\right) \tag{6}$$

and $(\hat{X})$ may be simply written as:

$$V(X) = N^2 \frac{\sigma^2}{n} [1 + (n-1)\, r_{FP}]. \tag{7}$$

The above formulas hold true, for all pps sampling procedures where Horvitz-Thompson estimators are employed. When the sizes are equal, σ^2 reduces to the classical formula $\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} (X_i - X/N)^2$ and in the case of simple random sampling without replacement, $r_{FP} = -1/(N-1)$, the classical finite population correlation. When sampling with pps with replacement, $r_{FP} = 0$ in (7).

In the case of systematic sampling with an equal step interval of N/n for units in a predetermined order, r_{FP} becomes the serial correlation whereby $\pi_{ij} = n/N$ for units N/n or a multiple of N/n apart and =0 for all other pairs. In the case of systematic sampling with pps among units in a fixed order of listing, r_{FP} may be defined as the serial correlation generalized to pps systematic whereby $\pi_{ij} = 0$ for those pairs of units that have no chance of entering the sample together.

A general formula for the joint probability of pairs of units being selected by systematic pps in a pre-determined order of listing and then the algorithm to calculate π_{ij} over all random arrangements of units will be described.

<u>Joint Probability of Selection of Units i and j when units in fixed order</u>

For a specified order of listing of the N units with units i and j separated by k units, π_{ij} will remain unchanged if we renumber the units so that i is the first unit, j is the (k+2)th unit and (i-1) is the last unit. π_{ij} will also remain unchanged if we interchange units i and j, if necessary, so that $np_i \leq np_j$ since the order of listing could be reversed without affecting the joint probabilities. If this is undertaken, Table 1 would be altered thus:

Table 2: Data as in Table 1 with renumbering as above

Unit No.	Relative Size	Probability of Selection	Accumulated Probabilities
i	p_i	np_i	$np_i = d_1$
	p_{i_2}	np_{i_2}	.
			.
k	.	np_{i_3}	.
units	.		.
a_k	.	.	.
	$p_{i_{k+1}}$	$np_{i_{k+1}}$	.
j	p_j	np_j	$I_{k+2} + d_{k+2}$
N-	.	.	.
(k+2)	.	.	.
remaining		.	.
units	.	.	.
a_{N-2-k}	p_{i_n}	np_{i_N}	n

Originally the cumulative probabilities for unit i was $I_i + d_i$ and for unit j, $I_{j+k+1} + d_{j+k+1}$ before renumbering the units. By deducting $c = I_i + d_i - np_i$ from the cumulative probabilities for units i to N and by adding n-c to the cumulative probabilities for the remaining units, we arrive at the new cumulative probabilities in Table 2. The original random number r may be adjusted to $r' = r - c + A$ for some integer A to ensure that $r' \varepsilon [0,1)$ and the random number r' will yield exactly the same sample in the renumbered list as r in Table 1.

If the original listing of units is simply reversed, the random number r may be adjusted to $r' = 1 - r$ with $r' = 0$ if $r = 0$ to ensure that $r' \varepsilon[0,1)$ instead of (0,1]. By employing r' in the reverse order the same selection will result as with random number r.

Units i and j are both selected with a fixed selection of k units or its complement of N-k-2 units lying between i and j when the adjusted random number $r' \leq np_i$ and $r' + a \in (I_{k+1} + d_{k+1}, I_{k+2} + d_{k+2}]$ for some integer a.

In FIG. I, the individual and joint probabilities of units i and j are illustrated by a series of lines parallel to the d-axis with C_1 A and its parallel lines in the shaded area representing $d_1 = np_i$ and BC representing np_j.

By observing the relative positions of B_iC_i (for different i) with respect to the shaded area, we can determine the value of π_{ij} for each position as follows: beginning with $d_{k+2} = 0$. If $\underline{a}_k$ represents a vector of k selected units between units i and j, $\pi_{ij|\underline{a}_k}$ represents the conditional joint probability for the given set.

Table 3: Value of $\pi_{ij|\underline{a}_k}$ according to range of values of d_{k+2} $(np_i + np_j \leq 1)$[2]

	Range of d_{k+2}	Value of $\pi_{ij\|\underline{a}_k}$
(3.1)	$[0, np_i)$	d_{k+2}
(3.2)	$[np_i, np_j)$	np_i
(3.3)	$[np_j, np_i + np_j)$	$np_i + np_j - d_{k+2}$
(3.4)	$[np_k + np_j, 1)$	0

By an argument similar to the above, using a figure such as the above for the case when $np_i + np_j > 1$, although each <1, we can obtain the probabilities $\pi_{ij|\underline{a}_k}$.

Table 4: Value of $\pi_{ij|\underline{a}_k}$ according to range of values of d_{k+2} $(np_i + np_j > 1)$[2]

	Range of d_{k+2}	Value of $\pi_{ij\|\underline{a}_k}$
(4.1)	$[0, np_i + np_j - 1)$	$np_i + np_j - 1$
(4.2)	$[np_i + np_j - 1, np_i)$	d_{k+2}
(4.3)	$[np_i, np_j)$	np_i
(4.4)	$[np_j, 1)$	$np_i + np_j - d_{k+2}$

[2] These results agree closely with those derived by W.S. Connor [1] in 1966.

Consider the two sets of units given below and the characteristic values.

Table 5: Set No. 1 (Horvitz-Thompson [4])

Unit No.	p_i	$2p_i$	$t_i = I_i + d_i$	X_i
1	18/394	18/197	18/197	19
2	9/394	9/197	27/197	9
3	14/394	14/197	41/197	17
4	12/394	12/197	53/197	14
5	24/394	24/197	77/197	21
6	25/394	25/197	102/197	22
7	23/394	23/197	125/197	27
8	24/394	24/197	149/197	35
9	17/394	17/197	166/197	20
10	14/394	14/197	180/197	15
11	18/394	18/197	1 + 1/197	18
12	40/394	40/197	1 + 41/197	37
13	12/394	12/197	1 + 53/197	12
14	30/394	30/197	1 + 83/197	47
15	27/394	27/197	1 110/197	27
16	26/394	26/197	1 + 136/197	25
17	21/394	21/197	1 + 157/197	13
19	19/394	19/197	1 + 185/197	19
20	12/394	12/197	2	12

p_i based on eye-estimated count of households in each of 20 blocks of Ames, Iowa X_i = actual count of households in the correcponding blocks.

Table 6: Set No. 2 (I.P. Fellegi[2])

Unit No.	p_i	$2p_i$	$t_i = I_i + d_i$	X_i
1	.10	.20	.20 = 0 + .20	.60
2	.14	.28	.48 = 0 + .48	.98
3	.17	.34	.82 = 0 + .82	1.53
4	.18	.36	1.18 = 1 + .18	2.16
5	.19	.38	1.56 = 1 + .56	2.85
6	.22	.44	2.00 = 2 + .0	4.18

To find the joint probability of units 2 and 5 being selected in Set No. 2, we note that unit 2 is the smaller so we renumber; thus:

Unit No.	p_i	$2p_i$	t_i	
2	.14	.28	.28	$d_1 = .28$
3	.17	.34	.62	
4	.18	.36	.98	
5	.19	.38	1.36	$d_1 = .36$
6	.22	.44	1.80	
1	.10	.20	2.00	

Here, k = 2, $d_1 = .28$, $d_4 = .36$, $2p_2 + 2p_5 = .66 < 1$ so we use Table 3 to find $\pi_{2,5|\underline{a}_2}$ where $\underline{a}_2 = (3,4)$.

Now $d_4 = .36$ lies between $2p_2 = .28$ and $2p_5 = .38$ and so by condition 3.2 of Table 3, $\pi_{2,5|\underline{a}_2} = .28$.

Proceeding in this manner for all pairs of units of Set 2 in the given order of listing, we find the joint probabilities, as follows:

Table 7: Value of π_{ij} (fixed order) as of Set 2 Unit i

Unit j	1	2	3	4	5	6
1	-	.0	.0	.18	.02	.0
2	.0	-	.0	.0	.28	.0
3	.0	.0	-	.0	.08	.26
4	.18	.0	.0	-	.0	.18
5	.02	.28	.08	.0	-	.0
6	.0	.0	.26	.18	.0	-
Sum	.20	.28	.34	.36	.38	.44

Turning now to the case of randomized order of units, we see that, for each given pair of units i and j, we need only find $\pi_{ij|\underline{a}_k}$ for $k = 0,1,2, \ldots, \frac{N-3}{2}$ if N odd and for $k = 0,1,2, \ldots, \frac{N-2}{2}$ if N even and only half of the sets for $k = \frac{N-2}{2}$ since each set of $\underline{a}_k$ yields a complement set with also N-2/2 units with the same joint probability. The arrangement of the k units in a given set $\underline{a}_k$ lying between i and j is immaterial since the cumulative probabilities remain unchanged for units i and j. Hence, we need only consider all possible selections of k units as k proceeds from k = 0 (when i and j are adjacent in the list) to k = N-3/2 or N-2/2. Or, we find $\pi_{ij|\underline{a}_k}$ for each set $\underline{a}_k$ and take an average over $\binom{N-2}{k}$ selections of k units and over (N-1)/2 distinct values of k, or for random ordering of units or

$$\pi_{ij} = \frac{1}{N-1} \sum_{k=0}^{N-2} \pi_{ij|k}$$

$$\pi_{ij} = \frac{2}{N-1} \sum_{k=0}^{(N-3)/2} \frac{1}{\binom{N-2}{k}} \sum_{\underline{a}_k} \pi_{ij|\underline{a}_k} \text{ for N odd}$$

$$= \frac{2}{N-1} \left[\sum_{k=0}^{(N-4)/2} \frac{1}{\binom{N-2}{k}} \sum_{\underline{a}_k} \pi_{ij|\underline{a}_k} + \frac{1}{\binom{N-2}{\frac{N-2}{2}}} \sum^*_{\underline{a}(N-2)/2} \pi_{ij|\underline{a}(N-2)/2} \right]^3 \qquad (8)$$

for N even.

3 This formula was also derived by W.S. Connor [1] who noted the symmetry in the joint probabilities for k and N-2-k units apart but not for each half of all the sets when k=(N-2)/2.

Σ* denotes summation over only 1/2 of the sets of N-2/2 units (easily accomplished by fixing a particular unit in each set when calculating the conditional joint probabilities).

In effecting the calculation of $\pi_{ij|k}$, the conditional joint probability of k units lie between i and j, the $\binom{N-2}{k}$ selections may be partitioned into four separate groups satisfying conditions 3.1, 3.2, 3.3, or 3.4 by examining the ranges of values of d_{k+2} over sub-sets of k selected units. In fact, $\pi_{ij|k}$ may be zero regardless of the selected units listed between i and j, a fact easily established by determining the range of values of $I_{k+2} + d_{k+2}$ between the smallest k units and the largest k units. Thus, the algorithm for calculating $\pi_{ij|k}$ for each value of k may be simplified and the calculations reduced by ordering the N-2 units in ascending order of size between i and j and examining the range of value of $I_{k+2} + d_{k+2}$.

Considering set 2 with 6 units, suppose now that $\pi_{2,5}$ is to be calculated over all possible random orderings of units.

Table 8: Re-ordering for calculation of $\pi_{2,5}$ (Set 2)

Unit	$2p_i$	Sum	Comments
2	.28		
5	.38	.66	Sum < 1, so use Table 3
remaining units ordered by size	$2p_i$		
1	.20	.20	
3	.34	.54	
4	.36	.90	
6	.44	1.34	

Calculations of $\pi_{2,5|k}$ for:

k=0 since d_2 = .66, condition 3.4 holds and so $\pi_{2,5|0} = 0$

k=1 Min. $I_3 + d_3$ = .86 and Max. $I_3 + d_3$ = 1.10 so for some sets $\underline{a}_1$, $\pi_{2,5|\underline{a}_1} \neq 0$ since d_3 = .10 for Max. $I_3 + d_3$, for which condition 3.1 holds. So proceeding thus; for

$\underline{a}_1$=(6) ,d_3=.10 ,$\pi_{2,5|(6)}$=.10 by 3.1,

$\underline{a}_1$=(4) ,d_3=.02 ,$\pi_{2,5|(5)}$=.02 by 3.1,

$\underline{a}_1$=(3) or (1),$d_3 \epsilon$[.66,1),$\pi_{2,5|\underline{a}_1}$ =0

Sum .12

Average $.12/\binom{4}{1}$= .12/4 = .03 = $\pi_{ij|1}$

k=2=(N-2)/2 so that we need consider only one-half of the $\binom{4}{1}$ sets, by fixing a particular unit, say no. 1 in each set.

Then, with unit no. 1 included in the 3 possible sets.

Min $I_4 + d_4$ = .66 + .20 + .34 = 1.20 or d_4 = .20

Min $I_4 + d_4$ = .66 + .20 + .44 = 1.30 or d_4 = .30

both yielding non-zero joint probabilities so for k=2, when

$\underline{a}_2$=(1,3),d_4=.20,$\pi_{2,5|(1,3)}$=.20 by 3.1,

$\underline{a}_2$=(1,4),d_4=.22,$\pi_{2,5|(1,4)}$=.22 by 3.1, and

$\underline{a}_2$=(1,6),d_4=.30,$\pi_{2,5|(1,6)}$=.28 by 3.2

Sum =.70

Average =.70/6=.116=$\pi_{2,5|2}$

Hence, $\pi_{2,5} = \frac{2}{5}\ (0+.0\dot{3} + .11\dot{6}) =$

$.4 \times .146 = .0586.$

Proceeding in like manner as above, for the remaining 14 pairs of units, we may derive the joint probability matrix in Table 9 and a similar matrix in Table 10 for the case of n=3 for the same population.

Table 9: Joint Probability for each pair of units of Set 2 (n=2) Unit i

Unit j	1	2	3	4	5	6
1	-	.0386	.0386	.0386	.0420	.0420
2	.0386	-	.0486	.0553	.0586	.0786
3	.0386	.0486	-	.0753	.0786	.0986
4	.0386	.0553	.0753	-	.0853	.1053
5	.0420	.0586	.0786	.0853	-	.1153
6	.0420	.0786	.0986	.1053	.1153	-
Sum	.2000	.2800	.3400	.3600	.3800	.4400

Table 10: Joint Probabilities between Pairs of Units (Population 2 when n=3)
Unit i

Unit j	1	2	3	4	5	6
1	-	.1080	.1163	.1213	.1263	.1280
2	.1080	-	.1630	.1680	.1730	.2280
3	.1163	.1630	-	.2096	.2346	.2963
4	.1213	.1680	.2096	-	.2596	.3213
5	.1263	.1730	.2346	.2596	-	.3463
6	.1280	.2280	.2963	.3213	.3463	-
Sum	.6000	.8400	1.0200	1.0800	1.1400	1.3200

Check = 2 x $3p_1$ = 2 x $3p_2$, etc.

and the calculations for say $\pi_{5,6}$ in Table 10:

i	$3p_i$	Sum				
5	.57					
6	.66	1.23 > 1 so $\pi_{5,6\|0}$ = .23 (by 4.1)				
-		k=1	$\pi_{5,6\|\underline{a}_1}$	$\underline{a}_2$	k=2	$\pi_{5,6\|\underline{a}_2}$
1	.30	1.53	.53 (by 4.2)	1,2	1.95	.28(by 4.4)
2	.42	1.65	.57 (by 4.3)	1,3	2.04	.23(by 4.1)
3	.51	1.74	.49 (by 4.4)	1,4	2.07	.23(by 4.1)
4	.54	1.77	.46 (by 4.4)			

$$\pi_{5,6|1} = \frac{2.05}{4} = .5125 \quad \pi_{5,6|2} = \frac{.74}{6} = .12\dot{3}$$

Hence, $\pi_{56} = .4\ (.2300 + .5125 + .123\dot{3}) =$

$.4 \times .8658\dot{3} = .346\dot{3}.$

For a population of 6 units, to calculate the joint probability of any pair of units requires only the calculation for 8 distinct arrangements of units and reading off the probabilities from Table 3 or 4 are required before a direct substitution in formula (8) can be made. However, for N units in general, 2^{N-3} distinct arrangements and readings from Table 3 or 4 must be obtained before substitution in formula (8). This is clearly impractical for manual calculations beyond say N=10 and perhaps even for a computer operation beyond say N=20 since 2^{17} = 131,072 and the running time on the computer could be extensive. The calculations may sometimes be considerably reduced by averaging over many possible selections of units between a given pair of units rather than by reading off the joint probability for each of the 2^{N-3} possible selections of units and then averaging the probabilities.

References

[1] Connor, W.S., "An exact formula for the probability that two specified sampling units will occur in a sample drawn with unequal probabilities and without replacement", Journal of the American Statistical Association, Vol. 61, (1966), pp. 384-390.

[2] Fellegi, I.P., "Sampling with varying probabilities without replacement: rotating and non-rotating samples", Journal of the American Statistical Association, Vol. 58, (1963), pp. 183-201.

[3] Hartley, H.O. and Rao, J.N.K., "Sampling with unequal probabilities without replacement from a finite universe", Journal of the American Statistical Association, Vol 47, (1952), pp. 663-685.

[5] Rao, J.N.K., "Sampling procedures involving unequal probability selection", (1961), unpublished Ph.D Thesis, Iowa State University, Ames, Iowa.

[6] Yates, F. and Grundy, P.M., "Selection without replacement from within strata and with probability proportional to size", Journal of the Royal Statistical Society, Series B, Vol. 15 (1953), pp. 253-261.

CHANGES IN THE OCCUPATION AND INDUSTRY CLASSIFICATION SYSTEMS BETWEEN THE 1960 AND 1970 CENSUS OF POPULATION

Stanley Greene and John Priebe
Social and Economic Statistics Administration, Bureau of the Census

I. INTRODUCTION

The world of work is of central concern in a person's life and the skills of its people a key to a country's development. The world of work comprises an extreme diversity of activities. The range of occupations extends from A. B. Seamen to Zoologist with 23,000 intervening discrete job titles in our Alphabetical Index of Industries and Occupations. To bring order and understanding to this extensive variety--to make it mentally manageable a system of classification or ordering is necessary. The simplest questions related to labor supply and demand, education, planning, and wage differentials require some systematic ordering of these information bits; thus the need for a classification system.

In recognition of this need for a systematic classification of occupations and industries the Federal Censuses have been developing classification systems in these fields since 1820. At the Census of 1820 the classification system comprised three major divisions, namely agriculture, commerce, and manufacturers. No occupational data were collected at the Census of 1830 but at the Census of 1840 the classification system had grown by over 100 percent constituting seven classes--mining; agriculture; commerce; manufacturers and trades; navigation of the ocean; navigation of canals, lakes, and rivers; and learned professions and engineers. The Census has been growing, changing, and adapting ever since.

It certainly was simpler to be a labor force analyst in the good old days. Whereas there were only three discrete "occupation" items in 1820, today there are three distinct classification systems--The occupational system comprising 441 items, the industrial system comprising 212 items, and the class of worker system comprising 6 items.

Although for historical and comparability purposes it is desirable to maintain consistency over time, the occupational data collected and presented in the several decennial censuses have seen changes introduced in the detail of occupation classification used. These changes are necessitated by a dynamic American economy--new industries, new methods, new occupations. The decade between 1960 and 1970 was no exception and many modifications were made.

These alterations are necessitated by a number of factors; Introduction of new jobs, change in work content, change in terminology, changes in schedule wording and enumeration techniques, and the general desire for improvement.

II. PURPOSE OF CHANGE

For the 1970 Census we had quite specific goals in mind in formulating our plans for the occupational classification system revisions. These were:

1. Reduction in the size of large categories.
2. More specificity relative to general categories.
3. More homogeneity among specific census occupation categories.
4. Identification of new and significant occupations.
5. Feasibility of coding actual census responses to the proposed categories.
6. Increased accuracy of terminology in titles and content of categories.

In recognition of the desirability of maintaining historical comparability, constraining guidelines were followed: introduce no category shifts between major groups except where compelling reasons existed in situations of evident misclassification and to introduce no new major groups nor eliminate any if comparability were affected.

With these goals and restrictions in mind, an organized and systematic review was undertaken of the job titles comprising each of the 297 occupation categories of 1960. Each of approximately 25,000 job titles was examined to determine whether it was to remain in a particular category or be shifted to an existing category or be included among those titles comprising a newly established category.

III. REVISION MATERIALS

Several important tools were used in revising the 1960 system. Among them were the following:

(1) The 1960 Classified Index of Occupations and Industries. This volume presents the job titles comprising the occupation categories established for the 1960 Census and was of primary importance to the undertaking.

(2) The Current Population Survey (CPS). This survey and other sources were used in identifying new and emerging occupation titles.

(3) The unpublished CPS annual average figures of the detailed occupation distribution. These figures served to identify growth occupations.

(4) The Dictionary of Occupational Titles (DOT).

(5) A special cross-tabulation of a CPS panel coded in both Census and DOT terms.

(6) A 100,000 card sample of the 1960 census occupation entries. This sample of the experienced civilian labor force in the 1960 census contains industry, occupation, and class of worker codes, and selected demographic characteristics. The sample card also includes the written occupation and industry responses clerically transcribed from the 1960 census schedules. This arrangement permitted the study of the job title composition of each specific occupation code category in quantitative terms. Since the sample was representative of the 1960 census, the card deck also allowed for the calculation of estimates of the size of the various revisions. One major result of these exercises was an approximate 50 percent increase in the number of occupations over the 1960 system: 441 compared with 297.

A significant aspect of the improved ability to make these additional distinctions stems from the introduction of two supplementary probing items for occupation in the 1970 census questionnaire.1/ The two probes ask for the major activities and job title in addition to "kind of work", providing more information for distinguishing the various occupations.

IV. DETAILED TYPES OF REVISIONS

A major effort was made to reduce the size and improve the identification of the "not elsewhere classified"(n.e.c.) categories, since some of these groups were too broadly descriptive and included a large number of workers. These categories were refined in the following fashion:

(1) New categories were established from components. For example, precision machine operatives were extracted from the "Operatives, n.e.c." and identified in four distinct occupations. In this fashion, 18 new occupation groups and large portions of four others were formed from the "Operatives, n.e.c." category. Moreover, a great number of job titles were shifted from this generalized group into more specifically identified categories.

(2) The "n.e.c." categories, where appropriate, were split among their integral components, i.e., generalized titles and minor occupations. For example, after extracting the significant and identifiable occupations and many job titles associated with existing occupations from the "Clerical workers, n.e.c." category, the remaining titles were assigned to two component categories: "Miscellaneous clerical workers," and "Not specified clerical workers."

(3) Many job titles were shifted about in order to improve the homogeneity of the occupation categories; for example, "impressario" was shifted from "Clerical workers (Agent, n.e.c.)" to "Managers and administrators." "Jailer" was moved from "Officials and administrators, n.e.c." to "Guards and watchmen."

In addition, three basic types of revisions were introduced:

(1) An existing category was separated into two or more new categories where the components were sufficiently large and occupationally distinctive. For example, "Automobile mechanics" was divided into two groups--"Automobile body repairmen" and "Automobile mechanics."

(2) In some cases, a portion of an existing category was subsumed by another category. For example, "marine engineers" was shifted from "ship officers" in "Managers" to "Mechanical engineers" in order to improve both categories in terms of uniform composition.

(3) Some categories were eliminated, and the components were placed elsewhere. For example, "Agents, n.e.c." represented a large heterogeneous "clerical" category in 1960. For 1970, all its component occupations were distributed to appropriate existing categories or newly established categories. In like manner, certain categories were combined in order to eliminate archaic or unnecessary divisions. For example, "Express messengers and railway mail clerks" was reclassified into two existing categories: "Railway mail clerks" were combined with "Postal clerks" and "Express messengers" with "Miscellaneous clerical workers." Another example, since "Baggage and cargo agents" were found to be doing like work as "Ticket, station, and express agents", these groups were combined.

Other aspects of the detailed revision process concerned the elimination of residual industrial overtones persisting in the occupational system. For example, such job titles as "carpenters" and "bulldozer operators" were removed from "Mine operatives and laborers, n.e.c." and "Lumbermen, raftsmen, and wood choppers" and placed in their respective occupation categories.

The system also addresses itself to growth occupations during the decade. For example, the computer field is represented by five distinct occupations in 1970. (In 1960 there were none.) In like manner, social welfare, health, and educational services have experienced a sizeable expansion in separately identified occupations.

Other aspects of the revised system have to do with a more flexible approach to computer processing of the data.

One such innovation is the inclusion of an unique "allocation" code for each major group. "Not reported" cases are assigned to one of these categories by the computer, based upon selected reported demographic characteristics. This results in the elimination of the single line "Occupation not reported" category used in 1960.

The second major innovation is that greater use is made in 1970 of "tabulation categories" which generally have not been published in printed census reports but have been reserved for special reports. In most other instances they will be combined with other groups for census publications. For example, medical and legal secretaries have been set up as tabulation categories. Another illustration--it was deemed appropriate to count the "apprentice" categories of 1960 presented then as "Operatives" with their counterpart craft occupations. The reasons behind this move were:

(1) To be consistent with the treatment of "trainees" which were classified with the craft.

(2) To recognize the fact that many apprentices often work at a skill level equivalent to the journeymen.

(3) The imperfect census count of apprentices, since many of them do not report they are in an apprenticeship program and thus are classified as in the craft. In some special reports, however, apprentices will be separately identified.

A third innovation is the use of the computer to supplement the hand coding activities. For example, it was deemed desirable to separate the large heterogeneous category of 1960 "Salesmen, n.e.c." into sales representatives, sales clerks, and salesmen of services. An examination of the responses to the expanded occupation item from the Census Pretest Program indicated that the information provided was insufficient to make this separation. However, a close correlation exists between certain detailed industries and the various types of sales personnel; therefore, although the coders will continue to classify by one broad code, the computer will be used to refine this code into five groups based upon the detailed industry code.

Another extended application of the computer capability was through computer edit of clerically assigned codes. For example, a new group "Restaurant, cafeteria, and bar managers" was established. This group was achieved by having the coders combine those clear-cut cases that are so designated verbally with those less specific entries that have an industry classification of "Eating and drinking places." That portion of this complex coding having to do with nonspecific manager entry coupled with the "Eating and drinking places" industry was then verified by computer to assure accuracy.

V. THE OCCUPATIONAL CLASSIFICATION SYSTEM

As noted, the new occupation system incorporates a sizable number of changes vis-à-vis the 1960 system. In general, many more detailed and more homogeneous categories were established: 441 compared with the 297 in 1960. For 1970, there are 12 major groups rather than the 11 of 1960. The new major group entitled "Transport equipment operatives" includes bus drivers, parking attendants, truck drivers, etc. The shift in categories comprising the new major group was confined to the 1960 major group "Operatives and kindred workers" to provide for historical comparability. A second revision (shown below) was the rearrangement of the major groups to reflect the traditional broad occupational areas--White collar workers, blue collar workers, farm workers, and service workers. The four worker divisions have been placed in their traditional sequence, but such sequence is not intended to imply that any division has a higher social or skill level than another.

1970 MAJOR OCCUPATION GROUPS

	Number of detailed categories comprising groups 1960	1970
White collar workers:		
Professional, technical, and kindred workers...........	84	124
Managers and administrators, except farm.................	13	24
Sales workers.................	9	15
Clerical and kindred workers..	28	48
Blue collar workers:		
Craftsmen and kindred workers.	61	96
Operatives, except transport..	53	54
Transport equipment operatives		12
Laborers, except farm.........	10	16
Farm workers:		
Farmers and farm managers.....	2	3
Farm laborers and farm foremen	4	5
Service workers:		
Service workers, exc. private household....................	28	38
Private household workers.....	4	6

Finally, rather than having the detailed categories with each major group listed alphabetically subgroupings or "families" have been established in several major groups. For example, the "Service workers" group, to clarify its content, has been recast into five subgroupings: Cleaning service, food service, health service, personal service, and protective service workers.

In addition to these conceptual revisions noted in the preceding statement, a great many detailed revisions were made to most of the 1960 categories having to do with sharpening homogeneity of the categories.

Quantification of the System

After placement of the job titles was completed, quantitative estimates of the 1970 classification changes were made for the 1960 experienced civilian labor force data published in the United States Census of Population: 1960. Detailed Characteristics, U.S. Summary. The following procedures were used to determine these levels.

The first step was to calculate on the basis of the 100,000 card sample (Item 6 of the source materials noted above) an estimate of the percent of a 1960 category which would be in a different category in the 1970 system. For each 1960 category then, cards containing job titles which were transferred to another category in the 1970 classification were counted separately for male and female. An adjustment percent was calculated individually for males and females by dividing the number of cards transferred to each 1970 category by the total number of cards in the 1960 category.

The second step was to apply these percents to the published data in the 1960 Detailed Characteristics, U.S. Summary table 201 to produce a numerical estimate of the change. For example, the 1960 category "lawyers and judges" was split into two distinct categories for 1970. There were 12 cards (male) for judges of the total 286 cards (male) in the category. Thus, judges formed 4.196 percent of the category. This percent was then applied to the 1960 total of 205,515 male "lawyers and judges" in the experienced civilian labor force to produce an estimate of 8,623 judges in 1960.2/

Table A presents the effect of the revisions at the major group level for the 1960 experienced civilian labor force.

The differences noted in Table A are the net result of changes across major occupational groups. The loss of a quarter million persons in the "Professional" category results mainly from the transfer of five categories out of the group. These were "School administrators", "Estimators and investigators", "Funeral directors", "Student nurses", and "Teacher's aides". A principle addition to the "Professional" major group was "craft teachers" added to "Adult education teachers" from "Craftsmen and kindred workers". Also, "not specified nurses" had been classified with "practical nurses", but our study indicated that many were more likely to be "registered nurses". Therefore, "not specified nurses" were moved to "Professional".

The addition of 218,000 "Managers and administrators" results mainly from the shifting of "School administrators" and "Funeral directors" out of "Professional" into this group, also important was the addition of titles from the "Agents, n.e.c." category in "clerical workers" to a number of "Manager" categories. Very little was moved out of this group.

The negligible decline in "Sales workers" (2,788) is accounted for by the removal of "Stock handlers" shifted to "Laborers, except farm". This was offset by a few smaller categories shifted into this group, primarily from "Agents, n.e.c." and "Professional, n.e.c.".

Table A.—SUMMARY OF CHANGES AMONG MAJOR OCCUPATION GROUPS - 1960 TO 1970 (1960 Experienced Civilian Labor Force)

MAJOR OCCUPATION GROUPS	TOTAL		
	1970 Classification System	1960 Classification System	Difference
EXPERIENCED CIVILIAN LABOR FORCE	67,990,073	67,990,073	-
PROFESSIONAL, TECHNICAL, AND KINDRED WORKERS	7,089,840	7,335,699	- 245,859
MANAGERS AND ADMINISTRATORS, EXCEPT FARM	5,708,247	5,489,489	+ 218,758
SALES WORKERS	4,798,553	4,801,341	- 2,788
CLERICAL AND KINDRED WORKERS	9,431,106	9,617,487	- 186,381
CRAFTSMEN AND KINDRED WORKERS	9,465,311	9,240,983	+ 224,328
OPERATIVES, EXCEPT TRANSPORT	9,580,810	12,846,044	- 591,764
TRANSPORT EQUIPMENT OPERATIVES	2,673,470		
LABORERS, EXCEPT FARM	3,755,237	3,530,022	+ 225,215
FARMERS AND FARM MANAGERS	2,527,755	2,525,907	+ 1,848
FARM LABORERS AND FARM FOREMEN	1,604,235	1,559,524	+ 44,711
SERVICE WORKERS, EXCEPT PRIVATE HOUSEHOLD	6,085,582	5,765,481	+ 320,101
PRIVATE HOUSEHOLD WORKERS	1,816,648	1,824,817	- 8,169
OCCUPATION NOT REPORTED	3,453,279	3,453,279	-

The decline of 186,000 in the "Clerical" group is due primarily to the breaking up and shifting out of such miscellaneous catchalls as "Agents, n.e.c." and "Clerical, n.e.c.". The major addition was the "Estimators and investigators" moved from "Professional".

The increase of 224,000 among the "Craftsmen" is largely accounted for by the shifting into the group of "Furniture and wood finishers" "caterpiller drivers operators", "Power station operators" and numerous job titles from "Operatives, n.e.c.".

The "Operatives" group declined by close to 600,000 and principally responsible for this shift was the loss of the additions to "Craftsmen" noted above; and work activities other than operatives in the "Mine operative" group, e.g., laborers and material handlers. Added to "Checkers, examiners, and inspectors; manufacturing" in the "Operatives" group were the "checkers and examiners" from "Clerical, n.e.c.". Also added to "Operatives" were "garage laborers" so as to be combined with "gas station attendents" and "sawyers" from "lumbermen". In addition the "Operative" group was divided in two--splitting out "transport equipment operatives".

The "Laborers, except farm" group was enlarged by 225,215. This is largely accounted for by "material handlers", "mine laborers", and certain other "n.e.c." titles from "Operatives". Other changes such as "sackers" from "Sales workers" to "Stock handlers" and the entire group of "Stock handlers" from "Operatives" to "Laborers" played a significant role. The "Farm workers" group saw little change. The increase of 320,000 in the "Service workers, except private household" group is caused by the shift there from "Craftsmen" of the "maintenance men", "medical and dental aides and assistants" from "Clerical", and "student nurses" and other "medical aides" from "Professional".

Table B quantifies all the shifts among the major groups. Table A provided the net shift for each major group. This net figure may be derived from Table B by subtraction of boxhead from stub line for corresponding groups and the table also charts the flow of changes of the '60 and '70 Classification System.

Table B.—SHIFTS AMONG MAJOR GROUPS BETWEEN 1960 AND 1970 OCCUPATION CLASSIFICATION SYSTEM (1960 Experienced Civilian Labor Force)

1970 major group gains	Total	1960 major group losses										
		Professional, technical, and kindred workers	Managers, officials, and proprietors, except farm	Sales workers	Clerical and kindred workers	Craftsmen, foremen, and kindred workers	Operatives and kindred workers	Laborers, except farm and mine	Farm and farm managers	Farm laborers and foremen	Service workers except private household	Private household workers
Total	1,837,406	-350,556	-26,897	-61,867	-260,354	-206,158	-748,432	-115,277	-1,496	-	-58,200	-8,169
Professional, technical, and kindred workers	+104,697	-	10,713	-	30,660	16,893	4,241	-	-	-	42,190	-
Managers and administrators, except farm	+245,655	178,112	-	-	64,132	674	-	2,737	-	-	-	-
Sales workers	+59,079	17,689	1,598	-	39,792	-	-	-	-	-	-	-
Clerical and kindred workers	+73,973	49,845	2,851	6,570	-	-	12,981	-	-	-	1,726	-
Craftsmen and kindred workers	+430,486	20,753	672	-	-	-	395,128	13,933	-	-	-	-
Operatives, except transport	+148,679	-	58	-	37,596	42,030	-	67,028	-	-	1,967	-
Transport equipment operatives	+7,989	-	-	-	5,226	1,457	-	1,306	-	-	-	-
Laborers, except farm	+340,492	3,466	2,208	55,297	11,8 5	2,754	251,079	-	1,496	-	12,317	-
Farm and farm managers	+3,344	-	-	-	-	-	3,344	-	-	-	-	-
Farm laborers and farm foremen	+44,711	-	-	-	-	-	44,075	636	-	-	-	-
Service workers, except private household	+378,301	80,691	8,797	-	71,073	142,350	37,584	29,637	-	-	-	8,169
Private household workers	-	-	-	-	-	-	-	-	-	-	-	-

VI. THE INDUSTRIAL CLASSIFICATION SYSTEM

The 1970 industrial classification system is patterned after the 1967 Standard Industrial Classification(SIC) established by the Office of Management and Budget. The revised Census classification contains 227 categories as opposed to 150 in 1960; again, like the occupation system about a 50 percent increase in detail. Most of the changes have resulted in establishing smaller,more homogeneous groups from larger groups. A major classification revision which was made had to do with the transfer of government welfare activities to "Welfare services" in the major group "Professional and related services" from "Public administration".

Although the Census system is developed from the Standard Industrial Classification, there is one difference which should be noted. In the 1967 SIC,all government workers are classified under a "government division" by level of government regardless of their activities. Within each level of government the SIC is further classified by industrial activity. One of these industrial activities is "regular government function--executive, legislative,and judicial". The Population Census, on the other hand, classifies all like industrial activity together without regard to government ownership. This classification system includes a major group "Public administration" which can be equated to the "regular government functions"category in the SIC. The Population Census identifies government ownership, by level of government, in its class of worker item. For example, the SIC would classify a person employed by the highway construction activity of the State highway department primarily as State government, and secondarily as "general contractor,except building". The Census would classify the person in the "general contractor, except building" industry,and identify the State government ownership in the independent class of worker item. It should be noted that the 1972 edition of the SIC has revised its system to eliminate the government primary sort. The following sources were used in revising the system:

1. The Standard Industrial Classification (1967).
2. The United States Bureau of the Census, County Business Patterns 1967, U.S. Summary CBP-67-1 giving first quarter 1967 employment data according to the SIC.
3. The 100,000 card sample of the 1960 census which was also used for revising the occupation classification.

The 100,000 card sample permitted the preparation of estimates of the effect of the revisions on the classification of 1960 data. The process of assigning weights to these cards was the same as used for the occupation revisions. Table 210 which presents detailed industry data in the Detailed Characteristics, U.S. Summary was used as a base for applying the industry adjustment percents to form the numerical estimates.

Presented in Table C is the 1960 experienced civilian labor force classified by the 1960 and 1970 systems showing differences at the **major** group level.

Table C.--SUMMARY OF CHANGES AMONG MAJOR INDUSTRY GROUPS - 1960 TO 1970
(1960 Experienced Civilian Labor Force)

MAJOR INDUSTRY GROUPS	TOTAL		
	1970 Classification System	1960 Classification System	Difference
EXPERIENCED CIVILIAN LABOR FORCE	67,990,078	67,990,078	-
AGRICULTURE, FORESTRY, AND FISHERIES	4,518,771	4,519,332	- 561
MINING	713,661	713,661	-
CONSTRUCTION	4,302,307	4,302,307	-
DURABLE GOODS MANUFACTURING	10,417,534	10,413,191	+ 4,343
NONDURABLE GOODS MANUFACTURING	8,118,930	8,122,712	- 3,782
TRANSPORTATION, COMMUNICATIONS, AND OTHER PUBLIC UTILITIES	4,633,016	4,633,016	-
WHOLESALE AND RETAIL TRADE	12,362,554	12,362,554	-
FINANCE, INSURANCE, AND REAL ESTATE	2,749,175	2,749,022	+ 153
BUSINESS AND REPAIR SERVICES	1,679,789	1,683,297	- 3,508
PERSONAL SERVICES	4,077,788	4,074,359	+ 3,429
ENTERTAINMENT AND RECREATION SERVICES	553,767	553,767	-
PROFESSIONAL AND RELATED SERVICES	7,834,278	7,714,244	+ 120,034
PUBLIC ADMINISTRATION	3,181,447	3,301,555	- 120,108
INDUSTRY NOT REPORTED	2,847,061	2,847,061	-

Since almost all the changes affecting the industry classification system were internal group splits, very little net change occured between groups. The 120,000 shift between "Public administration" and"Professional and related services" is caused, as noted earlier, by the shift of welfare activities from the former to the latter.

VII. EFFECT UPON "n.e.c." GROUPS

One of the major criticisms of the former occupational classification system centered on the "not elsewhere classified" groups. Those are the generalized residual categories,and the complaint was that they comprised a too large proportion of the work force. The attempt at resolution of this problem was a twofold approach. The first, as noted earlier, was to divide the n.e.c. groups into its two component parts, where possible, the not specified component, where the respondent entries were too vague to classify in a more refined manner,and the miscellaneous component which contains specific jobs but of limited significance or magnitude. This latter group was carefully examined to determine if new occupation groups could be established or job titles could be absorbed within existing groups. The measure of success this review met is shown in Table D.

It may be observed that whereas the 1960 system had 35 percent of its workers classified into 12 large n.e.c. groups, the reclassification reduced this to 23 percent for 15 categories. These 15 categories can be classified into three types. One type retains the label of "not elsewhere classified". Since the nature of the work performed by, e.g. "managers" and "salesmen" does not permit a determination of whether the entry is as detailed as possible. The other two, as noted earlier,are called "miscellaneous" and "not specified". The "miscellaneous" covers low frequency specific job titles;whereas the "not specified" categories include nondescriptive titles such as "office worker", "laborer", etc. Therefore, the 23 percent of the labor force included in these nonspecific occupation categories are distributed into these three types as follows:

The "not elsewhere classified" is 14 percent, the "miscellaneous" is 3.5 percent,and the "not specified"is 6 percent. Particularly noteworthy is the "Operatives" group where the "n.e.c." category was reduced from 7.3 percent of the ECLF to 3.8 percent, a decline of 48 percent. Moreover the 3.8 percent is distributed over four residual categories. The above figures were all standarized for the 1960 ECLF date base.

Prepublication data shows that these residual categories represent a lower proportion of the 1970 Census ECLF than the adjusted 1960 data presented here. In the 1970 Census, factors other than the reclassified occupation system play a role in the proportion these residual categories are to the whole. These factors would be the question wording and the quality of field and coding operations. The most important of these is the presence of two additional occupation questions on the 1970 Census questionnaire.

VIII. A STANDARD OCCUPATIONAL CLASSIFICATION SYSTEM?

The Interagency Occupational Classification Committee (of which the authors of this paper are members) under the auspices of the Office of Management and Budget played a strong role in providing an overview, guidance, and approval in the revision process. We look forward to the next quantum leap in the occupational classification work. This is a major attempt to establish a Standard Occupation Classification System. Such a system would parallel the Standard Industrial Classification System and would provide a uniform government-wide standard. This work has been underway for some time by the Interagency Committee. In fact, the revisions in the "Professional and technical"group of the 1970 Occupation system were largely the result of efforts in this direction.

During this experimental and evaluation work we hope to also introduce considerations relative to the major group classifications. These groupings have fallen under criticism for not being truly a socioeconomic grouping and for failing to be responsive to other concerns. We are ambitious enough to be considering a number of different sets of standard groupings based on a variety of criteria--earnings, training requirements, and status thay may serve to answer the different sets of questions asked of the data.

1/ A discussion of the background and development is given in "Some Recent Decennial Occupational Experimental Work" by Stanley Greene; Proceedings of the Social Statistics Section, 1966, American Statistical Association.

2/ For measures of the detailed movement of groupings among occupations and industries between the 1960 and 1970 Censuses see U.S. Bureau of the Census, 1970 Occupation and Industry Classification Systems in Terms of their 1960 Occupation and Industry Elements, Technical Paper No. 26, U.S. Government Printing Office.

TABLE D.—COMPARISON OF 1960 TO 1970 NOT ELSEWHERE CLASSIFIED (n.e.c.) CATEGORIES

(1960 Experienced Civilian Labor Force)

1960 Not Elsewhere Classified (n.e.c.) Categories			
Occupation	Total number	Percent major group	Percent of ECLF
Total ECLF	67,990,073	NA	100.0
Total of 12 n.e.c. categories	23,747,163	NA	34.9
Professional, technical, and kindred workers	7,335,699	100.0	10.8
Professional, technical, and kindred,n.e.c.	313,858	4.3	0.5
Managers, officials, and proprietors,exc.farm	5.489.489	100.0	8.1
Managers, officials, n.e.c.	4,586,035	83.5	6.7
Sales workers	4,801,341	100.0	7.1
Sales workers, n.e.c.	3,888,635	81.0	5.7
Clerical and kindred workers	9,617,487	100.0	14.2
Agents, n.e.c.	163,117	1.7	0.2
Clerical and kindred workers, n.e.c.	3,016,387	31.4	4.4
Craftsmen, foremen,and kindred workers	9,240,983	100.0	13.6
Foremen, n.e.c.	1,199,055	13.0	1.8
Mechanics, n.e.c.	1,237,064	13.4	1.8
Craftsmen, n.e.c.	112,225	1.2	0.2
Operatives and kindred workers	12,846,044	100.0	18.9
Operatives, n.e.c.	4,993,044	38.9	7.3
Laborers, except farm	3,530,022	100.0	5.2
Laborers, n.e.c.	2,762,824	78.3	4.1
Service workers, except private household	5,765,481	100.0	8.5
Service workers, n.e.c.	192,879	3.3	0.3
Private household workers	1,824,817	100.0	2.7
Private household workers, n.e.c.	1,281,740	70.2	1.9

1970 Not Elsewhere Classified (n.e.c.) Categories			
Occupation	Total number	Percent major group	Percent of ECLF
Total ECLF	67,990,073	NA	100.0
Total of 15 n.e.c. categories	15,846,713	NA	23.3
Professional, technical, and kindred workers	7,089,840	100.0	10.4
Research workers not specified	79,495	1.1	0.1
Managers and administrators, except farm	57,082,247	100.0	8.4
Managers and administrators, n.e.c.	4,268,389	74.8	6.3
Sales workers	4,798,553	100.0	7.1
Sales workers, n.e.c.	3,869,770	80.6	5.7
Clerical and kindred workers	9,431,106	100.0	13.9
Miscellaneous clerical workers	324,062	3.4	0.5
Not specified clerical workers	1,587,755	16.8	2.3
Craftsmen and kindred workers	9,465,311	100.0	13.9
Foremen, n.e.c.	1,185,586	12.5	1.7
Miscellaneous mechanics	179,599	1.9	0.3
Not specified mechanics	528,005	5.6	0.8
Craftsmen, n.e.c.	92,380	1.0	0.1
Operatives, including transport	12,254,280	100.0	18.0
Machine operatives,miscellaneous specified	762,954	6.2	1.1
Machine operatives, not specified	632,755	5.2	0.9
Miscellaneous operatives	799,178	6.5	1.2
Not specified operatives	371,715	3.0	0.6
Laborers, except farm	3,755,237	100.0	5.5
Miscellaneous laborers	309,018	8.2	0.5
Not specified laborers	856,052	22.8	1.3
Service workers,except private household	6,085,582	100.0	9.0
No n.e.c.	NA	NA	NA
Private household workers	1,816,648	100.0	2.7
No n.e.c.	NA	NA	NA

ESTIMATES FOR SAMPLES FROM FRAMES WHERE SOME UNITS HAVE MULTIPLE LISTINGS

Margaret Gurney and Maria Elena Gonzalez, U.S. Bureau of the Census

A. Introduction

It is often necessary to sample from a frame which is known to contain multiple listings for the same unit; this is particularly true for mail surveys. The estimation of totals and other statistics from such a sample presents some interesting and difficult problems.

In order to relate the subject matter of this paper to a real situation, we base our discussion chiefly on the sampling of small farms which was done as part of the 1969 Census of Agriculture. Units on the Census list which were expected to have Total Value of Products sold (TVP) less than $2,500 in 1969 were sampled at a rate of 1 in 2, and were sent a short questionnaire (form A2). Not all units on the mailing list represented farms according to the Census definition. In the interest of obtaining adequate coverage of farms, it was necessary to use several sources, not restricted to known "in-scope" farms, to construct the Census list. These lists were merged after unduplication on the basis of identical Social Security (SS) and Employer Identification (EI) numbers.

If anyone received more than one questionnaire, he was instructed to fill in any one of them, to mark the others "extra copy," and to return all of them in the same envelope. If any one of the forms returned by a respondent was an "A1," which was the regular form and was not subject to sampling, the form returned was assigned a weight of "1." If all of the questionnaires returned by a respondent were "A2's," the question of the best weighting procedure arose. Various estimation formulas can be used. Several are discussed below, including one based on a mathematical model for handling cases in which the respondent does not follow instructions.

B. Estimates for small farms when the respondent follows instructions

The properties of estimates discussed in this section depend on the following assumptions:

1. Respondents follow instructions, and return together all forms received by them.

2. The number of times a given unit is listed is uncorrelated with the characteristic being estimated. (For estimates of total number of units assumption 2 is not needed.)

Under these conditions we can make estimates which are unbiased, but have large sampling errors; estimates which are biased, but have smaller mean square errors, are also considered.

1. An unbiased estimate. An unbiased estimate, which is easy to apply, and which does not require that we know the total number of times a farm is on the listing, can be made by tabulating the data only for those farms for which the farm operator returned an odd number (1, 3, 5, ...) of questionnaires. These data are then multiplied by 2. This estimate is unbiased, which results from a property of the binomial expansion of $(1+1)^n$. To illustrate, consider farms which are on the list twice: approximately $\frac{1}{4}$ of them will not be selected, $\frac{1}{2}$ will be selected once, and $\frac{1}{4}$ will be selected twice. If we use only the farms which were selected once, and multiply the sum by 2, we shall have an unbiased estimate of the total of a characteristic for farms which are on the list twice. If we consider farms which are listed three times, the proportion which will not be selected at all is about 1/8; about 3/8 will be selected once; 3/8 will be selected twice; and 1/8 will be selected three times. The farms which are selected an odd number of times (3/8 plus 1/8 = 1/2) are used in making the estimate; multiplication by 2 will lead to unbiased estimates. The extension to higher amounts of duplication can be proved by induction.

This type of estimate is not a desirable one, since it is not using all of the data available: all of the farms which were selected an even number of times have a weight equal to zero. Consequently, the variance of this estimate may be larger than that of some other estimate.

2. An unbiased estimate for a list where every farm is listed k times. We consider, as the next step toward a better estimate, a list on which every listing is replicated exactly the same number of times, and ask for the optimum weighting scheme. If k is the number of replications, and N is the unduplicated number of farms, then the list contains k·N listings. If a sample of $\frac{1}{2}$ of the listings is selected, then approximately $[\binom{k}{1}/2^k]N$ will be selected once; $[\binom{k}{t}/2^k]N$ will be selected t times; etc. It can be shown that an unbiased estimate with minimum variance is obtained if all farms in the sample are given the same weight, and that this weight should be $2^k/(2^k-1)$. As k becomes large the optimum weight approaches unity.

Unfortunately, we do not know, when a farm operator returns one, two, or more copies of the A2 form, how many times his place was on the mailing list. We may have an approximate distribution, such as 80% of the farms are listed once, 10% are listed twice, 6% are listed three times, etc. But this doesn't help us to know how many times this particular farm was listed. Hence we go to another estimate, which is biased, but which can be expected to have a smaller mean square error for most items than the unbiased estimate described in section 1, above.

3. A reasonable biased estimate. Since the matching is presumed to have been reasonably successful, perhaps less than 20% of the farms are on

the listing more than once. This means that something more than 80% of the farms which respond with one questionnaire are actually singles, and should have a weight equal to 2. The remaining "single" responders should have weights of 4/3, 8/7, 16/15, etc., but we don't know which weight. Any farm which returned two or more questionnaires is on the list more than once, and with a sample of 1 in 2 has a high probability of being selected; hence it is reasonable to assign a weight = 1 to the "multiple" responders.

Since it is not possible to distinguish the "single" responders who are on the list more than once from those which are listed exactly once, we assign the weight "2" to them. The assignment of this weight to something less than 20% of the farms reporting once (the ones which are actually multiples) introduces an upward bias into the sample estimates. The size and importance of the bias will depend, of course, on the distribution on the list of single listings, doubles, triples, etc.; it will depend also on differences in the characteristics for farms which are listed once, as compared with farms which are listed more than once.

4. Illustration. Suppose that in a certain area there are 3200 farms, but some of them are on the mailing list more than once. Suppose the distribution by the number of replications is

Number of Replications	Frequency (P_i)	Number of Farms (N_i)	Number of Listings
1	.80	2560	2560
2	.10	320	640
3	.06	192	576
4	.03	96	384
5	.01	32	160
Total	1.00	3200	4320

The number on the listing is about 36% larger than the actual number of farms.

Now take a 50% sample of the listings; 2160 listings will be selected representing 1600 farms. The number of farms in the sample is a random variable, and will not be exactly equal to 1600.

	Expected Distribution of the Sample						
	Not selected	Number of times selected					
Number of replications	0	1	2	3	4	5	Number of farms selected
1	1280	1280					1280
2	80	160	80				240
3	24	72	72	24			168
4	6	24	36	24	6		90
5	1	5	10	10	5	1	31
Total	1391	1541	198	58	11	1	1809

If we use the unbiased estimate from section 1. above, the expected value of the estimate of the number of farms is 2(1541 + 58 + 1) = 3200.

The biased estimate of section 3. leads to an estimated number of farms: n' = 2(1541) + 1(198 + 58 + 11 + 1) = 3350, which is an overestimate by 150 farms, or about 4.8%. This can be shown to have a smaller MSE than the unbiased estimate.

5. Another biased estimate. If we have any information which will allow us to estimate the approximate magnitude of P_1, the proportion of unduplicated units, it is possible to obtain a closer estimate of the number of farms, which will usually be an underestimate (depending largely on how well we estimate P_1). Notice that the proportion of farms which look like "singles" (1541/1809 = .852) is considerably larger than the true proportion of singles in the population, which is .80. If the respondent fails to follow instructions and returns only one form when he receives several, or returns more than one but in different envelopes, the proportion which look like "singles" will be even greater than .85. Consequently, it may be desirable to assign a weight of "1" to some small proportion of the apparent singles. A rough guess of the proportion of true singles can lead to a smaller mean square error; for example, if we guess the proportion of singles which are true "singles" to be .84, then the estimated number of farms is

$$n'' = 2(1541 \times .84) + 1(1541 \times .16 + 198 + 58 + 11 + 1) = 2(1294) + 1(515) = 3103$$

which is closer to 3200, the number of farms in the area, than the estimate of 3350, in section 3. In this illustration, a guess as low as 81%, or as high as 99% will lead to a smaller bias in the estimate of the number of farms than the simple procedure of section 3.

The discussion of the preceding sections can be easily extended to other sampling fractions. If the sampling fraction is less than $\frac{1}{2}$, the variance of the unbiased estimate (section 1.) is usually even larger, relative to the biased estimate of the latter part of section 3., and would not normally be considered a candidate for practical use.

C. Estimates when the respondent does not follow instructions

Respondents in a mail survey may fail to follow instructions in a number of ways. A respondent receiving t agricultural questionnaires may return any number from 0 to t of them. We have no way of knowing how many questionnaires a respondent received. Since the questionnaires he does not return are not known to be duplicates, they will probably be considered "nonresponses," and further followup will be done, and imputation will be applied to them if the followup is not successful.

It is possible to construct models which take into account the sampling fraction (f) and the probability that a respondent who receives t questionnaires returns r of them. Such a set of probabilities ($p_{t,r}$) can be applied to various sets of probabilities of multiple listings (P_i) to get some notion of the efficiency of various estimators. The models developed below start (section 1.) with a sampling frame consisting of k lists, each containing all N elements in the population (see section B2., above). Two methods are compared (section 2.): the method of section B1., and the latter part of section B3. The extension is then made (section 3.) to variable numbers of duplications.

1. Model[1/]

Let a population consist of N elements, and let the sampling frame consist of k lists, each containing all N elements. In each list, each element is given a probability f of being selected for a sample. The number of times it is selected is a random variable t_i with values in the range $t_i = 0, 1, \ldots, k$, and $\Pr(t_i=t) = \binom{k}{t} f^t(1-f)^{k-t}$. If an element is selected t times, it may respond (return the questionnaire) from 0 to t of them. Let $p_{t_i,r}$ be the probability that an element responds r times, given that it has been selected t_i times. Let X_i be a value associated with the i-th element, and consider a statistic

$$x = \sum_{i=1}^{N} u_i X_i.$$

Two methods have been investigated for the choice of the random variable u_i.

a. Method 1

$$\text{Let } u_i = \begin{cases} 2 \text{ if } r \text{ is odd } (1, 3, 5, \ldots) \\ 0 \text{ if } r \text{ is even } (0, 2, 4, \ldots) \end{cases}$$

Then

$$E\, u_i = E\, E\,(u_i|t_i) = 2\, E\, P_{t_i}$$

where $P_{t_i} = \sum_{j=0}^{\left[\frac{t_i-1}{2}\right]} p_{t_i,\, 2j+1}$ and the upper limit is the greatest integer less than or equal to $\frac{t_i-1}{2}$.

Then

$$E\, u_i = 2\sum_{t=0}^{k} P_t \binom{k}{t} f^t(1-f)^{k-t} = 2E(P_t) \qquad 1)$$

$$\begin{aligned} \text{Var } u_i &= E \text{ Var } (u_i|t_i) + \text{Var } E\,(u_i|t_i) \\ &= E\left[4P_{t_i}(1-P_{t_i})\right] + \text{Var } (2P_{t_i}) \\ &= 4\sum_{t=0}^{k} P_t \binom{k}{t} f^t(1-f)^{k-t} \\ &\quad - 4\sum_{t=0}^{k} P_t^2 \binom{k}{t} f^t(1-f)^{k-t} \\ &\quad + 4\left\{\sum_{t=0}^{k} P_t^2\binom{k}{t} f^t(1-f)^{k-t} - \left[\sum_{t=0}^{k} P_t\binom{k}{t} f^t(1-f)^{k-t}\right]^2\right\} \\ &= 4E(P_t) - 4E(P_t)^2 = E\,(u_i)[2-E(u_i)] \end{aligned} \qquad 2)$$

b. Method 2

$$\text{Let } u_i = \begin{cases} 2 \text{ if } r = 1 \\ 1 \text{ if } r > 1 \\ 0 \text{ if } r = 0 \end{cases}$$

$$\begin{aligned} \text{Then } Eu_i &= E\, E\,(u_i|t_i) \\ &= E\,(2p_{t_i,1} + 1 - p_{t_i,o} - p_{t_i,1}) \\ &= E\,(1 + p_{t_i,1} - p_{t_i,o}) \\ &= 1 + \sum_{t=0}^{k} (p_{t,1} - p_{t,o})\binom{k}{t} f^t(1-f)^{k-t} \end{aligned} \qquad 3)$$

Var u_i

$$\begin{aligned} &= E \text{ Var } (u_i|t_i) + \text{Var } E\,(u_i|t_i) \\ &= E\,\{E(u^2|t) - [E(u|t)]^2\} + \text{Var } (1 + p_{t,1} - p_{t,o}) \\ &= E\,[4p_{t,1} + (1-p_{t,1} - p_{t,o}) - (1+p_{t,1} - p_{t,o})^2] \\ &\qquad + \text{Var } (1+p_{t,1} - p_{t,o}) \\ &= E\,[p_{t,1} + p_{t,o} - (p_{t,1} - p_{t,o})^2] \\ &\qquad + \text{Var } (p_{t,1} - p_{t,o}) \\ &= E\,(p_{t,1} + p_{t,o}) - [E\,(p_{t,1} - p_{t,o})]^2 \end{aligned}$$

$$= \sum_{t=o}^{k} (p_{t,1} + p_{t,o}) \binom{k}{t} f^t (1-f)^{k-t} - [\sum_{t=o}^{k} (p_{t,1}-p_{t,o})\binom{k}{t}f^t(1-f)^{k-t}]^2 \quad 4)$$

2. Comparison of Methods 1 and 2

We note that

$$Ex = \sum_{i=1}^{N} X_i \; Eu_i = (Eu_i) \sum_{i=1}^{N} X_i$$

$$Var\ x = \sum_{i=1}^{N} X_i^2 \; Var\ (u_i) = Var\ (u_i)\ \Sigma\ X_i^2$$

Hence

$$MSE(x) = Var\ x + (Ex - X)^2$$

$$= Var\ (u_i) \sum^{N} X_i^2 + (1-Eu_i)^2 \ (\Sigma\ X_i)^2$$

$$= N\sigma_u^2 \ (\bar{X}^2 + \sigma_X^2) + N^2 \ (1-Eu)^2 \ \bar{X}^2$$

$$= N\sigma_u^2 \ \sigma_X^2 + [\sigma_u^2 + N\ (1-Eu)^2]\ N\bar{X}^2 \quad 5)$$

This provides a means of comparing the mean square errors of the two methods for hypothetical sets of values $p_{t,r}$.

3. Extension to lists with variable number of duplications.

Let the population be divided in N_1 elements that are on only 1 list, N_2 elements that are on exactly 2 lists, etc. Then, if x_j denotes the estimate for the j-th set

$$x = \sum_{j=1}^{L} x_j = \sum_{j=1}^{L} \sum_{i=1}^{N_j} X_{ji} \quad 6)$$

$$Ex = \Sigma_j\ Ex_j = \sum_{j=1}^{L} E(u|j)\ X_j \quad 7)$$

$$Var\ x = \sum_j Var\ x_j = \sum_j Var\ (u|j) \sum_{i=1}^{N_j} X_{ji}^2 \quad 8)$$

$$MSE(x) = Var\ x + (Ex - X)^2$$

$$= \sum_j Var(u|j) \sum_{i=1}^{N_j} X_{ji}^2 + \{\sum_j [E(u|j)-1]X_j\}^2$$

$$= \sum_j \sigma_{uj}^2\ N_j\ (\sigma_{xj}^2 + \bar{X}_j^2) + \{\sum_j [E(u|j)-1]\ N_j\bar{X}_j\}^2 \quad 9)$$

Formula 9 can be used to compare the efficiency of the two methods.

4. Examples

Two examples are shown in Exhibit I, comparing the two methods for hypothetical populations.

Exhibit I

Data for Example A

Five lists
Sampling fraction = 1/2
Mean per stratum = 1.0
Variance per stratum = 1.0

Stratum: Number of times unit is listed in population	Number of listings	Unduplicated number of units
1	10,900	10,900
2	2,528	1,264
3	477	159
4	108	27
5	90	18
Total	14,103	12,368

Number of times unit is selected (t)	Expected number of units selected	Proportion of times unit responds (r)					
		0	1	2	3	4	5
0	5778.1	1.00					
1	6151.2	0.05	0.95				
2	391.4	0.04	0.50	0.46			
3	32.2	0.03	0.45	0.47	0.05		
4	4.5	0.03	0.40	0.40	0.10	0.07	
5	.6	0.03	0.40	0.35	0.07	0.12	0.03
Total	12,368.0						

Data for Example B

Three lists
Sampling Fraction = 1/2
Mean per stratum = 1.0
Variance per stratum = 1.0

Stratum: Number of times unit is listed in population	Number of listings	Unduplicated number of units
1	139	139
2	84	42
3	12	4
Total	235	185

Number of times unit is selected (t)	Expected number of units selected	Proportion of times unit responds (r)			
		0	1	2	3
0	80.5	1.00			
1	92.0	0.60	0.40		
2	12.0	0.20	0.50	0.30	
3	.5	0.10	0.45	0.40	0.05
Total	185.0				

(continued)

Exhibit 1, continued

Comparisons of Methods 1 and 2
for Examples A and B

	Method 1				
Example	Estimate	Variance	Bias	Mean square error	Relative root mean square error
A	12,116	24,564	-252	88,068	0.024
B	86	259	-99	10,060	0.54

	Method 2				
Example	Estimate	Variance	Bias	Mean square error	Relative root mean square error
A	12,311	23,951	-57	27,200	0.013
B	90	256	-95	9,281	0.52

For both of the examples presented in Exhibit I, method 2 gives a smaller mean square error than method 1.

The relative root mean square error of method 2 for example A is 0.013, while for method 1 it is 0.024. Example B represents an extreme case in many respects. It assumes that:

1. Only about 60 percent of the list corresponds to single units.

2. Only about 40 percent of units selected once responded.

The relative root mean square error for this example is about 0.5 for both methods, because of the extremely biased results obtained. In this example again, method 2 is slightly better than method 1.

D. Reducing the amount of duplication

The unduplication of the various lists for the 1969 Census of Agriculture was done by matching on EI and SS numbers. It was felt that most of the remaining duplication would be removed as a result of the instruction to the respondent who received more than one questionnaire that he should fill out only one, but return them all in the same envelope. This system was adopted because of budget and time considerations and because it was believed that the amount of duplication was small. However, the mailings were made at different times. About 90 percent of the mailing pieces were sent in January 1970, and 10 percent in May and July. Therefore, not all forms were received by respondents at the same time; so respondents were sometimes unable to mail back all forms in the same envelope. In addition, some respondents who should have followed this instruction failed to do so.

An idea of the reasons for duplication can be obtained from Exhibit II-B, which is based on a small sample of "births," which were names added to the mailing list in July 1970. Potential "births" were matched against the original Census list on the basis of SS and EI numbers. While the distributions of duplicates described in Exhibit II-B apply only to the "birth" match, they may provide a general indication of the kinds of duplication problems that are not adequately taken care of by a straight match on identification numbers.

The data in Exhibit II, and results from some other studies of duplication in selected geographic areas, indicate that, if additional characteristics, such as name and address are used in the matching, the amount of duplication may be reduced considerably.

Further unduplication of the mailing lists is necessary for several reasons:

1. To reduce the reporting burden on respondents, and costs to the Census Bureau.

2. To the extent that respondents fail to return duplicate questionnaires, the application of imputation procedures for nonrespondents to these cases may produce a significant upward bias in farm counts and related statistics.

3. Respondents may fill out and return more than one report for the same operation. To the extent that it is not possible to identify these as duplicates, there is an upward bias in all statistics.

In preparation for the 1974 Census of Agriculture, we are investigating various methods of linking records to devise an unduplication procedure which is more effective than the one used for the 1969 Census of Agriculture.

FOOTNOTE

1/ This analysis was suggested by Benjamin J. Tepping of the U.S. Bureau of the Census.

Exhibit II

A. SAMPLE FROM "BIRTH" MAILING LIST: NUMBERS OF DUPLICATED LISTINGS AND OF TOTAL DUPLICATION, BY REGION

Characteristic of the sample	United States	Northeast	North Central	South	West
Total sample size	535	126	134	137	138
Sample listings duplicated one or more times	175	44	44	49	38
Total duplications	267	68	63	79	57

B. PERCENT DISTRIBUTION OF WEIGHTED NUMBER OF DUPLICATES, BY REASON FOR DUPLICATION AND BY REGION

(Based on sample of 267 duplications to a sample of birth listings)

Reason for duplication	United States	Northeast	North Central	South	West
Total	100.0	100.0	100.0	100.0	100.0
Incorrect punching or reporting of SS or EI numbers*	12.3	22.8	9.7	9.2	12.4
SS number on one listing, EI number on other, same name on each	34.1	30.8	28.3	36.6	39.0
Different SS number reported for same name or one listing missing SS number	11.0	4.8	8.0	20.4	5.7
Different EI number reported for same name or one listing missing EI number	10.6	13.5	3.5	16.9	7.6
Other members of partnership or operation	30.6	24.8	46.9	16.9	35.2
No apparent reason	1.5	3.4	3.5	-	-

* SS stands for Social Security and EI stands for Employer Identification.

THE STUDY DESIGN OF AN EXPERIMENT TO MEASURE THE EFFECTS OF USING PROXY RESPONSES IN THE NATIONAL HEALTH INTERVIEW SURVEY

Kenneth W. Haase and Ronald W. Wilson, National Center for Health Statistics

Although the first census was taken in this country in 1790, the utilization of an ongoing national survey procedure designed to obtain a large amount of detailed information about a specific area of interest, such as health, is still a rather new phenomenon. The Current Population Survey is only 32 years old. The Health Interview Survey conducted by the National Center for Health Statistics, although first implemented on an ad hoc basis in 1935, has only been an ongoing operation since 1957. For this reason there are still many things we do not know about the techniques of collecting valid and reliable information from a household respondent on topics which often can be quite complex.

Important public health decisions made at the national and local levels are often influenced by data obtained in the Health Interview Survey. These decisions can affect the lives of large segments of the population. In the process of improving both the validity and reliability of the data collection techniques, we must first try to better understand the types and magnitude of the errors and biases that exist within these procedures.

In assessing the validity of the statistics derived from the Health Interview Survey, of major concern has been the effect of the respondent rules which allow adult respondents to report for other family members. It is generally accepted that for most types of personal information, such as health, the best informant will, with a few exceptions, be the person to whom that information is related. The adoption of the present respondent rules was based on the concept that, while the validity of the data may be improved by implementing a "self-respondent only" procedure where all adults are interviewed for themselves, the greatly increased cost and decreased efficiency resulting from such a change would more than offset any gain.

Over the years, a number of special studies have been conducted to evaluate the validity of specific types of information obtained in the Health Interview Survey.[1/] These studies often involve the validation of information obtained in the interview with records from selected health facilities where the respondents were patients (such as hospitals or clinics). Most of these studies tend to indicate that the most important problem is related to underreporting. There is a tendency for some people not to report all of their experiences either because of memory loss or for some other reason. Further analysis of these studies also indicate that the degree of underreporting for selected health variables tends to be more severe when the information is obtained through a proxy respondent rather than from the person himself.[2/]

The evidence on reporting differences resulting from self or proxy respondents derived from these and other studies[3/] is far from conclusive. This is due to a number of reasons, including the small sample size used in some of these studies, the fact that most of these studies were not specifically designed to investigate this specific phenomenon and the fact that these studies often included very select population groups which made inference to the general population difficult.

From April to June 1972, the Division of Health Interview Statistics conducted a study designed to measure the degree to which the use of proxy respondents affects the national statistics. The findings from this study could have a wide range of implications far beyond the National Health Survey and could affect the Census Bureau and other survey activities.

First, it might be helpful to look at some of the survey data which are related to this respondent problem. Data on approximately 40 percent of all adults in the Health Interview Survey sample are obtained from proxy respondents, that is, the information about them is obtained from another family member. If we look at the rates of selected health indices we find the following differences between data obtained from self and proxy respondents.

Table 1 presents the annual rate of restricted activity days for persons 17 years of age and over. The restricted activity statistic is derived from questions that ask about cutting down on the things a person usually does.

There are more than 17 days of restricted activity per year for each adult in the United States. However, there is considerable variation in these rates depending upon whether the information was obtained directly from the person himself or through another family member. Self-respondents report a rate of almost 20 days of restricted activity per year as compared to only 13 days when data are obtained by proxy respondents. The higher rates for self-respondents are found entirely among persons age 17-64.

For the 65 year and over age group, there is a reversal in the rates according to respondent status. The 29 days reported for the self-respondents is lower than the 35 days reported by the proxy respondents.

Table 2 shows the number of days of bed disability per person 17 years of age and over per year according to respondent type and age. Since by definition all bed days are also restricted activity days, we would, therefore, expect a somewhat similar pattern for these two sets of data. The overall rate for self-respondents is higher than the rate for proxy respondents, although this difference is not quite as great as that observed for total restricted activity. For each of the age groups shown, a similar pattern is observed; the self-respondent rates for persons under 65 years of age are higher than the rates obtained by proxy respondents. Also, the same reversal is observed for persons over 65 years of age, with more than twice as many days of bed disability reported by proxy respondents as compared to self-respondents.

Table 3 presents the number of physician visits per person per year by respondent status and age. During 1970, there were five physician visits for each adult in the United States; whereas the self-respondent rate is 5.6 visits per person per year, the proxy respondents report only 4 visits per year. The self-proxy difference is most pronounced for persons under 65 years of age. The reversed pattern for persons 65 years and over that was observed for the disability data is not present in the physician visit data.

Estimates from the Health Interview Survey indicated that during 1970 there were 1.6 acute conditions for each adult in the United States. The same pattern for the two respondent groups emerges in Table 4 that was observed for physician visits.

Table 5 presents the annual number of short-stay hospital discharges per 100 persons 17 years of age and over. Again, as in all previous tables a higher rate for self-respondents than proxy respondents and that the younger age groups are the major contributors to this difference.

Summarizing the findings for these selected health indices, the rates for adult self-respondents are higher than those obtained by proxy respondents. The pattern held consistently for persons under 65 years of age. The pattern for persons over 65 was not consistent, with proxy respondents reporting higher rates for several variables. Although not shown here, there were also sex differences in the pattern for the aged. While these variations in the general pattern of higher rates for self-respondents are interesting in and of themselves, an indepth analysis of these variations are not essential to the main purpose of this paper.

What can be inferred from these data? The three most obvious hypotheses that can be developed to account for these differences are:

1. These differences are due to different degrees of error in reporting by proxy as compared to self-respondents.
2. These differences are due to differences within these two population groups and, therefore, reflect true differences and not reporting error.
3. These differences are due to a combination of both factors, reporting error and the select nature of the population groups involved.

The possibility that true differences exist between self-respondents and persons for whom data is obtained from another family member can be elaborated further. Whether data is gathered from a self or proxy respondent depends on whether or not the subject is available at the time of that interview. However, the person's availability may well be a consequence of his health status. Since most of the interviews are conducted during the work week and within the daylight hours, a person who might otherwise be expected to be at work or away from home for other activities but who is at home at the time of interview might well have a health problem and can conceivably inflate these health statistics for the self-respondent groups. The observed higher rates for persons over 65 for whom data were obtained by proxy could be explained by the fact that when these older persons are ill they are more likely to be bed-ridden or hospitalized and, therefore, data must be obtained by proxy.

To know how much of the differences in the rates for these health indices are due to reporting error as compared to true population differences, the following questions must be answered: First, what would be the rates of these selected health indices for persons responded for by proxy if they had responded for themselves? If this question is answered adequately, it is necessary to know how much of these differences in health indices are due to the use of proxy respondents and how much of these are due to differences in health characteristics between the persons at home and those not at home at the time of the interview.

If the difference in reporting bias between self and proxy respondents is significantly large, there would be a second question to be answered. Is the proxy respondent's relationship to the person for whom he is responding an important factor in these differences? Are there certain groups of respondents that can more adequately respond for another family member? For example, is a woman responding for her husband a better respondent than a woman responding for her father-in-law?

A third question that this study was designed to answer is how much will a change in the respondent rules cost, both in terms of dollars and other factors such as the nonresponse rate?

Finally, there is one important question this study will not answer, that is, "What type of respondent reports the most accurate information?" We would need to tie the present study to health records in order to determine the accuracy of reporting.

The remainder of this paper will describe the study design. During the second quarter of the 1972 data collection year, 6 of the 13 weeks were selected for the use of special respondent rules. Under the regular respondent rules, any responsible adult can respond for another related household member.

During the 6 weeks of the test, interviewers were instructed to follow their regular schedule in making their initial contact at a household. They then determined which adults were at home and obtained information from them as a self-respondent. Appointments were made to return and interview as self-respondents all adults not home at the time of the initial interview. However, a special notation was made to indicate that these persons would have been responded for by proxy under the normal survey rules. Thus, under the regular rules we have persons responded for by self and proxy determined by whoever is at home, while under the special rules we have all self-respondents, but some who would normally have been responded for by proxy under the old rules. If it was impossible to obtain a self-interview, rather than lose information, data were obtained by proxy, but this residual group will be deleted from the major analysis.

Table 6 further illustrates the study design and the proposed analysis plans.

The first two columns represent respondent status under the usual respondent rules. Column 3 represents those adults at home at the time of the initial interview. They would have been self-respondents under the usual procedure and will also be self-respondents under the experimental procedure. Column 4 represents those adults who were not at home during the initial interview and, therefore, would have been responded for by proxy under the usual procedure but will be interviewed for themselves during the experimental procedure.

This table illustrates how the various hypotheses discussed above can be tested. If the earlier observed differences between self and proxy respondents resulted entirely from reporting error, then Column 4 would approximate Column 1 and 3 and Column 4 would be greater than Column 2. If the different rates reflect true differences in the two groups, then Column 4 would approximate Column 2 and be less than Columns 1 and 3. If, however, both these factors account for the different rates, which is probably the case, then the effect of each can be calculated. Column 4 minus Column 2 represent the difference due to reporting and Column 3 minus Column 4 would represent the difference due to population difference. If a significant part of these self-proxy differences are a result of reporting error on the part of proxy respondents, it is necessary to determine if certain types of proxy respondents report more adequately than others. For this analysis the proxy columns (2) and (4) will be divided into spouse and other respondents.

What are the implications of this study? In terms of the Health Interview Survey, the first benefit that can be derived from this study is a much better understanding of how the national statistics are being affected by the utilization of the present respondent rules. Regardless of whether any changes are made in the respondent rules, this knowledge would greatly benefit all future analysis of Health Interview Survey statistics. If it is found that the self-proxy differences are primarily the result of true population differences, then probably no change in the respondent rules would be necessary.

If, however, it is found that the present respondent rules create a very significant bias in the statistical product, alternative procedures must be developed. These procedures could include being more discriminating in who can be a proxy respondent or a decision that only self-respondents are permissible. If the latter course of action is taken, this could possibly affect the entire household interview procedure. For example, if a self-respondent procedure is adopted it might be necessary to interview only a sample of the persons in the assigned households rather than all household members as is presently done.

REFERENCES

(1) National Center for Health Statistics, Vital and Health Statistics, PHS Pub. 1000, Series 2, Public Health Service, Washington, U. S. Government Printing Office: "Measurement of Personal Health Expenditures," No. 2, (June 1963); "Health Interview Responses Compared With Medical Records," No. 7, (July 1965); "Interview Responses on Health Insurance Compared With Insurance Records," No. 18, (August 1966); "Interview Data on Chronic Conditions Compared With Information Derived from Medical Records," No. 23, (May 1967); "Development and Evaluation of an Expanded Hearing Loss Scale Questionnaire," No. 37, (April 1970); "Optimum Recall Period for Reporting Persons Injured in Motor Vehicle Accidents," No. 50, (April 1972).

(2) Series 2, Nos. 6 and 7

(3) Ballweg, John A., "Husband-Wife Response Similarities on Evaluative and Non-Evaluative Survey Questions," Public Opinion Quarterly, 33 (2) (Summer 1969), 249-254.

Cartwright, Ann, "The Effect of Obtaining Information from Different Informants on a Family Morbidity Inquiry," Applied Statistics 6 (1), (March 1957), 18-25.

Elinson, Jack, and Trussell, Ray E., "Some Factors Relating to Degree of Correspondence for Diagnostic Information as Obtained by Household Interview and Clinical Examinations," American Journal of Public Health, 49, (March 1957), 311-321.

Enterline, Philip E., and Capt, Katherine G., "A Validation of Information Provided by Household Respondents in Health Surveys," American Journal of Public Health, 49, (Jan-June 1959), 205-212.

Feldman, Jacob J., "The Household Interview as a Technique for the Collection of Morbidity Data," Journal of Chronic Diseases, 11, (May 1960), 534-554.

Fisher, Gordan, "A Discriminate Analysis of Reporting Errors in Health Interviews," Applied Statistics, 11 (3), (November 1962), 148-163.

Nisselson, Harold, and Woolsey, Theodore D., "Some Problems of the Household Interview Design for the National Health Survey," Journal of the American Statistical Association, 54, (March 1959), 69-87.

Scanzoni, John, "A Note on the Sufficiency of Wife Responses in Family Research," Pacific Sociological Review, 8, (Fall 1965), 109-115.

Weiss, David J., and Dawis, Rene V., "An Objective Validation of Factual Interviewing Data," Journal of Applied Psychology, 44 (6), (1960), 381-385.

Table 1. Number of days of restricted activity per person 17 years of age and over per year by age and respondent type: United States, 1970

Age of Subject	Respondent Type		
	Total	Self Respondent	Proxy Respondent
	Number of days per person per year		
All persons 17+	17.2	19.6	12.8
17-44 years	12.0	14.4	9.0
45-64 years	20.0	23.3	14.0
65+ years	30.7	29.2	35.2

Table 2. Number of days of bed disability per person 17 years of age and over per year by age and respondent status: United States, 1970

Age of Subject	Respondent Type		
	Total	Self Respondent	Proxy Respondent
	Number of days per person per year		
All persons 17+	7.0	7.5	6.2
17-44 years	4.9	5.6	4.0
45-64 years	7.5	8.5	5.8
65+ years	13.8	10.9	23.3

Table 3. Number of physician visits per person 17 years of age and over per year by age and respondent status: United States, 1970

Age of Subject	Respondent Type		
	Total	Self Respondent	Proxy Respondent
	Number of visits per person per year		
All persons 17+	5.0	5.6	4.0
17-44 years	4.6	5.4	3.5
45-64 years	5.2	5.6	4.6
65+ years	6.3	6.3	6.0

Table 4. Incidence of acute conditions per person 17 years of age and over per year by age and respondent status: United States, 1970

Age of Subject	Respondent Type		
	Total	Self Respondent	Proxy Respondent
	Number of conditions per person per year		
All persons 17+	1.6	1.7	1.5
17-44 years	1.9	2.1	1.7
45-64 years	1.3	1.4	1.1
65+ years	1.0	1.1	0.9

Table 5. Number of discharges from short-stay hospitals per 100 persons 17 years of age and over per year by age and respondent status: United States 1970

Age of Subject	Respondent Type		
	Total	Self Respondent	Proxy Respondent
	Number of discharges per 100 persons per year		
All persons 17+	16.5	19.6	11.6
17-44 years	15.7	20.7	9.3
45-64 years	14.7	15.9	12.6
65+ years	23.4	23.1	24.4

Table 6. Example of the type of table format that can be used to analyze the findings from the Respondent Rules Study

Health Indice	Present Respondent Rule Procedure		Experimental Respondent Rule Procedure	
	Self	Proxy	Self Self	Proxy Self
	(1)	(2)	(3)	(4)

HO_1: Self-proxy differences entirely due to reporting bias
Then: Col. (4) should equal Cols. (1) and (3)
Col. (4) should be greater than Col. (2)

HO_2: Self-proxy differences entirely due to population differences
Then: Col. (4) should equal Col. (2)
Col. (4) should be less than Cols. (1) and (3)

HO_3: Both factors involved
Then: Col. (4) minus Col. (2) should equal the difference due to reporting
Col. (3) minus Col. (4) should equal the true differences in population

EMPLOYMENT PROFILES OF SELECTED LOW-INCOME AREAS, 1970

Harvey R. Hamel, Bureau of Labor Statistics

For many years, researchers, policy makers, and program administrators have been trying to deal with the many and diverse manpower problems in low-income areas of our major cities without having adequate information on exactly what these problems were and how widespread or serious they were. A demand for comprehensive, timely, and accurate information necessary to deal with these problems has come from State and local governments, universities, private research organizations and others. This demand has been somewhat satisfied by several BLS and Census Bureau efforts completed in recent years.

Some early urban area surveys were conducted by the Department of Labor in November 1966 in 8 cities. This was followed by several methodological and experiemental surveys in 1967, and by the Urban Employment Survey (UES) from 1968 to 1970, covering the poverty areas of 6 major cities -- Atlanta, Chicago, Detroit, Houston, Los Angeles and New York City. The findings and results of the UES demonstrated that it was possible to collect detailed and extensive socioeconomic data in low-income areas. In view of the need for these same kind of data for a great many more areas, plans were made to conduct the Census Employment Survey (CES) by tying it into the 1970 Decennial Census. However, the survey was conducted separately from the Census as it would have been impossible to include on the Census questionnaire the detailed information in the fields of labor, housing, education, income and other subject areas which were needed.

The data were collected in low-income areas of 51 cities, 7 rural areas and one Indian reservation. Interviewing began in some areas in August 1970 and was completed for the last area in March 1971. Survey results in the form of statistical reports were published by the Census Bureau in series PHC(3) of the 1970 Census reports. All of the 76 separate reports in the series are now available from the Census Bureau in Washington and its field offices.

The purpose of this paper is to examine each of the major subject matter areas on which information was obtained in the CES. It focuses on reasons why information on specific subjects was collected, how the findings are presented, and notes particularly interesting and useful measures which have been tabulated. Specific areas for possible analysis by those with responsibilities for the application and use of the data are also pointed out. The paper also includes information on potentially significant data which was collected but not tabulated in the CES reports. These unpublished data, which are available through special arrangement with the Census Bureau, vary in amount from one subject area to another and, in some cases, are a substantial part of the overall findings on a particular subject.

Demographic and Social Characteristics

The examination of the CES begins with the most basic information on the group being surveyed -- their personal and social characteristics. The Census Employment Survey provides essentially the same detailed information on demographic and social characteristics for residents of the selected low-income areas as is available for the total United States, individual States, SMSA's and cities from the Decennial Census and the Current Population Survey. Demographic or personal characteristics are essential for a complete profile of any group of persons being studied. They serve as a framework to a fuller understanding of the comprehensive employment related subject matter on which the survey is focused. Available from this survey are detailed data on age, sex, school status, veteran status, race and ethnic origin, marital status, family size and household composition.

CES data for many of the subject matter characteristics are published with data disaggregations for particular age groups; this includes such age breaks as 16-19, 16-21, 20-24, and 10 year breaks for persons 25 years and over, with the amount of detail varying by table. The detailed age categories allow for a greater degree of flexibility in the application of the survey data as labor force and related characteristics of men and women in various age groups differ significantly. For example, labor force participation rates (the proportion of the population in the labor force) vary sharply by age and sex. Men 35 to 44 years usually represent maximum labor force participation and are apt to hold or seek full-time jobs. In contrast, because of family responsibilities, labor force participation of women in that as well as other age groups is significantly lower than that of men, and women workers are more likely than men to hold or seek part-time jobs. Trends show however, that during the 1960's, an increasing proportion of women entered the work force to supplement family income and for other reasons, an occurence which has narrowed the gap in participation rates for men and women.

Separate figures for Vietnam veterans 22 to 34 years old are presented on a number of tables. The fact that many returning veterans have had serious difficulties in locating a satisfactory job is of great concern at all levels of government -- national, State, and local. The cross-classification of employment status of veterans by other detailed characteristics should shed some light on whether the problems of veterans living in the low-income areas are principally the result of lack of civilian skills, lack of previous civilian work experience, and/or low educational attainment.

Detailed information on race and ethnic origin is included on each table in the CES reports in order to better focus on the particular problems of minority-group residents in the poverty areas. There is a particular need for information on the employment and other characteristics of low-income area residents with diverse racial and national origin backgrounds. Because of differences in customs, language barriers, low educational attainment, lack of skills, recent migration into the area, discrimination, or other factors, members of minority groups are more likely to have employment problems than their non-minority neighbors. The CES data will enable the user to determine the extent of and reasons for these differences among racial groups in particular low-income areas.

The extent of the race and ethnic origin detail presented in individual CES reports depends directly on the relative make-up of the population in each area. Detail is presented separately for the total area population, the majority and for the largest minority race and ethnic groupings in some combination of the following -- Negro, white, white Spanish-American, other white, and/or other races. The summary tables (A through N) include information on any three of the above race or ethnic groups which comprise at least 5 percent of the population while the detailed tables (1 through 54) present separate data on those groups which make-up at least 20 percent of the population. In addition, there is a special report on the combined New York City low-income area presenting information on the Puerto Rican and the non-Puerto Rican population.

Education and Training

The demand for higher educational and skill requirements to fill most of the new job opportunities in the economy has been rising rapidly. Moreover, in most existing occupations, there has been an increasing demand for workers with more education and training. Educational attainment classifies persons by years of schooling completed in "regular" schools, which include graded public, private, and parochial schools, whether day schools or night schools. Detailed data on educational attainment in the CES poverty areas provide the opportunity to analyze the relationship between levels of schooling and its association with other important socio-economic indicators, including labor force status, earnings, job training, and annual income.

Also relating to the question of how people prepare themselves for jobs, is information on the extent and sources of formal job training programs completed by low-income area residents. Along with educational attainment, this information was collected in order to help determine the relationship between job preparation and job success. For the purpose of this survey, job training includes completion of any formal job training program in a high school, trade school or junior college, the Armed Forces, or apprenticeship programs. It also includes training programs operated under the Manpower Development and Training Act, the Neighborhood Youth Corps, and other special manpower programs. Job success is measured by current occupational status and hourly earnings of employed persons -- both those who have and those who have not completed training programs. For example, data is available on the occupational distribution and average hourly earnings of both high school and non-high school graduates who either did or did not complete job training programs. The information is tabulated according to whether the training was completed prior to or since January 1961 in order to focus principally on "recent" job training experience (that completed since January 1961) which is assumed to be more closely related to current work activity than training completed more than ten years ago.

Most of the data obtained on job training is published in the CES reports. However, an important survey question, results of which might bear upon the efficacy or success of job training pro -- "have you ever used any of this training on any of your jobs" -- was not tabulated. These unpublished data would indicate the frequency of use of each source of job training by current occupation of the training recipient and by other variables. This would help in measuring the impact of training programs completed by area residents and could be of major assistance to policy makers and program administrators in the design and operation of both skill development and supportive training programs.

Employment and Labor Force Status

Current employment status is perhaps the most widely used and quoted manpower measure. The proportion of the population in the labor force (labor force participation rate), the number of employed and unemployed persons, and the unemployment rate provide quick measures of the economic well-being of residents of any area. They can be particularly useful in pointing to the differences in employment success between residents of poverty areas and that of the general population.

The CES contains unusually detailed data upon which a profile of the unemployed in the poverty areas can be established. This includes separate information on those who are looking for full-time employment and those who want a part-time job. It also includes a number of other important variables such as duration of unemployment, the reasons for unemployment (lost last job, left last job, and new entrants or reentrants to the labor force) problems in finding a job, methods of seeking work, availability of transportation, extent of unemployment for part-year workers, and main reasons for less than full-year work. Since most of these variables are cross-classified by other important characteristics, the data will identify whether lack of appropriate education, experience or skill, or other factors are the most serious problems for unemployed persons living in these areas. The data should be of assistance in the planning of programs to aid the unemployed in their job search.

In using the CES data, one objective of the analyst might be to provide a measure of the number of additional jobs needed to absorb the unemployed in poverty areas. Another objective might be to measure the number of jobs needed to insure full-time employment for the component most easily discernible as underemployed -- those on part time for economic reasons. The additional jobs needed for residents of the poverty areas must be linked to available or potential worker skills before fully or partially unemployed workers can be matched with job vacancies, or the demand for workers in the area.

Since the skill level of the worker is closely aligned with his occupation, an analysis of both the occupational structure and the educational, earnings, and other characteristics of workers in each occupational group is essential to obtain a comprehensive picture of the unemployed, the underemployed, and the full-time employed. For poverty areas such as those surveyed in the CES where low-skill levels contribute heavily to the employment problems and low earnings of area residents, occupational data are of primary importance. For this reason, specific focus in the CES was aimed at the occupational composition of the work force.

Three of the CES tables present data for subdivisions commonly referred to as "detailed occupation groups" for current job, first regular full-time job, and longest job since leaving school. These data should be of primary interest in the analysis of worker skills available to the community. On the remaining CES occupational tables, major occupation group data are cross-classified by several important variables which should add considerably to knowledge of characteristics of workers at general levels of skill. Information on possible occupational upgrading is also presented by a cross-tabulation of occupation of current job with occupation of longest job.

Problems Affecting Jobholding and Jobseeking; Desire for Work Among Persons Not in the Labor Force

Because of interest in information on barriers to satisfactory employment faced by low-income area residents, the CES obtained specific data on the problems or responsibilities that prevent or affect jobholding or jobseeking and on the desire to work of persons outside the labor force. This information is central to the purposes of the CES. Hopefully, it will assist in the design of effective manpower programs to alleviate these problems and enable low-income persons to find and keep good jobs.

The list of specific problems or responsibilities keeping some low-income area workers from searching for or from finding or holding a good job is quite extensive. The specific reasons identified in the CES included several categories. Family responsibilities; health problems; lack of experience, skills, or education; and employers think too young or too old were tabulated for employed and unemployed persons as well as those not in the labor force. In addition to these four categories, persons not in the labor force were also classified as to whether their nonparticipation status was due to one of the following reasons: transportation problems, looked but couldn't find work, or retirement.

Each problem category may include a variety of aspects. Family responsibilities might be housekeeping duties, care of children or care of sick relatives. Health problems include general poor health, physical handicaps, specific illness, and other disabilities. Persons indicating lack of experience, education and skill include those who are unable to find jobs because of limited education, experience or skills. Those who indicated that employers thought they were too young or too old consist of those too young to get work permits, as well as those who have been refused a job because either the job requires someone more "mature" or someone younger who can be expected to remain on the job for a longer time.

Among the additional reasons identified for those not in the labor force, the retirement category includes persons who report old age as their reason for not looking for work or wanting a job. It consists of persons who think of themselves as beyond working age, as contrasted to those who report that employers regard them as too old to work. Transportation problems include poor or non-existent public transportation or no car or other private transportation being available. The category looked but couldn't find work is for an individual who is not presently looking for work because he had looked unsuccessfully at a previous time. The presumption here is that he believes that there are still no jobs available for him, so he is not bothering to look.

A great deal of attention has recently been focused on "discouraged workers" -- persons not in the labor force who want a job but are not looking because they think they cannot get work. In the CES, a count of these persons could include persons whose primary reason for not seeking work was one of the following: they previously had looked but couldn't find work; employers thought them to be too young or too old; they lacked necessary skills, experience, or education; or they had transportation problems preventing them from getting to a job.

For persons out of the labor force, information also was obtained and published on the extent of desire for a job in addition to the information on reasons for not seeking work. Persons classified as employed or unemployed may be assumed to have some desire for a job -- they are either actually working or actively seeking work. Thus, any reason they perceive as preventing them from finding a job or getting a better one may be considered a barrier to satisfactory employment. On the other hand, persons not in the labor force may or may not want to work and any attempt to gather and evaluate information on reasons for their nonparticipation must first ascertain their desire for work and degree of labor force attachment. Those who indicate a desire for a job and who identify specific reasons for not looking for work also may be considered as facing a barrier to satisfactory employment.

Thus, in terms of desire for a job, the CES classifies persons not in the labor force as follows: those who want a job now, those who may want a job, those who would want a job if it weren't for their problem or responsibility, and those who do not want a job. Nonparticipants were first asked if they wanted a job. Those who wanted or might have wanted a job were asked for their reasons for not seeking work, the remainder were asked why they did not want to work. If a member of the latter group indicated a specific reason for not wanting a job, he was asked if he would want a job if it were not for the specific problem or responsibility keeping him out of the labor force.

Additional data for some of the problem categories faced by both workers and potential workers were collected but not published. For persons with health problems, the specific illness or disability and its duration were also obtained from the respondent. Persons indicating that their family responsibilities included child care were asked additional questions about their problems in arranging for the care of their children during work hours and specifically for their attitudes toward child care centers. Persons indicating a lack of skill, education or experience were asked about their interest in returning to school for additional training. In addition, persons not in the labor force because they looked but couldn't find work were asked a series of questions about the jobseeking methods which they used and the job training programs which they might have taken.

Work Experience

In addition to the wealth of information on the current employment status of poverty area residents, the CES collected some important data on their employment experience over the 12 months prior to the survey week. These data are conceptually the same as those collected annually for the Nation as a whole in the March supplement to the Current Population Survey and published by the Bureau of Labor Statistics in its Special Labor Force Report series.

Work experience data provide a more comprehensive insight into the overall employment experience of workers by obtaining information on the number of weeks a person was employed or unemployed during the entire year rather than his status as of the survey week, which may or may not be typical of his usual status. Thus, the work experience data provide a more useful tool for indicating patterns of labor force participation by measuring the components of both stability (the proportion of year-round workers) and of movement (part-year workers).

The work experience data are particularly useful for an analysis of the employment situation of persons likely to have multiple employment problems, such as residents of low-income areas. The year-long concept permits the development of measures that supplement the more often used information on employment and unemployment at a point in time. For example, it provides information on the number of persons who were successful in obtaining only intermittent or part-year employment, the number working year-round full-time, but earning low wages; the number who experienced several spells of unemployment during the year; and other measures of the lack of employment success.

Most of the work-experience data are published in the CES reports. However, there are some non-tabulated data relating to the work history of the population that would shed additional light on the extent of unemployment during the year among two specific worker groups. The first group is those persons who unsuccessfully looked for work during the year but never worked. In addition, data are not shown on the number of year-round workers who could have been unemployed 1 or 2 weeks during the year. Although part-year workers make up the vast majority of the unemployed over the year, the missing data would provide a complete measure of the extent of unemployment among the low-income area population.

Principal Jobseeking Methods

One of the reasons sometimes cited for the lack of success that workers in low-income areas have in finding jobs or getting better ones is that they have a limited knowledge of the available job opportunities, and their job search techniques are inefficient. How do poverty area residents seek jobs? Do they generally use formal or informal means? Do methods differ substantially among occupational groups? In order to shed some light on these questions, the CES provides information on the principal job seeking methods used by workers who looked for a job during the previous 12 months cross-tabulated by occupation, industry and spells of unemployment.

Most of the information obtained on jobseeking methods is published in the CES reports. However, the data were tabulated only for the principal method of job search used by part-year workers. In addition, information from the question -- "Which way of looking for work got you your present (or most recent) job?" -- was not tabulated. The collation of these data would permit a much better measure of the impact of effectiveness of the various jobseeking activities and could provide valuable information for the planning and structuring of employment programs aimed at more effective matching of job seekers with available jobs.

Transportation

The absence of quick, inexpensive, and adequate transportation to the workplace, particularly those located in the fast growing job markets outside central cities, can be a major barrier to the finding and holding of good jobs by inner city residents and may create a sense of isolation from socio-economic opportunities. More important, it can perpetuate the marginal existence of low-income area residents of inner cities by exacting daily transportation costs (in terms of time and money) large enough to make working at a regular job not a viable way out of poverty conditions. The CES tabulations relating to transportation provide information on the usual place of work, method and cost of transportation, and time required to get to work for the employed and the availability of transportation to work for unemployed poverty area residents. The interrelationship of these data with several key variables should answer many questions about the commuting problems of inner city residents, particularly those who must daily go to jobs outside the city. Hopefully, they will permit a clearer basis for policy decisions in this area.

Most of the data collected on transportation characteristics is published in the CES reports. However, information was collected, but not published, from the following question asked of employed workers: "What time do you usually get to work" and "Do you use any other way at least once a week?" In addition, responses to questions on the place unemployed workers usually went to look for work and the time required and cost of transportation used to get there have not been tabulated. Collation of these data would allow a fuller profile of the transportation to work characteristics, patterns and problems of low-income area workers.

Earnings and Income

The CES reports include an abundance of data on both the earnings and income of low-income area families. Earnings and income data are among the most connotative indicators of the well-being of workers and their families. The data reflect the labor force status of the population and often are direct measures of employment problems; they are particularly useful in measuring the employment difficulties of low-income area residents. For example, high hourly or weekly pay coupled with low annual earnings indicate the incidence of part-year or part-time employment (or both). Data on annual earnings below poverty levels for full-time year-round work provide still another measure of lack of employment success. Other revealing measures are the percent of families below the poverty level by source of income or occupation of the family head, and levels of family income by number of family earners.

When coupled with the large number of demographic and employment variables tabulated in the CES reports, the earnings and income data provide a quite comprehensive economic profile of the survey area worker and his family. They are also useful as a measure of the well-being of persons and families living in the target areas compared to those living in the Nation as a whole.

The annual income data are quite diverse and cross-tabulated with a number of important variables. In addition, most of the income tables include information on the number and percent of persons and families below the poverty line. The data on annual income and poverty level are tabulated by size of family, number of children under 18 years, fixed monthly expenses, monthly charges for mortgages, rent, and utilities, age and education of family head, age and education of unrelated individual, number of earners and nonearners, source of income, detailed source of income (without poverty level information), work experience of family head in last 12 months, and major occupation group of family head's longest job in last 12 months.

Lowest Acceptable Weekly Pay

One subject matter for which little information is available for the general population or for persons living in specific areas is that of the lowest pay acceptable to persons when they were looking for work. One well-known hypothesis which has been advanced to explain why the unemployed, particularly jobless youth, are unsuccessful in the job search is that their wage expectations are unrealistic, that is, they are not willing to accept the prevailing wage for the type of job which is available to them. Information is available in the CES reports on the lowest weekly pay acceptable to persons who looked for work at any time during a 12 month period prior to the survey. This information is presented in extensive detail by age and sex, family status, educational attainment, major occupation and industry group, and duration of unemployment of part-year workers. The lowest acceptable pay data could be evaluated in conjunction with actual weekly earnings data from the survey. Together these data would provide an indication of whether low-income and workers in these cities have unrealistic wage demands, or whether they expect realistic pay, consistent with prevailing local wage rates.

Mobility and Migration Patterns

The CES provides important information on mobility and migration of poverty area residents. One of the most significant factors affecting the job market in large American cities in recent years has been the migration of large population groups. The exodus of Negroes from the rural South to large northern cities, the large number of Puerto Ricans moving from Puerto Rico to New York City and other eastern cities, the migration of Cuban residents to cities in Florida, and the movement of Mexican residents to many southwestern and western U. S. cities are examples of this pattern. Unfortunately, little is known about the condition of the newly arrived compared to that of the long-time resident.

It is necessary to seek the answers to a number of questions concerning the residents of these areas. For example, who are the people living in the Nation's low-income areas? Are the majority of the residents members of a particular ethnic group? Where did these residents come from? Are those who have been lifetime residents of the area better off than the newcomers? Are those who were born in farm or rural areas as well off as those who were born in the city in which they now reside? Are newcomers to the area younger than longtime residents? Is educational attainment higher for residents born in the area than it is for migrants? Substantive answers to these questions should indicate more clearly the direction that programs and action should take in order to solve the employment, income and related problems of the residents of these areas. The CES provides information relative to several of these questions based on place of birth and place of residence 5 years ago cross-tabulated by annual income, size and composition of family, and educational attainment of family head.

Additional information on residential mobility was collected but not included in the CES reports. These data include the number of years the respondent has lived at his present address, how many times he moved during the last year, where he lived at age 16, and how many years (in detail) he has lived within his present city limits.

****

The CES reports contain a substantial amount of detailed data on these as well as other subject matter characteristics. The data are presented in each urban area report according to the formats described in this paper. However, there are several differences in the subject matter content of the rural area surveys. Thus, many statements in this paper do not apply to the CES rural area data.

Although the Census Employment Survey will obviously not answer all our questions concerning the complex socio-economic problems of low-income area residents, it does provide more information for more specific areas than was previously available. When viewed and used together with other survey data such as that from the Decennial Census and the Current Population Survey, it will hopefully help all of us to at least better understand the extent and kinds of employment difficulties faced by a large number of our fellow citizens. Once those problems are accurately measured however, there still remains the infinitely difficult job of solving them.

ESTIMATING THE DISTRIBUTION OF THE FEDERAL INCOME AND SOCIAL SECURITY TAXES - AN INTERIM REPORT ON A SIMULATION APPROACH

by

Roger A. Herriot*
Social and Economic Statistics Administration, Bureau of the Census

I. Introduction

In this paper I shall present some preliminary findings from a continuing research project at the Census Bureau on the distribution of taxes. In past studies,1/ we have presented estimates of the average amount of various taxes paid by families and unrelated individuals in each adjusted total money income class.2/ However, more work is needed to disaggregate and prepare estimates for various subgroups of the population. For example, how much tax is paid by people with "low income," or those over 65 years old?

In order to fulfill this need we have been developing a simulation procedure wherein we could simulate the amount of tax paid by each person or household covered in the Current Population Survey (CPS). Thus far we have simulated information for only the Federal income tax and the Personal Contributions for Social Security. While I had hoped to present some estimates of these taxes by race, age, and "low income" status, these estimates are not yet finished. We have however, completed some estimates of these taxes by household and family status. The next section of this paper will discuss the simulation methodology. This will be followed by a summary evaluation of the procedure. In the final section, a brief discussion of some of the results based on income for 1968 will be presented.

II. Methodology

A. Income Tax

Determination of Adjusted Gross Income

Although it is not possible to estimate, with precision, the Internal Revenue Service (IRS) concept of Adjusted Gross Income (AGI) from the money income information reported in the CPS, our approximation of their concept is believed to be sufficient for most purposes. This approximation of AGI will be referred to in this paper as "AGY". To approximate AGI, the following transfer payments were subtracted from Total Money Income (TMY): (1) Social Security and Railroad Retirement payments, (2) Public Assistance or Welfare, (3) Veterans' benefits, (4) Unemployment Compensation, (5) Money Workmen's Compensation, and (6) Government Employee Pensions. Although the last item should not have been excluded, it was reported in total money income with other transfers and thus had to be removed with them. The most serious faults of AGY as an approximation of AGI is in the underreporting and nonreporting of property income (i.e. interest, dividends, and net rent) in the CPS, and the fact that AGY does not include any "income" from capital gains. In addition, it is not possible to code in the CPS, an amount for an individual type of income in excess of $99,999. Although these shortcomings of the CPS permeate throughout the income distribution, they only become serious (for our purpose) at the higher income levels.

Determination of Filing Status and Type of Return

For each potential filing unit (a filing unit consists of either a married couple or a person), a series of tests was performed to determine their filing status. It was assumed that any potential filing unit whose AGY was greater than the legal limit required for filing in 1968 would file a return. The legal limits were $1,200 for a married couple in which both the husband and wife were under 65 years of age, $1,800 if either husband or wife (but not both) was 65 years or over, and $2,400 if both husband and wife were 65 years or over.

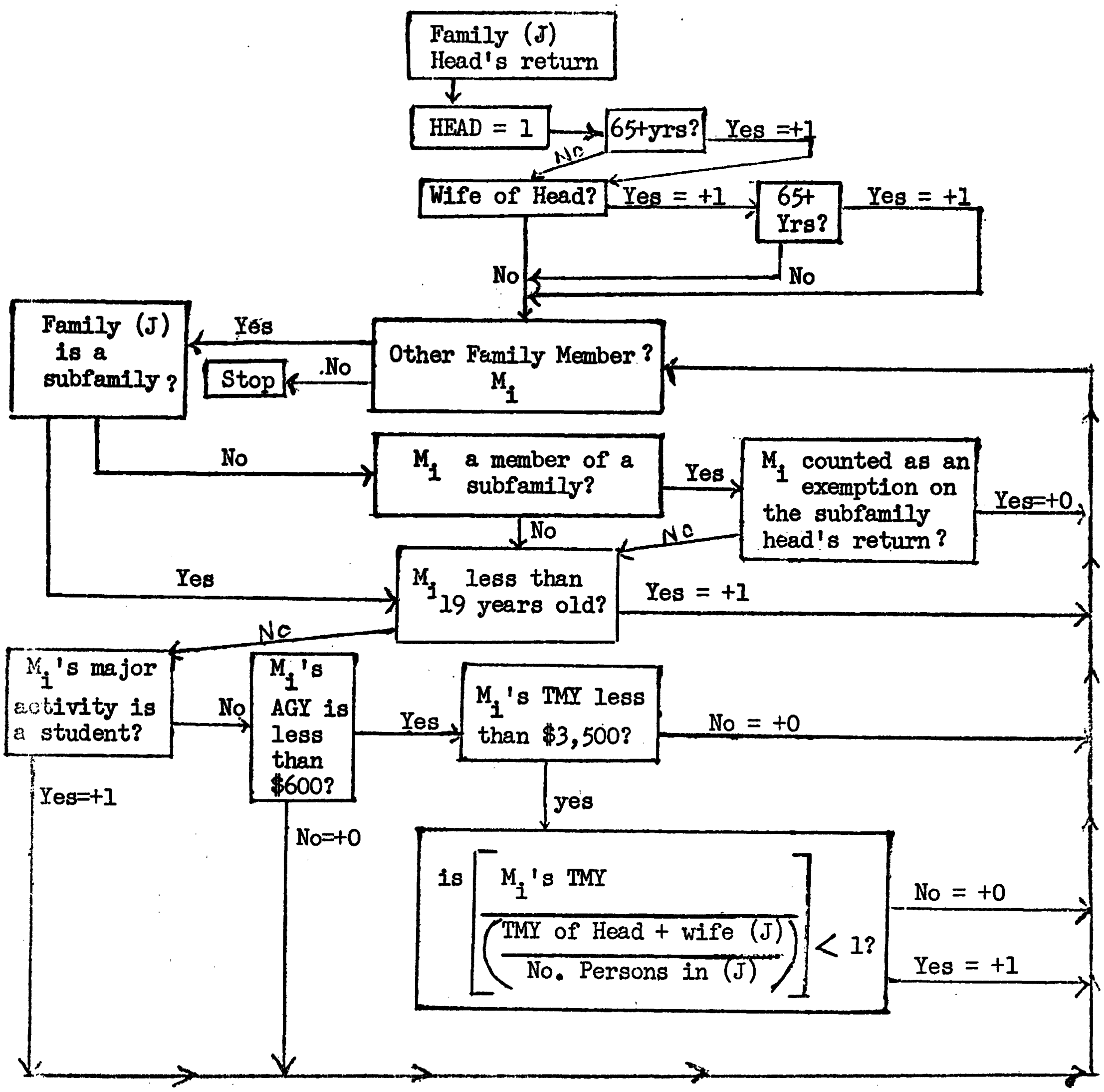
All other persons were assumed to be filers if they were less than 65 years old and had AGY of more than $600 or were 65 years or over and had AGY of more than $1,200. In addition, any potential filing unit which failed to meet the above test, but which received wage and salary income during the year was assumed to file a return to get its withholdings refunded.

To determine the type of return filed for each filing unit we made the following assumptions:

1) Family head returns:
   a) Married Couples.--all married couples were assumed to file joint returns.
   b) Qualified Heads.--these returns, which approximate IRS's "Heads of Household" and "Surviving Spouse" returns, were for heads of other than husband-wife families who met the qualification test. To qualify, the head's TMY had to be greater than the sum of the TMY of all other family members. All who qualified were assumed to file "Heads of Household" returns.
   c) Nonqualifying Heads.--returns of heads of other than husband-wife families who failed to meet the above qualification test. These persons were assumed to have filed "Single person not head of household or surviving spouse" returns.

Determination of Number of Exemptions per Return

To estimate the number of exemptions on family head returns, we used the procedure shown in the following flow chart:



A family unit can consist of a "primary family," including one or more "subfamilies" or a "secondary family." A subfamily is a married couple with or without children, or one parent with one or more single children under 18 years old living in a household and related to, but not including the head of the household or his wife. As members of a subfamily are also members of a primary family, the number of subfamilies is not included in the number of "families." A "secondary family" is similar to a subfamily except that the members of the secondary family are not related to the primary family members. Since subfamily members could be counted as exemptions on either a subfamily head's return or on a return of the head of the primary family, we begin with subfamily members. An exemption is granted for the head and additional exemptions are granted for a wife and for age (i.e., 65+ years). An additional exemption is also granted for other subfamily members if they are under 19 years of age, or their major activity is a student. In addition, if the person fails to qualify as an exemption on the above grounds, they can qualify if their AGY is less than $600 and they pass our crude "support" test. We assumed financial dependency if their TMY is less than $3,500 and in addition, their TMY is not greater than the TMY of the filing unit spread equally among all members of the family. After we have completed the process for all subfamily members, the process is repeated for all primary family members (including members of any subfamilies). The same set of procedures applies to the returns of primary and secondary family heads with the exception that members of the primary family who are also members of a subfamily cannot be counted as exemptions on the primary family head's return if they have already been counted as an exemption on the subfamily head's return.

For all returns other than those for family heads, one exemption was granted for the person and an additional one was granted if he or she was 65 years or over.

Determination of Deduction, Taxable Income, and Tax

To estimate the deductions (DED), taxable income (TY), and income tax (YT), for each return, the following procedures were used:

Assume a return (r) of type (i) having \$AGY which is in the (jth) AGI interval.

Then:

$$\$DED_r = \frac{IRS\ \$DED}{IRS\ \$AGI_{ij}} \times (\$AGI)_r$$

$$\$TY_r = (\$AGY)_r - (\$DED)_r - (EXEMPT)_r \times (\$600)$$

$$\$YT_r = \frac{IRS\ \$YT}{IRS\ \$YT_{ij}} \times (\$TY)_r$$

Where:
r = type of return, i.e.

1. Joint return, head less than 65 years old.
2. Joint return, head 65 years or over.
3. Head of Household return.
4. Return of Single persons (not H. of H. or S.S.), less than 65 years
5. Return of Single persons (not H. of H. or S.S.), 65 years or over.

B. Social Security Contributions

For all persons who were estimated to be covered under the Social Security programs, we assessed a "contribution" of 4.4 percent of the first \$7,800 of their wage and salary earnings and 6.4 percent of the first \$7,800 of their self-employment earnings. However, if a person received both wage or salary and self-employment earnings, the contribution was limited to a maximum of 6.4 percent of the first \$7,800 of their total earnings.

All persons receiving wage and salary income were assumed to be "covered" with the exception of those whose occupation of longest job was a government worker, a farm laborer, or a domestic. Since only a portion of the people in these occupations were covered, a probability algorithm was used to assign "covered" status. The proportion for each occupation which was covered was estimated from unpublished Social Security administrative data.

All persons whose class-of-worker of longest job was self-employed in agriculture or nonagricultural industries and who has \$400 or more in net self-employment earnings were assumed to be covered. Actually, there are other options open to farmers allowing them to be covered even if they have less than \$400 in net earnings, but we did not have sufficient information available in the CPS to allow for these options.

III. Summary Evaluation of the Simulation Procedure

Although many of our procedures only roughly approximated IRS concepts, the overall results appear to be adequate when compared with published IRS statistics. For example, in Table 1, the CPS estimated number of filing units, by type of return are compared with those reported by IRS.

Table 1.--Comparison of CPS Estimated Number of Filing Units Filing Each Type of Return With Those Reported by IRS for 1968 (In millions)

CPS				IRS	
Potential Filing Unit Description	Total	Estimated Filers		Filers	Filing Unit Description
Family Heads1/	51.4	46.1		NA	-
Married Couples1/	44.3	41.1		42.7	Joint Returns + ½ of Separate Returns of Husband and Wives
Qualified Heads1/	4.5	3.5		2.6	Returns of Head of Household or Surviving Spouse
Nonqualifying Heads1/	2.6	1.5			
Other Persons (14+) in Family	32.8	18.4	29.2	27.0	Returns of Single Persons not Head of Household or Surviving Spouse
Unrelated Individuals	13.8	9.4			
Total	98.1	73.9		72.3	Total

1/ Includes primary, secondary, and subfamily heads.

NA Not Available

Of the 51.4 million family heads (counting subfamily heads as the head of a family), it was estimated that 46.1 million were filers. Married couple filers accounted for about 41.1 million of these; correspondingly, IRS shows about 42.7 million husband and wife filing units. Of the remaining 5 million family head returns, we estimated that about 3.5 million qualified to be taxed at "Head of Household" rates. This is considerably more than the 2.6 million returns IRS shows filing as either "Head of Household" or "Surviving Spouse." The remaining returns consisted of 1.5 million Nonqualifying Head returns, 18.4 million returns from Other Persons in Families, and 9.4 million Unrelated Individual returns. These 29.2 million returns were assumed to file single person returns. The number of "single person" returns according to IRS was about 27.0 million returns. Much of the differences between these CPS and IRS numbers is believed to be accounted for by the way we handle the approximately 2 million persons who were married, not separated, but having a spouse who was absent. While many of these people might have actually filed joint returns to IRS, they are excluded from the married couple filing units in our procedure. As we had no information regarding income or other characteristics of the absent spouse, we treated these people as an "other family head" or an "other person in families" depending on their current living arrangement. We suspect that most of those to whom we gave filing status, and had children were counted in the "Qualified Head" category. If this is true it would account for much of the differences from the IRS figures for Married Couple and Qualified Head returns.

In Table 2, when we compared the CPS estimated number of returns, amount of income tax, and mean income tax per return with corresponding IRS figures by adjusted gross income intervals, we again found that overall, the estimates correspond reasonably well with those of IRS. The major exception to this general conclusion is in the CPS returns having more than $50,000 AGY. The $50,000 and over interval is seriously deficient in both numbers of returns and mean AGY and thus, in the mean income tax per return. This interval accounts for almost all of the differences between the CPS aggregate income tax and that reported by IRS. This can be seen by comparing the aggregates in the "Total" row with those in the "Total under $50,000" row.

Table 2 Comparison of CPS Estimated Number of Returns, Amount of Income Tax and Mean Income Tax Per Return by Adjusted Gross Income Interval with IRS Data, for 1968.

Adjusted Gross Income Intervals	CPS					IRS				
	Estimated Number of Returns		Estimated Amount of Income Tax		Mean Income Tax per Return	Number of Returns		Amount of Income Tax		Mean Income Return
	No.	Cum. No.	Amount	Cum. Amount		No.	Cum. No.	Amount	Cum. Amount	
	mil.	mil.	bil.$	bil.$	dollars	mil.	mil.	bil.$	bil.$	dollars
TOTAL	73.9	-	$63.7	-	$ 862	73.7	-	76.6	-	$1,039
Total Under $50,000	73.9	-	61.8	-	836	73.3	-	63.4	-	864
Under 1,000	11.4	11.4	0.0	0.0	0	7.7	7.7	X	X	1
1,000 - 1,999	5.8	17.2	0.2	0.2	41	7.5	15.2	0.4	0.4	51
2,000 - 2,999	4.6	21.8	0.6	0.9	134	5.9	21.1	0.9	1.2	146
3,000 - 3,999	5.2	27.0	1.3	2.2	252	5.6	26.7	1.5	2.7	264
4,000 - 4,999	4.9	32.0	1.8	4.0	370	5.3	31.9	2.0	4.7	381
5,000 - 5,999	5.1	37.1	2.5	6.5	490	5.0	36.9	2.5	7.3	509
6,000 - 6,999	5.1	42.2	3.2	9.7	529	5.0	41.9	3.1	10.4	635
7,000 - 7,999	5.2	47.4	4.0	13.7	758	4.7	46.6	3.7	14.1	781
8,000 - 8,999	4.5	52.0	4.1	17.8	892	4.6	51.3	4.2	18.3	913
9,000 - 9,999	3.9	55.9	4.1	21.9	1,044	4.0	55.3	4.3	22.6	1,068
10,000 - 14,999	12.2	68.1	18.4	40.2	1,509	12.0	67.3	18.3	40.9	1,527
15,000 - 19,999	3.5	71.6	8.8	49.0	2,506	3.7	70.9	9.3	50.2	2,535
20,000 - 24,999	1.1	72.7	4.2	53.2	3,695	1.2	72.1	4.4	54.6	3,730
25,000 - 50,000	1.2	73.9	8.6	61.8	7,428	1.2	73.3	8.8	63.4	7,084
50,000 +	Z	73.9	1.9	63.7	28,031	0.4	73.4	13.2	76.6	34,479

X = less than $50,000
Z = less than 50,000 returns

In addition, although it has little effect on the aggregate amount of tax, there is a less than satisfactory distribution of CPS returns among the lower intervals. For example, while the total number of CPS returns with less than $3,000 AGT is roughly comparable with the number of IRS returns under $3,000 AGI, there is considerable difference between the CPS and IRS numbers below $1,000.

A final anomaly is the mean tax per return in the $25,000 to $50,000 AGY interval. Although the estimated mean tax is a few dollars lower than the IRS counterpart for all other intervals (except the $50,000+ interval already discussed), the mean for this interval is several hundred dollars higher than the corresponding IRS amount. This condition is thought to result from the relatively greater concentration of returns in the upper half of this interval in the CPS compared to that of IRS. This would cause the CPS average income and hence the tax amount to be greater than the corresponding IRS figure. The relatively greater concentration of returns in the upper half of this interval in the CPS could result from the coding of people with $50,000+ AGI in the $25,000 - $49,999 AGY interval because of the exclusion of capital gains and underreporting of income in CPS. Both of these factors are particularly pronounced among very high income families.

Regarding the personal contribution to the Social Security program, there is not much outside data with which to compare. Our estimate of the aggregate amount of personal contributions for Old-Age Survivors, Disability, and Hospital Insurance was $16,873 million which was surprisingly close to the Bureau of Economic Analysis estimate of $16,859 million.

IV. Results by Total Money Income Class

Income Tax

Of the estimated 64.5 million returns filed by family members in 1968, about 45.1 million or 70 percent were filed by the heads of primary and secondary families (see table 3). These returns accounted for roughly 91 percent of total family income tax. The remaining 30 percent of family returns were filed by other family members (including the heads of subfamilies) and accounted for about 9 percent of family income tax. The proportion of family returns accounted for by returns of the family heads, tends to be highest for the middle income groups, however, the proportion of family's income tax bill picked up by the head increased with income but tends to be roughly constant among the middle income groups. For families above $4,000 total money income, the head becomes increasingly the primary taxpayer. This is reflected in the increasing divergence in the mean tax per return of the head's return and other family member's returns.

If one were to attempt to describe the Federal income tax situation of the "average" American family in 1968, one would probably not be far off if they looked at the families in the "medianth" income group (i.e., the income group which contained the median family). This would be the $8,000 - $8,999 income group in 1968. The roughly 4 million families in this group filed about 5.1 million returns in 1968, or about 1.3 returns per family. They paid around 3.1 billion dollars in Federal income tax, which works out to be about $613 per return and $789 per family. Virtually all, about 98 percent, of the family heads filed a return - paying, on average, $730 in income tax per return. The returns of these family heads accounted for about 77 percent of the returns and 91 percent of the income tax paid by these families. The remaining 23 percent of the returns were filed by other members of the family (excluding wives) and they paid, on average, $226 in income tax per return.

Social Security Personal Contributions

The estimates of the distribution of Personal Contributions for Social Security are presented in Table 4. There is little to say about these results since they pretty much speak for themselves. The average amount contributed per earner in families increases up to the $10,000-$14,999 Total Money Income interval. The average amount per family increases to at least the $50,000 money income class due to both the increasing average contribution per worker and the increasing number of workers per family.

1/ Herriot, Roger A. and Herman P. Miller, "The Taxes We Pay" The Conference Board Record, May 1971 and "Changes in the Distribution of Taxes: 1962-1968," Proceedings of the Business and Economic Statistics Section, American Statistical Association, 1971.

2/ The amounts of money income reported in the CPS were not adjusted for this study as they were in the above studies. While the adjustment procedures in our earlier studies appear adequate when applied to all families and unrelated individuals, they were judged to be inadequate when applied to certain subgroups of the population.

* I would like to thank my colleague, John Coder, who did all of the programming in this and our earlier studies.

Table 3.--CPS Estimate of Returns and Mean Federal Income Tax per Return by Family Status and Total Money Income Class for 1968

Unadjusted Total Money Income Intervals	Families									Unrelated Individuals		
	Total Number Families	Total Family Returns		Returns of Primary and Secondary Family Heads			Returns of Other Family Members			Total Number of Unrel. Indiv.	Total Unrel. Indiv. Returns	Mean Tax per Return
		No.	Agg. Tax	Percent of Total Fam. Ret.	Percent of Total Fam. Tax	Mean Tax per Return	Percent of Total Fam. Ret.	Percent of Total Fam. Tax	Mean Tax per Return			
	mil.	mil.	bil.$	percent	percent	dollars	percent	percent	dollars	mil.	mil.	dollars
TOTAL	50.0	64.5	54.4	70.0	91.2	$1,139	30.0	8.8	$258	13.8	9.4	$ 776
Under $1,000	0.9	0.6	0	57.8	-	0	42.2	-	0	2.0	0.6	0
1,000-1,999	1.8	1.2	Z	64.2	14.3	1	35.8	85.7	14	3.4	1.1	46
2,000-2,999	2.6	2.0	0.1	69.9	55.2	23	30.1	44.8	44	1.8	1.2	132
3,000-3,999	3.1	3.0	0.3	74.1	74.1	89	25.9	25.9	89	1.5	1.4	295
4,000-4,999	3.1	3.4	0.6	75.1	78.9	171	24.9	21.1	137	1.1	1.1	487
5,000-5,999	3.5	4.2	1.1	76.6	84.8	282	23.4	15.2	166	0.9	0.9	683
6,000-6,999	3.8	4.7	1.8	77.6	87.8	433	22.4	12.2	207	0.7	0.7	902
7,000-7,999	4.1	5.1	2.6	77.2	89.8	587	22.8	10.2	226	0.6	0.6	1,102
8,000-8,999	4.0	5.1	3.1	76.6	91.4	731	23.4	8.6	226	0.4	0.4	1,320
9,000-9,999	3.7	4.8	3.5	74.4	90.7	878	25.6	9.3	263	0.3	0.3	1,498
10,000-14,999	12.5	18.0	17.5	68.7	91.8	1,296	31.3	8.2	253	0.7	0.7	2,148
15,000-19,999	4.5	7.6	10.4	59.0	90.1	2,087	41.0	9.9	331	0.1	0.1	3,656
20,000-24,999	1.5	2.7	4.9	55.4	89.3	3,040	44.6	10.7	452	0.1	0.1	5,283
25,000-49,999	1.3	2.2	9.0	59.4	94.0	6,621	40.6	6.0	619	0.1	0.1	9,839
50,000+	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Z = less than $500,000

NA = Not Available

Table 4 Amount of Personal Contributions for Social Security by Family Status - Total, per Family with Earnings, and per Workmen - by Total Money Income Intervals for 1968

Unadjusted Total Money Income Interval	Personal Contributions for Social Security				
	Contributions of Family Members			Contributions of Unrelated Individuals	
	Total Amount	Amount per Family with Earnings	Amount per Earner	Total Amount	Amount per Earner
	mil.$	dollars	dollars	mil.$	dollars
TOTAL	$15,442	$332	$182	$1,432	$167
Under $1,000	10	15	15	10	16
1,000 - 1,999	44	44	35	44	40
2,000 - 2,999	134	85	70	79	76
3,000 - 3,999	309	132	105	146	122
4,000 - 4,999	458	174	130	154	166
5,000 - 5,999	703	219	151	175	210
6,000 - 6,999	968	263	174	165	238
7,000 - 7,999	1,226	304	198	180	284
8,000 - 8,999	1,315	334	203	125	321
9,000 - 9,999	1,297	359	207	88	297
10,000 - 14,999	5,213	418	222	201	317
15,000 - 24,999	3,075	517	221	48	338
25,000 - 49,999	675	531	225	15	326
50,000 +	NA	NA	NA	NA	NA

NA = Not Available

MEASURING SCHOOL AND COLLEGE ENROLLMENT: A COMPARISON OF CENSUS BUREAU AND OFFICE OF EDUCATION DATA

Charles E. Johnson, Jr. and Larry E. Suter
Bureau of the Census

The data on school and college enrollment collected by the Bureau of the Census in the 1970 Census are generally comparable with the data collected by the Bureau in the October 1969 Current Population Survey (CPS) and with the data collected in the fall of 1969 by the Office of Education (OE). These three series of data on school and college enrollment are collected for different purposes and present different geographic detail and different characteristics concerning the student population, but they do have similar rim totals on the numbers in elementary school, high school, and college. The Census data present information on the demographic, soial, and economic characteristics of the enrolled and not enrolled population for State and local areas once every 10 years. The CPS data present annual National information on the demographic and socioeconomic characteristics of the enrolled and not enrolled population collected in the household. The OE data present information on the characteristics of the schools, teachers, and school finance, as well as some student characteristics, collected from the institutions themselves or from State and local officials.

Although generally comparable, there are some interesting differences among the three series of data. Most of the diversity among the data appears to be due to the differing collection dates, methods, and definition of terms. The Current Population Survey data were collected during the first semester of school (October 1969) as were the Office of Education data while the Census data were collected during the second semester of school (April 1970). The fall and spring collection dates caused differences in level of enrollment (primarily because of attrition between the first and second semester) and differences in enrollment rates by age. The 1970 Census data were collected from a 15 percent sample of the Nation's population by both enumerators and self-response. The CPS data were collected from a sample of 50,000 households by direction enumeration of any qualified respondent. The Office of Education data were collected from a virtually complete enumeration of institutions or State and local officials.

Level of School

1970 Census and 1969 CPS.--Overall the number of students 3 to 34 years old, enrolled in school in the Census for April 1970 was 2 percent smaller than the number in the CPS of October 1969 (Table 1). Attrition between semesters probably accounts for some of this difference in total enrollment. The number of persons 3 to 34 years old who were not enrolled in school in the Census was 9 percent larger than the number in the CPS. The universe dissimilarity between the Census and the CPS probably accounts for this difference since the Census universe included the total resident population but the CPS was restricted to the civilian noninstitutional population.

At the nursery school level, the number of pupils enrolled was 11 percent larger in the Census (spring) than in the CPS (fall). This increase is contrary to the attrition hypothesis. However, since nursery school enrollment is increasing, and since nursery school enrollment does not necessarily have to begin in the fall, this increase might be reasonable. At the kindergarten level, the number of pupils enrolled was 8 percent smaller in the Census than in the CPS. Since some attrition is not unexpected, and since kindergarten enrollment is not compulsory in all States, this decrease might not be unreasonable. At the elementary school level, there were 2 percent fewer pupils enrolled in the Census than in the CPS; at the high school level, there were 0.5 percent fewer students enrolled in the Census than in the CPS; and at the college level, there were 6 percent fewer students enrolled in the Census than in the CPS.

Census and Office of Education. At the elementary and high school levels there were 4 percent fewer enrolled in the Census than in the Office of Education figures for the fall of 1968 (OE collected data on only public school enrollment below the college level in 1969), and at the college level there were 7 percent fewer enrolled in the Census than in the OE data for the fall of 1969.

CPS and Office of Education. The CPS and OE estimates of total degree-credit college enrollment in the fall of 1969 were quite similar (table 3). The CPS estimate was less than 1 percent smaller than the OE figure--not a statistically significant difference because of a sampling error of about 120,000 in the Current Population Survey. However, there were differences by level of college. The CPS estimate of undergraduate enrollment was 6 percent lower than the OE estimate. And the CPS estimate of graduate enrollment was 27 percent higher than the OE figure. The larger number of graduate students in the CPS was somewhat surprising, since the CPS universe has an upper age limit of 34 years whereas the OE universe has no upper age limit. This finding suggests that there may be some overreporting of enrollment in the fifth year of college or more in the CPS.

Office of Education and CPS estimates of college enrollment are shown in table 3 for each year from 1960 to 1970. In all but three years the OE total degree-credit college enrollment level was higher than the CPS estimate of college enrollment of 14 to 34 year olds. The yearly variations in enrollment are greater in the CPS than the OE figures probably because the CPS is more influenced by changes in enrollment rates of persons 18 to 24 years old than are the OE figures which include persons over age 34. There have been significant shifts in enrollment rates for men of draft age since 1964; enrollment rates increased between 1964 and 1966 and decreased between 1967 and 1970.

The 1969 CPS estimate of part-time undergraduate degree-credit enrollment was 64 percent lower than the OE estimate whereas the CPS estimate of full-time enrollment was 6 percent higher (table 4). One possible reason for the difference between the CPS and OE data on part-time enrollment may be the CPS questionnaire format. Several questions relating to major activity and employment status are asked prior to the question on school enrollment in the CPS and this may lead some persons who are attending school part-time, but working full-time, to respond "no" when asked if they are attending or enrolled in school. If the CPS question asked if the persons were enrolled or attending school either part-time or full-time, the CPS might uncover more persons attending college on a part-time basis. Another reason is coverage. The larger number of part-time students counted by OE may be due to the restricted age limit of the CPS, which is cut off at age 34.

Control of School

There was a considerable amount of variation between the Census and the CPS measurement of level of enrollment in public and private schools (table 5).

At the nursery school level, there were more children enrolled in both public and private schools in the Census than in the CPS. In public kindergartens there were 138,000 fewer children enrolled in the Census than in the CPS and 114,000 fewer children enrolled in private kindergartens in the Census than in the CPS. (One standard error on the CPS estimate would be around 17,000 persons). The proportion of all nursery school pupils who were in private schools was about the same.

At the elementary school level, the Census and the CPS data both show the same proportion of elementary school pupils enrolled in private schools--11.5 percent in the Census and 11.7 percent in the CPS. The Census data show fewer children enrolled in parochial elementary schools (in the Census this would include all children in any school supported by a religious organization) than the Official Catholic Directory shows enrolled in Catholic elementary schools in the fall of 1969. The Census shows 3.4 million children enrolled in parochial elementary schools (which would include Catholic and other church-supported schools) whereas the Official Catholic Directory shows 3.7 million children enrolled in Catholic elementary schools in the fall of 1969. The Census figure on the percent of private elementary school children enrolled in parochial schools is lower than the Office of Education estimate for the fall of 1968. (OE did not collect similar data in the fall of 1969). In the Census, 88 percent of the students enrolled in private elementary schools were in parochial schools as compared with 95 percent of the students included in the Office of Education survey. The Office of Education proportion might be somewhat high, since they obtained more complete coverage of the Catholic parochial schools than of the secular private schools. And the Census estimate may be low, if some students enrolled in private Catholic schools not under the control of the parish and some students enrolled in non-Catholic but church-supported schools classified themselves as enrolled in "other private" schools rather than in parochial schools.

At the high school level, there were fewer students enrolled in public schools but more students enrolled in private schools in the Census than in the CPS. The 1960 Census also showed more persons enrolled in private high schools than did the October 1959 CPS, but the increase was only 3 percent as compared with the 23 percent increase between the October 1969 CPS and the 1970 Census. (The 1960 Census also showed more persons enrolled in public high schools than did the 1959 CPS).

Preliminary Office of Education data for the fall of 1970 show 1.2 million students enrolled in private high schools (grades 9 to 12) as compared with 1.4 million students in the 1970 Census. It has been suggested that the opening of private academies in some States after the school year began may have increased private school enrollment between the October 1969 CPS and the April 1970 Census. However, the CPS data for October 1970 show still fewer students enrolled in private high schools than in October 1969, as the decline in private school enrollment which began in 1966 apparently continued. It is possible that the Census residence rule which counted students living away from home at a private boarding school while attending school as a part of their parental household may have led to some double counting of these students, if the instruction not to count these students at their schools was not completely adhered to.

The Census data on enrollment in parochial high schools are roughly comparable with the data in the Official Catholic Directory on enrollment in Catholic high schools in the fall of 1969. The Census shows 1.0 million students enrolled in parochial high schools and the Official Catholic Directory shows 1.1 million enrolled in Catholic high schools. The Catholic high school enrollment was comprised of 667,000 students enrolled in schools under the control of the parish or diocese and 387,000 enrolled in schools under private Catholic control. Except for attrition between semesters, one would expect more students to be enrolled in parochial schools in the Census than just the number enrolled in Catholic schools, since the Census parochial definition (which was read by enumerators, but not respondents) included any private schools controlled or supported primarily by a religious organization. The New Haven pretest of the 1970 Census questions indicated that children enrolled in Catholic schools, whether under parish or private control, were being classified as enrolled in parochial schools. The comparability of Census data with data from the Official Catholic Directory appears to indicate that, in general, they were so classified in the Census. One surprising finding is the relatively low proportion of private high school enrollees in the Census who were in parochial (church-controlled or supported) schools when compared with the Office of Education data for the fall of 1968. In the Census, only 70 percent of the students enrolled in private high schools were in parochial schools as compared with 88 percent of the students included in the Office of Education survey in 1968. The Office of Education proportion might be high, since they obtained more complete coverage of Catholic and other church-related schools than of secular private high schools. On the other hand, the category "parochial" may not have been reported for some students enrolled in private Catholic schools, that is, those not under the control of the parish; moreover, the category might not have been reported for students enrolled in schools supported by non-Catholic religions who may have thought that "parochial" referred only to Catholic schools. In future censuses and surveys this category should be fully explained in the instructions provided to the respondents and, perhaps, the wording of the category should be changed to read "parochial and other church-related schools" or just "church-related schools".

At the college level, there were fewer students enrolled in public colleges and more students enrolled in private colleges in the Census than in the CPS. In public colleges, there were 813,000 fewer students enrolled in the Census than in the CPS--a 15 percent difference. (One standard error on the CPS estimate is around 70,000 persons). Assuming some attrition, this may not be an unreasonable change. But in private colleges, there were 345,000 more students enrolled in the Census than in the CPS--a 17 percent increase. (One standard error on the CPS estimate is around 45,000). It was suggested in the 1960 Census Monograph, Education of the American Population, by Folger and Nam, that the students in college--a larger proportion of whom report for themselves in the Census than in the CPS--may make a better determination of the control of their colleges in the Census than their parents do in the CPS. This explanation was buttressed with a comparison of Census, CPS, and Office of Education data which showed a similar percentage of college students enrolled in private colleges in the Census and Office of Education data, but a smaller percentage in the CPS data. However, the 1970 Census, CPS, and Office of Education comparison does not support continuation of this reasoning. The CPS and OE data show approximately the same proportion of students in private colleges, around 27 percent, whereas the Census shows 34 percent in private colleges. The data in the 1970 Census and 1960 Census are consistent in that both showed a larger proportion of college students in private colleges than did the preceding CPS. Apparently, self-reporting of control of school varies significantly.

Conclusion

The 1970 Census data on school enrollment appear to be valid and reliable. The major unaccountable differences are the larger number of students enrolled in private high schools and colleges in the Census than in the CPS, and the relatively low proportion of private high school students enrolled in parochial schools in the Census when compared with Office of Education data. It may be that the Census data on parochial school enrollment in the Census relate primarily to Catholic school enrollment and that students in other church-related schools are included in the "other private" category. The CPS data and OE data also are reasonably comparable with the major difference being the lower number of part-time undergraduate students in the CPS. Because of different collection methods and dates, the Census data, the CPS data, and the Office of Education data should not be facsimiles of each other, but should approximate each other, and they appear to do this reasonably well.

Table 1.--SCHOOL ENROLLMENT OF PERSONS 3 TO 34 YEARS OLD BY LEVEL OF SCHOOL IN THE 1970 CENSUS AND THE OCTOBER 1969 CURRENT POPULATION SURVEY

(Numbers in thousands. CPS data for civilian noninstitutional population)

Level and control of school	Census April 1970	CPS October 1969	Difference	
			Number	Percent
Total, 3 to 34 years old	107,983	105,145	2,838	2.7
Enrolled in school	58,635	59,913	-1,278	-2.1
Nursery school	954	860	94	10.9
Kindergarten	3,024	3,276	-252	-7.7
Elementary school (1 to 8 years)	33,210	33,788	-578	-1.7
High school (1 to 4 years)	14,481	14,553	-72	0.5
College	6,966	7,435	-469	-6.3
Not enrolled in school	49,348	45,232	4,116	9.1

Note: Minus sign denotes a larger CPS figure.

Table 2.--COLLEGE ENROLLMENT BY LEVEL IN CURRENT POPULATION SURVEY AND IN THE OFFICE OF EDUCATION SURVEYS: 1967 TO 1969

(Numbers in thousands. CPS data for civilian noninstitutional population)

College enrollment	CPS October 1/	Office of Education 2/	Difference	
			Number	Percent
Total, undergraduate enrollment				
1967......................	5,437	5,653	-216	-4.0
1968......................	5,858	6,101	-243	-4.1
1969......................	6,296	6,656	360	5.7
1970......................	6,273	7,020	-747	-11.9
Total, graduate enrollment 2/				
1967......................	963	753	210	21.8
1968......................	944	797	147	15.6
1969......................	1,138	828	311	27.3
1970......................	1,140	900	240	21.1

1/ Graduate enrollment is defined as enrolled in the 5th year of college or higher.

2/ Estimated resident graduate enrollment. About 130,000 in 1969 and 1970 were extension graduate students, Projections of Educational Statistics to 1980-81, U.S.O.E., National Center for Education Statistics, 1971 edition, table 18.

Note: Minus sign denotes larger OE figure.

Table 3.--COLLEGE ENROLLMENT IN THE CURRENT POPULATION SURVEY AND OFFICE OF EDUCATION SURVEY: 1960 TO 1970

(Numbers in thousands)

	College enrollment			Full-time college enrollment		
	OE 1/	Census	Census - OE	OE 1/	Census	Census - OE
1960.......	3,583	3,570	-13	2,466	2,681	215
1961.......	3,861	3,731	-130	2,714	2,902	188
1962.......	4,175	4,208	33	2,902	3,237	335
1963.......	4,495	4,336	-159	3,068	3,260	192
1964.......	4,950	4,643	-307	3,418	3,556	138
1965.......	5,526	5,675	149	3,910	4,414	504
1966.......	5,928	6,085	157	4,225	4,847	622
1967.......	6,406	6,401	-5	4,556	4,976	420
1968.......	6,928	6,801	-127	4,937	5,357	420
1969.......	7,484	7,435	-49	5,254	5,810	556
1970.......	7,920	7,413	-507	5,489	5,763	274

1/ Fall degree-credit enrollment.

Table 4.--**GRADUATE AND UNDERGRADUATE COLLEGE ENROLLMENT IN THE CURRENT POPULATION SURVEY AND OFFICE OF EDUCATION SURVEYS:** 1967 TO 1969

(Numbers in thousands. CPS data for civilian noninstitutional population under 35)

Year and full-time part-time attendance	CPS October	Office of Education	Difference	
			Number	Percent
Undergraduate				
Part-time				
1967	952	1,414	-462	-48.5
1968	994	1,496	-502	-50.5
1969	1,077	1,766	-492	-64.0
1970	1,066	1,910	-844	-79.2
Full-time				
1967	4,485	4,239	246	5.5
1968	4,864	4,606	258	5.3
1969	5,219	4,890	329	6.3
1970	5,207	5,110	97	1.9
Graduate				
Part-time				
1967	472	436	36	7.6
1968	451	455	-4	-0.9
1969	545	464	81	14.9
1970	585	521	64	10.9
Full-time				
1967	491	317	174	35.4
1968	493	342	151	30.6
1969	593	364	229	38.6
1970	555	379	176	31.7

See notes on table 2.

Table 5.--SCHOOL ENROLLMENT OF PERSONS 3 TO 34 YEARS OLD BY LEVEL AND CONTROL IN THE 1970 CENSUS AND THE OCTOBER 1969 CURRENT POPULATION SURVEY

(Numbers in thousands. CPS data for civilian noninstitutional population)

Level and control of school	Census April 1970	CPS October 1969	Percent	
			Census	CPS
Nursery school.........................			100.0	100.0
Public.............................	303	245	31.8	28.5
Private............................	650	615	68.1	71.5
Kindergarten..........................			100.0	100.0
Public.............................	2,544	2,682	84.1	81.9
Private............................	480	594	15.9	18.1
Elementary school (1 to 8 years).......			100.0	100.0
Public.............................	29,375	29,825	88.5	88.3
Private............................	3,835	3,964	11.5	11.7
High school (1 to 4 years).............			100.0	100.0
Public.............................	13,063	13,400	90.2	92.1
Private............................	1,417	1,153	9.8	7.9
College...............................			100.0	100.0
Public.............................	4,626	5,439	66.4	73.2
Private............................	2,340	1,995	33.6	26.8

Note: Minus sign denotes a larger CPS figure.

Table 6.--PERCENT OF COLLEGE STUDENTS IN PUBLIC AND PRIVATE COLLEGES IN THE 1960 AND 1970 CENSUS, THE OCTOBER 1959 AND 1969 CPS, AND IN THE OFFICE OF EDUCATION FALL 1959 AND 1969 SURVEYS

Control of College	Census April 1970	CPS October 1969	Office of Education Fall 1969 1/	Census April 1960	CPS October 1959	Office of Education Fall 1959 1/
Total enrolled.....	100.0	100.0	100.0	100.0	100.0	100.0
Public.................	66.4	73.2	72.4	58.9	63.5	58.9
Private................	33.6	26.8	27.6	41.1	36.5	41.4

1/ Degree-credit enrollment.

THE DURATION OF UNEMPLOYMENT

Hyman B. Kaitz, Bureau of Labor Statistics

Unemployment is a complex form of human behavior, and there is probably no single analysis which attempts to cover all of its facets. One aspect of this behavior has had little understanding in the press and in public discussions, but has been exposed to an increasing amount of study among economists in connection particularly with job search theory, and that is the phenomenon of turnover among the unemployed and unemployment duration. At the most primitive levels of belief, the fact that unemployment has been at a level of 5 million people for well over a year is translated into a visual image of 5 million people each with over a year of unemployment. It is not known generally--and we share the blame for inadequate communication on this score--that a level of unemployment of 5 million for a year corresponds to at least one week of unemployment experience for perhaps 20 million people in 30 million or more spells, and that the vast majority of these spells--three-quarters or more--are less than five weeks in total duration. Only a small proportion of the people who experience unemployment have extended periods of joblessness. The public policy issues are obviously quite different with respect to the short-term unemployed as compared with the long-term unemployed. However, the nature of turnover among the unemployed is not explicitly given in the data customarily published. Published data present the cross section distribution of the unemployed, i.e., the distribution of the unemployed by the duration of their unemployment through the end of the reference week each month (the week including the 12th). Only a small fraction of these people are actually completing spells of unemployment in the reference week; the remainder will go on to experience additional week(s) of unemployment. Because reference weeks are four or five weeks apart, the intervening movements into and out of unemployment are not evident, and the full nature of unemployment turnover has remained somewhat obscure. A first attempt to deduce the patterns of completed spells of unemployment and of turnover is presented in an article in the Monthly Labor Review for November 1970.

The present paper continues this work by developing a time series of completed spells of unemployment. The details of the estimating procedure are given in the first part below. In the second part certain shortcomings of the under lying data are examined, and a method for eliminating them to a substantial extent is developed. Since the techniques used here are somewhat different from those presented in the earlier paper, some comparisons between the results in the two papers are made, and the validity of certain assumptions in the earlier article is examined. Finally, a time series of completed spells and their average duration from mid-1967 to mid-1972 is presented.

Derivation of Monthly Series on Completed Spells

Let g_1 = number of people who report they have been unemployed for 1 week through the end of the reference week.

Let N = the number of people completing spells of unemployment in the week prior to the reference week.

Let F = the total number of people unemployed in the week prior to the reference week.

Finally, let G = the total number of people unemployed in the reference week.

It may be readily shown that $N = F - G + g_1$. 1/

In words rather than symbols, the number of completed spells in the week prior to the reference week is equal to the number of unemployed in that week minus the number of unemployed in the following week (the reference week) plus the number of unemployed who had completed only one week of unemployment in the reference week).

G and g_1 are available from the regular current population survey data. (There is a basic problem in the measurement of g_1 which must be faced; this is the burden of the next section of this paper.)

F is not available from the current population survey since it refers to a nonsurvey week. However, it can be estimated by interpolation as described a little later. It must be emphasized that the calculation of the number and duration of completed spells requires knowledge of the indicated statistics in adjacent weeks.

In order to estimate the average duration of completed spells in the week prior to the reference week, we need several additional symbols.

Let A_f = the cumulated number of weeks of unemployment of those unemployed in the week prior to the reference week.

Let A_g = the corresponding quantity for those unemployed in the reference week.

Let A_s = the cumulative number of weeks of unemployment of those completing their spells of unemployment in the week prior to the reference week.

If we let $\bar{x}_f$ and $\bar{x}_g$ represent the cross section average durations of unemployment in the week prior to the reference week and in the reference week respectively, and $\bar{x}_s$ represent the average duration of unemployment of those completing spells in the week prior to the reference week, we have the following relationships by definition:

$$A_f = \bar{x}_f F; \quad A_g = \bar{x}_g G, \quad \text{and} \quad A_s = \bar{x}_s N$$

It can be readily shown that

$$A_s = A_f - A_g + G. \quad 2/$$

A_g is obtained directly from current population survey data for the reference week. As will be indicated below A_f is estimated by interpolation as is F.

The average duration of completed spells in the week prior to the reference week can therefore be written as:

$$\bar{x}_s = (A_f - A_g + G) / (F - G + g_1).$$

Under conditions of stability or equilibrium, we will have $A_f = A_g$, and $F = G$, so that the expression for average duration simplifies to

$\bar{x}_s = G / g_1$, a result previously given in the cited article in the Monthly Labor Review.

Even under conditions of instability, if we let G be the average annual total unemployment, g_1 be the average annual number of unemployed with one week of unemployment (entrants), the average duration of all spells completed in the year is given by

$$\bar{x}_s = \frac{(A_1 - A_{53}) / 52 + G}{(G_1 - G_{53}) / 52 + g_1}$$

with the subscripts in the parentheses referring to the first and fifty-third weeks. Since A is about 10 times the size of G, and G is about 6 times the size of g_1, and the differences, $A_1 - A_{53}$, and $G_1 - G_{53}$ are, in most cases, fractions of the sizes of A and G respectively, the nature of the approximation in the formula $\bar{x}_s = G/g_1$ is pretty good on an average annual basis, although it is by no means perfect.

The exposition in the preceding paragraph is intended to serve as a rationale for the derivation of completed spell distributions from average annual cross section data in the Monthly Labor Review article. It may also be shown with the same reasoning that the derivation of the distribution of completed spells by intervals of weeks is also generally a reasonable first approximation.

The present paper does not go on however to estimate a time series for the distribution of completed spells by duration intervals for reasons which are briefly discussed in the Monthly Labor Review article, and may profitably be repeated here. The recorded cross section distributions of the unemployed by single weeks of duration are subject to a number of irregularities, most prominently the rounding effect in the process of recall on the respondent's part, which introduces local modes into the data at 4, 8, 13, 26, etc. weeks of duration, because of the tendency to report unemployment in terms of months, quarters, half years and so on. Another irregularity, which may not superficially appear to be one is the local mode at a duration of two weeks. This irregularity is of fundamental importance to this paper and is discussed in more detail in the next section. Until these irregularities can be smoothed in an appropriate way, the derivation of the time series of the distribution of completed spells by duration intervals must be held in abeyance.

One other point will conclude this section. As mentioned earlier, it is necessary to estimate A_f and F, since they do not apply to the reference week. The simplest thing to do is to use linear interpolation for F between the two G values in the adjacent month reference weeks. In the same way linear interpolation between adjacent month reference week values of A_g will yield A_f approximations.

A Bias in the Duration Distribution of Unemployment and its Correction

As noted earlier, a major problem in the derivation of the duration distribution of completed spells is in the biased reporting of those unemployed in the reference week who indicate that they have had one, two or three elapsed weeks of unemployment. The bias is evident in the data below for the average annual cross section duration distribution for 1971, but it appears in the data for almost all individual reference weeks, and in all years.

Duration in weeks	Number of Unemployed
0	12
1	445
2	656
3	532
4	583
5	177

The 12 (000) people who have zero weeks of unemployment need not concern us particularly. Conceptually they can only be people who were out of the labor force for more than half of the reference week before they began looking for a job. For convenience they are included with the number identified with one week of unemployment.

On the other hand, it is conceptually and empirically impossible for those unemployed with two weeks of unemployment to regularly exceed, and usually by a substantial amount, those with one week. This is the phenomenon studied in this section of the paper.

Finally, the local mode at 4 weeks is obviously due to rounding by the respondent or the interviewer (4 weeks equals one month) so that suitable smoothing techniques must be developed to remove this mode. Fortunately this effort is not required for this paper, so it is not considered further here.

A hypothesis for the aberrant pattern among duration groups of 1, 2 and 3 weeks is explored below. In order to develop a reasonable basis for this hypothesis, we must first look at the actual questionnaire used in the Current Population Survey.

The week including the 12th of the month has been called the reference week in the discussion thus far. The following week (which includes the 19th of the month) is the survey week. On each working day of that week the interviewer visits households in his or her part of the sample. On the CPS questionnaire, questions 19, 20 and 21 ask about activities in the preceding week with emphasis on the phrase "LAST WEEK" (the capitals are on the questionnaire form itself). The appropriate responses to questions

20, 21, and 22 classify a person as unemployed during the reference week.

However, the phrasing in question 22C may be subject to some ambiguity in response. It is as follows:

22C. 1) How many weeks has . . . . been looking for work?
2) How many weeks ago did . . . . start looking for work?
3) How many weeks ago was . . . . laid off?

These three questions are alternate versions to be used as appropriate, depending on replies to earlier questions.

The hypothesis advanced here is as follows: In answering question 22C, the respondent calculates the length of time he (or she) has been looking for work up to the time of the interview, and not up to the end of the preceding week (the reference week). Since answers are recorded in whole weeks (rounding is called for), it is likely that persons in households which are interviewed in the second half of the survey week tend to round their answers up one week. For example, suppose the respondent has only been unemployed in the reference week (one week duration), but is still unemployed in the survey week. The interviewer visits his household on Thursday of that week and asks: "How many weeks have you been looking for work?", and he replies: "About two", which is an appropriate reply from his perspective. Now the interviewer's instruction manual says: "In computing the weeks a person has been looking for work, count the number of weeks from the time he started looking for work through the end of the reference* week for the current month." (Emphasis supplied in original) However, this is the instruction to the interviewer alone. Question 22C does not in its own phrasing pin down the actual duration through the end of the reference week. The hypothesis offered here is that respondents in fact are likely to use the date of the interview as the end of the elapsed duration period.

The balance of this section examines the effect on the data if this hypothesis is true and estimates how this bias may be removed.

Let the "true" number of those unemployed people with an elapsed duration of 1 week through the end of the reference period be denoted by g_1, those with 2 weeks by g_2 and those with 3 weeks by g_3. Let the corresponding measured quantities be denoted by the same symbols with primes.

Let the symbol f with appropriate subscripts refer to the corresponding values for the week prior to the reference week and the symbol h apply to the week after the reference week.

Now let r be the fraction of those with one elapsed week of unemployment to the end of the reference period, who are still unemployed in the survey week and who are interviewed in the latter part of that week and consequently report themselves with two elapsed weeks of unemployment. Let the same fraction hold for those with two elapsed weeks who report themselves with three weeks (and for those with three elapsed weeks who report four). We will therefore find the following relationships between the true and measured quantities:

$$f_1' = f_1 - r g_2$$

$$f_2' = f_2 + r g_2 - r g_3$$

$$f_3' = f_3 + r g_3 - r g_4$$

Of the f_1 people who experience one week of unemployment by the end of the reference week, g_2 go on to experience a second week of unemployment. When these g_2 people are visited in their households, $r g_2$ of them report themselves with two weeks of unemployment up to the end of the reference week (the preceding week). The reasoning is the same for the other groups to whom the second and third equations above refer.

In the same way relationships for the following week may be specified:

$$g_1' = g_1 - r h_2$$

$$g_2' = g_2 + r h_2 - r h_3$$

$$g_3' = g_3 + r h_3 - r h_4$$

We have assumed that the fraction of people (r) who round their length of unemployment up to the next week is the same in the two adjacent weeks as it is for the first few duration categories.

The relationships which follow from this assumption are still indeterminate, because there are more unknowns than relationships. We need therefore to introduce an additional assumption in order to permit a solution. This assumption is that the continuation (survival) rate is also constant over the first several weeks of duration, and in the two adjacent weeks.

The continuation rate is defines as $p_{i+1} = g_{i+1}/f_i$, i.e., the proportion of people unemployed i weeks who go on to have at least one additional week of unemployment. p has a subscript to identify its position in the duration scale, and is not, as a rule, constant over the entire duration (see MLR article, op. cit.) but it does not usually change very much within a duration range of two or three weeks. The assumption used here is that, within our area of consideration, the continuation rate is constant.

In other words,

$$p = g_2/f_1 = g_3/f_2 = g_4/f_3 = h_2/g_1 = h_3/g_2 = h_4/g_3$$

* The manual uses the word "survey" here to designate what I have called the "reference" week.

An earlier equation may be rewritten as

$$g'_3 = pf_2 + rp\, g_2 - rp\, g_3$$

We also have the earlier relationship

$$f'_2 = f_2 + r\, g_2 - r\, g_3 \,.$$

If we divide the first of these by the second we get $g'_3 / f_2 = p$, a direct estimate of the continuation rate.

Another of the earlier relationships may be rewritten as

$$g'_2 = p\, (f_1 - rp\, f_1 + r\, g_1)$$

Two other equations may be expressed as follows:

$$f'_1 = f_1\, (1 - rp)\ ;\ g'_1 = g_1\, (1-rp) \quad \text{or}$$

$$f'_1 / f_1 = 1 - rp = g'_1/g_1, \text{ whence } f'_1/g'_1 = f_1/g_1$$

In other words, the ratio of the observed number of people with one week of unemployment in two adjacent weeks is equal to the ratio of the "true" values of these numbers. Using this result in the equation for g'_2, we arrive at a solution for r:

$$r = \frac{(g'_2/f'_1)/p - 1}{g'_2/f'_1 + g'_1/f'_1 - p}$$

The observed number of people entering the unemployed in the reference week can be corrected to the "true" number by use of

$$g_1 = g'_1 / (1 - rp)$$

For 1969 annual average data, letting $f'_1 = g'_1$, and estimating p by use of g'_3 / g'_2, we have $p = 0.733$, and $r = 0.459$. The estimated continuation rate of 0.73 for low duration values in 1969 compares with estimates in the .70 - .80 range for 1969 estimated by a different approach (MLR article, p. 13, Table 4). The r value of 0.46 is consistent with the rather simple and plausible notion that about half the g_2 people round their responses on duration to one week, and the other half round to two weeks.

As indicated in the last formula, we now have a way of estimating g_1 values month by month, which are essential to the calculation of completed spell statistics.

Because the monthly series on completed spells constitutes only between 15 and 20 percent of the total level of unemployment, it is subject to much greater irregularity. Consequently the derived time series shown in the table below is given in terms of quarterly averages only.

Time Series of Completed Spells of Unemployment
(Weekly Averages Per Quarter)

Year-Quarter	Spells Completed (000)	Spells Begun (000)	Average Duration (in weeks) of Completed Spells
1967 - III	746	686	4.4
IV	511	492	5.3
1968 - I	480	488	6.0
II	559	616	6.0
III	703	627	4.6
IV	504	487	4.8
1969 - I	493	513	5.2
II	665	724	4.5
III	709	678	4.2
IV	515	489	5.3
1970 - I	500	580	5.3
II	685	766	5.8
III	871	843	5.0
IV	705	734	5.0
1971 - I	711	747	6.0
II	761	769	7.3
III	796	745	6.6
IV	733	713	6.2
1972 - I	660	694	6.2
II	735	735	8.1

At this point it may be useful to compare the average annual data on completed spells estimated directly from the average annual cross section duration distributions of the unemployed (A) with the average annual data derived from the twelve monthly observations in each year (M). The table below presents these comparisons.

Year	Type of Estimate	Average Weekly No. of Completed Spells (000)	Average Duration of Completed Spells
1968	M	562	5.3
	A	554	5.1
1969	M	596	4.7
	A	601	4.7
1970	M	690	5.3
	A	734	5.6
1971	M	750	6.6
	A	742	6.7

The difference between the results obtained from time series versus those obtained from cross section data are the greatest for the year 1970, when the seasonally adjusted rate of unemployment rose from 3.9 percent in January to 6.1 percent in December. Even here, the estimates based on cross section data may be considered to be reasonable approximations of the estimates based on time series. In the other three years the two estimates for both number of completed spells and average duration are very close to each other and the approximation is excellent.

Two additional comments about these estimates may be made in closing:

1. The average duration of completed spell is substantially and uniformily lower than the average duration of cross section data regularly published. The table below compares these two averages in annual average form.

Average Duration of Unemployment

Year	Cross Section	Completed Spell
1968	8.5	5.3
1969	8.0	4.7
1970	8.8	5.3
1971	11.4	6.6

A discussion of the reasons for these differences is found in the MLR article, op. cit., and will not be repeated here.

2. Average weekly additions to unemployment for new entrants as a percent of the average level of unemployment for these four years is as follows:

Year	New Entrants/Total Unemployment
1968	.20
1969	.21
1970	.18
1971	.15

No extensive analysis of this phenomenon will be undertaken here, but it is interesting to note that the rise in unemployment in a recession partakes of two elements:

1) a rise in the number of entrants to unemployment and
2) a decline in the rate of mobility into and out of the ranks of the unemployed.

Again, there is some discussion of this phenomenon in the reference cited, but a full study of it remains to be made.

Footnotes

1. Let C = number of unemployed with unemployment in both the reference week and the prior week.

   Then $F = C + N$
   $G = C + g_1$

   By subtraction we get $N = F - G + g_1$

2. Let A_c = cumulative number of manweeks of unemployment of the C people through the end of the week prior to the reference week. In the following week they will have $A_c + C$ weeks.

   Then $A_f = A_c + A_s$; $A_g = A_c + C + g_1$,
   but $G = C + g_1$ so $A_g = A_c + G$.
   By subtracting this last equation from the first, we get the desired result:

   $$A_s = A_f - A_g + G.$$

Reference

Kaitz, Hyman B. - "Analyzing the length of Spells of Unemployment" Monthly Labor Review, November 1970, pp. 11-20.

APPLICATION OF LIFE TABLE TECHNIQUES TO CRIMINAL STATISTICS: PAROLE VIOLATION

Nathan Kantrowitz, Kent State University

I wish to outline a way to study criminal statistics which involve follow-up over a period of time. Essentially, it is a simple application of ideas from life tables which restricts itself to an analogy with the "probability of death", or q_x column. We shall call this analogy the "probability of parole violation", or for brevity, "PV_x". My purpose is to propose what I hope may be a useful idea, illustrated with the data I have available, of all 1,325 men (or more exactly, adult male felons) paroled from the prisons of Illinois during 1960. The need to present several applications and also to be brief unfortunately makes for some superficiality and sets some limits. Although the PV_x idea can be applied to other populations such as probationers, or even men entering prison, we shall discuss only parolees. Also, our definitions of failure, Parole Violation Warrants, are complex administrative documents which can be equated with only rough justice to criminality. Technical ommissions are that life table functions other than q_x will not be discussed, and our time periods will be limited to unsmoothed monthly rates.

We shall proceed by considering first, usual measures of parole success or failure; second, the idea of the PV_x; third, the relationship of PV_x to social indicators such as unemployment; and fourth, a suggestion of how PV_x may be of use in evaluating some kinds of prison programs.

I. THE MEASURE OF PAROLE VIOLATION

This paper claims no great originality, for measures equivalent in concept to PV_x do exist.[1] However, the usual measure of parole success or failure is the percentage of parolees who are violators; that is, the number of persons initially released on parole divided into the number of those who violated during some time interval. Thus, of our 1,325 parolees in 1960, 508 or 38.3 percent violated parole. Of course our parolees could also be classified by various attributes; for example, by race, 353, or 41.8 percent of the 844 white parolees violated parole, as did 155, or 32.2 percent of the 481 Negro parolees. This oversimplifies existing scholarship, but the percent violators is by far the accepted measure.

An example of the manner in which the number of months actually on parole constitutes a problem can be seen in our own sample. Here, because the cut-off date for follow-up was July 1962, men paroled in January 1960 were followed for 24 months, while those paroled in December 1960 were followed for 18 months. The problem is compounded by the fact that some men were legally freed from parole during those time periods. We think PV_x allows for attrition of the original sample and also is useful in dealing with the problems of time in follow up. Breaking time periods into small units (as for example, the month) permits more equal measurement of time in follow-up. Moreover, as a practical matter, it may reduce the waiting period for results from years to months, no small matter to administrators who support research.

II. "PROBABILITY OF PAROLE VIOLATION"

The analogy to q_x is spelled out in Table 1 for the 100 men paroled during January 1960. Column 1 shows the monthly time periods used. Column 2, analagous to l_x, begins with the 100 men initially released, and decreases this number in each succeeding month by the numbers in Column 3, 4, 5. These columns show the number of parolees who, during each month, were legally freed from parole supervision (Column 3), died of natural causes while on parole (Column 4), or who were considered parole violators (Column 5). I have shown all three columns -- especially Column 4 which has no entries -- to illustrate the possibility of creating multiple decrement tables. For the present however, we shall focus on Column 5, the parole violators; when we divide the numbers in Column 5 by those of Column 2, the result in Column 6 is the number of parolees who violated parole during each month as a percent of those who began the month. This is our "probability of parole violation", or PV_x.

In another illustration of PV_x, Table 2 considers the entire year's parolees, according to each of the 18 months of supervision; that is, a January parolee who violated parole in two weeks and a December parolee who violated parole in three weeks would both be included in the 23 parole violators of Column 3, line 1. Column 4, the PV_x for all parolees illustrates one of several kinds of curves which may be characteristics of the PV_x statistic: in this case, a biomodal distribution with peaks around the third and twelfth months. Other different PV_x patterns are illustrated for whites and Negroes in Columns 7 and 10, and graphed in Figure 1.

III. THE PV_x IN RELATION TO UNEMPLOYMENT

If we wish to consider whether such social indicators as unemployment cause crime, and we wish to go beyond cross-sectional, survey data, the most precise procedure is to repeatedly interview a cohort of men over a period of time, something rarely attempted because it is difficult and expensive. A less precise, and cheaper, indirect way is to correlate unemployment rates with crime statistics. PV_x is a useful supplement to indirect correlation for two reasons: first, it is a more precisely defined rate; and second, it uses a shorter time period.

In the first case, precision of rate, crime statistics typically consist of the annual number of crimes reported by the police, divided by an estimated mid-year population of whatever geographical area is being investigated. The numerator of official crime statistics is grievously deficient; for example, not all crimes are reported. In addition, the denominator for these rates, the population at risk, is a usually imprecise mid-year population estimate for a State, (or city or metropolis). The PV_x is much more precisely defined.

In addition to this first problem, precision of definition as a rate, the PV_x makes it possible to compare monthly rates rather than annual ones. This is illustrated in Table 3 which shows the probability of parole violation during specific months in the follow-up period, classified by the month of initial release during 1960. Column 1, for January (that is, men released on parole during January 1960), is the PV_x already illustrated in Column 6 of Table 1; we have graphed these PV_x in Figure 2. We can see in Figure 2 that there is no simple way to summarize the patterns traced out by each of these succeeding month's cohorts - although using moving averages might help. Our immediate concern however is the relationship of PV_x to unemployment as shown in the graphs of Figures 2 and 3.

Figure 3 illustrates the pattern for unemployment, just one of many possible social indicators, in Illinois and the Chicago area by month during 1960-1962. The evidence of Figures 2 and 3 suggest the hypothesis that there is no correlation between unemployment and parole violation. For example, in Figure 3, the volume of unemployment steadily rises for six months between September 1960 and February 1961; and thereafter falls through October 1961. Yet neither of these six and eight month periods is reflected in the PV_x of Figure 2. For example, if we consider the men released in September or October 1960, we can see that these two month's patterns differ both from each other as well as from Figure 3.

IV. THE EVALUATION OF PRISON PROGRAMS

The way in which the PV_x might be a useful supplement to other means of evaluating correctional programs becomes apparent if we consider an oversimplified example of a research study. Here, a sample of men released during the calendar year are assigned to experimental and control groups and then are followed for another calendar year or longer; this sample is cross-classified by variables such as age, race, criminal history, or personality. Then the parole violation rate is calculated after the end of a year; that is, the number of violators is divided by the number initially released.

The end result may be interesting, but it is also a long time in coming, often two or three years. Some of the possible usefulness of PV_x may be illustrated if we turn back to Figure 1. But now instead of its actual tabulation by race, let us imagine that we have graphed the PV_x of experimental and control groups of parolees released during some month.[2] Suppose further that the broken (that is, Negro) line was the experimental group, while the solid (that is, white) line was the control group; in this case one could see that there was a great impact by the treatment (however defined) used in the experimental group in the early months of parole, diminishing after the first half-year. On the other hand, if we reversed our imaginary study, so the experimental group was the solid line, and the control group was the broken one, there would be an immediate feedback that the treatment accorded the experimental group was disastrous. Obviously, one does not expect an experimental group to show such extreme results, but it may not be difficult to work out a program of sequential analysis in which an experiment could be terminated if the experimental group did not consistently show some benefit, month by month.

CONCLUSION

In this brief presentation, I have been able only to sketch an approach which may (or may not) be a useful addition to ways of studying parole and similar programs. The PV_x is no replacement for other statistics, and has its own limitations. For example, among its disadvantages are the relatively large numbers needed during a short time period for the initial cohort. On the other hand, it appears to have some advantages; for example, it may enable us to create a finer, month-by-month evaluation of follow-up. Fortunately, a closer look at its usefulness should require little expense. All that is necessary is the secondary analysis of existing studies of parole violation or equivalent populations with three issues in mind. First, as in our own Figure 1 by race, do different groups show different characteristic curves; if they do, how is this related to the erraticness of Figure 2. Second, what is the correlation with monthly social indicators such as unemployment. Third, is the re-analysis of experimental and control groups. These secondary analyses should indicate whether we have an interesting but useless idea, or whether there is some practical wisdom in considering it further.

FOOTNOTES

*The data for 1960 Illinois State Peninteniary parolees were originally gathered to update Parole Prediction Tables; however, all statements made are the responsibility of the author. Illinois unemployment statistics were provided by Virginia Peyton, Chicago Association of Commerce and Industry.

[1]For example: Joan Havel and Elaine Sulka, Special Intensive Parole Unit, Phase III, Research Report No. 3, Research Division, California Department of Corrections (1962), Figure 2; Nathan Kantrowitz, "Joliet-Menard 1960 Parolees", Illinois Department of Public Safety, Bulletin of the Sociologist-Actuary, Number 2 (April 5, 1963), Table 7; Frances H. Simon, Prediction Methods in Criminology, Home Office Research Studies Number 7, H.M.S.O. London (1971), Appendix I.

[2]Havel and Sulka illustrate what appears to be this approach.

Table 1. Probability of Parole Violation For Men Released in January 1960 from Illinois Prisons

(1) Number of Months Actually Under Parole Supervision in the Free Community		(2) Total Number of Parolees Remaining Under Supervision at the beginning of the Month	(3) Number of Parolees, who During this month: were legally freed from Parole	(4) died while on Parole	(5) violated Parole	(6) Number of Parolees who violated Parole during a month as a Percent of those who began the month (Col 5/Col 2)
at least	but less than					
0	1	100	0	0	2	2.0
1	2	98	0	0	1	1.0
2	3	97	1	0	7	7.2
3	4	89	0	0	1	1.1
4	5	88	0	0	5	5.7
5	6	83	0	0	4	4.8
6	7	79	3	0	2	2.5
7	8	74	0	0	3	4.1
8	9	71	1	0	3	4.2
9	10	67	1	0	1	1.5
10	11	65	1	0	2	3.1
11	12	62	0	0	3	4.8
12	13	59	0	0	1	1.7
13	14	58	2	0	2	3.4
14	15	54	3	0	2	3.7
15	16	49	0	0	1	2.0
16	17	48	3	0	1	2.1
17	18	44	5	0	1	2.3
18	29	38	10	0	4	-
30 and over		24	-	-	-	-

Table 2. Probability of Parole Violation (per 100) By Race, Illinois Parolees, 1960

(1) Midpoint of Number of Months Successful on Parole	(2) All Parolees: Total who Began the Month	(3) Number who violated During the Month	(4) PV (3)/(2)	(5) White: Total who Began the Month	(6) Number who violated During the Month	(7) PV (6)/(5)	(8) Negro: Total who Began the Month	(9) Number who violated During the Month	(10) PV (9)/(8)
0.5	1325	23	1.7	844	21	2.5	481	2	0.4
1.5	1302	34	2.6	823	22	2.7	479	12	2.5
2.5	1267	54	4.3	801	40	5.0	466	14	3.0
3.5	1209	48	4.0	759	34	4.5	450	14	3.1
4.5	1155	49	4.2	722	38	5.3	433	11	2.5
5.5	1100	40	3.6	682	30	4.4	418	10	2.4
6.5	1045	33	3.2	644	22	3.4	401	11	2.7
7.5	991	30	3.0	615	18	2.9	376	12	3.2
8.5	930	27	2.9	585	22	3.8	345	5	1.4
9.5	867	20	2.3	538	11	2.0	329	9	2.7
10.5	835	23	2.8	522	19	3.6	313	4	1.3
11.5	786	28	3.6	486	18	3.7	300	10	3.3
12.5	738	19	2.6	456	9	2.0	282	10	3.5
13.5	700	15	2.1	440	9	2.0	260	6	2.3
14.5	655	12	1.8	410	7	1.7	245	5	2.0
15.5	614	9	1.5	389	4	1.0	225	5	2.2
16.5	582	8	1.4	370	3	0.8	212	5	2.3
17.5	543	5	0.9	344	3	0.9	199	2	1.0

Table 3

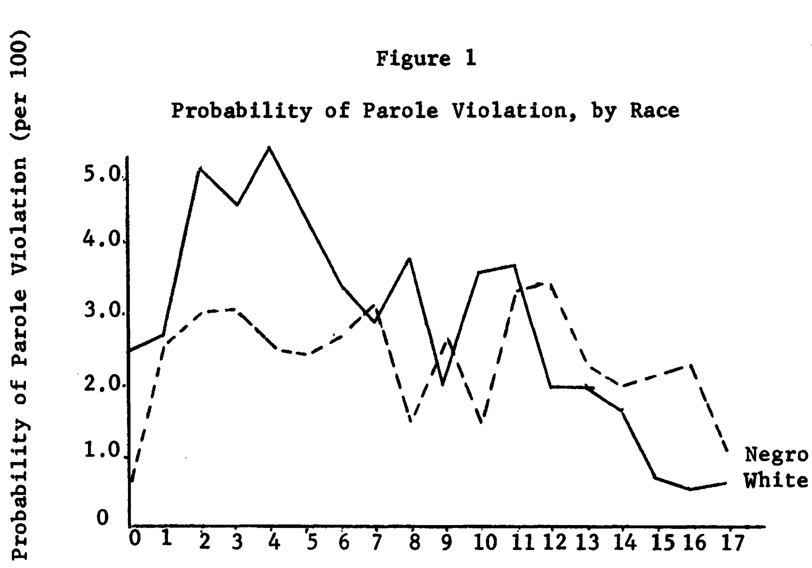
Probability of Parole Violation for Specific Months, by Month of Initial Release (per 100)

Prob. of PV for Month of:	Probability of Parole Violation, By Cohort Released during the following month of 1960:											
	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
1960 Jan.	2.0											
Feb.	1.0	0.8										
Mar.	7.2	0.0	1.7									
April	1.1	4.2	2.6	4.0								
May	5.7	2.6	5.3	3.1	0.0							
June	4.8	2.7	4.7	6.5	0.8	5.7						
July	2.5	8.4	4.9	4.7	4.0	4.3	0.9					
Aug.	4.1*	3.1	1.0	4.9	4.2	1.8	2.8	2.0				
Sept.	4.2	1.1	3.2	1.3	4.4	1.8	1.0	2.0	0.0*			
Oct.	1.5	1.1	4.6	2.7	3.8	1.9	1.0	5.2	3.5	0.7		
Nov.	3.1	2.2	0.0	5.6*	3.0	1.9	5.9	5.5	3.6	5.1	1.8	
Dec.	4.8	4.8	1.3	1.5	4.2	4.0	2.1	2.4	5.7	4.6	2.7	0.8
1961 Jan.	1.7	3.9	4.1*	6.3	4.7	4.3	7.8	7.2	2.0	2.4	3.7	3.4
Feb.	3.4	4.1*	1.4	3.3	2.7*	3.5	0.0	2.6	4.2#	2.5	7.8	4.4
Mar.	3.7	2.9	0.0	3.7	2.8	1.2	3.9	1.4*	4.3	3.4	4.2	7.3
April	2.0#	3.2	1.6	3.9	1.5	1.3	4.3*	0.0	6.8	3.5	2.2	8.9
May	2.1	0.0	1.7	2.0#	1.6	1.4	0.0	0.0	5.0	1.9	1.1	3.3
June	2.3	0.0	0.0	2.2	1.6	2.8	3.1	3.2	0.0	2.9	1.2	0.0
July		0.0	1.9	2.3	1.7	1.5	4.8	3.4	2.8	1.1	4.9	3.5
Aug.			0.0	5.3	0.0	0.0	3.4	5.6	8.8	3.3	2.6	3.8
Sept.				0.0	0.0	4.8	1.9	2.1#	0.0	5.7	0.0*	4.0
Oct.					0.0	1.7	0.0	4.3	7.1	2.5	4.2	4.2*
Nov.						0.0	2.0#	0.0	0.0	2.6	1.5	2.9
Dec.							4.4	0.0	0.0	1.4*	0.0	1.6
1962 Jan.								0.0	0.0	1.6	1.6	0.0
Feb.									0.0	1.7	1.7	0.0
Mar.										1.8	1.8	3.8
Apr.											1.9	0.0
May												0.0#
Number beginning Cohort	100	121	119	100	127	123	107	100	57	139	113	119

* = N_x fell below 75 for this month
# = N_x fell below 50 for this month

Figure 1

Probability of Parole Violation, by Race



Months Duration of Parole

Source: Table 2

Figure 3. Volume of Unemployment (in 000), Chicago Area and State of Illinois By Month, 1960-1962

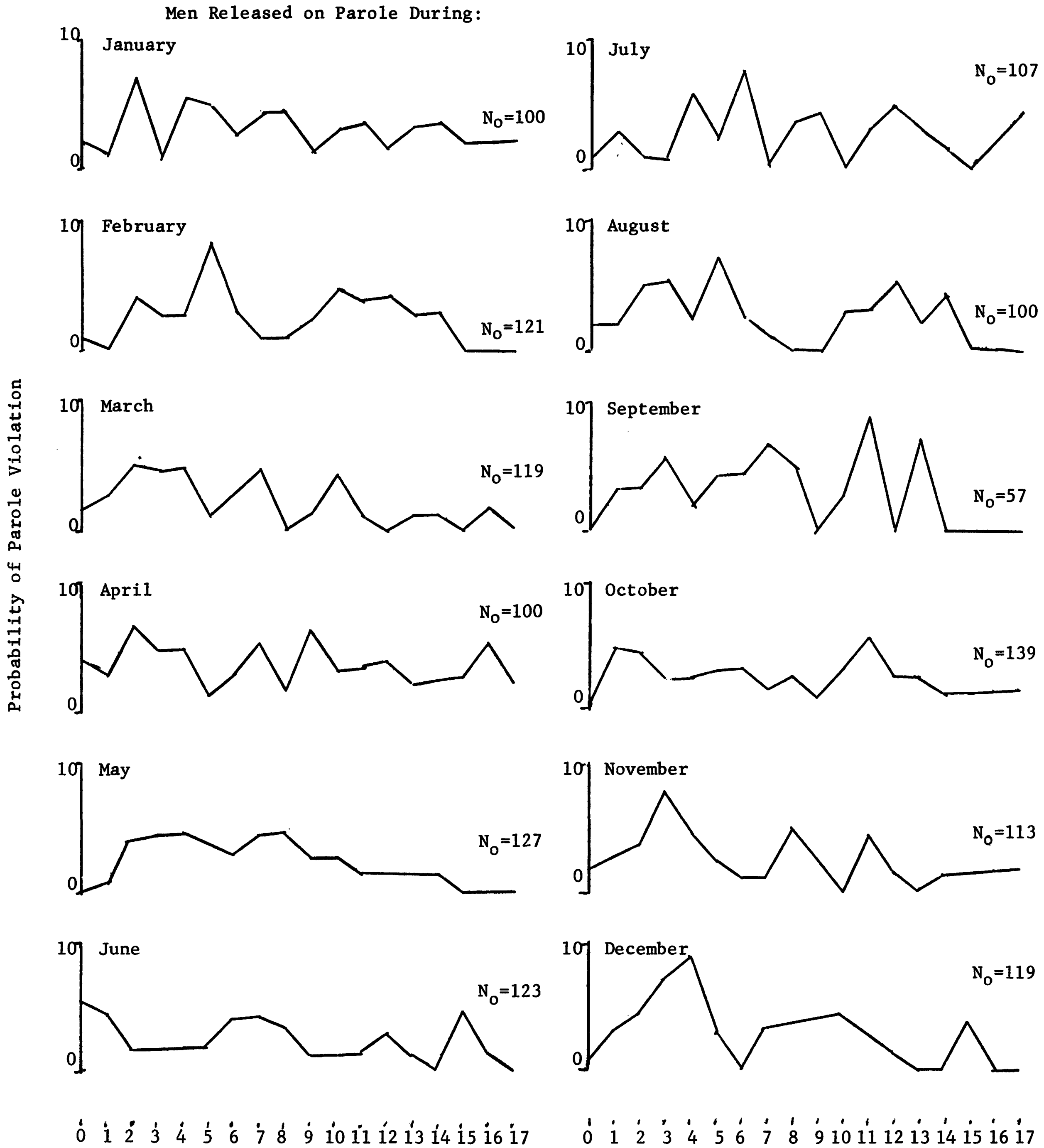
a. Total Unemployment, Cook and DuPage Counties (Illinois) and Lake County (Indiana)
b. Insured Unemployment, State of Illinois
c. Insured Unemployment, Cook and DuPage Counties, Illinois

Source: Research and Statistics Division, Chicago Association of Commerce and Industry (from data provided by U.S. Dept. of Labor-Bureau of Labor Statistics, Illinois State Bureau of Employment Security, and Indiana Employment Security Division)

Figure 2

Probability of Parole Violation, By Month Released and Duration (per 100)

of Successful Parole Time



Source: Table 3

AN ANALYSIS OF THE EFFECT OF INCOME ROUNDING
IN THE CURRENT POPULATION SURVEY

by

Joseph J. Knott
Social and Economic Statistics Administration, Bureau of the Census

This paper analyzes a form of income reporting bias caused by income sample survey respondents reporting income in rounded or approximate amounts.

This paper makes 7 major points:

1. An income rounding bias exists.

2. Responses from "proxy" respondents show more bias than responses from persons who report their own income to the interviewer.

3. The propensity to round is directly related to income.

4. Female persons have more tendency to round income data than male persons.

5. The bias is found at all levels of income aggregation with respect to persons income.

6. The bias may affect the usefulness of the CPS income data for some very specialized uses.

7. There are several changes in procedures which will reduce the bias.

Income Rounding Bias

It is well known that in any sample survey involving voluntary responses, such as the CPS, respondents do not always report exact income data. In fact, interviewers are specifically instructed to accept a reasonable estimate rather than report the income item as a nonresponse or "NA" (not available). This estimation bias has never been to my knowledge specifically quantified, although its presence has been recognized by many researchers. As a result of this bias, when income is tabulated by intervals which are less than $1,000 wide, the interval containing an exact thousand dollar amount or an exact $500 amount shows on extraordinarily high reporting frequency. This bias is most easily illustrated by a bar graph by $100 intervals. (see graph 1) this rather peculiar distribution is at variance with any of the usual income size distribution models (normal, lognormal, log-log).

Proxy vs Self Reporting

In the CPS interview process interviewers are not required to interview every member of the household, but may accept information from a responsible and knowledgeable "proxy" respondent. There were 40,046 proxy respondents or 43.0 percent of the 93,193 persons 14 years old and over who responded to the March, 1970 CPS survey. As the horizontal percentages in table 1 show, over half of the responses for male persons were provided by proxy respondents, but only about 30 percent of the responses for females persons were made by proxy respondents.

The percent of total persons reporting income in the $n,000 - $n,099 (n=1 to .14) intervals who had proxy respondents is higher in all 14 income intervals than the percent reporting for themselves who reported in that interval. Above $4,000, the percent is higher than 25 percent in all intervals tabulated (see table 2).

Propensity to Round Income Amounts

The propensity to round income amounts is in general directly related to the amount of income in the $1,000 to $15,000 range. As table 2 shows, the propensity to report income in the lowest $100 interval containing the exact thousand dollar amount increases as income increases. The tendency to round is very pronounced in the $10,000, $11,000, and $12,000 intervals. In several of the $1,000 intervals over 50 percent of the units reported the $n,000 to $n,099 interval (see table 2).

Reporting by Male vs Female Persons

In general, female self respondents have more of a tendency to report in the $n,000 - $n,099 than male self respondents. In only 4 out of the 14 intervals are the percent of male higher or equal to the percent reporting in the lowest $100 of the $1,000 interval. (see table 2).

Levels of Aggregation

The income rounding bias is present at all levels of aggregation, from a single type of income for a person to total family income. As Table 3 shows, the percent of persons reporting in the lowest $100 of an income class is higher for the individual income types than for total money income. This is true for both male and female respondents. The presence of the bias in total family income has been reported by this author in another paper. 1/

The degree in which the bias is present in persons total money income for persons and for total family income is the surprising aspect of the bias. It had been the assumption of this author and other researchers that as the eight individual income types were aggregated to total persons income and as persons income was aggregated to total family income, that the effect of the income bias would be greatly reduced if not elimated. However, it seem clear that the degree of the income bias exceeds the expected amount.

Impact of the Rounding Bias

In general the rounding bias does not affect the March income data greatly. The March income supplement to the CPS is designed to collect general statistics on the National level. In this sense, I doubt if the cost of collecting more precise data would be justified by the benefit derived. However, for some specialized researchers, this may not be the case. More and more the March CPS file is being used as a microdata file and for these users the rounding bias may case problems. The possible impact of the income rounding bias on the analysis of the low income population and on econometric modeling is pointed out below. Also the impact on income interval means and interpolation is discussed.

Economic Models

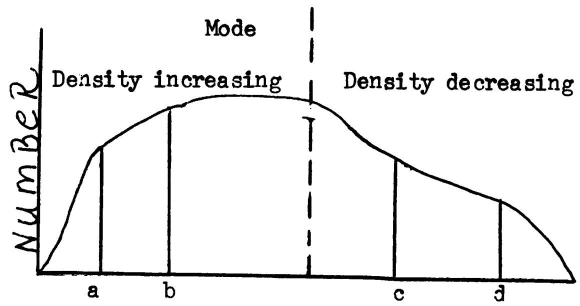
One use of the CPS March file is as input to economic models to evaluate the cost of effects of various programs. This is usually done by assigning different elegibility is dependent on income, the income reporting bias could be reflected by unexpected jumps in the number of eligibles as the qualification income threshold level is increases to include these discontinuous "Lumps" at thousand and five hundred dollar levels. I doubt that the income reporting bias would seriously effect the overall conclusion, but for specialized types of analysis it may.

Possible Impact on the Analysing Low Income Population

Since the low income thresholds are adjusted by the Consumer Price Index (CPI) each year, the rounding bias could cause abberations in the numbers of families classified as low income. For example, an increase in a low income threshold from 3,899 to 3,999 will increase the number of low income families less than an increase from 3,999 to 4,099. However, this aspect of the income rounding bias requires more analysis before its effect on changing the number of lower income families can be understood.

Income Interval Means

Another consequence of rounding reported income values is that the income class interval means are depressed. It is the "usual" practice to tabulate income in $1,000 or $500 intervals expect for high income values. For the intervals close to the class containing the mode income value, one can assume that the actual income values are distributed in the manner shown below. For the income class intervals below the mode (class ab for example), it is expected that the class means would fall above the midpoint ($\overline{ab} < \frac{b+2}{2}$). For the income class intervals above the mode (cd for example), it is expected that the class mean would fall below the midpoint ($\overline{cd} > \frac{c+d}{2}$).



However, this is not the case for class interval below the mode. As table 4 clearly shows the class intervals below the median (assumed to be close to the mode) the tabulated class falls well below the midpoint in the March Current Population Survey (CPS) income data. The Internal Revenue Service class means (table 4) falls slightly below the class midpoint but are much closer to the expected value than the CPS. The IRS data may also indicate some slight tendency for money to be paid in round amounts.

Interpolation for Median Income

It is standard practice to assume a linear distribution within a income class for the computations of the median income value. As can be seen from the $100 interval bar graph this appears to be wrong assumption (see graph 1-3). A negative exponentialfunction ($y = 1n\ e\ -x$) would be a better fit for the observed $100 interval values within a $1000 income class. However, since I believe the observed reporting pattern to be the result of a rounding bias, the assumption of a linear distribution may more closely conform with the true income distribution in the real world.

Reduction of the Income Bias

The income rounding bias could be reduced by several modifications in the instructions covering the field work on collecting the data. These are:

1. To the extent possible, interview the respondent personally.

2. To the extent possible, encourage the respondent to consult records for exact amounts rather than relying on memory.

3. Interviewers should be less willing to accept estimates.

While these modifications will reduce the income rounding bias they will also increase cost of field work and probably increase the income nonresponse rate. These tradeoffs would have to be considered implementing the proposed modifications.

Tabulation Modification

The importance of the bias can be reduced by tabulating data in $1000 intervals which start with the $500 level. For example, under $500, $500 to $1499, $1500 to $2499, $2500 to $3499, etc. Also it must be recognized that tabulation by intervals less than $1000 in size will result in some distortion of the data.

Conclusions

The data presented in this paper clearly show the presence of an income rounding bias. The hypothesis that the bias is caused by respondents reporting in exact thousand dollar amount and to a lesser extent in exact five hundred dollar amount has been established but not proven. Further investigation of this subject is being planned. This will require tabulation of persons reporting exactly an exact thousand dollar amount and file hundred dollar amount.

1/ "The Index of Income Concentration in the 1970 Census of Population and Housing", Joseph J. Knott, ASA Proceedings of the Social Science Section 1971.

Table 1.--Respondents 14 Years Old and Over by Sex and Self vs. Proxy Reporting Status in 1969

	Total	Self	Proxy
Total	93,192	53,146	40,046
Male	42,619	17,614	25,005
Female	50,573	35,532	15,041
Horizontal Percent			
Total	100.0	57.0	43.0
Male	100.0	41.3	58.7
Female	100.0	70.3	29.7
Vertical Percent			
Total	100.0	100.0	100.0
Male	45.7	33.1	62.4
Female	54.3	66.9	37.6

Table 2.--Number of Units Reporting Total Money in the $n,000 - $n,099 Interval as a Percent of the Units in the $n,000 - $n,099 Interval by Self and Proxy

Total Money Income	Percent of Units in $n,000 - $n,099 Intervals					
	Total		Male		Female	
	Self	Proxy	Self	Proxy	Self	Proxy
$1,000 - $1,999	14.0	16.1	11.6	15.1	15.1	18.4
$2,000 - $2,999	16.8	20.1	14.3	18.7	18.1	22.2
$3,000 - $3,999	18.7	20.7	17.4	19.3	19.4	22.7
$4,000 - $4,999	20.2	22.0	18.9	20.7	20.9	24.1
$5,000 - $5,999	22.1	25.3	21.1	22.9	23.0	31.3
$6,000 - $6,999	23.6	26.1	21.7	24.7	25.7	32.8
$7,000 - $7,999	18.9	24.2	19.0	22.8	18.7	35.1
$8,000 - $8,999	24.0	27.7	23.4	27.1	25.6	35.2
$9,000 - $9,999	22.7	28.5	22.5	27.9	23.8	40.7
$10,000 - $10,999	29.8	35.3	29.8	35.4	29.8	54.2
$11,000 - $11,999	27.5	31.1	28.3	30.7	23.2	42.4
$12,000 - $12,999	32.7	40.3	32.1	39.9	36.2	52.0
$13,000 - $13,999	26.6	31.6	26.0	30.9	31.2	58.3
$14,000 - $14,999	26.7	28.5	27.0	28.4	24.3	40.0

NOTE: See Table 3 for details.

TABLE 3.-- Percent of Persons 14 Years Old or Over who Reported Income of The Specified Type and Reported in The Lowest $100 of The Specified $1000 Interval in 1969 by Sex and Reporting Status (Self vs. Proxy)

Income Level $n,000	Self: Total Money Income	Self: Wages and Salaries	Self: Nonfarm Self-Employed	Self: Farm Self-Employed	Self: Social Security + Railroad Retirement	Self: Dividends Interest Net Rent ect.	Self: Public Assistance	Self: Public Transfers	Self: Private Pensions, Annuities, Alimony, ect	Proxy: Total Money Income	Proxy: Wages and Salaries	Proxy: Nonfarm Self-Employed	Proxy: Farm Self-Employed	Proxy: Social Security + Railroad Retirement	Proxy: Dividends Interest Net Rent ect	Proxy: Public Assistance	Proxy: Public Transfers	Proxy: Private Pensions Annuities
MALE																		
1000 - 1999	11.6	12.8	25.6	22.0	8.1	24.3	25.6	24.6	14.6	15.1	17.1	37.1	30.4	9.8	24.3	29.9	16.9	18.1
2000 - 2999	14.3	22.1	38.3	37.4	12.4	29.1	13.3	10.2	13.3	18.7	22.9	38.1	35.8	14.6	24.9	11.1	12.7	18.2
3000 - 3999	17.4	21.9	40.0	37.5	20.3	42.3	17.6	15.1	26.9	19.3	20.8	54.4	43.1	12.5	39.0	9.1	28.3	40.6
4000 - 4999	18.9	23.4	42.1	50.0	20.0	37.2	16.7	20.3	25.0	20.7	23.3	38.0	48.9	7.7	37.7	20.1	15.4	42.1
5000 - 5999	21.1	24.1	61.1	44.6	(B)	52.3	-	15.0	45.2	22.9	26.5	48.1	40.0	(B)	50.0	(B)	20.0	44.4
6000 - 6999	21.7	23.9	62.0	40.5	(B)	54.5	(B)	36.4	36.4	24.7	27.6	56.6	36.7	(B)	40.0	(B)	33.3	(B)
7000 - 7999	19.0	21.4	43.8	65.2	(B)	28.6	-	22.2	33.3	22.8	25.2	49.3	51.9	-	25.0	(B)	28.6	(B)
8000 - 8999	23.4	27.3	55.8	45.8	-	40.0	(B)	37.5	40.0	27.1	32.7	69.4	36.8	-	37.5	-	14.3	(B)
9000 - 9999	22.5	28.7	62.2	50.0	(B)	46.7	-	36.4	(B)	27.9	32.0	61.5	66.7	-	14.3	-	57.1	(B)
10000 - 10999	29.8	39.2	75.7	72.7	-	68.4	-	(B)	40.0	35.4	43.3	83.8	76.5	-	56.3	-	(B)	(B)
11000 - 11999	28.3	42.3	72.7	83.3	-	57.1	-	(B)	(B)	30.7	41.3	54.5	50.0	-	(B)	-	-	-
12000 - 12999	32.1	52.4	84.6	77.8	-	50.0	-	40.0	(B)	39.9	54.9	76.8	83.3	-	(B)	-	(B)	-
13000 - 13999	26.0	45.5	72.2	60.0	-	33.0	-	-	(B)	30.9	42.5	46.2	(B)	-	(B)	-	(B)	(B)
14000 - 14999	27.0	39.0	70.0	33.3	-	(B)	-	-	-	28.4	38.6	75.0	(B)	-	0.0	-	-	(B)
FEMALE																		
1000 - 1999	15.1	16.9	24.4	25.7	16.5	22.3	22.2	17.0	14.1	18.4	18.2	31.8	(B)	21.8	24.6	26.4	21.1	20.2
2000 - 2999	18.1	19.9	37.7	30.0	13.4	27.1	14.8	20.5	17.5	22.2	22.5	29.4	(B)	36.4	32.4	21.4	11.1	22.2
3000 - 3999	19.4	21.0	32.6	28.6	13.2	30.3	15.8	14.3	24.5	22.7	23.7	46.2	(B)	(B)	25.0	0.0	(B)	66.7
4000 - 4999	20.9	23.6	41.4	(B)	26.5	28.1	8.0	29.2	19.5	24.1	25.6	12.5	-	(B)	33.3	(B)	(B)	33.3
5000 - 5999	23.0	26.1	47.2	42.9	(B)	37.9	(B)	30.0	29.4	31.3	32.9	(B)	-	-	66.7	(B)	(B)	-
6000 - 6999	25.7	27.9	45.5	33.3	(B)	37.5	(B)	(B)	50.0	32.8	34.2	(B)	(B)	-	(B)	-	(B)	(B)
7000 - 7999	18.7	19.8	38.5	-	(B)	30.8	(B)	(B)	0.0	35.1	37.7	-	-	-	62.5	-	-	(B)
8000 - 8999	25.6	31.0	40.0	-	-	69.2	-	(B)	(B)	35.2	37.0	(B)	-	(B)	(B)	-	-	-
9000 - 9999	23.8	30.6	(B)	-	(B)	36.4	-	(B)	40.0	40.7	37.9	(B)	-	(B)	(B)	-	(B)	(B)
10000 - 10999	29.8	33.3	63.6	(B)	-	81.8	-	-	(B)	54.2	57.1	(B)	-	-	(B)	-	-	(B)
11000 - 11999	23.2	36.9	(B)	-	-	40.0	-	(B)	(B)	42.4	57.7	(B)	-	-	(B)	-	-	-
12000 - 12999	36.2	49.3	(B)	-	-	80.0	-	-	(B)	52.0	55.6	(B)	-	-	(B)	-	-	-
13000 - 13999	31.2	51.7	(B)	-	-	16.7	-	-	(B)	58.3	60.0	(B)	-	-	-	-	-	-
14000 - 14999	24.3	35.0	(B)	-	-	-	-	-	-	40.0	66.7	-	-	-	-	-	-	-

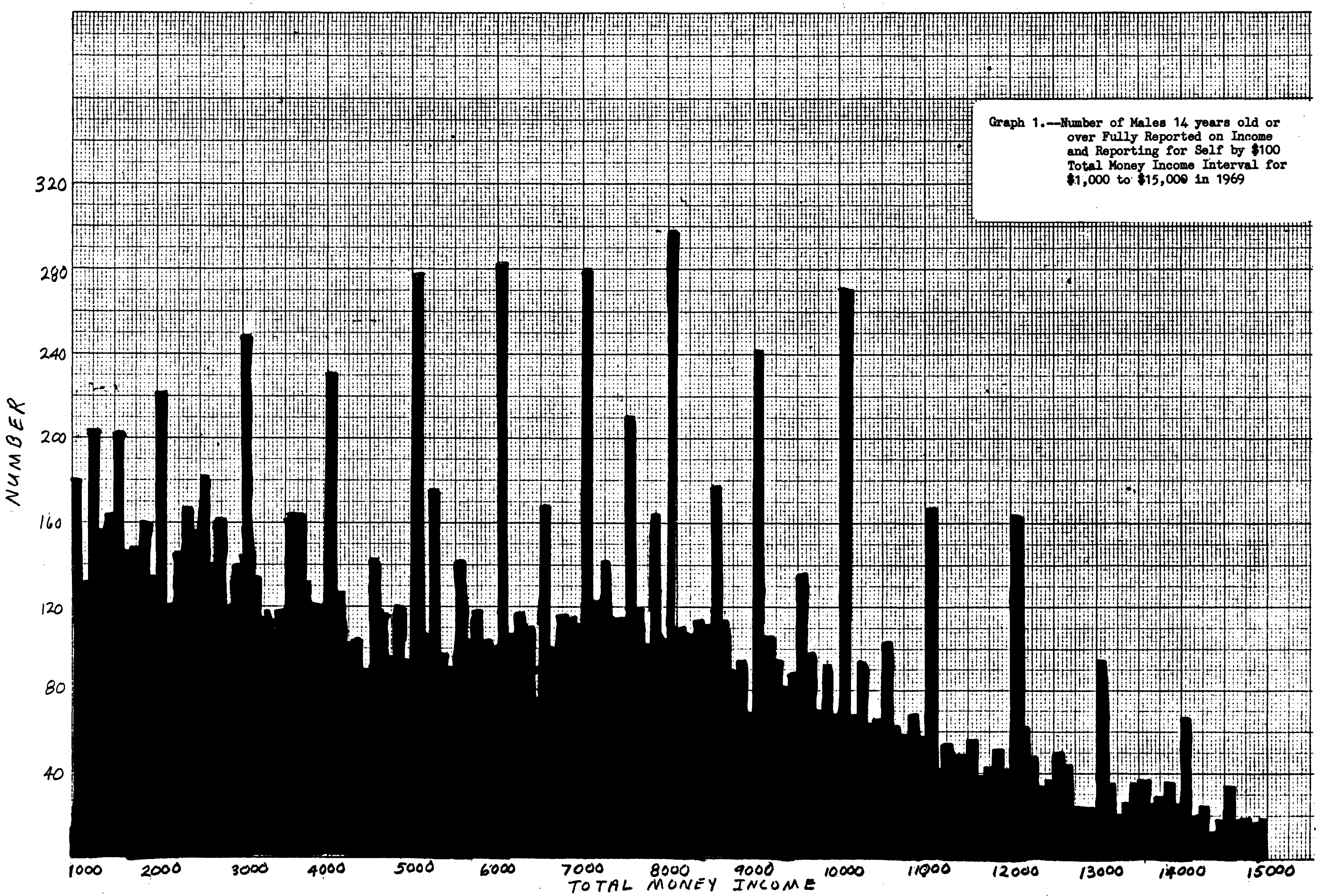
(B) - ~~Base~~ less than 5   - NONE

Percent = $\frac{\text{number } \$n,000 - \$n,099}{\text{number } \$n,000 - \$n,999} \times 100$

TABLE 4.--Mean Income of Specified Type by Size Class for Males 14 years Old and Over Reporting for Self in 1969 and All Individual Income Tax Returns in 1968

Income of Specified Type	MEAN INCOME OF SPECIFIED TYPE: CPS Males Reporting for Self with Specified Type of Income: Total	Wages and Salaries	Nonfarm Self-Employed	Farm Self-Employed	Social Security Railroad Retirement	Dividends Interest Net Rent Ect.	Public Assistance	Public Transfers	Private Pensions, Annuities Ect.	All Returns with Specified Type of Income: Adjusted Gross Income	Wages and Salaries	Business or Professional Net Profit	Farm Net Profit	Dividends	Interest
TOTAL	NA	NA	NA	NA	NA	NA	NA	NA	NA	$ 8,783	$6,839	$6403	$2993	$1355	$540
Under $1000	NA	NA	NA	NA	NA	NA	NA	NA	NA	938	502	449	413	191	196
$1000 to 1999	$1475	$1438	$1393	$1378	$1486	$1332	$1373	$1331	$1436	1483	1472	1462	1434	1402	1408
2000 to 2999	2465	2403	2324	2313	2395	2330	2470	2476	2416	2492	2486	2473	2468	2446	2433
3000 to 3999	3452	3424	3296	3325	3395	3283	3364	3451	3416	3502	3498	3478	3464	3450	3433
4000 to 4999	4433	4426	4287	4273	4402	4315	4216	4457	4401	4497	4493	4494	4484	4483	4437
5000 to 5999	5410	5386	5180	5248	(B)	5234	-	5392	5230	5500	5493	5471	5483		
6000 to 6999	6424	6408	6204	6354	(B)	6256	(B)	6341	6367	6500	6493	6476	6504		
7000 to 7999	7432	7410	7277	7185	-	7391	-	7391	7375	7560	7508	7479	7490	6921	6,647
8000 to 8999	8390	8366	8210	8282	(B)	8313	(B)	8356	8434	8489	8495	8476	8447		
9000 to 9999	9401	9367	9195	9295	-	9260	-	9366	(B)	9489	9480	9475	9478		
10000 to 10999	10358	10295	10105	10109	-	10142	-	(B)	10360						
11000 to 11999	11376	11279	11104	11033	-	11257	-	(B)	(B)						
12000 to 12999	12315	12217	12071	12113	-	12210	-	12220	(B)						
13000 to 13999	13393	13279	13122	13160	-	13266	-	-	(B)	16836	14,690	22,427	16,185	31,971	19,003
14000 to 14999	14386	14301	14090	(B)	-	(B)	-	-	(B)						
15000 and over	NA	NA	NA	NA	NA	NA	NA	NA	NA						

NA- NOT AVAILABLE (B) - BASE LESS THAN 5



Graph 1.--Number of Males 14 years old or over Fully Reported on Income and Reporting for Self by $100 Total Money Income Interval for $1,000 to $15,000 in 1969

CAUSES OF DIFFERENTIALS IN EARLY LABOR MARKET SUCCESS AMONG YOUNG WOMEN*

Andrew I. Kohen and Roger D. Roderick
Center for Human Resource Research, The Ohio State University

I INTRODUCTION

This paper is part of a larger research effort to identify the causes of differentials in early labor market success among youth. Here, we are primarily interested in (a) the effect of education, independent of ability, on early labor market success, and (b) white-black differentials in the determinants of early labor market success, where our measure of "success" is hourly earnings. For the most part, the extensive literature on women's wages consists of descriptions of male-female pay differences and attempts to answer questions concerning discrimination (1, 3, 6, 10, 11, 12, 14).[1] Our results are both relevant to those issues--indeed, such issues are central to our larger project--and of interest in their own right. Furthermore, we believe that the youth dimension in our study is of special import, for just as it has been shown that early encounters with unemployment have consequences for subsequent labor force experience (8), we contend that the wage rates received early in one's work history have an impact upon later experience.

II CONCEPTUAL FRAMEWORK

Our examination of differentials in early labor market success is by means of a two-equation, recursive model, with educational attainment as the dependent variable in the first equation and labor market success as the dependent variable in the second. Our explanatory variables are race, family background, ability, health, quality of education, and (for the second equation) quantity of education. Being black is expected to have a negative effect upon both educational attainment and labor market success, in that it is expected to reflect any existent racial discrimination (current or past) and any resultant intercolor variations in norms and attitudes. For reasons discussed below, each equation is run separately for whites and for blacks, rather than entering race as a dummy explanatory variable. Family background is expected to be positively related to both dependent variables, as a result of heredity and financial and environmental support. Ability is also presumed to bear a direct relationship to both education and labor market success. While we would anticipate good health to be associated positively with labor market success, it is not clear what sign to expect in its association with years of education completed. On the one hand, youngsters with severe health problems are often unable to continue in school. On the other hand, the foregone earnings cost of continuing school is probably considerably lower for them than for youth without disabilities.[2] Quality of schooling is hypothesized to be positively related to both dependent variables. Products of "better" schools should be more likely to be both motivated to seek and able to attain additional education, and should be better equipped for success in the labor market. Similarly, because we anticipate positive returns from investment in education, we expect the coefficient of education to be positive in the labor market success equation.

III THE DATA BASE

The universe under consideration here is a subset of a larger, national sample of women who were 14 to 24 years of age when interviewed in January/February 1968. The sample is a multi-stage probability sample selected from the civilian noninstitutionalized population of women within the prescribed age limits, and is a part of the National Longitudinal Surveys of labor market and educational experiences.[3] Data for all variables other than ability and quality of schooling were obtained through personal interviews conducted by the Bureau of the Census. Data on ability and quality of schooling were obtained through a mailed survey of the high schools attended by all 14- to 24-year-old male and female respondents in the National Longitudinal Surveys panels.

The characteristics of the available data base are such that we are both permitted and constrained to focus on a rather narrow universe of young women. This universe is specified as follows: females 18 to 24 years of age in 1968 who were employed as wage and salary workers and who were not enrolled in school at the time of the survey. The universe is further confined to those who had attended at least the first year of high school (the 9th grade) because the sources of our measure of mental ability were the records of secondary schools. The minimum age and enrollment status criteria were chosen so as to permit a reasonable range of occupations to be represented and to exclude those teenagers whose principal labor force activity is the rather casual occupation of babysitting. The employment status restriction was imposed after some experimentation with the data. We concluded that the value of a larger sample size would be outweighed by the costs of having to adjust the wage rates of the nonemployed for the effects of inflation. That is, for those unemployed or out of the labor force, strict comparability of measurement would have required dating their most recent job and adjusting the wage rate on that job for price level changes between that date and January/February 1968. In some cases the most recent job might have been held as much as five years prior to the survey.

IV METHODOLOGY

<u>Specification and Estimation of the Model</u>

As a first approximation to the complex nexus of variables which causes "early labor market success" among young women we employ a two-equation, recursive model. It can be stated as follows:

(1) EDU = F (RACE, SEL, SIBS, IQ, HLTH, QUAL)

(2) LMS = G (RACE, SEL, SIBS, IQ, HLTH, QUAL, EDU)

where each functional relationship indicates a linear structural equation. The acronyms are defined below in the discussion of the measurement of the concepts. This format was prompted by several factors. First, there seems to be a growing recognition among economists that single-equation models are inadequate to describe and explain the effects of schooling upon earnings (5, 9). Second, sociologists interested in similar hypothetical causal structures have found at least a modicum of success with the recursive structure, especially when interested in examining direct and indirect effects (2). Third, because of the recursive-structure hypothesis, the parameters of the structural equations can be estimated using ordinary least-squares regression analysis (7), which is a low-cost, convenient starting point. Finally, the multiple-equation format permits us to examine both direct and indirect effects of some variables.[4] It should be recalled that the parameters of the model are estimated and shown separately for whites and blacks.[5] This estimation procedure reflects our belief that race interacts with the other determinants of success.[6]

Measurement of the Concepts[7]

Labor market success (LMS) The measure of success used here is hourly rate of pay on current job. It differs from a "contract wage" because the time unit in which a respondent reported her earnings was at her discretion. If she reported it in other than hourly terms, the figure was converted to weekly units and divided by the usual number of hours she worked per week. Although the measure ignores psychological dimensions of success, it seems to us to be the most appropriate measure of economic success. Annual earnings is a more commonly used measure, but seems inferior to us because it is "contaminated" by the effects of annual hours of work which may or may not be positively related to success. The young woman who seeks and obtains part-year or part-time employment may be considered as "successful" as her counterpart who is employed full time all year. The social stigma which is applied to out-of-school males who are out of the labor force a significant part of the year does not seem to be applied to comparable females, even if they are unmarried.

Amount of schooling (EDU) The amount of schooling completed is measured as a continuous variable in units of single years. Although nonlinearities in the effects of schooling (e.g., "sheepskin" effects) may exist, we do not test for them here.

Quality of schooling (QUAL) We operationalize this concept in an index based on the following characteristics of the last secondary school attended by the respondent: per-pupil availability of library facilities, guidance counselors/100 pupils, pupils/full-time teachers, and starting salary of an inexperienced teacher with a bachelor's degree. The last variable is not school-specific. Rather, it refers to district-wide schedules for the 1967-1968 academic year and was adjusted for inter-area differences in price levels in 1967. The quality index is conceived purely in ordinal terms and was constructed to correspond to the assumption that schools from which the data were obtained are normally distributed with respect to quality. The final scaling of the index assigned a range of values of 1-11 where: 1 = lowest 1 percent of the quality distribution, 2 = next higher 4 percent, 3 = next higher 7 percent, 4 = next higher 12 percent, 5 = next higher 16 percent, 6 = middle 20 percent, etc.

Family background (SEL and SIBS) Family background is operationalized in two variables, i.e., an index of the socioeconomic level of the parent family (SEL) and number of siblings of the respondent (SIBS). The index is a simple average (mean = 10.0, s.d. = 3.0) of the linearly transformed values of the following measures: father's education, mother's education, education of oldest older sibling, father's occupation, and availability of reading material in the home of orientation. In order to preserve data cases, we permit the index to be computed for any respondent who provided information on at least three of the five components. For the sake of computational convenience, the number of siblings is excluded from the index but included separately in the estimating equation.

Mental ability (IQ) Our measure of mental ability is a standardized measure derived from the score on one of many tests of mental capacity reported by the secondary schools. For the entire sample of young people (male and female) for whom the data were collected, the variable was scaled to a metric which is conventional in educational testing--i.e., mean = 100, s.d. = 16. A lengthy analysis of pooling scores from different tests has been performed which establishes the legitimacy of the technique for this type of empirical research (9).

Health (HLTH) The measure of health condition used here is a dummy variable (1 = "unhealthy") based on self-reported physical limitations on current work activity. The measure is less than ideal as a proxy for health conditions which might have affected school attendance, and in future work we intend to attempt a refinement of it based on the duration of the professed limitation.

V DISCUSSION OF FINDINGS

In the educational attainment equation for whites, all of the variables except HLTH are significant and all have the hypothesized signs (Table 1). In sum, the equation explains 24 percent of the variance in educational attainment. The nonsignificance of the health variable should be interpreted with caution. First, there are the measurement problems mentioned above. We are unable to relate the timing of the reported health limitation to the respondent's period of school enrollment. Furthermore, we have no knowledge of the nature of the reported health impairment. Second, nonsignificance may result from two offsetting impacts of poor health: some of the unhealthy young women may not be physically or mentally capable of continuing in school, while others may be induced by the relatively lower opportunity costs to stay in school.

For blacks, the explanatory power of the equation is substantially less than for whites, i.e., it accounts for only 10 percent of the variance in educational attainment. There are also other major intercolor differences in the equation. For blacks, the only variable which is significant at conventionally accepted levels is IQ. Examination of (a) the zero-order correlation among the regressors and (b) several different sequences of entrance of the variables into the equation indicates that the nonsignificance of family background[8] and school quality[9] is "real" rather than a product of collinearity.

In the wage equation for whites the $\bar{R}^2$ is .167 and only the coefficients for EDU and SIBS are significant. The direction of effect is as hypothesized in both instances. For the blacks, EDU, IQ and QUAL all have significant coefficients, again in the hypothesized

Table 1 Regression Results for Model F-1: Out-of-School Females 18 to 24 Years of Age in 1968 Employed as Wage and Salary Workers

(t-ratios)

Explanatory variable[b] \ Dependent variable[a]	EDU (years)		LMS = WAGE (cents/hour)	
	WHITES	BLACKS	WHITES	BLACKS
EDU	c	c	+16.6 (9.41)	+19.1 (5.00)
IQ	+.04 (7.94)	+ .03 (3.12)	+ .1 (0.30)	+ 1.1 (2.50)
SIBS	-.07 (2.65)	-.05 (1.04)	- 2.1 (1.78)	+ .1 (0.50)
HLTH	+.01 (0.02)	-.47 (0.92)	+ 2.9 (0.25)	+15.9 (0.71)
QUAL	+.07 (2.56)	-.09 (1.57)	+ .4 (0.31)	+ 5.9 (2.26)
SEL	+.25 (7.29)	+.08 (1.33)	- .9 (0.58)	- .8 (0.31)
CONSTANT	+6.12	9.68	- .4	-175.1
$\bar{R}^2$	.237	.099	.167	.219
F-ratio/S.E.E.	38.4/1.39	4.06/1.38	21.0/59.9	7.56/61.4
N	602	141	602	141

a See text for definition of dependent variables.
b See text for definition of explanatory variables.
c Variable did not enter equation as explanatory factor.

directions. Here, the $\bar{R}^2$ is somewhat higher, .219. Of particular interest is the intercolor differential in the effect of ability on earnings. It is substantially more important for young black women; evidence, perhaps, of an element of racial discrimination in the labor market. For example, the fact that a white job applicant has received a high school diploma might alone be sufficient to secure employment for her, whereas only the most able of the black graduates might be accepted for that same job.[10]

It should be noted at this point that the explanatory powers of the above equations differ as between whites and blacks. In the education equation we are able to explain more of the variance among whites than among blacks, whereas in the wage equation the reverse is true. Additionally, the intercolor difference is far more striking in the education than in the wage equation.

The notable intercolor difference in the behavior of the measure of mental ability prompts us to examine it in somewhat greater detail. Given the causal ordering embodied in our model, the "total" effect of IQ on WAGE net of all prior factors can be decomposed into a direct effect (i.e., the regression coefficient shown in Table 1) and an indirect effect.[11] The latter is obtained by computing the effect of ability on hourly earnings through education. If we categorize the variables in our model into I(IQ), E(EDU), W(WAGE) and O(OTHER), the total effect of I on W can be seen as $b_{WI.O}$. The decomposition of this total effect implied by our model is as follows:

$$b_{WI.O} = b_{WI.EO} + (b_{WE.IO})(b_{EI.O}).$$

The first term on the right-hand side of the equation is the net, direct effect and the second term is the indirect effect.

Using the estimated regression coefficients to compute the total effect of IQ on WAGE for whites and blacks yields values for $b_{WI.O}$ of .6¢ and 1.6¢, respectively. This suggests that ability is about three times as important in the determination of the wages of young black women as in the determination of the wages of young white women. While this difference is by no means trivial, it is certainly more reasonable than the relative difference of 24:1 (if $b_{WI.EO}$ for whites is assumed $\neq 0$) or of ∞ (if $b_{WI.EO}$ is assumed = 0). Furthermore, it indicates that the size of the indirect effect of mental ability, through schooling, is about the same for white and black women.

Finally, because in general women's wages are affected by their concentration in lower paying occupations, we have begun to explore the effect of occupational assignment upon our estimates. To examine this effect, we first employed a dummy variable representing "atypicality of occupation,"[12] because there is some evidence that women in occupations where the proportion of female employment is low are likely to receive higher wages than those in occupations where women make up a large proportion of the work force (3, 4, 11, 12, 14). When this atypicality variable was added to the wage equation we found that, among whites,[13] those employed in atypical occupations earn approximately 13 cents per hour more than do their counterparts holding more traditional female jobs.

We then expanded our exploration to include a full set of interaction terms between atypicality of occupation and the other determinants of hourly rate of pay. Previous work (13) had suggested that the background characteristics of young women in atypical occupations were sufficiently different from those of the respondents in typical jobs to warrant testing for the existence of these interactions. The final equation containing only significant interaction terms is shown below (values in parentheses are t-ratios).

$$\underset{}{\mathrm{WAGE}} = .7 + \underset{(10.2)}{17.5\ \mathrm{EDU}} - \underset{(0.69)}{.2\ \mathrm{IQ}} - \underset{(1.84)}{2.1\ \mathrm{SIBS}} + \underset{(2.62)}{1.3(\mathrm{IQ})(\mathrm{ATYP})}$$

$$- \underset{(2.38)}{10.0\ (\mathrm{EDU})(\mathrm{ATYP})}$$

This equation shows that, in terms of labor market success, education is less important for the atypicals than for the typicals, while ability is more important. More so than in traditional "women's work," the labor market success of a young woman who enters an atypical occupation is a function of her ability rather than merely the amount of schooling she completes.

VI SUMMARY

We have hypothesized and tested a two-equation, recursive model to identify the determinants of early labor market success among young women, where success is operationalized as hourly earnings. For the most part, the explanatory variables included in the model performed in accordance with our expectations. The principal exception was the consistent nonsignificance of the measure of health condition, although for reasons discussed earlier we caution against strict interpretation of our results for this variable. Another unexpected finding was that family background does not play a significant role in determining the educational attainment of young black women. It should be recalled that this applies only to those young women who at least enter high school, and that nearly one of every eight black women in the age range of our study did not go beyond the eighth grade (15). Each of our measures of mental ability, amount of schooling and quality of schooling exhibit significant direct effects on hourly earnings for one or both of the color groups.

The several black/white differences revealed by the regression analysis provide *prima facie*, if not statistically rigorous, support for our belief that race interacts with other determinants of labor market success. One of the intercolor differences is in the explanatory power of the model, i.e., we are able to explain more of the variance in success among blacks than among whites, but just the reverse is true for the educational attainment equation. Our results indicate also that mental ability has both direct and indirect effects on the early success of young black women, but only an indirect impact for their white counterparts. We have suggested that finding a significant direct effect of ability for blacks but not for whites may be evidence of racial discrimination in the labor market--i.e., the "creaming" of only the most able blacks.

Our brief exploration of the role of occupational assignment in the determination of hourly earnings indicates that young white women in atypical occupations receive higher wages and that their labor market success is more likely to be dependent upon ability than is the success of their counterparts in traditional women's occupations. Finally, the net effect of years of schooling is significantly lower for young women in atypical occupations than for those holding more conventional occupational assignments.

FOOTNOTES

[*]This paper is based on data from the National Longitudinal Surveys, a project sponsored by the Manpower Administration, U.S. Department of Labor, under the authority of the Manpower Development and Training Act. Researchers undertaking such projects are encouraged to express their own judgment, thus interpretations or viewpoints stated in this document do not necessarily represent the official position or policy of the Department of Labor.

[1]All item numbers in text and footnotes refer to numbered reference items.

[2]The sign of the estimated coefficient can be used to test the relative power of these competing hypotheses. Of course, a nonsignificant coefficient does not resolve the ambiguity.

[3]For a detailed description of the sampling design and the complete interview schedule see (15).

[4]Among the alternative specifications of the model which will be examined in the larger study of which this is a part are multiplicative equations and the hypothesis that the system is more appropriately viewed as simultaneous.

[5]The term "blacks" refers here exclusively to Negroes. The nonwhite nonblacks in the sample were purposely excluded in order to focus on Negro-white differences.

[6]Unfortunately, we are unable to subject this belief to rigorous statistical testing because the computer technology available to us at the present time does not permit precise consideration of the fact that blacks were oversampled relative to whites in the ratio of 3:1. We expect, however, that this technological bottleneck will be eliminated in the near future.

[7]For means, standard deviations, and zero-order correlations for the variables described below, see Appendix Table 1.

[8] Previous tabular analysis of the relationship between family background and educational attainment, using slightly different measures of the former, produced similar results (15, pp. 30-32).

[9] In an earlier version of this paper we had suggested that a black-white differential in graduation from vocational high schools might be responsible for the unexpected sign of QUAL in the education equation for blacks. However, re-estimation of the equations including variables representing vocational school enrollment did not confirm this speculation.

[10] Because we believed that both statistical and "real" interrelationships among the regressors could be leading our interpretations astray, we re-ran both equations, entering only the variables which were significant in the model specified above. For both the wage and education equations, the results produced by the re-estimation were virtually identical to those from the original run.

[11] This decomposition is analogous to those performed using beta weights or path coefficients and can also be derived by applying the conventional "omitted-variable" formula (5).

[12] The categorization of occupations as "atypical" or "typical" for women was based on the occupational distribution of women in the labor market as of the 1960 Census. At that time women comprised 32.8 percent of the U.S. labor force; thus those occupations for which the proportion of female incumbents was less than that percent were classified as "atypical" for women. For a complete description of the variable and examples of atypical occupations see (14, pp. 3-4).

[13] We did this only for whites because the number of sample cases of blacks in atypical occupations was too small for confident analysis.

REFERENCES

(1) Ashenfelter, Orley. "Changes in Labor Market Discrimination Over Time." Journal of Human Resources, V (Fall 1970), 403-30.

(2) Duncan, Otis Dudley, David L. Featherman, and Beverly Duncan, Socioeconomic Background and Occupational Achievement: Extensions of a Basic Model (Washington, D.C.: U.S. Department of Health, Education, and Welfare, Office of Education, 1968).

(3) Fuchs, Victor R. "Differences in Hourly Earnings Between Men and Women." Monthly Labor Review, XCIV (May 1971), 9-15.

(4) Gordon, Margaret S. "Women in the Labor Force." Industrial Relations, VII (May 1968), 187-92.

(5) Griliches, Zvi, and William Mason. "Education, Ability, and Earnings." Discussion Paper No. 207, Harvard Bureau of Economic Research, Cambridge, Massachusetts, October, 1971.

(6) Gwartney, James, and Richard Stroup. "Sex Discrimination in Employment: The Man or the Myth." Paper presented at Western Economic Association Meetings, August, 1971.

(7) Johnston, J. Econometric Methods (New York: McGraw-Hill Book Co., Inc., 1963).

(8) Kalachek, Edward. The Youth Labor Market, Policy Papers in Human Resources and Industrial Relations, No. 12 (Ann Arbor, Michigan: University of Michigan, Institute of Labor and Industrial Relations, 1969).

(9) Kohen, Andrew I. "Determinants of Early Labor Market Success Among Young Men: Race, Ability, Quantity and Quality of Schooling" (Ph.D. Dissertation, The Ohio State University, 1972).

(10) Kreps, Juanita. Sex in the Marketplace: American Women at Work (Baltimore, Maryland: Johns Hopkins Press, 1971).

(11) McNulty, Donald J. "Differences in Pay Between Men and Women Workers." Monthly Labor Review, XC (December 1967), 40-43.

(12) Oppenheimer, Valerie Kincade. "The Sex-Labelling of Jobs." Industrial Relations, VII (May 1968), 219-34.

(13) Roderick, Roger D. and Joseph M. Davis. "Correlates of Atypical Occupational Assignment" (Columbus, Ohio: The Ohio State University, Center for Human Resource Research, 1972).

(14) Sanborn, Henry. "Pay Differences Between Women and Men." Industrial and Labor Relations Review, XVII (July 1964), 534-50.

(15) Shea, John R., Roger D. Roderick, Frederick A. Zeller, and Andrew I. Kohen. Years for Decision, Vol. I, U.S. Department of Labor, Manpower Administration, Manpower Research Monograph No. 24 (Washington, D.C.: U.S. Government Printing Office, 1971).

Appendix Table 1 Means, Standard Deviations and Zero-Order Correlation Coefficients[a] for Variables[b] in Labor Market Success Model, by Color: Employed Nonstudent Females 18 to 24 Years of Age in 1968 Who Had Completed at Least 8 Years of Schooling

	Whites									Blacks	
	Mean	S.D.	EDU	IQ	SEL	SIBS	HLTH	QUAL	WAGE	Mean	S.D.
EDU (years)	12.6	1.6		30	16	-17	-06	-13	44	12.3	1.4
IQ (index)	104.6	12.3	38		17	-20	09	-07	30	91.4	12.5
SEL (index)	9.0	1.8	38	25		-22	-03	14	10	8.3	2.1
SIBS (persons)	2.7	2.1	-19	-11	-20		05	01	-10	4.1	2.8
HLTH (percent)	4.7	21.2	-04	-08	-06	01		-02	-06	5.7	23.2
QUAL (index)	5.9	2.0	10	01	03	-01	02		11	6.0	2.0
WAGE (cents/hour)	200.6	65.6	41	16	15	-14	-01	05		190.3	69.5

a Correlation coefficients for whites are below the main diagonal and those for blacks are above the main diagonal. Decimal points are omitted from the coefficients.

b For definition of the index units of measurement see text.

ON THE ESTIMATION OF FUNCTIONAL RELATIONSHIPS FOR A FINITE UNIVERSE

J. C. Koop, Research Triangle Institute, N. C.

1. Introduction

In the modern theory of least squares, and the theories from which it evolved [12], the notion of residuals or the somewhat kindred notion of errors estimation, both of which have a strong empirical basis, play a basic role [2, 3]. Either one of these notions, in their relevent context, with the twin principles of unbiasedness and minimum variance, yield the solutions to the problems of fitting curves or surfaces, or adjustment of observations.

In this connection, in the theory of regression for sample surveys presented by Konijn [6], the notion of residuals also plays a basic role, plus either of the following assumptions:

> (i) "the population...itself constitutes a proportionate stratified sample from a (conceptually) infinitely large population of individuals with similar behavior. This makes the residual z" (equal to $y - a - \beta x$ in his notation) "an ordinary random variable, in the sense of a drawing from an infinite population, uncorrelated with each x and with zero mean," and (ii) "for any given individual in the finite population the conditional mean of y given x is linear in x, and . . . the deviation z corresponding to a given individual and for a fixed value of x represents one realization among a class of potential fluctuations about this individual's conditional mean y for any fixed x",

each leading to two fifferent linear models proposed by him.

Recently Godambe and Thompson [5] introduced with much ingenuity the notion of "error" vector for their "attempt to show how the regression analysis generally used for hypothetical populations can also be validated for the actual populations commonly dealt with in statistical surveys" by starting their arguments with the assumption which they stated as follows:

> "Suppose that for every unit $i(i=1,..,N)$ of the population we have knowledge *a priori* (i.e. before any sampling is done) of the (real) value y_i associated with the unit $i(i=1,..,N)$ of some auxiliary variate y".

Their arguments for arriving at the "error" vector are different from those of Konijn of whose work they do not appear to be aware.

However, in the theory presented in section 2, these restrictive concepts of residuals or errors (restrictive, because other than the dependent random variable other random variables pertaining to the units of the sample would have to be known or to remain fixed) do not play any role. Rather, the notion of the ascribed (or hypothesized) line, surface, or hypersurface, with variables that are not random variables, passing through the centroid,* the statistics of which have desirable properties in the context of sample survey theory, plays an important role. The latter notion immediately leads to the random function, which plays a key role as illustrated in the example of section 3. Further the present theory has a different empirical basis from that of least squares. In sequel it will appear that the logical basis of this theory, characterized by an economy of principles and assumptions, is different from those of Konijn, Godambe, and Thompson.

In regard to principles, what is common between the theory presented in section 2 and least squares is the Gauss-Laplace principle of minimum variance. Because of this circumstance, in the matter of fitting a straight line (or any plane), for the case of a single stratum or universe, despite the difference in approach, the end results agree, when the line is made to pass through the estimated centroid. This agreement is evidenced by equation (27).

Thus so far in the literature of statistical theory and methods, a theory of functional relationship which directly recognizes the circumstance that all values observed on the units of a probability sample from a finite universe are random variables does not exist. The theory presented in section 2, which is fairly general, is an attempt to fill this gap, and is an elaboration of the two concise statements [9, 10] made by the writer more than eight years ago.

One illustration of the theory given in section 3 has obvious applications, for example in studying relations between income and expenditure in consumer expenditure surveys carried out by methods of statistical sampling, and doubtless other sociological surveys similarly carried out.

2. Theory

There is a finite universe U composed of N identifiable elements $\{u_i: i=1,2,...,N\}$ which are the ultimate units of sampling. These elements may be directly identifiable as single units, or identifiable indirectly as distinct members of larger units. The ℓ (≥ 2) measurable characteristics (variate values) of unit u_i are $(x_i, y_i, z_i...)$, and for the purpose of

---

* This idea was used by Boscovich as early as 1757 in fitting a line by his method of minimizing the sum of the absolute deviations [2]. I must thank Dr. Churchill Eisenhart for the gift of his paper on the work of Roger Joseph Boscovich.

this paper we assume that they can be measured without error.

A de facto functional relationship exists between the ℓ conjoint values of the respective characteristics of each element u_i as soon as the corresponding sets of finite values are defined as ordered ℓ-tuples, and the whole (finite) collection of such ℓ-tuples

(1) $\{(y_i,x_i,z_i,\ldots): i=1,2,\ldots,N\}$

defines this functional relationship completely in a mathematical sense. The way of expressing the functional relationship at (1) may be interpreted to mean that the first coordinate y, in some sense, depends on the remaining ℓ-1 corrdinates (x,z,...); this formulation is sufficient to initiate the next statement of the problem and also to suggest the possibility of casual relationships in the context of problems in the real world.

Obviously such a function is discontinuous, but the theoretical possibility that it can be satisfactorily approximated by some continuous (real) function remains open.

Generally such a (hypothetical) function approximating the de facto function, and purporting to show the relationship between the variate values of the elements of the universe may be expressed as a multivariate polynomial (of some prescribed degree with $p(\geq \ell)$ parameters, β_o, β_1, β_{p-1}, but less than the number of observations, viz,

(2) $Y=\beta_o+f(X,Z,\ldots;\beta_1,\ldots,\beta_{p-1})$,

where the domains of the ℓ-1 arguments, X,Z,..., whatever they may be, are such that each includes the respective domains of x,z,... . The ℓ-1 domains of x,z,... are defined by the extreme values relevant to each marginal distribution. The domain of Y will be determined by $\beta_o + f$. We stress here that since the (Y, X, Z,...)'s are not in one-to-one correspondence with any probability measure, they are ordinary variables despite the fact that their domains cover the domains of (y,x,z,...) which later are found to be strictly random variables. The form of the polynomial relationship between Y and X,Z,... may be suggested by previous information and/or that of the sample drawn for its estimation. The reader may now see that any of the ℓ variables may be chosen as the dependent variable so that in general ℓ different formulations similar to (2) are possible. The user of the theory will have to decide which of the ℓ formulations is revelant in the context of his problem. It may very well be that more than one formulation is revelant. Obviously the question of relevance, in problems of specific applications of the theory, is something outside the scope of a general mathematical-statistical theory. What, in sequel, will appear important for the development of the theory is that f is linear in the ℓ-1 unknown parameters.

To estimate the parameters of the functional relationship (2) we have a sample of distinct units, s, from U which is realized with positive probability p(s). To elaborate further, s is a member of a collection S, i.e. S = {s}, and p(s), a rational positive number less than unity, is defined for all s. For the sake of generality the sample design leading to s (with all its appropriate randomization procedures for the selection of units) is not specifically defined. It may be simple random sample or it may be a multi-stage sample where units at one or more stages are drawn with varying probabilities, with or without replacement. What we have are the variate values of the distinct units of s and its corresponding probability of realization, viz.

(3) $\{(y_i,x_i,z_i,\ldots): i\varepsilon s\}$ and p(s).

We conclude from the statement at (3) and the preceding explanations of this paragraph that y, x, z,... are all random variables.

It is important to note that p(s) (unlike, for example, the multivariate normal probability for correlated variates) cannot tell us anything about the nature of the relationship between the random variables y, x, z, etc., simply because it is a numerical value. However, it is through p(s) that the values of the random variables incident to the distinct units of s are revealed, and which in turn suggest the form of the (approximate) functional relationship. Logically therefore we cannot ignore p(s), or equivalently the sample design, when we attempt to infer about (1).

Now (2), as it stands, is just a hypothetical relationship between the ℓ variables, as yet unrelated to the N discrete points

(4) $(y_i,x_i,z_i,\ldots)$, $(i=1,2,\ldots,N)$

in the same ℓ-dimensional Euclidean space, which define the de facto functional relationship (1). The statistical data given at (3) in some sense represents (4). To arrive at a statistically meaningful functional relationship a conjunction of (3) and (2) needs to be effected in some way.

Heuristically we feel that some condensation of the data at (3) is required as a first step to effect conjunction with (2). It is natural to think of the centroid of these points for the purpose of condensation.

Apart from the problem of functional relationships, it is desirable from the point of a view of the known estimation theory for sample surveys, to have the true centroid of (4) and the centroid of (3), which will be an estimate of the true centroid, as close as possible in some sense.

The true centroid is given by

(5) $(\bar{\bar{y}}, \bar{\bar{x}}, \bar{\bar{z}},\ldots)$

where

$$\bar{\bar{y}} = \sum_{i=1}^{N} y_i/N,$$

$$\bar{\bar{x}} = \sum_{i=1}^{N} x_i/N,$$

$$\bar{\bar{z}} = \sum_{i=1}^{N} z_i/N, \text{ etc.}$$

On the basis of the data at (3) let

(7) $(\hat{y}, \hat{x}, \hat{z},\ldots)$

be estimate of (5). The estimating formulas for $\hat{y}, \hat{x}, \hat{z}$, etc. will depend on the underlying sample design. We shall require these estimates to be admissible and consistent, and more desirably to be unbiased. Because there are no best estimates in the context of sample surveys (see [4] and [7]) this is all we can do to ensure that (7) shall be "close" to (5).

To effect conjunction let the polynomial (hypersurface) (2) pass through the point (7), which is composed of estimates (based on data (3)) which are also random variables. Then we have

(8) $\hat{Y}=\hat{y}+f(X,Z,..;\beta_1,..,\beta_{p-1})-f(\hat{x},\hat{z},..;\beta_1,.,\beta_{p-1})$

as the equation of the hypersurface passing through (7). Hence $\hat{Y}$ is a function of random variables or a random function for short. It is important to note that $f(X,Z,\ldots;\beta_1\ldots,\beta_{p-1})$ in (8) is just a real function without probabilistic character since X,Z,... are not defined as random variables.

Because $\hat{Y}$ is a random function there is no way to apply the classical method of least squares for the estimation of $\beta_1,\ldots,\beta_{p-1}$. However, one of the twin principles on which least squares is founded, the Laplace-Gauss principle of minimum variance, is applicable in the sense that it allows us to determine the form of the β's which minimize the variance of the random function (8). We proceed to apply this principle. We have

(9) $V(\hat{Y})=V(\hat{y})-2\,\mathrm{Cov}\,[\hat{y},f(\hat{x},\hat{z},\ldots;\beta_1,\ldots,\beta_{p-1})]$
$+V[f(\hat{x},\hat{z},\ldots;\beta_1,\ldots,\beta_{p-1})]$.

In determining $V(\hat{Y})$, $f(X,Z,\ldots)$ behaves as a constant by virtue of the fact that X, Z, etc., are not defined as random variables. Equation (9) reveals the interesting result that the variance of $\hat{Y}$ is constant for all (X, Z,...).

We note that f is a polynomial linear in the β's and also that the β's are implicit in $V[f]$ and $\mathrm{Cov}\,[\hat{y}, f]$, and both these functions which are respectively of degree two and degree one in the β's, involve the variances of $\hat{x}, \hat{z}$, etc., and also the relevant covariances. The value of the β's which minimize $V(\hat{Y})$ will be given by the solution of the following p-1 equations

(10) $$\frac{\partial V[(\hat{Y})}{\partial\beta_a} = \frac{\partial V[f]}{\partial\beta_a} - 2\frac{\partial\ \mathrm{Cov}[\hat{y}, f]}{\partial\beta_a} = 0,$$

$(a = 1, 2,\ldots p-1)$.

It is not difficult to see that the solution of the set of p-1 simultaneous equations given at (10) minimizes $V(\hat{Y})$. From (9), since the covariance function is linear in the β's, the Hessian $\left|\frac{\partial^2 V(\hat{Y})}{\partial\beta_a\partial\beta_b}\right| = \left|\frac{\partial^2 V(f)}{\partial\beta_a\partial\beta_b}\right|$, $(a=b=1,2,\ldots,p-1)$

will obtained with variances of $\hat{x}, \hat{z}$, etc., making up the diagonal elements and the corresponding covariance expressions making up the off-diagonal elements. This Hessian and all its principal minors, starting with $\frac{\partial^2 V(f)}{\partial\beta_1{}^2}$,

are all positive so that $V(\hat{Y})$ is minimized.

The simultaneous equations (10) involve (unknown) universe values of the variances and the covariances. Hence the solution for each β will be a ratio of functions of variances and covariances. For the sake of argument, if these variances and covariances were known, the β's obtained by solving (10) would really minimize $V(\hat{Y})$. But this cannot be realized. What can be done, as proposed in the foregoing account, is to substitute the unbiased or consistent estimates in the formal solution for each β_a.

The expressions for the estimated β's will be in the form of ratios, the respective numerators and denominators of which involve estimates of the variances and covariances. If $\hat{\beta}_a$ is an estimate of β_a, then because it is a ratio estimate

$$E(\hat{\beta}_a) \neq \beta_a,\ (a = 1, 2,\ldots, p-1),$$

so that it is a biased estimate. We shall recur briefly to this problem in section 3. At any rate these estimated β's, despite their bias, are in some sense near-minimum values, since the values implied by their parent expressions are minimum values.

In summary, and ascribed functional relationship is estimated in the following steps:

(i) estimate the centroid of the universe,
(ii) determine the random function, i.e., the equation of the given surface

(or line) passing through this estimated centroid,

(iii) determine the variance of this random function,

(iv) determine the parameters involved so as to minimize the variance of the random function, and

(v) estimate these parameters by substituting in their expressions unbiased or consistent estimates for the variances and covariances involved, or functions thereof.

With these estimates of the parameters and the centroid, an estimate of the functional relationship is

$$(11)\quad Y'=\hat{y}+f(X,Z,..;\hat{\beta}_1,..;\hat{\beta}_{p-1})-f(\hat{x},\hat{z},..;\hat{\beta}_1,..,\hat{\beta}_{p-1})$$

The expression for the variance of Y' will be extremely complex. Because of this circumstance the estimation of the variance of Y' for given (X, Z,...) for any specified sample design and functional relationship, will be difficult. A way of circumventing this difficulty through the well known technique of independent replicated samples, initiated by the late Professor Mahalanobis in the thirties in India, is discussed in section 3.

3. Example for Stratified Random Sampling

In this section we shall consider the estimation of a linear functional relationship for the case of stratified random sampling. Examples for more ramified sample designs are available but have not been included here to save space.

For the case of linear relationship with simple random sampling a brief discussion of

(i) the bias of the estimated parameter,

(ii) the estimation of the variance of a Y-value, for given X, through the technique of independent replication, and

(iii) a type of probability statement regarding the median Y-value for a given X

will be given. This discussion is intended to suggest solutions for similar problems incident to functional relationships considered with sample designs other than simple random sampling.

The universe U is subdivided into L strata each containing N_h elements (h=1,2,...,L). From stratum h, n_h (h=1,2,...;L) elements are selected with equal probabilities and without replacement at each draw, and two measurable characteristics are observed. Let (y_{hi}, x_{hi}) denote the variate values of the characteristics for the i^{th} element $(i=1,2,\dots,n_h)$ of the h^{th} stratum. We wish to estimate a linear functional relationship

$$(12)\quad Y = \beta_o + \beta_1 X.$$

We proceed as indicated in the summary paragraph of section 2. The centroid $(\bar{\bar{y}}, \bar{\bar{x}})$ is given by

$$(13)\quad \begin{cases} \bar{\bar{y}} = \sum_{h=1}^{L} \sum_{i=1}^{N_h} y_{hi} / \sum_{h=1}^{L} N_h \\ \bar{\bar{y}} = \sum_{h=1}^{L} \sum_{i=1}^{N_h} x_{hi} / \sum_{h=1}^{L} N_h \end{cases}$$

The unbiased estimates of $\bar{\bar{y}}$ and $\bar{\bar{x}}$ are

$$(14)\quad \begin{cases} \hat{y} = \sum_{h=1}^{L} \frac{W_h}{n_h} \sum_{i=1}^{n_h} y_h \\ \hat{x} = \sum_{h=1}^{L} \frac{W_h}{n_h} \sum_{i=1}^{n_h} x_n \end{cases}$$

where $W_h = N_h / \sum_{h=1}^{L} N_h$ (h=1, 2,...L). Thus the equation of the line passing through the estimated centroid $(\hat{y}, \hat{x})$ will be

$$(15)\quad \hat{Y} = \hat{y} + \beta_1 (X - \hat{x}).$$

By virtue of $(\hat{y}, \hat{x})$, (15) is a random function, with variance

$$(16)\quad V(\hat{Y}) = V(\hat{y}) + \beta_1^2 V(\hat{x}) - 2\beta_1 \operatorname{Cov}(\hat{y}, \hat{x}).$$

The value of β_1 which minimizes $V(\hat{Y})$ is given by the solution of the equation $\partial V(\hat{Y})/\partial \beta_1 = 0$, that is

$$(17)\quad \beta_1 = \operatorname{Cov}(\hat{y}, \hat{x})/ V(\hat{x}).$$

This is the minimum value, since $\frac{\partial^2 V}{\partial \beta_1^2} = 2V(\hat{x}) > 0$.

Now according to standard theory

$$(18)\quad \operatorname{Cov}(\hat{y},\hat{x}) = \sum_{1}^{L} W_h^2 \, S_{hxy} \left(\frac{1}{n_h} - \frac{1}{N_h}\right)$$

and

$$(19)\quad V(\hat{x}) = \sum_{1}^{L} W_h^2 \, S_{hx}^2 \left(\frac{1}{n_h} - \frac{1}{N_h}\right),$$

where

$$(20)\quad S_{hxy} = \sum_{i=1}^{N_h} (x_{hi} - \bar{\bar{x}}_h)(y_{hi} - \bar{\bar{y}}_h)/(N_h - 1)$$

and

$$(21)\quad S_{hx}^2 = \sum_{i=1}^{N_h} (x_{hi} - \bar{\bar{x}}_h)^2/(N_h - 1)$$

in which $\bar{\bar{x}}_h$ and $\bar{\bar{y}}_h$ are the usual stratum means. As n_h increases, the variances of the estimates given by (14) decrease. In this connection the reader might say "what happens to β_1 if $n_h = N_h$ for all h?" Mathematically, of course, we have a situation when β_1 as expressed in terms of values given by (18) and (19) assumes an indeterminate value $\frac{0}{0}$. But here logic, the premises for which are already embedded in the basis of the problem comes to the rescue. The reply is that the need for estimating the hypothesized relationship, along the lines stated in the paper, would vanish, simply because complete knowledge of the functional relationship is already given by the de facto relationship (1), applicable to the case of two variables. The same kind of argument holds if $n_h = N_h$ for some strata. The need for determining a functional relationship would then be restricted to the remaining strata.

An estimate of β_1, is obtained by substituting unbiased estimates of Cov $(\hat{y}, \hat{x})$ and $V(\hat{x})$ in (17) and their estimation entails unbiased estimates of S_{hxy} and S^2_{hx}. For a given h, these estimates are

$$(22)\quad s_{hxy} = \sum_{i=1}^{n_h} (x_{hi}-\bar{x}_h)(y_{hi}-\bar{y}_h)/(n_h - 1)$$

and

$$(23)\quad s^2_{hx} = \sum_{i=1}^{n_h} (x_{hi}-\bar{x})^2/(n_h - 1),$$

where $\bar{x}_h$ and $\bar{y}_h$ are the usual unbiased estimates of $\bar{\bar{x}}_h$ and $\bar{\bar{y}}_h$.

Thus an estimate of the linear relationship will be

$$(24)\quad Y' = \hat{y} + \hat{\beta}_1 (X - \hat{x}).$$

where

$$(25)\quad \hat{\beta}_1 = \sum_1^L W_h^2 \frac{s_{hxy}}{n_h} \left(1 - \frac{n_h}{N_h}\right) \Big/ \sum_1^L W_h^2 \frac{s^2_{hx}}{n_h} \cdot \left(1 - \frac{n_h}{N_h}\right).$$

The formula (25) speaks for itself. We interpret the multiplying factor W_h^2 in both numerator and denominator as expressing the importance of the strata involved.

When there is proportional allocation, i.e. $n_h/\sum_1^L n_h = W_h$ for all h, then

$$(26)\quad \hat{\beta}'_1 = \frac{\sum_1^L W_h\, s_{hxy}}{\sum_1^L W_h\, s^2_{hx}}$$

a result independent of sample size. Further when $W_1 = 1$, and $W_h = 0$ for $h = 2,\dots L$, i.e. there is only one stratum, then, dropping the stratum identification subscript, we find

$$(27)\quad \hat{\beta}''_1 = \sum_{i=1}^{n} (x_i-\bar{x})(y_i-\bar{y}) \Big/ \sum_{i=1}^{n} (x_i - \bar{x})^2,$$

which is a familiar estimate obtained in the least squares method. The identity of this result with that of classical least squares regression is certainly reassuring.

By the use of the generalized k-statistics [1] it can be shown that

$$(28)\quad \frac{E(\hat{\beta}''_1) - \beta}{\beta} = \left(\frac{K_{40}}{K_{20}^2} - \frac{K_{31}}{K_{20}K_{11}}\right) \cdot \frac{(N-n)(Nn-n-N-1)}{n(n-1)N(N+1)} + 0(n^{-2}),$$

where $\beta = \sum_1^N (x_i - \bar{\bar{x}})(y_i - \bar{\bar{y}}) \Big/ \sum_1^N (x_i - \bar{\bar{x}})^2$ may be defined as the true regression coefficient and K_{40}, K_{31}, K_{20}, and K_{11} are parent bivariate k-statistics for a finite universe. Thus the bias of $\hat{\beta}''_1$ relative to the true β is of order 1/n. The proof is omitted to save space. Similar results hold for $\hat{\beta}_1$ and $\hat{\beta}'_1$ but are complicated by the weights, and additionally the sampling fractions involved in the case of $\hat{\beta}_1$.

The variance of Y', for any given X, is

$$(29)\quad V(Y') = V(\hat{y})+X^2\, V(\hat{\beta}_1)+V(\hat{\beta}_1\, \hat{x})-2X\, \mathrm{Cov}(\hat{\beta}_1, \hat{x}) - 2\, \mathrm{Cov}\, (\hat{\beta}_1\, \hat{x}, \hat{y})+2X\, \mathrm{Cov}\, (\hat{y}, \hat{\beta}_1),$$

Other than the expression for $V(\hat{y})$, which is similar to (19), the expressions for the remaining terms will be extremely complex and will not be exact when determined by series expansion methods. Thus, estimation of V(Y') for a given X poses a problem.

The difficulty of variance estimation can be circumvented by the technique of independent replication. Suppose the universe was not stratified and suppose we draw k independent replicated simple random samples (see [11] and [8]), each of the same size, and estimate k statistically independent lines

$$(30)\quad Y'_a = \hat{y}_a + \hat{\beta}_{1a} (X - \hat{x}_a) \quad (a = 1,2,\dots,k)$$

according to the foregoing theory. The additional subscript in (30) identifies the k different lines. Then an estimate of the Y-ordinate, for a given X, is simply

$$(31)\ \bar{Y}' = \frac{1}{k} \sum_{a=1}^{k} Y'_a$$

with variance

$$(32)\ V(\bar{Y}') = V(\bar{Y}')/k$$

an unbiased estimate of which is

$$(33)\ v(\bar{Y}') = \sum_{a=1}^{k} (Y'_a - \bar{Y}')^2 / \{k(k-1)\},$$

despite the complexity of the expression for $V(Y')$.

The method of independent replication also leads to a method for making probability statements about the median. If the sample size for each independent replicate is n, then we can have $\binom{N}{n}$ possible estimated lines. For a given X, the Y-ordinate will cut these lines in $\binom{N}{n}$ points, assuming that the slope of every estimated line is at angle other than 90^0. If $\binom{N}{n}$ is an even number we may define the median Y as the point midway between the $\{\binom{N}{n}/2\}^{th}$ Y-value and the $\{\binom{N}{n}/2 + 1\}^{th}$ Y-value. Recalling the results of classical nonparametric theory or arguing from first principles, we find with independent replicates, for a given X, that

$$(34)\ P\{Y'_{least} < Y_{median} < Y'_{largest}\} = 1-(\tfrac{1}{2})^{k-1}.$$

Order statistics other than the extremes may also be used in making probability statements.

The reasons for restricting a statement of probability to the median of the Y's are as follows:

(A) The de facto functional relationship is restricted only to the N discrete points so that the given X (corresponding to which inference about Y is desired) may not relate to any of the true X's.

(B) Nonetheless on the supposition that X may be a possible value we wish to know the probability limits of the corresponding Y-value. For given X, the only probability distribution of the Y-values is those of the corresponding Y-ordinates of the $\binom{N}{n}$ lines generated by the randomization procedure for selecting the $\binom{N}{n}$ possible samples. The statment of probability, of necessity, must therefore be restricted to this set of Y-values (for given X), and the only exact statement of probability we can make is about the median Y-value of this distribution.

REFERENCES

[1] David, F. N., Kendall, M. G. and Barton, D. E. Symmetric Function and Allied Tables, Cambridge: Cambridge University Press 1966.

[2] Eisenhart, C., "Roger Joseph Boscovich and the Combination of Observations," Actes des Symposium International R. J. Boscovich 1961(1962), 19-25.

[3] ____________ "The Meaning of "Least" in Least Squares", Journal of the Washington Academy of Sciences, 54 (1964), 24-32.

[4] Godambe, V. P., "A Unified Theory of Sampling Finite Populations", Journal of the Royal Statistical Society B, 17 (1955), 268-78.

[5] Godambe, V. P. and Thompson, M. E., "Bayes, Fiducical and Frequency Aspects of Statistical Inference in Regression Analysis in Survey-sampling", Journal of the Royal Statistical Society B, 33 (1971), 361-90.

[6] Konijn, H. S., "Regression Anaylsis in Sample Surveys", Journal of the American Statistical Association,57 (1962), 590-606.

[7] Koop, J. C., "Contributions to the General Theory of Sampling Finite Populations Without Replacement and with Unequal Probabilities", Ph.D. Thesis, North Carolina State University Library (1957). (Also in N. C. Institute of Statistics. Mimeo. Series 296, 1961).

[8] ___________ On Theoretical Questions Underlying the Technique of Replicated or Interpenetrating Samples. Procedings of the Social Statistics Section, American Statistical Association (1960), 196-205.

[9] _________ Bulletin of the International Statistical Institute, 40 (1963), 1062-64, (Contributions to the discussion on T. Cunia's paper).

[10] ___________ "On the Estimation of Functional Relationships for a Finite Universe". (Abstract) Annals of Mathematical Statistics 35 (1964), 930.

[11] Lahiri, D. B. "National Sample Survey No. 5. Technical Paper on Some Aspects of the Sample Design", Sankhyā B, 25 (1954) 268-316.

[12] Sheynin, O. B., "Origin of the Theory of Errors", Nature, 211 (1966), 1003-4.

A GENERALIZED APPROACH TO SAMPLING ON SEVERAL OCCASIONS

Carl F. Kossack, University of Georgia

In many sample survey situations interest centers on the behavior in a finite population of the mean of some response variable over several time periods. Often in such situations there is associated with each unit of the population one or more descriptive variables whose values are known for each unit of the population.

In considering the type of survey design that would be most efficient in the above circumstances certain concepts seem to be such as to contribute to improving the effectiveness of any design. These include:

(a) Stratification should be used such that the design allocates to each stratum but one sample observation.

(b) In designing the sample over time a rotational system should be followed and the estimate should make use of the correlation that exists over time.

(c) The available descriptive variables should be used so as to provide the most efficient stratification possible.

(d) Regression estimates should be used whenever possible.

The technique to be presented herein involves all four of these features but its complexity is such that formal analysis of the procedure has as yet to be accomplished. However the technique has many characteristics which from a heuristic point of view seem most desirable. In order to keep the basic nature of the procedure from being obscured by the algebraic notation an actual numerical illustration will be presented. In addition to providing a vehicle by which the technique can be described the numerical example has also been used to study in a very limited manner the operational characteristics of the techniques using a Monte Carlo approach.

In considering the stratifications that are evolved over two time periods, t and t + 1, the rotational feature will result in some strata being "matched" in as much as the same sample unit will be used in each time period. The other strata will be "unmatch" since a different unit will have been selected for time period t + 1. To use the estimation procedure we require two estimates of the stratum mean for each time period. These are acquired following the formulae given below and then they are combined into a single estimate using the variances as weights.

For an "unmatched" stratum the following estimates are generated.

$$\hat{Y}'_{h,\ t+1} = r_{t+1}\ \hat{Y}_{h,t}$$

where $$r_{t+1} = \frac{\sum_{h_m} Y_{h,\ t+1}}{\sum_{h_m} Y_{h,\ t}}$$

($\sum_{h_m}$ indicates summation over "matched" stratum)

$\hat{Y}''_{h,\ t+1} = y_{h,\ t+1}$, the new sample observation

then $$\hat{Y}_h = \frac{1}{\frac{1}{V'} + \frac{1}{V''}}\left(\frac{\hat{Y}'_{h,\ t+1}}{V'} + \frac{\hat{Y}''_{h,\ t+1}}{V''}\right)$$

where $V' = r^2\hat{S}^2_{h,t}$ and $V'' = \hat{S}^2_{h,\ t+1}$

and $\hat{S}^2$ is obtained from the estimated Y's

thus $$V(\hat{\bar{Y}}_{h,t+1}) = \frac{V'V''}{V'+V''}$$

For "matched" stratum the approach used is the standard regression estimate but applied separately to each stratum and thus

$$\hat{\bar{Y}}_{h,t+1} = y_{h,t+1} + \hat{b}\ (\hat{\bar{Y}}_{h,t} - y_{h,t})$$

where $$\hat{b} = \frac{\sum_{h_m} y'_{h,t}\ y'_{h,\ t+1}}{\sum_{h_m} y^{2'}_{h,t}}$$

and
$$V\ (\bar{Y}_{h,t+1}) = (1 - \frac{1}{N_h})\ \hat{S}^2_h\ (1 - \hat{\rho}^2)$$

where $$\hat{\rho}^2 = \frac{\sum_{h_m} y'_{h,t}\ y'_{h,t+1}}{\sqrt{\sum_{h_m} y'^2_{h,t} \sum_{h_m} y'^2_{h,t+1}}}$$

(primed letter indicates deviation from mean)

To develop the illustrative example a finite population of N=80 units was developed and only three time periods, t=0, 1, and 2, considered. Births and deaths were not allowed and it was assumed that the descriptive variables, which were limited to two, did not change values during the time periods. The population data were generated using random normal deviates and assumed regression models. The descriptive variables X_1 and X_2 were assumed to be independent and were generated as random samples for normal populations with mean and standard deviations equal to 10, 5 and 30, 10 respectively. The three response variables Y_0, Y_1, and Y_2 were then generated from the descriptive variables using the regressions

$$Y_0 = 100+.4x_1+.6x_2+\varepsilon_0$$

with $\varepsilon_0 \sim$ NID (0,10)

$$Y_1 = 120+.5x_1+.5x_2+\varepsilon_1$$

with $\varepsilon_1 \sim$ NID (0,15)

$$Y_2 = \varepsilon_3 Y_1$$

with $\varepsilon_3 \sim$ uniformly (1.00 to 1.20)

Typical population values were

i	X_1	X_2	Y_0	Y_1	Y_2
1	3	11	75	82	90
2	9	27	124	142	153
..	...	....	.....	.....	.....
80	21	30	149	175	196

The two descriptive variables were used to design a sample for estimating the mean level of Y for the three time periods. The computational steps are described below assuming a given sample size, n, constant over time and a fixed rotation fraction, λ. In the illustrative example the sample size was taken to be 8 and the rotational fraction 3/8.

Step 1. Generate the estimates of Y_0 for each unit of the population using an assumed regression of Y_0 on X_1 and X_2.

In our illustration we assumed that the actual regression was known and thus

$$\hat{y}_{0i} = 100+.4x_{1i}+.6x_{2i}$$

Step 2. Stratify the population by ordering by $\hat{Y}_0$ and evolve optimum unit sample allocation by determining strata division points so that $W_h \hat{S}_h$ are all relatively constant.

In our illustration we used the $\hat{Y}_0$'s to obtain the $\hat{S}_h$'s and used an interactive computer program for the strata division determination. The analyst used his judgement in varying the stratification. It took on the average of eight iterations to acheive optimum stratification starting from stratum of equal size.

Step 3. For the initial time period a double sample was used to provide the required stratum mean estimates and an estimate of its variance.

The estimate of the population mean is then obtained as $\Sigma W_h \bar{Y}_h$.

Step 4. For the new time period, t=1, use the assumed regression of Y_1 on X_1, X_2 and Y_0 to obtain $\hat{Y}_1$'s and repeat stratification step (2).

In our example we assumed that

$$\hat{y}_{1i} = 120 + 0.5\ x_{1i} + 0.5\ x_{2i}$$

Step 5. Identify the stratum for which a new sample unit is to be selected (unmatched) by first including all stratum for which the units in the t=0 sample transferred out of the stratum and no sampled units transferred in. If additional stratum are needed drop other stratum out randomly to select them. In resulting unmatched strata randomly select a new sample unit.

Step 6. In the remaining strata randomly reduce the double sample to a single unit and survey that unit as the stratum observation for t=1.

Once the rotational process has been initiated the selection mechanism can utilize the concept of rotating out the observation that has been in the sampling plan the longest time and select for the new sampled unit that unit which has been out of the sample the longest time. Random selection would only be resorted to if ties occur.

The results of the initial Monte Carlo study can be summarized as given below:

MONTE CARLO STUDY RESULTS

Time Period	t=0	t=1	t=2
Population Mean	122.06	141.33	152.61
$\sigma\ (\bar{Y}_{srs})$	5.75	9.29	10.92
Mean of Estimates	122.3	142.4	155.43
Computed Standard Error	2.68	5.90	9.26
Efficiency, $V(srs)/V(\hat{\bar{Y}})$	460%	248%	139%

In considering the results obtained in this initial Monte Carlo study there are several observations that can be made.

(1) The loss of efficiency over time can generally be attributed to the loss of information contained in the two descriptive variable relative to the re-

sponse variable. This loss is accummulating as far as the data generation model is concerned and should not be encountered in practice.

(2) The increase in the apparent bias of the estimate over time may also be reflecting the data generation model used but may only be reflecting the influence of one or two large strata estimators obtained in the limited sampling program.

(3) Since the estimated Y's evolved from the assumed regression fail to reflect the error variance in the original population data, the use of $\hat{S}_h$ as an estimate of the within strata standard deviation can involve a significant negative bias. This underestimation may be of modest importance when the standard deviation estimates are being used for weights, but if they are used to estimate standard errors the bias results in a critical underestimation.

In conclusion there are several areas that demand attention if one is to obtain a better appreciation of the effectiveness of this new approach to sample survey design. These include:

(1) A critical study of how the efficiency of the estimator is affected by the several population parameters and in particular the error variance of the assumed regression.

(2) An examination of the appropriateness of the finite correction term, $1 - \frac{1}{N_h}$, used in the formulation since preliminary studies indicate that this correction may be subject to large errors when the stratum is small.

(3) Methods of estimate the standard error from the available data need to be evolved. One possible approach now under study is to evolve appropriate finite population distribution models from which such estimates can be made.

(4) There is need for a method of determining the two critical design characteristics n and λ if designs are required to meet error variance criteria.

(5) The development of an algorithm that can be used with the computer to evolve the optimum stratification since this would eliminate the need for the interactive computer approach now being followed.

RACE-SEX DISCRIMINATION INDICES FOR NON-METROPOLITAN STATE ECONOMIC AREAS

G. K. Kripalani, Western Michigan University

INTRODUCTION

One major group of factors influencing internal migration decisions stems from the "socio-cultural environment" of the areas of origin of potential migrants and their anticipated evaluation of corresponding elements in the areas of prospective in migration. Differences in internal net migration patterns may be analyzed by non-linear iterative least squares estimation procedure to isolate a component which reflects the impact of forces of relative race-sex discrimination in an area vis-a-vis the rest of the nation and which represents the net migration rate that would occur if relative economic opportunity factors in this area were as good as in the rest of the nation. This component of net migration is used as a basis of the definition of the race-sex discrimination index of a state.

The principal premise that underlies this study is that there are at least a few major independent variables affecting net migration and that some of these are non-measurable or non-observable, and that valid data series for such variables do not exist for use in empirical investigations. The method of analyses used is, therefore, designed to recognize and take into account this problem of nonobservability of some of the major explanatory variables. It is further recognized that net migration behavior patterns vary between the races, between the sexes and between age groups within each race-sex category. Consequently, there is need for stratification of an area's population into reasonably small homogeneous age, sex and race groups.

It is hypothesized that factors influencing internal net migration decisions of an age-sex-race group are of three categories:

1. Sub-area-related relative opportunity factors. These factors are the same for all age groups within a race-sex category. These relative opportunity factors are represented by an omnibus variable Z_t which is an index representing all relevant sub-area related relative opportunity factors. It is assumed that Z_t, which is the independent variable, is non-observable.

2. Age-related relative opportunity factors. These factors do not vary over a sub-area t in cross-section analyses, but vary between age groups within a race-sex category. Such age-related relative opportunity factors are denoted by a nonobservable index m_i where i refers to age group.

3. Race-sex-related relative opportunity factors. These factors do not vary over a sub-area in cross-section analyses or between age groups within a race-sex category. But these factors vary between race-sex categories and they reflect the impact of relative race-sex discrimination elements of the socio-cultural environment of the state in question.

In practical language the model separates net migration into three components: race-sex discrimination effect 'autonomous' component α' which is the same for all age groups within a race-sex category. This component would reflect the amount of net migration that would occur if $Z_t = 1$ and $m_i = 1$, that is, if net migration induced by relative opportunity and age-related factors were zero. It is this component which is defined to reflect "race-sex discrimination effect." There are two induced effects, one representing response to area-related omnibus independent variable Z_t representing relative opportunity factors and the other to age-related factors m_i.

For the purpose of this study, we may define the "race-sex discrimination" index of an area as the race-sex related component of internal net migration of that area (component α'). It is, however, recognized that the subset S_1 of elements of a socio-cultural environment S giving rise to what is called "race-sex discrimination" may consist of two types of elements--subset S_{11} consisting of elements which are the same for all age groups within a race-sex category (component α') and a subset S_{12} consisting of elements which vary between age groups within a race-sex category (component m_i). The latter component may reasonably be thought of, in given situations of being the result of "race-sex discrimination" and should appropriately be attributed to it.[1]

A real difficulty comes in the interpretation of the significance of the forces represented by m_i. Some of the forces underlying m_i may stem from those elements of the "socio-cultural environmental" complex as may be said to represent "race-sex discrimination," while it may legitimately be argued that some of these age-related factors stem from the fact that the assumption of a common index of relative opportunity facing all age-groups is unrealistic and that the index of relative opportunity is a function of both t and i. In such a situation,

Z_t would represent an average index of relative opportunity and a part of m_i would represent departures of the omnibus variable for the age group from the average Z_t for the category. Under these conditions, it would be necessary to identify the two subsets of the elements underlying m_i; those that relate to race-sex discrimination and those that reflect the situation that the index of relative opportunity is both age-related and time-related.

The "race-sex discrimination" index of an area may be viewed as a measure of the net migration effects of factors other than age and area related factors. Viewed thus, a comparative analyses of α''s may enable us to answer questions such as, (1) Are females "potentially" more migratory than males when the influences of age-time related factors are eliminated or equalized out or are Southern nonwhite males potentially more migratory than the Southern nonwhite females? and (2) Does the socio-cultural environment of a state discriminate against females or against nonwhites?

The significance of the positive or negative sign of α' may be clearly understood. Since total internal net migration of a color-sex category for the nation as a whole must be zero, it is easy to see that for each race-sex category:

$$\sum_s \alpha'_{cs} \cdot W_{cs} = 0 \ (c = WM,\ WF,\ NM,\ NF)$$

where α'_{cs} equals race-sex discrimination index of category c in state s, and W_{cs} equals proportion of category c population in state s (as proportion of the total category population in the nation).

Consequently, index α' is an index of relative "discrimination," in relation to the average for the nation which is zero. A positive α' does not signify that "discrimination" however defined, is absent in that state; it only signifies that "discrimination," if any in this state, is less than the average for the nation as a whole.

The results reported in this paper are with reference to the third component of internal net migration which stems from forces which are constant over age groups within a race-sex category, but which vary between the four race-sex categories, namely, white males (WM), white females (WF), nonwhite males (NM), and nonwhite females (NF). These results are based on the analyses of nonmetropolitan state economic areas (NSEA) data for 1950-60 decade (Table 1). Results based on the analyses of 1950-60 net migration data for metropolitan state economic areas (MSEA) as units of study were reported in a paper read at the 1970 Detroit meetings.

The nonlinear iterative least squares estimation procedure developed by Johnston and Tolley [1] in their study "Supply of Farm Operators" was used to estimate values of model parameters α', and m_i and the nonobservable variable Z_t. The basic properties of this model were, however, crucially different in some respects from the properties of Johnston-Tolley model and consequently necessary modifications were introduced in evaluation procedures. Estimation method is dealt with in Section II.

Empirical Results

Some interesting results were: (a) Inter-sex comparisons showed that NSEA's which are predominantly rural areas are relatively more favorable to white males than white females. For nonwhites the evidence was not clear. (b) Inter-racial comparisons between white and nonwhite females provided no clear evidence; for males, however, there was some evidence that NSEA's are relatively favorable to white males than to nonwhite males. (c) The indices of relative race-sex discrimination were generally negative except for some interesting cases, viz., (i) white females for North Carolina, (ii) both nonwhite males and nonwhite females for Virginia and Mississippi; and (iii) nonwhite females for Florida. These exceptional cases indicate that there would be positive net inmigration of these categories into the NSEA's of these states, if the opportunity factors in these areas were as good as in the rest of the nation.

Section II

Model and the Method of Estimation

Consider the model:

$$Y'_{it} = a'_i + \beta_i Z'_t + e'_{it} \qquad (1)$$

where primed variables represent logarithms of the original variables. The subscript i refers to the age group and subscript t refers to the sub-area in cross-section analyses. In this model the value of the dependent variable Y'_{it}, which depends upon i and t, is known while the independent variable on the right-hand side Z'_t which is nonmeasurable and hence unknown, is independent of i and is a function of t alone. Thus, the analysis of net migration data by the use of the above model would imply a critical assumption viz., that the net migration rates for different age groups i = 1, 2, ... i in a given nonmetropolitan state economic area (NSEA) are all functions of the same variable Z_t. This means that all the age groups in a given NSEA face the same index of relative opportunity. While the response coefficients for different age groups will be different, the critical assumption is that the independent non-observable omnibus variable to which these age groups are responding is the same.

Johnston and Tolley (1968) in their analysis of the supply of farm operators investigated a model of the above form using the NILES iterative procedure. An essential property of this iterative procedure is that the sequence of parameter estimates obtained at various iteration stages converges to the underlying parameter value not in the absolute sense, but in a relative sense. Hence, the character of parameter estimates obtained at the final stage is not cardinal but ordinal.

Consider the hypothesis: The number of persons of age i who will be staying in a nonmetropolitan state economic area t at the end of a decade (when the only cause of decrement or increment operating on the group is net migration) is a function of the index of relative opportunity Z_{it}, the exposed to risk of net migration E_{it} being the proportionality factor. This hypothesis gives rise to the multiplicative model of the form:

$$P_{it} = a_i \, E_{it} \, Z_{it}^{\beta_i} e_{it} \qquad (2)$$

where $P_{it} = E_{it} + M_{it}$ (3)

The notation is:

E_{it} = Population of age group i exposed to risk of net migration in sub-area t during the decade, i.e., the population of age group i which would be in area t in the absence of any net migration.

M_{it} = Net migration of age group i into or out of sub-area t during the decade. M_{it} is positive when there is net inmigration and negative when there is net outmigration.

P_{it} = $E_{it} + M_{it}$ = Population of age group i in sub-area t at the end of the decade (if the only cause of decrement or increment operating on the group was net migration).

Y_{it} = $P_{it}/E_{it} = 1 + M_{it}/E_{it}$ = 'Survival' rate against net migration where P_{it} is the quantity of supply of population of age group i in sub-area t and E_{it} is the supply shifter.

Z_{it} = Nonobservable independent variable representing the index of relative opportunity.

β_i = Age group i's net migration response coefficient (elasticity) to relative opportunity index Z_{it} facing it.

a_i = Constant term for age group i.

e_{it} = Disturbance term.

(2) may be written as

$$Y_{it}' = a_i' + \beta_i \, Z_{it}' + e_{it}' \qquad (4)$$

where primed variables represent logarithms of the original variables. Note that (4) will become (1) if it is assumed that $Z_{it} = Z_t$ for all i.

Consider a given sex-color group in a NSEA in a state, say white males in a particular NSEA t in state s. This group is subdivided into nine age groups, 0-9, 10-14, 15-19, 20-24, 25-34, 35-44, 45-54, 55-64 and 65+ at the start of the decade 1950-60. The data regarding the number of net migrants (M_{it}) for each age group for each of the several NSEA's in each state for 1950-60 decade with the "appropriate" exposed to risk of net migration (E_{it}) were taken from the statistics published by the Economic Research Service of the U.S. Department of Agriculture (1965) in their Population Migration Report giving net migration numbers and rates by age, sex and color separately for metropolitan and nonmetropolitan state economic areas. The values of $Y_{it} = (M_{it} + E_{it})$ are given for $i = 1, 2, \ldots, 9$ and $t = 1, 2, 3, \ldots,$ depending upon the number of NSEA's in a state.

The procedure consists of taking logarithm of the functional relationship $Y_{it} = a_i Z_t^{\beta_i} e_{it}$ and minimizing the sum of squares of the random term e_{it}' ($e_{it}' = \log e_{it}$), the summation being a double summation over i and t. The process starts with an arbitrarily selected set of values for Z_t, $t = 1, 2, \ldots, t$, ($N_t = t$).

$$(5) \quad SSE = \sum_i \sum_t e_{it}'^2 = \sum_i \sum_t (Y_{it}' - a_i' - \beta_i Z_t')^2$$

To obtain least squares estimates for parameters a_i' and β_i, we have the usual normal equations by taking partial derivatives of relation (5) with respect to a_i' and β_i, and solving these we have:

$$(6) \quad \hat{a}_i' = \frac{\sum_t Y_{it}' - \beta_i \sum_t Z_t'}{N_t}$$

$$(7)\quad \hat{\beta}_i = \frac{N_t \sum_t Y'_{it} Z'_t - \sum_t Y'_{it} \sum_t Z'_t}{N_t \sum_t Z'^2_t - \left(\sum_t Z'_t\right)^2}$$

Setting the partial derivative of SSE with respect to Z'_t equal to zero, we obtain an additional normal equation which when solved for Z'_t gives the least squares estimate of Z'_t in terms of a'_i, β_i and Y'_{it}.

$$(8)\quad \hat{Z}'_t = \left(\sum_i \hat{\beta}_i Y'_{it} - \sum_i \hat{a}_i \hat{\beta}_i\right) \Big/ \sum_i \hat{\beta}_i^2$$

Thus, the estimate of the value of Z'_t is obtained in terms of the estimated values of β_i and a'_i. To summarize, the process consists in starting with an assumed arbitrary set of values for Z'_t. Using this set, we arrive in the usual way, via normal equations, at the least squares estimates of a'_i and β_i. Then using these estimates of a'_i and β_i and the given values of Y'_{it}, we obtain the new set of estimates for Z'_t from equation (8). This completes the first iteration and we are at the second "stage" of the iteration process having given values of Y'_{it} as before but a new set of values of Z'_t. The process is repeated and the new set of Z'_t used to obtain a new set of a''s and β's in the first step of the second iteration and the process is repeated until estimates are approximately equal from the K^{th} and $(K+1)^{th}$ iteration. We will assume, for our purpose, certain convergence properties of this iterative procedure and the following results demonstrated by Johnston and Tolley (1968). The final estimates of a'_i, β_i and Z'_t depend upon the initial set of Z'_t's arbitrarily chosen as the starting point of the iterative procedure. It can, however, be shown that the ratios of β's, the ratio of differences for the Z'_t's and a certain linear function of a''s and β's have the property of convergence to the underlying value. Thus,

$$(9)\quad \lim_{s\to\infty} \frac{\beta_i^s}{\beta_j^s} = \frac{\beta_i}{\beta_j}$$

$$(10)\quad \lim_{s\to\infty} \frac{Z'^s_1 - Z'^s_j}{Z'^s_1 - Z'^s_2} = \frac{Z'_1 - Z'_j}{Z'_1 - Z'_2}$$

$$(11)\quad \lim_{s\to\infty} \frac{a'^s_\ell \beta^s_k - a'^s_k \beta^s_\ell}{Y'_{\ell t} \beta^s_k - Y'_{kt} \beta^s_\ell} = \frac{a'_\ell \beta_k - a'_k \beta_\ell}{Y'_{\ell t} \beta_k - Y'_{kt} \beta_\ell}$$

where $Y'_{\ell t}$ and Y'_{kt} are the known values of the dependent variable and s refers to the parameter estimates at the s^{th} iterative stage.

An important assumption underlying Johnston's model ($Y'_{it} = a'_i + \beta_i Z'_t + e'_{it}$) was that the variable Z is independent of i and depends on t only; that is, all age groups within a color-sex category in a NSEA faced the same index of relative opportunity. The index is thus assumed to change over NSEA's in a cross-section analysis, but it does not vary from age group to age group within a color-sex category in a given NSEA. It is proposed to relax this assumption and to regard the variable Z as a function of both i and t, and to replace it by a less exacting assumption that the index of relative opportunity varies from age group to age group but the ratio of any two Z's is fixed over NSEA's, i.e., it does not change over NSEA's in a cross-section analysis (or over time in a time series analysis). Mathematically, this is equivalent to assuming

$$(4.1)\quad Z_{it} = k_i Z_t \quad (i = 1, 2, \ldots i, t = 1, 2, \ldots t)$$

Hence,

$$(4.2)\quad \frac{Z_{it}}{Z_{jt}} = \frac{k_i}{k_j} = f(i, j)$$

It is important to note that k_i's are relative and it will be valid to regard $k_i = 1$ for $i = i_0$ and express all other k_i's ($i = 0, 1, \ldots i, i \neq i_0$) in relation to $k_i = 1$ for $i = i_0$. The condition that k_i's are relative enables us to put a constraint on k_i's, e.g. $\prod_{i=1}^{i} k_i = 1$. Z_t is a sort of average of individual Z_{it}'s depending upon the constraint imposed on k_i's. If $\prod_{i=1}^{i} k_i = 1$, $\prod_{i=1}^{i} Z_{it} = (Z_t)^i$ or Z_t is the unweighted geometric average of Z_{it}'s.

The model may be written as:

$$(4.3)\quad Y_{it} = \alpha_i Z_{it}^{\beta_i} e_{it}$$

Substituting $Z_{it} = k_i Z_t$ and taking logarithm, we have

$$(4.4)\quad Y'_{it} = \alpha'_i + \beta_i (k'_i + Z'_t) + e'_{it}$$

$$(4.5)\quad = (\alpha'_i + \beta_i k'_i) + \beta_i Z'_t + e'_{it}$$

The problem now is to minimize $S = \sum\sum_{it} e'_{it}2$.

$$(4.6)\quad S = \sum\sum_{it} (Y'_{it} - \alpha'_i - \beta_i k'_i - \beta_i Z'_t)$$

Taking partial derivatives with respect to the unknown parameters α'_i, β_i and k'_i, (assuming the initial arbitrary set of values for Z'_t) and equating the expressions to zero to obtain the normal equations, we have:

$$(4.7)\quad \frac{\partial S}{\partial \alpha'_i} = \sum_t (Y'_{it} - \alpha'_i - \beta_i k'_i - \beta_i Z'_t) = 0$$

$$(4.8)\quad \frac{\partial S}{\partial \beta_i} = \sum_t (Y'_{it} - \alpha'_i - \beta_i k'_i - \beta_i Z'_t)(k'_i + Z'_t) = 0$$

$$(4.9)\quad \frac{\partial S}{\partial k'_i} = \sum_t (Y'_{it} - \alpha'_i - \beta_i k'_i - \beta_i Z'_t)\beta_i = 0$$

Now a difficulty arises in the solubility of this system. Equations (4.9) are the same as equations (4.7); and equations (4.8) simplify to:

$$\sum_t (Y'_{it} - \alpha'_i - \beta_i k'_i - \beta_i Z'_t) Z'_t = 0$$

because the terms $k'_i \Sigma_t (Y'_{it} - \alpha'_i - \beta_i k'_i - \beta_i Z'_t)$ vanish by virtue of (4.7). We are thus left with the following normal equations:

(4.10) $\alpha'_i \sum_t 1 + \beta_i k'_i \sum_t 1 + \beta_i \sum_t Z'_t = \sum_t Y'_{it}$

(4.11) $\alpha'_i \sum_t Z'_t + \beta_i k'_i \sum_t Z'_t + \beta_i \Sigma Z'^2_t = \sum_t Y'_{it} Z'_t$

that is,

(4.12) $(\alpha'_i + \beta_i k'_i) \sum_t 1 + \beta_i \Sigma Z'_t = \sum_t Y'_{it}$

(4.13) $(\alpha'_i + \beta_i k'_i) \sum_t Z'_t + \beta_i \Sigma Z'^2_t = \sum_t Y'_{it} Z'_t$

The solutions are:

(4.14) $\alpha'_i + \beta_i k'_i = \frac{1}{\Delta}\left[\sum_t Z'^2_t \cdot \sum_t Y'_{it} - \sum_t Z'_t \sum_t Y'_{it} Z'_t\right]$

(4.15) $\beta_i = \frac{1}{\Delta}\left\{-\sum_t Z'_t \sum_t Y'_{it} + \sum_t 1 \cdot \sum_t Y'_{it} Z'_t\right\}$

where

(4.16) $\Delta = \sum_t 1 \cdot \sum_t Z'^2_t - \left(\sum_t Z'_t\right)^2$

Denote

(4.17) $a'_i = \alpha'_i + \beta_i k'_i$

Then

(4.18) $k'_i = (a'_i - \alpha'_i)/\beta_i$

After the iterative procedure has run to termination, the estimates for a'_i and β_i are obtained. To solve this model, we have to solve system (4.18) containing i equations and 2i unknowns k'_i and α'_i.

To make the system solvable, two alternatives may be considered by imposing constraints on the model:

(4.19) Case I: $\alpha'_1 = \alpha'_2 = \ldots = \alpha'_i = 0$

(4.20) Case II: $\alpha'_1 = \alpha'_2 = \ldots = \alpha'_i = \alpha'$ and

$$k'_1 + k'_2 + \ldots + k'_i = 0$$

The rationale for a constraint on α'_i is dictated by an important consideration arising from the arbitrary nature of the estimates of these parameters. As Johnston has shown, the estimated values of α'_i's depend upon the arbitrary set of values chosen as the starting Z_t's for the iterative process. The regression line gives the intercept α'_i on the Y'_{it} axis when $Z'_{it} = 0$. On theoretical grounds, however, it may be argued that when $Z'_{it} = 0$, the component of net migration induced by relative opportunity factors will be zero; but an autonomous component of net migration may, however, still occur on account of factors other than relative opportunity factors which are assumed to have the same non-zero impact on all age groups within a color-sex category in a state. In the terminology of this study, this autonomous component represents the index of race-sex discrimination. Consequently, it is Case II which conforms to the hypothesis developed in this study.

The rationale for Case II may be argued this way. The quantities k'_i's represent relative magnitudes and hence one constraint can be imposed on k'_i's. Let $\sum_i k'_i = 0$ be assumed. Theoretical considerations in a study may warrant provision in the model for a constant term which will be the same for all age groups within a color-sex category in a state and which would reflect the impact of race-sex discrimination. Such an effect can be provided for in the model by the introduction of the parameter α' independent of i and t. This is equivalent to the assumption $\alpha'_1 = \alpha'_2 = \ldots = \alpha'_i = \alpha'$ though the significance of common α' is different from the significance of different a'_i's when they were included in the model. The alteration of the model reduces the number of parameters to be estimated to (i + 1). The number of equations i and the constraint $\sum_i m'_i = 0^2$ give i + 1 equations, thus making the system solvable. The new model designated as the Race-Sex Discrimination Effect Autonomous Component Model is:

(4.25) $Y_{it} = \alpha Z_{it}^{\beta_i} e_{it}$

Putting $Z_{it} = m_i Z_t$, taking logarithm and simplifying this assumes the form

(4.26) $Y'_{it} = \mu'_i + \beta_i Z'_t + e'_{it}$

where

(4.27) $\mu'_i = \alpha' + \beta_i m'_i$.

The model thus has the same basic form as that of Johnston's model. The iterative procedure yields estimates of μ'_i, β_i and Z'_t where the value of μ'_i will be the same as that of a'_i in Johnston's model. Using the constraint $\prod_i m_i = 1$ or $\Sigma m'_i = o$, and simplifying, we have:

(4.28) $m'_i = (a'_i - \alpha')/\beta_i$

and

(4.29) $\alpha' = \sum_i (a'_i/\beta_i) / \sum_i (1/\beta_i)$

(4.30) $= \sum_i a'_i c_i / \sum_i c_i$

where

$$c_i = 1/\beta_i$$

α' is thus the weighted arithmetic mean of a'_i's the system of weights being $c_i = 1/\beta_i$, the reciprocal of β_i. α' may be written as $\alpha' = \bar{a}'_i(c_i)$ to denote the weighted average of a'_i's, the weights being c_i.

Now

$$(4.31) \qquad m_i' = (a_i' - \alpha')/\beta_i$$

$$(4.32) \qquad = c_i \; (a_i' - \alpha')$$

$$(4.33) \qquad = k_i' - c_i \alpha'$$

Relation (4.33) is the relationship between m_i' of Case II and k_i' of Case I.

In terms of the parameters of the initial multiplicative model, viz.

$Y_{it} = \alpha m_i^{\beta_i} Z_t^{\beta_i} e_{it}$, we have $\mu_i = \alpha m_i^{\beta_i}$

that is,

$$(4.34) \qquad m_i = \left(\frac{\mu_i}{\alpha}\right)^{1/\beta_i} = \left(\frac{\mu_i}{\alpha}\right)^{c_i}$$

The constraint $\sum_i m_i' = 0$ is equivalent to $\prod_i m_i = 1$. Therefore, $1 = \prod_i m_i =$

$$\prod_i \left(\frac{\mu_i}{\alpha}\right)^{c_i} \quad \text{or} \quad \alpha^{\sum_i c_i} = \prod_i \mu_i^{c_i} \quad \text{or}$$

$$(4.35) \qquad \alpha = \left[\prod_i \mu_i^{c_i}\right]^{1/\sum_i c_i}$$

i.e. α is the weighted geometric mean of μ_i's (or a_i') the system of weights being c_i, reciprocal of β_i. α being known, m_i may be estimated by the relation $m_i = (\mu_i/\alpha)^{c_i}$.

Our object is to know the estimates of α (or α') and m (or m') in the final iteration which serves as the solution. It is not necessary to estimate α and m_i at intermediate stages in the iterative procedure. The iterative procedure for the Race-Sex Discrimination Effect Autonomous Component Model is the same as that for Johnston-Tolley model. It is only when the estimates of β_i and a_i' (or μ_i in Case II) are obtained in the final iteration of the procedure that values of α' and m_i' may be estimated by the relationships developed.

Like β and Z parameters, m' can be estimated in relative terms only. In the case of m''s, the invariant quantity is given by

$$(4.38) \qquad \frac{(Y_{jt}'/\eta_{ji} - Y_{it}') - \alpha\,(1/\eta_{ji} - 1)}{Y_{\ell t}'/\eta_{\ell i} - Y_{it}') - \alpha\,(1/\eta_{\ell i} - 1)} = \frac{m_j' - m_i'}{m_\ell' - m_i'}$$

Hence under certain assumptions when either η_{ji} and $\eta_{\ell i}$ are nearly equal or the quantities $\alpha\,(1/\eta_{ji} - 1)$ and $\alpha\,(1/\eta_{\ell i} - 1)$ are small in relation to the respective terms in the numerator and the denominator of (4.38), the invariant property of $\dfrac{m_j' - m_i'}{m_\ell' - m_i'}$ would remain and m_j' and m_ℓ' could be compared in the ordinal sense.

Strictly speaking, α's for two color-sex categories are not comparable even in the ordinal sense, due to change of origin in each case, since the estimated values come from two different regressions and each α' has its own specific invariant function. Under certain conditions, however, if the Z's between states are in fact similar, the comparisons between α's for two color-sex categories are valid.

Appendix I. Race-Sex Discrimination index, by State, based on cross-section analyses of net outmigration data for Non-metropolitan State Economic Areas, 1950-60 decade.

Region/State	White: Male	White: Female	Nonwhite: Male	Nonwhite: Female	$H_{1.} \overset{D}{=}$ (2)-(3)	$H_{2.} \overset{D}{=}$ (4)-(5)	$H_{.1} \overset{D}{=}$ (2)-(4)	$H_{.2} \overset{D}{=}$ (3)-(5)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
New England					(2, 0)			
Maine	.0088	-.0316			1			
New Hampshire	.0226	-.0272			1			
Middle Atlantic					(1, 1)			
New York	-.0140	-.0163			1			
Pennsylvania	-.0115	-.0042			0			
East No. Central					(2, 1)			
Indiana	-.0216	-.0269			1			
Illinois	-.0249	-.0155			0			
Wisconsin	-.0002	-.0137			1			

Appendix I. (Continued)

Region/State	White		Nonwhite		$H_{1.}^{D} =$ (2)-(3)	$H_{2.}^{D} =$ (4)-(5)	$H_{.1}^{D} =$ (2)-(4)	$H_{.2}^{D} =$ (3)-(5)
	Male	Female	Male	Female				
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
West No. Central					(2, 4)	(2, 0)	(0, 1)	(1, 1)
Minnesota	-.0050	-.0100	.0400	-.1003	1	1	0	1
Iowa	-.0120	-.0060			0			
Missouri		-.0133	.0913	-.0031		1		0
North Dakota	-.0321	-.0205			0			
South Dakota	-.0146	-.0109			0			
Nebraska	-.0244	-.0008			0			
Kansas	-.0057	-.0067			1			
South Atlantic and D. C.					(2, 3)	(3, 3)	(3, 2)	(3, 2)
Virginia	-.0100	-.0056	.0031	.0015	0	1	0	0
West Virginia	-.0040	-.0051	-.0147	-.0172	1	1	1	1
North Carolina	-.0153	.0071	-.0170	-.0132	0	0	1	1
South Carolina	-.0240	-.0113	-.0450	-.0351	0	0	1	1
Georgia	-.0097	-.0121	-.0021	-.0064	1	1	0	0
Florida			-.0161	.0064		0		
East So. Central					(1, 1)	(2, 2)	(1, 1)	(1, 2)
Kentucky			-.0161	-.0080		0		
Tennessee		-.0085	-.0157	-.0785	1			
Alabama	-.0149	-.0155	-.0710	-.0047	1	0	1	0
Mississippi	-.0236	-.0069	.0091	.0071	0	1	0	0
West So. Central					(1, 3)	(1, 3)	(2, 2)	(3, 1)
Arkansas	-.0363	-.0362	-.0288	-.0200	0	0	0	0
Louisiana	-.0110	-.0115	-.0425	-.0316	1	0	1	1
Oklahoma	-.0255	-.0104	-.1548	-.0405	0	0	1	1
Texas	-.0453	-.0095	-.0222	-.0577	0	1	0	1
Mountain					(1, 3)			(1,0)
Montana	-.0212	-.0196		-.0680	0			1
Idaho	.0275	.0057			1			
Wyoming	-.0605	-.0135			0			
Colorado	-.0350	.0579			0			
Pacific					(0, 1)			
Washington	-.0166	-.0111			0			
Oregon	4.5990							
United States Total					(12, 17)	(8, 8)	(6, 6)	(8, 7)

Notes: An explanation is necessary for columns (6) through (9). Let a'_{11}, a'_{12}, a'_{21}, and a'_{22} refer to the values of race-sex discrimination index a', relating to white male, white female, nonwhite male and nonwhite female categories, respectively. (The first subscript refers to race, 1 for whites and 2 for nonwhites; the second subscript refers to sex, 1 for males and 2 for females.) Let us define:

$$H_{1.}\underset{=}{D} \quad a'_{11} - a'_{12} = 1 \text{ if } a'_{11} > a'_{12}$$
$$= 0 \text{ if } a'_{11} < a'_{12}$$
$$H_{2.}\underset{=}{D} \quad a'_{21} - a'_{22} = 1 \text{ if } a'_{21} > a'_{22}$$
$$= 0 \text{ if } a'_{22} < a'_{22}$$
$$H_{.1}\underset{=}{D} \quad a'_{11} - a'_{21} = 1 \text{ if } a'_{11} > a'_{21}$$
$$= 0 \text{ if } a'_{11} < a'_{21}$$
$$H_{.2}\underset{=}{D} \quad a'_{12} - a'_{22} = 1 \text{ if } a'_{12} > a'_{22}$$
$$= 0 \text{ if } a'_{12} < a'_{22}$$

It will be observed that $H_{1.}$ involves a comparison between the sexes among whites. $H_{1.}$ gets value 1 if a' for white males is greater than a' for white females. Value 1 for $H_{1.}$ signifies that the autonomous part of net migration pull or push, as the case may be, is greater for white males than for white females in the particular state. Similarly, $H_{2.}$ involves a comparison between the sexes in the case of nonwhites; $H_{.1}$ involves a comparison between males among the races and $H_{.2}$ involves a comparison between white females and nonwhite females.

Footnotes

*Important results based on data for metropolitan state economic areas (MSEA) as the units of study were reported in a paper read at the 1970 Detroit meetings and published in the Proceedings Volume of the Social Statistics Section.

Acknowledgment

**Main work on this research was done while the author worked at the N. C. State University on a National Science Foundation Project. "Area Population Adjustments in Relation to Economic Activity." The author gratefully acknowledges the help he received from C. E. Bishop, G. S. Tolley, C. H. Hamilton, and F. E. McVay. Results reported here are contained in author's report to NSF. However, work on the interpretation and significance of race-sex discrimination effect component of internal net migration has been done during author's stay at Western Michigan University.

[1] I am indebted to my colleague Louis Junker for suggesting that the variable m_i may represent and capture age-related factors of "race-sex discrimination" and that "race-sex discrimination" index need not necessarily be totally described by component α'.

[2] Since the estimated values of k_i' in Cases I and II leading to modified versions of Johnston's model be different, the problem of distinguishing as between them will arise. Hence, in Case II, we denote the k parameters by the letter m.

Principal References

1. Johnston, W. E. and Tolley, G. S., 1968. "The Supply of Farm Operators." Econometrica Vol. 36, No. 2 (April, 1968), p. 365-382.

2. Lyttkens, Ejnar, 1965. 'On the Fixed-Point Property of Wold's Iterative Estimation Method for Principal Components,"-- a paper presented at the International Symposium on Multivariate Analysis at University of Dayton, Dayton, Ohio, June 14-19, 1965.

3. Wold, Herman, 1965. "Nonlinear Estimation by Iterative Least Squares;" unpublished paper, Institute of Statistics, University of Uppsala, Sweden.

4. U. S. Department of Agriculture, 1965. States, Counties, Economic Areas and Metropolitan Areas, Volume I. Population-Migration Report; Net Migration of the Population, 1950-60, by Age, Sex, and Color. U. S. Government Printing Office, Washington, D. C.

THE STABILITY OF VOTER PREFERENCES IN TWO SOUTHERN STATES: AN ANALYSIS OF PANEL DATA WITH ATTRITION

Robert G. Lehnen and Gary G. Koch, University of North Carolina

Every four years millions of Americans spend considerable amounts of their time, energy, and money to influence the choice of the next President. The cost of this choice is enormous, especially during the three or so months of campaigning following the selection of candidates at the major party conventions. Given the social and political importance accorded these collective decisions, one needs to ask an obvious question: are a sufficient number of preferences changed over the course of the campaign to make a difference in the outcome?

The research pertaining to the problem of changing voter preferences during election campaigns is relatively sparse. The only major study by Berelson, Lazarsfeld, and McPhee is based on a sample of a single community in upstate New York during the 1948 election. [1] Its findings are based on a multi-wave panel design, and the analysis in most cases uses only part of the total data. Such panel designs have several attractive features, including the ability to analyze gross (individual) as well as net (marginal) change and the added power of inference derived from repeated measurements on the same respondent, but their major statistical drawback is one of uncontrolled attrition. Even the best planned and executed design cannot obtain complete information on all individuals for each time period. Such factors as cost, time, the mobility of the respondent, and his overall cooperativeness operate to produce incomplete data on some individuals for selected time periods. The more common analytical choice when confronting incomplete panel data is to analyze only a subset of respondents with more or less complete data, although a unified approach using the total information available would be more desirable. We shall present such an approach based on the general linear analysis of categorical data originally introduced by Grizzle, Starmer, and Koch called the "supplemented marginals" approach. [4] This technique is designed to test hypotheses about net (marginal) change across time periods even where data may be incomplete for some respondents. The example presented below is based on two three-wave panel designs administered during the 1968 presidential election campaign in Florida and North Carolina.

The Supplemented Marginals Approach

In a three-wave panel design seven combinations of data are possible. One type represents the situation of complete information while the other six types have some form of incomplete data. In most applications, the supplemented marginal approach treats each of the seven types of data as distinct subpopulations, within which certain relevant functions can be defined. Next, one constructs appropriate estimates of these functions and their associated variance-covariance matrix. Finally, statistical tests are undertaken to determine whether the estimates of such functions from the respective subpopulations may be regarded as having come from the same underlying population. In effect, one is asking whether the factors influencing the occurrence of missing data for some respondents and not others is related to the functions of interest.

For the sake of brevity, the approach will be illustrated for the sample data used in the example. Because of the design of the survey, only three patterns of data were present: respondents interviewed (1) at all three time periods, (2) at the first two time periods, and (3) at the first period only. However, it will be apparent from the generality of the discussion that this same methodology is applicable to those situations where other patterns of missing data are present. For additional details in this respect, the reader is referred to Koch, Imrey, and Reinfurt. [5]

The data for each state may be arrayed into a complex contingency table having 3 rows (subpopulations) and 64 columns. The first subpopulation has 64 possible combinations of response, resulting from a four level response (Humphrey, Nixon, Wallace, and Don't Know) measured at three time periods ($4^3 = 64$). The second subpopulation has only 16 possible combinations resulting from measurement only at times 1 and 2 ($4^2 = 16$). Finally, the third subpopulation has four combinations ($4^1 = 4$).

Three functions are of interest for characterizing the net level of preference for each candidate at each time period. These functions are as follows:

$$f_{t,1} = \log_e \left\{\frac{H}{N}\right\}$$

$$f_{t,2} = \log_e \left\{\frac{W}{H + N}\right\} \qquad (2.1)$$

$$f_{t,3} = \log_e \left\{\frac{D}{H + N + W}\right\}$$

where t = (1, 2, 3) and H = the proportion preferring Humphrey, N = the proportion preferring Nixon, W = the proportion preferring Wallace, and D = the proportion undecided.

These log-linear functions of the response data may be constructed by application of the categorical data analysis approach by the selection of the appropriate $\underset{\sim}{A}$ and $\underset{\sim}{K}$ matrixes. These matrices define 18 functions of interest, 9 for the subpopulation with complete data, 6 for the second subpopulation, and 3 for the third. For the second subpopulation, the 3 functions for the third time period (November) are undefined; for the third subpopulation, the 6 functions for October and November are undefined.

A supplemented marginals model was used to determine whether the 18 **functions could** be accounted for by 9 parameters (2.2). These nine

$$
\underset{(18\tilde{x}9)}{\tilde{X}} = \begin{bmatrix}
1&0&0&0&0&0&0&0&0\\
0&1&0&0&0&0&0&0&0\\
0&0&1&0&0&0&0&0&0\\
0&0&0&1&0&0&0&0&0\\
0&0&0&0&1&0&0&0&0\\
0&0&0&0&0&1&0&0&0\\
0&0&0&0&0&0&1&0&0\\
0&0&0&0&0&0&0&1&0\\
0&0&0&0&0&0&0&0&1\\
1&0&0&0&0&0&0&0&0\\
0&1&0&0&0&0&0&0&0\\
0&0&1&0&0&0&0&0&0\\
0&0&0&1&0&0&0&0&0\\
0&0&0&0&1&0&0&0&0\\
0&0&0&0&0&1&0&0&0\\
1&0&0&0&0&0&0&0&0\\
0&1&0&0&0&0&0&0&0\\
0&0&1&0&0&0&0&0&0
\end{bmatrix}
\quad \text{and} \quad
\underset{(9\tilde{x}1)}{b} = \begin{bmatrix}
b_{1,1}\\ b_{1,2}\\ b_{1,3}\\ b_{2,1}\\ b_{2,2}\\ b_{2,3}\\ b_{3,1}\\ b_{3,2}\\ b_{3,3}
\end{bmatrix}
\qquad (2.2)
$$

parameters represent the three candidate comparisons of interest for each of the three time periods. The remaining 9 degrees of freedom excluded from the supplemented marginals model provide means for testing the goodness-of-fit (GOF) of the model (2.2). This GOF test provides a comparison between the 18 observed values with the 18 values predicted by (2.2). When this test is not significant, one may interpret the results to mean that the model adequately characterized the data, or in other words, that the factors affecting the occurrence of missing data, in general, have not operated to make the estimated values of the functions from the various subpopulations different. When the test is significant however, it is evidence that the subpopulations are not homogeneous--that is, the estimates of the functions for each of the three subpopulations are statistically different.

When the model described in (2.2) was fitted to the statewide sample data for Florida and North Carolina, the GOF tests were significant in both cases, indicating that the factors affecting whether a respondent was interviewed once, twice, or three times were in some way related to candidate preferences.

When such a situation arises, the best strategy is to define subpopulations of interest that may adequately characterize the factors affecting the occurrence of missing data. Since the first wave of the sample was measured by a household interview while the second and third waves were completed by telephone, we suspected that occupation of the head of the household would be an efficient indicator for characterizing the likelihood of being reached by telephone for a reinterview. Also, we hypothesized that the race of the respondent would significantly affect the chance for a reinterview. These two factors together, no doubt, efficiently characterize such factors as social class, place of residence, and income that would affect one's chance for reinterview. As a result, each state sample was further stratified by race and occupation. Because the number of black respondents in each sample was small, only white respondents were further classified by six levels of occupation as follows:

WC	white; white collar and professional
SBM	white; small businessmen
SBC	white; skilled blue collar worker
UBC	white; unskilled blue collar worker
F	white; farmer or tenant worker
R	white; retired
B	black; --

These seven race x occupation subpopulations were defined for each of two states creating a total of 14 subpopulations. Within these 14 subpopulations were three further divisions--interviewed three times, interviewed twice, and interviewed once. Thus, a total of 42 distinct subpopulations were created. The initial task was then to fit the supplemented marginals model (2.2) to each of the 14 subpopulations defined by state, race, and occupation.

Only one of the 14 GOF tests was significant at the $\alpha = .10$ level (that is, white professionals in North Carolina), and we concluded that these supplemented marginal models provided an appropriate description of the data from the 14 subpopulations. Some further justification for this decision will be given in the next section. Otherwise, it should be recognized that those subpopulations where the fit is judged not to be satisfactory could be subdivided into smaller, and more homogeneous, aggregates within which better fits of the supplemented marginals model could be anticipated. Because of sample size requirements however, there are certain limits on the amount of subpopulation partitioning that may be done. Hence, a point is eventually reached where goodness-of-it and sample size criteria must be compromised. In sum, one must formulate a set of subpopulations that are of interest and make a decision as to whether the supplemented marginal model is appropriate to them.

Combining the Supplemented Marginals Models

Given the acceptable GOF test for the 14 state x race x occupation subpopulations, one may be interested in the importance of these three effects on the level of candidate preferences. We shall discuss the approach whereby state and time effects may be examined, although the method is general, and therefore, may be used to study race and occupation effects as well.

Each of the 14 models reported in the previous section has a 9-vector of parameters (p_{ij}) and a corresponding variance-covariance matrix ($\tilde{V}_{ij}$ where i = 1, 7; j = 1, 2). These 14 sets of parameters and variance matrices may be used to fit further models designed to characterize the importance of state, time, race, or occupation. In the case of state and time effects, the corresponding sets of parameters and variance matrices for each race x occupation group i may be combined in the following manner:

$F_i \hat{=} X_2 b_2$ where "$\hat{=}$" means "is estimated by",

$$\underset{(18 \times 1)}{F_i} = \begin{bmatrix} p_{i1} \\ p_{i2} \end{bmatrix}, \quad \underset{(18 \times 18)}{V_i} = \begin{bmatrix} V_{i1} & 0 \\ 0 & V_{i2} \end{bmatrix}$$

with 0 being a 9 x 9 matrix of zeroes.

Thus, F_i is the vector of parameters for race x occupation class i for both Florida and North Carolina; V_i is the appropriate variance matrix for F_i. The corresponding model X_2 includes state effects, time effects, and state x time interaction effects for each of the three candidate comparison functions.

The resulting GOF and test statistics provide criteria for decisions reducing the original model X_2 in a manner that efficiently characterizes the important sources of variation.

The models fitted for each of the seven race x occupation classifications produce predicted values for the functions defined in (2.1). These values may be transformed into the proportion preferring each candidate at the three time periods since

$$a = \log_e \{\frac{H}{N}\},$$

$$b = \log_e \{\frac{W}{H+N}\}, \text{ and}$$

$$c = \log_e \{\frac{D}{H+N+W}\} \text{ imply that}$$

$$P(\text{Nixon}) = 1 / k,$$

$$P(\text{Humphrey}) = e^a / k,$$

$$P(\text{Wallace}) = e^b (1 + e^a) / k,$$

$$P(\text{Undecided}) = e^c (1 + e^b) (1 + e^a) / k, \text{ and}$$

$$k = (1 + e^a)(1 + e^b)(1 + e^c).$$

The estimates of the level of preference for the three candidates are reported in Table 1.

TABLE 1

PROBABILITIES OF PRESIDENTIAL CANDIDATE PREFERENCE (1968) DERIVED FROM REDUCED MODELS BY RACE AND OCCUPATION SUBPOPULATIONS

		WC	SBM	SBC	UBC	F	R	B
FLORIDA								
SEPTEMBER	Nixon	.383	.249	.217	.189	.285	.367	.070
	Humphrey	.147	.166	.123	.111	.082	.216	.726
	Wallace	.290	.457	.488	.556	.462	.234	.037
	Undecided	.180	.128	.172	.144	.171	.183	.167
OCTOBER	Nixon	.402	.249	.263	.204	.285	.367	.070
	Humphrey	.159	.166	.148	.112	.082	.216	.726
	Wallace	.260	.457	.460	.531	.462	.234	.037
	Undecided	.180	.128	.129	.144	.171	.183	.167
NOVEMBER	Nixon	.359	.249	.310	.189	.285	.367	.070
	Humphrey	.230	.166	.175	.111	.082	.216	.726
	Wallace	.231	.457	.420	.556	.462	.234	.037
	Undecided	.180	.128	.095	.144	.171	.183	.167
NORTH CAROLINA								
SEPTEMBER	Nixon	.383	.232	.217	.254	.230	.367	.070
	Humphrey	.147	.156	.140	.150	.167	.216	.726
	Wallace	.290	.484	.514	.492	.432	.234	.037
	Undecided	.180	.128	.129	.104	.171	.183	.167
OCTOBER	Nixon	.402	.249	.288	.289	.275	.367	.070
	Humphrey	.159	.166	.124	.170	.035	.216	.726
	Wallace	.260	.457	.460	.397	.519	.234	.037
	Undecided	.180	.128	.129	.144	.171	.183	.167
NOVEMBER	Nixon	.359	.265	.284	.313	.315	.367	.070
	Humphrey	.230	.177	.183	.184	.081	.216	.726
	Wallace	.231	.429	.405	.306	.432	.234	.037
	Undecided	.180	.128	.129	.198	.171	.183	.167

Some Comments on the Effects of 1968 Campaign

These sample data provide evidence that the election campaign as applied to the electorate of these two southern states had no significant political effect. In fact, one can be impressed by the durability of net candidate preferences among almost all sectors of the voters. Except for Unskilled Blue Collar Workers in North Carolina, the candidate who enjoyed a plurality in September at the formal beginning of the campaign maintained his plurality through the campaign to election day. Although one can find statistically significant time effects among several classes and states, these time effects had no important consequences. The general trend of these campaign effects, however, is a slight loss of support for Wallace and some gain for Nixon. When one compares the magnitude of the X^2 associated with the time parameters with other effects though, one can conclude that the amount of variation associated with these parameters is relatively small. We should emphasize, however, that when we stress the lack of campaign (time) effects on these two state electorates, we are speaking about the net effect of possibly many individual changes. It is in this sense that we can say that the campaign made no political difference.

The election in Florida and North Carolina was between Nixon and Wallace. Although black voters overwhelmingly supported Humphrey, white voters supported either Nixon or Wallace. Two other studies support these findings--one in Tennessee [2] and the other nationally [3]. Nixon gained his support from the White Collar, Professional, and Retired workers, while Wallace held a broad base of support among all other classes--Small Businessmen; Blue Collar-Workers, skilled and unskilled; and Farmers. In the one case where the campaign made a politically significant difference (Unskilled Blue Collar Workers in North Carolina), Wallace lost his plurality and Nixon gained. Humphrey was never a serious contender in these two states, for his only sizable basis of support was among blacks.

Finally, one finds little evidence for the importance of different state "political cultures" on the stability of voter preferences [6]. One might anticipate "state" effects to occur from the many cultural and historical differences that make Florida different from North Carolina with respect to how people decide their choice for President. These state effects though unspecified would be characterized by significant "main effects" for state. No statistically significant state effects were found however. Thus, our findings give little support to the theory that state electorates differ markedly in their political processes (with respect to the level and stability of candidate preferences).

There is a more limited sense, however, in which state differences do appear. The relatively minor time effects are, in fact, different for each state within certain race x occupational classifications. The most dramatic example is for Unskilled Blue Collar workers: although Wallace maintained his plurality in Florida, he lost it in North Carolina. Other lesser examples of state x time effects may also be found for Small Businessmen, Skilled Blue Collar Workers, and Farmers. White Collar Workers and Professionals, Retired Workers, and Blacks, however, show no variations based on time within state.

In conclusion, let us stress that our analysis has focused on net political preferences and not political behavior. We have determined the level of choice for each of the major candidates in the 1968 election and not the proportion of the actual vote. Clearly, different race and occupational groups vote at different rates, and there is no direct relationship between preference and voting. Thus, one can anticipate slight discrepancies between the predicted levels of preference and election outcomes.

REFERENCES

We wish to thank Herbert M. Kritzer, who worked as research assistant on this paper.

[1] Bernard R. Berelson, Paul F. Lazarfeld, and William N. McPhee, Voting (University of Chicago Press, 1954).

[2] William M. Berenson, Robert D. Bond, and J. Leiper Freeman, "The Wallace Vote and Political Change in Tennessee," Journal of Politics 33 (May, 1971), 515-20.

[3] Philip E. Converse, Warren E. Miller, Jerome G. Rusk, and Arthur C. Wolfe, "Continuity and Change in American Politics: Parties and Issues in the 1968 Election," American Political Science Review 63 (December, 1969), 1083-1105.

[4] J. E. Grizzle, C. F. Starmer, and G. G. Koch, "Analysis of Categorical Data by Linear Models," Biometrics 25 (September, 1969), 489-504.

[5] G. G. Koch, P. B. Imrey, and D. W. Reinfurt, "Linear Model Analysis of Categorical Data with Incomplete Response Vectors," Biometrics 28 (September, 1972), forthcoming.

[6] Samuel C. Patterson, "The Political Culture of the American States," Journal of Politics 30 (February, 1969), 187-209.

QUALITY CONTROL OF STATISTICAL REPORTS

Paul S. Levy and Monroe G. Sirken, National Center for Health Statistics

1. Introduction

The National Center for Health Statistics (NCHS) develops and maintains systems capable of providing reliable, general purpose, national, descriptive health statistics on a continuing basis and publishes these statistics for the use of the health industry and related industries, both public and private [1]. Examples of such data systems are the national vital statistics of births, deaths, fetal deaths, marriages and divorces; sample surveys linked to birth and death records; a continuing nationwide survey of households by means of interviews; a series of national surveys based on physical examinations of samples of the population; surveys of hospitals and other health care facilities; as well as other periodic and ongoing surveys.

The principal form of output of the Center's work is published statistical reports. These come out in several series, one of which is entitled Vital and Health Statistics. This series often is referred to as the "rainbow series" since each data system has its own series of reports with distinctively colored jackets. In this series, some 35 to 40 substantive statistical reports are produced every year covering various aspects of the data collected in the systems mentioned above. Each substantive report contains a text which analyzes the data presented in a set of statistical summary tables, which are subject to sampling and measurement errors, estimates of which are present in the appendix. In order to ensure that the technical material presented in these reports meet specified standards and that the statistical statements made in the text are valid, a surveillance program in the form of a standardized procedure for reviewing such reports has been developed within the Office of Statistical Methods (OSM), which is the primary statistical support group within NCHS. The purpose of this article is to describe some of the main features of the quality control program for published statistical reports.

2. General Format of the Program

It should be mentioned that this surveillance program is not yet fully operational, and our experience is derived primarily from pilot projects. We have begun this program on a limited basis and expect it to become fully operational within the next several months.

At present, most of the reviews are being performed *ex post facto*, i.e., after the reports have been published. While this is not optimal, publication schedules do not allow these reports to be delayed too long for a statistical review. It is hoped, however, that as we gain experience and familiarity with the review procedures, and as we recruit additional personnel to do the reviewing, we can undertake to perform speedy reviews of each report prior to its publication.

As it stands now, however, a published report would be assigned for review to a junior or mid-level mathematical statistician within OSM soon after it was published. The review procedure which will be described below in more detail consists of two general parts. The first part deals with the review of the technical material, tables and figures in the report, while the second part deals with the review of statistical statements made in the report. An instruction manual has been drafted which guides the reviewer, step by step, in the procedure and ensures that all reviews are performed according to protocol [2].

When the procedure is completed, the findings are first reviewed by the senior statistician responsible for this surveillance program, and then they are discussed with the author of the report and/or the director of the subject matter division responsible for the report.

3. Review of the Technical Material, Tables and Figures

The purpose of this phase of the procedure is to ensure that (1) the report describes the essential design and estimation features of the data collection system, and (2) all statistics presented in the substantive tables meet prescribed standards of accuracy and precision, and have sampling errors available for them in the report.

In NCHS reports, the technical material describing the design and estimation features of the data collection system often is presented in a technical appendix but also may appear in the main part of the text. Specific items comprising this technical material which we feel should be adequately described in the report are the universe, frame, number of primary sampling units (PSU's), stratification, clustering, data collection and processing procedures, quality control procedures, estimation methods, methods of obtaining sampling and measurement errors, etc.

Standards for these items are given in the instruction manual, and the reviewer using a checklist systematically goes through the report and notes whether each technical item on the checklist meets the standards given in the instruction manual. For example, the description of the sampling frame in a substantive report would meet the standard set for it if it clearly states what are the enumeration units and elements, and if it makes reference to the NCHS publication describing the frame in detail (should such a publication exist).

Each substantive table and figure is checked for purposes of determining whether the statistics meet specified standards of precision and accuracy, whether the reader has sufficient information in the report to determine the sampling error of every statistic presented in the substantive table and whether the technical terms used in the table titles are defined clearly and accurately in the report.

4. Review of Statistical Statements

4.1 Overall Objective

The second major component of the review procedure is the review of the statistical statements made in the text of the report. Its purpose is to ensure that the inferences made in the statistical statements are based on sound statistical methodology and judgment.

The primary objective is to estimate the proportion of statistical statements made in the report that are valid, invalid or untestable. In order to achieve this objective, it was necessary to develop a methodology which will be described in the remainder of this section.

4.2 Statistical Statements

4.2.1 Definitions

We use the following as a working definition of a statistical statement:

A statistical statement is any phrase, clause, sentence or combination of words that makes an inference from sample observations about characteristics of a population. A statistical statement is valid if the inference made in the statement is justified statistically. If the inference is not justified, the statement is said to be invalid. If neither of these decisions can be made, it is said to be untestable.

A statement that is untestable may be adequate or inadequate. It is considered adequate if the inference made in it is clear but there is no usable statistical procedure for validating the statement. It is considered inadequate if the inference made in the statement is unclear.

For these definitions to be applied with any reproducibility, some objective criteria were decided upon for identifying statistical statements within the text of a report and for declaring a specific statistical statement valid, invalid, or untestable.

4.2.2 The identification and classification in a text

The main criterion for deciding whether a sentence, phrase, clause or group of words constitutes a statistical statement is whether an inference is drawn from a sample to a population. If more than one such inference is drawn, then the group of words would constitute more than one statistical statement. Using this criterion, a reviewer with a little instruction and practice can identify statistical statements in a report with reasonable reliability.

Since the basic characteristic of a statistical statement is that it draws a statistical inference, and since the tools that the reviewer uses to judge the validity of the inference depend on the type of inference implied in the statement, we have attempted to classify statements according to type of inference and to specify methods for testing each type of statement. After some effort, we have arrived at the following taxonomy of types of statistical statements:

Type 1. Quotation of Estimates. These are statements which characteristically involve only one subdomain and are generally, as the name implies, simple quotations of estimates. For example, the statement "On the basis of examinations, approximately 24 million U.S. children aged 6-11 years averaged an estimated 1.4 DMF per child" [3]. A Type 1 statistical statement is considered valid if its coefficient of variation is below the tolerance set by the appropriate subject matter division.

Type 2. Simple Comparisons. These are statistical statements in which comparisons are made between two domains. These statements are divided into two subtypes given below:

Type 2A. Simple Comparisons of Equality. These are statements which draw the inference that the level of a characteristic in one domain is different from that in another domain. For example, "White children had somewhat better levels of oral hygiene than Negro children. As a result, the average OHI-S (Simplified Oral Hygiene Index) was 1.41 for the former and 1.66 for the latter"[4]. A statement of this type is considered valid for our purposes if the difference between the two domains with respect to the quoted statistics is significant at the 5 percent level of significance as determined by the usual test of a difference between two means or proportions.

Type 2B. Simple Comparisons Involving Magnitude. This type of statistical statement is of the form "X_1'" is r times as large (or as small) as "X_2'" where X_1' is the estimated level of a characteristic in one domain, X_2' is the estimated level of the same characteristic in another domain and r the estimated ratio of X_1' to X_2'. This type of statement is considered valid if (1) r' meets the standards of precision given for Type 1 statistical statements and (2) the difference between X_1' and X_2' is significantly different from zero, as defined in the discussion of Type 2A statements.

Type 3. Comparisons Involving More Than Two Domains.

Type 3A. Comparisons Among Several Subdomains Within a Domain. These statements have the general form "within domain A, there were differences in the level of characteristic X among the subdomains A_1, A_2, and A_k." In judging the validity of such statements, a multiple comparison test based on the Bonferroni inequality is used [5].

Type 3B. Comparisons Between the Corresponding Subdomains of Two Domains. The inference in this type of statement is of the form "within each subdomain, the level of characteristic X is higher in domain A than in domain B. To judge the validity of this type of statement, differences between domain A and B in the level of characteristic X are tested for each of the k subdomains implied in the statement. If all k of these differences are statistically significant, the statement is considered valid. If r (< k) of these differences are significant, the number, r, is examined in a sign test table based on k signs. If r is greater than or equal to the critical value of the sign test statistic (for the 5 percent level of significance), then the statement is considered valid. Otherwise, it is considered invalid.

Type 4. Statements of Statistical Relationship.

Type 4A. Statements of "Trend". This type of statement makes an inference that there is an association between two variables whose domains of definition are on interval or ordinal scales. For example, "the relationship is an inverse one with the proportion of men and women of all races who need to see their dentist at an early date decreasing sharply with rising levels of yearly income [6]. The validity of this type of statement is judged by testing whether the linear component representing the relationship is significantly different from zero. Because of the complexities of data collected from complex surveys, a modified estimate of the linear relationship and a modified significance test was devised for this purpose [3].

Type 4B. Statements of Association Other Than Trend. This category would include any statement implying a relationship between two or more characteristics that cannot be interpreted as a Type 4A statement. Most of these statements imply a relationship between two variables, one or both of which are attribute or categorical variables. Other statements of this type express rather complex types of association. An example of such a statement is "The difference in the direction of the trend (in need for dental care) between Negro men and the other sex-race groups is not due to differences in the age composition of the various educational attainment groups." [6] This type of statement is often difficult to interpret and we have no standard protocol for testing statements of this type. Often what appears to be a Type 4B statement can be interpreted as another type or broken up into one or more types for which we have standard tests. However, many of these Type 4B statements elude interpretation or validation.

These are the main categories of statements that we have classified. From preliminary studies, we found that these four types account for almost 75 percent of the statements found in NCHS statistical reports, and that approximately two-thirds of all statistical statements made are testable. As we gain experience with this procedure, we expect to develop a more refined taxonomy of statistical statements and to develop further methodology for testing these statements.

4.2.3 Sampling of statistical statements. A main objective of the review of statistical statements is to obtain an unbiased estimate of the proportion of valid, invalid, and untestable statistical statements made in the report. Since resources do not permit the testing of every statistical statement, our estimates are based on a probability sample. The choice of what kind of sampling design to use presented us with an interesting statistical problem which is the subject of a separate article [7]. In summary, after experimentation with several types of sampling designs, we have chosen for reasons of logistics and cost efficiency, one which uses lines of text as the enumeration units, statistical statements as the elementary units and a conventional enumeration rule to link the enumeration units with the elementary units.

Once the required sample size, ℓ, of enumeration units is specified, a systematic sample of approximately ℓ lines of text is taken using a random start. After these are chosen, the reviewer examines each sample line and reads the entire paragraph overlapping the line. If one or more statistical statements appear on the line, each entire statistical statement is enclosed in brackets. A statistical statement may begin on a previous line and/or continue on subsequent lines. In such a case, the enclosed statement may overlap several lines in addition to the sample line. Once the statistical statements overlapping the sample line are identified all statistical statements which begin on the sample line are included in the sample of statements to be tested.

4.3 Estimation of the proportion of valid, invalid, and untestable statistical statements.

After each statistical statement has been tested and found to be either valid, invalid or untestable, inflation estimates are obtained of the total number of statistical statements in the report, and the total number of valid, invalid, and untestable statements. Finally, ratio estimates are obtained for the proportion of statistical statements which are valid, invalid, and untestable; and the estimated variances of these estimated proportions determined.

5. Comments

In the management of large government statistical systems, a considerable effort in the way of time and money goes into the quality control of the data collection, data preparation and data processing operations for the system. Effort, however, is needed also to assess the quality of the end product of the system, namely the substantive report. Since the responsibility for preparing substantive statistical reports generally lies in the hands of subject matter persons who are not professional mathematical statisticians, we feel that a surveillance of the technical items and statistical statements in these reports is necessary to ensure a high quality product. The procedure described above is a first attempt to formulate such a systematic quality control program.

Such a program, even done ex post facto, gives us information on the types of errors which analysts are making in their inferences and on inadequacies in their presentations of technical material. This information can serve as a useful resource in our planning of intramural training programs which would have as an objective the teaching of analysts to make clear, accurate and testable statistical statements in their reports. We feel that ultimately this would result in improved statistical reports.

One of the spinoffs which we hope to obtain from this program is in the analysis of the statements found by reviewers to be untestable. While some of these are untestable because of the lack of clarity, others are untestable because no methodology exists for making a test which is appropriate for data collected from complex surveys. It is hoped that as we catalogue these statements, we can encourage research in the development of methodology for testing hypotheses in data collected from complex sample surveys.

References

[1] U.S. Department of Health, Education and Welfare; The Mission and Policies of the National Center for Health Statistics.

[2] Levy, Paul S. and Sirken, M. G.; A Manual for the Review of Statistical Reports. Unpublished document.

[3] National Center for Health Statistics; Decayed, Missing and Filled Teeth among Children, United States. Vital and Health Statistics PHS Pub. No. 1000 - Series 11 No. 106. Public Health Service, Washington. U.S. Government Printing Office, August, 1971.

[4] National Center for Health Statistics; Periodontal Disease and Oral Hygiene Among Children - United States - 1963-65. (A forthcoming NCHS report).

[5] Miller, Rubert G. (1966). Simultaneous Statistical Inference. McGraw-Hill, New York.

[6] National Center for Health Statistics; Need for Dental Care Among Adults, United States, 1960-62. Vital and Health Statistics PHS Pub. No. 1000 - Series 11 - No. 36. Public Health Service, Washington. U.S. Government Printing Office, March, 1970.

[7] Sirken, M. G. and Levy, P.S.; Multiplicity estimation of proportions based on ratios of random variables. Unpublished manuscript (1972).

THE RELATIVE EFFECTIVENESS OF DENSITY-RELATED MEASURES FOR PREDICTING ATTITUDES AND BEHAVIORAL VARIABLES[1]

Robert W. Marans and Lewis Mandell, The University of Michigan

Over the past decade the social science literature has contained with increasing regularity reports of studies which deal entirely or in part with the effects of population density on human behavior and attitudes. While numerous studies have based their measures of population density on census or other aggregate data, others have considered the size of place of residence as a measure of how close people live to each other. In other words, lacking the means of obtaining an objective measure of how close people live together, many researchers have used size of place as a surrogate measure of residential density. In order to see if, in fact, size of place is a reliable measure of residential density, this paper will consider the relationship of size of place to the most objective density measure we could obtain - the actual number of dwellings per given unit of land. Next, less objective density-related measures which are easier to obtain will be considered in relation to our objective density measure. Finally, in an exploratory effort, we will examine the relative effects of each of our measures in predicting density-related dependent variables.

Background and Methodology

Prior to discussing relationships among our density-related measures and their relative effect in predicting two responses, a brief discussion of the sources of data including the procedures used in obtaining density measures is in order.

In the fall of 1971, approximately 1300 interviews were taken with the heads of a sample of households throughout the United States. The survey was one of a series conducted quarterly by the Survey Research Center and included several questions dealing with the respondents' recreation behavior and residential environment.

Types of Questions

With respect to recreation behavior, respondents were asked several questions about their level of participation in a number of outdoor activities. One question dealt with family participation in hunting or fishing, activities thought to be more closely associated with people living in small towns, rural or other sparsley populated areas.

Among the questions related to the residential environment, respondents were asked to assess their immediate neighborhoods on several semantic differential type dimensions. Three of these (attractive-unattractive, pleasant-unpleasant, great place to live-poor place to live) were used to create a general scale of neighborhood satisfaction. Another specific dimension which the respondent assessed was the extent to which he viewed the neighborhood as crowded. Although it can be argued that subjective assessments of neighborhood crowding should, by their very nature, be correlated with neighborhood satisfaction measures, crowding was intentionally omitted as an item in our satisfaction scale. Our intent was, first, to see its relative importance among a number of subjective assessments of specific neighborhood attributes and second, to see how crowding relates to an objective measure of density. With considerable discussion in the literature as to whether "crowding" and "density" are interchangeable terms, (Stokels, 1972; Hawley, 1972), the analysis in this paper will enable us to see if, in fact, these two variables can be substituted for one another.[2]

Interviewer Observations

In addition to conducting interviews with the respondents, interviewers were asked to assess a number of characteristics of the respondents' dwellings and the area around it. While interviewers have identified type of structure and land use in the surrounding area in previous studies, we know of no attempts to have interviewers assess the extent to which the immediate environment around the respondents dwelling was built up. In an exploratory effort, interviewers in this survey were asked to make a judgement about the amount of open space around the respondent's dwelling. Specifically, they were asked,

> Is the area that can be seen within 100 feet of the front door of the dwelling (or 100 feet of the front door of the building in which the respondents' dwelling is located) best described as:
>
> 1. Entirely built up with buildings, no open space present except that associated with privately owned buildings.
> 2. Mostly built up with buildings; some open space present such as parks, school yards, open fields, vacant lots, body of water, etc.
> 3. Sparsely built up with buildings; mostly open space present such as parks, school yards, open fields, vacant lots, bodies of water, etc.
> 4. All open space; respondent's dwelling is the only building within the area.

To date, little effort has been made to check the reliability of our interviewer's assessments. Nevertheless, we feel this variable may offer another measure of how close people live to each other. Subsequent work with this item in predicting to a number of attitudes and behaviors should reveal its relative importance as a variable.

Objective Density Measure

While the respondent's perception of crowding and the interviewer's assessment of open space are relatively inexpensive measures for any survey, measures of the actual density at which people in a sample live are more difficult and costly to obtain, particularly when the sample of households is distributed among four hundred and sixty-five (465) locations throughout the United States. Within the context of a methodological study, a procedure and cost estimates were developed for measuring residential density for clusters of households falling in our national sample.[3] In developing this work, residential density was defined as the number of households within a geographic area known as the "micro-neighborhood". Operationally, a micro-neighborhood includes the structure containing the respondent's dwelling and the five or six dwellings or buildings that can be seen from the respondent's dwelling entrance or from the entrance of the building within which the respondent lives. It should be remembered that this closely approximates the environment that the respondent and the interviewer were asked to assess as to the extent of crowding and open space, respectively.

For our calculation of actual density, we had hoped to follow a procedure similar to that used in a study of planned residential environments (Lansing, Marans and Zehner, 1970). Briefly, this procedure involved the use of a transparent grid of a 2-acre square which would be used in conjunction with scaled maps or aerial photographs containing buildings and dwelling units falling in our sample. The number of dwellings within a square would be counted and divided by two in order to get a measure of the number of dwellings per acre for any cluster of dwellings.[4] Unfortunately, detailed maps or photographs were not available for the 465 clusters of dwellings in our national sample. We viewed the time and cost involved in obtaining them as prohibitive. Instead, we relied on "sketch maps" prepared by our interviewers in enumerating dwellings and blocks for sampling purposes. These "sketch maps" contained block outlines, structures and the number of dwellings in each structure if it were devoted to residential use. In some instances, other physical features such as railroad lines, rivers and streams were shown. If these "sketch maps" had been drawn with consistent accuracy and to a specified scale, our task of measuring the land area within which our sample dwellings were situated would have been easy. However, in many instances, structures were not accurately located on the map nor were maps drawn to a specified scale. Furthermore, structures which may have existed across the street from the sample dwellings were not shown. With these limitations, our measurements were not as precise as we had wanted. Nevertheless, we were able to make reasonably accurate estimates of land area based on a set of definitional criteria and reference maps covering all our sample points. These were either census tract maps, highway department or road commission maps, U.S.G.S. maps or other maps available from local planning agencies. We are now in the process of validating our measures by obtaining and using aerial photos and detailed plat maps covering a number of sample sites.[5]

After completing our calculations of density for each sample cluster, these values were assigned to each dwelling within the cluster. Seven density classifications were created ranging from "less than one DU per acre" to "40 DU's per acre and over" with the modal category being 1 to 3 DU's per acre.[6]

We have developed an operational procedure for determining residential densities for small areas within the framework of national samples. Although these measures are not precise, we believe they are sufficiently accurate and can be used in studies where residential density is considered to be an important explanatory variable. The procedure for obtaining the density measure, however, is both expensive and time consuming. A basic question is whether the cost is warranted based on the benefits derived from having this objective density measure.

Correlations Between Objective Density and Other Measures

Thus far, we have discussed four of the density measures used in this paper. Ranging in order from the most objective to the most subjective, they are: the objective density measure, size of place, interviewer's assessment of open space and respondent's perception of crowding. A fifth measure, urbanicity, was added as a possible improvement over the traditional size of place measure. Like size of place, urbanicity can be easily obtained from existing data, but it has a more detailed set of ordered classes ranging from central cities to rural areas.

As we indicated earlier, one objective of this paper is to consider the relationship between the objective density measure and size of place of residence, a measure which has often been used to approximate population density. The product-moment correlation shown in Table 1 suggests

Table 1

Product-Movement Correlations Between Objective Density Measures and Other Density-Related Variables

	Objective Density Measure
Size of place of residence	.70
Urbanicity Scale	.72
Interviewer's assessment of open space	.47
Respondent's perception of crowding	.47

that the relationship between these variables is fairly strong ($r = .70$). With urbanicity, the product-moment correlation is only slightly stronger than that between density and size of place.[7] While these relationships are fairly strong, the question remains whether either of these variables representing the urban-rural continuum is a good substitute measure for actual density when predicting attitudes and behaviors. In the next section of this paper we will attempt to shed additional light on this question when we consider the marginal contributions of each of these variables in two regression analyses.

The extent to which our two less objective density-related measures are related to the objective density measure are also shown in Table 1. The product-moment correlations between the interviewers' assessment of open space and density and the respondents' perception of crowding and density are identical ($r = .47$). However, these relationships are not as strong as the relationships between the density measure and size of place and urbanicity. As a result of these correlation analyses, we can tentatively conclude that, within the context of national samples, size of place (or urbanicity) may be appropriate substitute measures for residential density. Furthermore, the interviewer assessment of open space and the respondents' perceptions of crowding do not appear to be appropriate substitutes for actual residential density. In fact, they may be measuring other dimensions of density to which people will respond.

Multivariate Analysis

To find out if size of place (or urbanicity) are appropriate substitutes for the objective density measure and if interviewer assessments and respondents' perceptions are measuring other dimensions of density, we considered two series of regression analyses - ones with the respondents' level of satisfaction with the neighborhood as a dependent variable and the other with an observable behavior - whether the family participates in hunting and fishing.

Satisfaction with Neighborhood

In the analysis of neighborhood satisfaction, several multiple classification analyses were run. We included in each equation a base set of independent variables -- housing type, income, race, life cycle and education. The first equation included only these variables. In subsequent equations we added to the base set of independent variables each density-related variable, one at a time, to observe its marginal effects. Table 2 shows the results of adding each of our five density-related measures to the base set of predictors. In the equation predicting to neighborhood satisfaction using only the base set of predictors, 20.6 percent of the total variance is explained. When adding the objective density measure the proportion of variance explained increases to 23.2 percent with a partial R^2 of 2.6 percent. Similarly, the additions of size of place, urbanicity, and the interviewers assessment of open space produce partial R^2's ranging from 2.6 to 2.9 percent.[7] However, when the respondents' perception of crowding is added to the base set of predictors, the proportion of variance explained is 34.0 percent, an increase of 13.4 percent. While it appears that the objective density measure and the interviewers' assessment of open space are no more useful in predicting peoples' satisfaction with their neighborhoods than the size of place (or the urbanicity) measure, their perception of crowding, contributes much more to the proportion of variance explained than the other measures.

Table 2

Determinants of Neighborhood Satisfaction

Independent Variables	R^2	Partial R^2
Base Set (Income, Education, Race, Life Cycle, Housing Type)	20.6	---
Base Set, Objective Density	23.6	2.6
Base Set, Size of Place	23.2	2.6
Base Set, Urbanicity	23.4	2.8
Base Set, Interviewer's Assessment of Open Space	23.5	2.9
Base Set, Respondent's Perception of Crowding	34.0	13.4

Participation in Hunting and Fishing

Having tested the various density measures as predictors of a subjective variable -- namely the respondent's satisfaction with his neighborhood -- we next tested the measures against a more objective dependent variable, the family's participation in hunting and fishing. This variable was chosen because it is observable behavior which has been found to be strongly related to a correlate of density -- size of place (Mandell and Marans, 1972).

First, we regressed a dummy variable indicating whether the family participates in hunting or fishing against the most important non-density related independent variables, namely income, education and age. Together, this base set of variables explained 10.3 percent of the total variance.

Table 3 summarizes the results of adding each of the five density measures to the base set of independent variables in order to see the marginal effects of each. Again, as in the prediction of neighborhood satisfaction, the objective density measure does not appear to be more useful in predicting the family's proclivity to hunt or fish than the urbanicity or size of place variables. However, all three of these more objective measures predict much better than the more subjective measures of the interviewer or the respondent.

Table 3

Determinants of Hunting and Fishing

Independent Variables	R^2	Partial R^2
Income, Education Age	10.3	---
Income, Education, Age, Objective Density	15.0	4.7
Income, Education, Age, Size of Place	14.0	3.7
Income, Education, Age, Urbanicity	15.2	4.9
Income, Education, Age, Interviewer Assessment of Open Space	12.0	1.7
Income, Education, Age, Respondents' Perception of Crowding	10.9	0.6

Summary and Conclusions

In this paper we have considered the relationship of residential density as an objective measure of how close people live to each other and size of place of residence - an often used surrogate density measure. We have also investigated relationships between the objective density measure and the assessments and perceptions of interviewers and respondents. Methods and procedures for obtaining data from a national sample of households were discussed along with their relative costs. Analyses of two dependent variables -- satisfaction with neighborhood and participation in hunting and fishing -- were presented showing the relative value of each density-related variable as a predictor.

As a result of these analyses, a number of conclusions can be drawn. First, size of place of residence and urbanicity are much better proxies for our objective density measure than the interviewer's assessment of open space or the respondents' perception of crowding. Second, these more subjective measures (interviewer open space assessment and respondent perception of crowding) appear to be measuring things other than density per se. This finding with respect to crowding tends to support the notion that the psychological experience should be distinguished from the physical condition of density. Third, the traditionally used size of place of residence seems to be as good a measure as the urbanicity measure consisting of a more detailed and ordered set of classes. Finally, after controlling for a number of socio-economic and demographic variables, the objective density measure does not appear to be better than size of place of residence in predicting density-related dependent variables.

APPENDIX A

CONSTRUCTED VARIABLES

Size of Place (1960 Census Classification)

1. Central cities of 12 largest SMSA's (including Consolidated Areas)
2. Cities of 50,000 and over, exclusive of (1)
3. Urban Places, 10,000-49,999
4. Urban places 2,500-9,999 and other urbanized areas not included in above codes
5. Rural, in an SMSA psu
6. Rural, not in an SMSA psu

Urbanicity

1. Central cities of 12 largest SMSA's (exclude Long Beach and Jersey City)
2. Cities with population (1960) of over 100,000 excluding those coded 1 in this variable
3. Suburbs, population (1960) 2,500 to 100,000 within the twelve largest SMSA's
4. Cities with population (1960) of 10,000 to 100,000, excluding those coded 3
5. Places with population (1960) of 2,500 to 10,000 excluding those coded 3
6. Rural places (population of less than 2,500 (1960) in an SMSA)
7. Rural, not in an SMSA; adjacent areas
8. Rural, not in SMSA; outlying areas

APPENDIX A

CONSTRUCTED VARIABLES (continued)

Density

1. Less than one DU per acre
2. 1.00 - 3.49 DU's per acre
3. 3.50 - 6.49 DU's per acre
4. 6.50 - 11.49 DU's per acre
5. 11.50 - 19.49 DU's per acre
6. 19.50 - 39.49 DU's per acre
7. 39.50 DU's per acre and over

FOOTNOTES

[1]The study which formed the basis of this paper was sponsored by the Bureau of Outdoor Recreation of the U. S. Department of the Interior. Work on the objective density measure was sponsored by a grant to the Institute for Social Research from the National Science Foundation - RANN Division.

[2]The analyses of the relative importance of assessments of neighborhood attributes in predicting overall neighborhood satisfaction is in process.

[3]The methodological study is one of several being funded at the Institute for Social Research through a grant from the National Science Foundation.

[4]See Lansing et. al, p. 107 for a detailed discussion of the procedures for calculating residential density.

[5]The procedure and cost estimates used in obtaining measures of residential density are described in a Survey Research Center staff working paper "A Methodology for Measuring Residential Density in National Samples" by Robert W. Marans and Jean Wineman.

[6]See Appendix A for categories of this constructed variable.

[7]See Appendix A for the classes used in the size of place of residence and urbanicity variable.

[8]When urbanicity and the objective density measure were added to the base set of predictors, the partial R^2 was 3.2 percent.

REFERENCES

[1] Stokols, Daniel, "A Social-Psychological Model of Human Crowding Phenomena," Journal of the American Institute of Planners, 38 (March 1972), 72-83.

[2] Hawley, Amos H., "Population Density and City Life," Presidential Address of the Population Association of America, April, 1972.

[3] Lansing, John B., Robert W. Marans and Robert B. Zehner, Planned Residential Environments, Ann Arbor: Institute for Social Research, The University of Michigan, 1970.

[4] Mandell, Lewis and Robert W. Marans, "Participation in Outdoor Recreation: A National Perspective," Ann Arbor: Institute for Social Research, The University of Michigan, 1972.

SURVEY MEASURES OF CHANGES IN HOUSEHOLD WEALTH AND BUYING EXPECTATIONS: A NOTE ON METHODOLOGY

John McNeil, Bureau of the Census

Since 1959, the Bureau of the Census has collected quarterly household survey data which are intended to help predict changes in consumer spending on durables.

In the portion of the survey dealing with expected automobile purchases, consumers are asked to estimate their chances of buying a car during the next 6, 12, and 24 months. Changes in responses to these questions are then used as indicators of probable changes in actual spending. If, for example, the seasonally adjusted mean probability for all households of buying a car within 6 months rises from one quarter to the next, then, the assumption is that, other things being equal, the actual level of car purchases will rise.

We determine whether a change in car buying chances has occured by measuring the level of probabilities in the first quarter, measuring them again in the second quarter, and subtracting one from the other. An alternative method of measuring change would be to ask during the second quarter if a household's chances of buying a car had increased or decreased since the first quarter.

In fact, both these methods, the subtraction of one level from another and the single change question, were used in interviews with a panel of households who participated in a special Census Bureau study of consumer anticipations.

From 1968 through 1970, the Census Bureau conducted a series of five interviews with a panel of approximately 3,500 households. The sample was not representative of U.S. households; it was a convenience sample of generally middle income, suburban households from three widely dispersed geographic areas.

The purpose of the operation was first, to test alternative methods of collecting data on consumer buying expectations, secondly, to test the feasibility of collecting anticipations data in areas other than that of cars and household durables, and thirdly, to provide data which would help in the interpretation of the national data which was being collected regularly each quarter.

This paper reports on the consistency of the two methods of measuring change. The alternative methods were used in measuring the change in the household's chances of buying a car and the changes in the amounts held by the household in savings accounts and U.S. Government bonds. The tests were built into the questionnaires used for the first two visits. During the first visit, in May 1968, households were asked to estimate their chances of buying a car within 6, 12, and 24 months. The procedure was the same as that used on the national sample - respondents were asked to select their answers from an 11 point scale with the numbers 0, 10, 20, and so on up to 100. Respondents were also asked about ownership of certain assets, amoung them, savings accounts, and U.S. Government bonds. If the respondent reported ownership of an asset, he or she was asked about the amount. The respondent was asked to indicate the amount by selecting a category from an answer sheet containing 15 categories of unequal intervals - under $250, $250 to $499, $500 to $999, $1,000 to $1,999, and on up to $50,000 and over. During the second visit, in November 1968 households were asked to estimate their chances of buying a car within 6 and 12 months by selecting an answer from the same 11 point scale as had been used previously. They were then asked if their chances of buying a car had increased or decreased since our May 1968 visit. Households were again asked the asset ownership and amount questions. If they reported ownership in November, they were asked if the amount was greater or less than it had been in May. If a respondent reported a change, he was asked to name the amount by using the 15 point scale.

The result of these sequences of questions was that we had two estimates each of the May to November change in three variables - (1) the subjective probability of buying a car, (2) the amount in savings accounts, and (3) the amount in U.S. bonds. As we have mentioned, the estimates of levels and changes in asset amounts were somewhat crude in that they were in terms of intervals rather than dollar amounts, but it was still possible to test for consistency in the two methods.

Tables 1 and 2 show the consistency of responses for the first variable - the household's subjective probability of buying a car. The November 1968 question on whether the probability of buying a car had increased or decreased was somewhat ambiguous because it did not specify a time or horizon. It is conceivable that a respondent's short-term probability would move in a different direction from his long-term prospects. We will review the consistency of the November 1968 change question with both the computed change in 6-month probabilities from May to November and the computed change in 12-month probabilities.

Table 1 shows that 803 households reported in November 1968 that their chances of buying a car had increased from May to November. The table further classifies these households by the alternative measure of change - the November probability of buying within 6 months minus the May probability of buying within 6 months. Only 461 of the households show an increase by the second measure; 187 show no change, and 155 show a decrease. Three hundred and forty two households out of the 803, or 43 percent, gave answers that were not consistent. The inconsistency rate for households which reported in November 1968 that their chances had not changed was 36 percent, the rate for those who had reported that their chances had decreased was 49 percent. The total number of households in table 1 is 3,489 and 1,383, or 40 percent, of those households gave inconsistent replies to questions about their chances of buying a car.

Table 2 repeats table 1 except that the alternative measure of change is defined as the difference between the November and May probabilities of buying a car within 12 months. The proportion of inconsistent replies rises to 48 percent using this definition.

The findings suggest that considerable measurement errors exist in at least one of the methods used to measure the change in the car buying probabilities of households. Since the single question method involves asking the respondent to compare his current chances with the chances of six months earlier, it is the most suspect method. Presumably, the shorter the recall period, the more accurate the response. It is possible to judge the relative accuracy of the two methods by looking at the actual car purchase behavior of the households.

Tables 3 through 5 show the percent of households buying a car between November 1968 and May 1969 for households classified by their response to the November 1968 change question and their responses to the May 1968 and November 1968 question on probabilities. Table 3 indicates that both the November change question and the computed May to November change succeed in placing respondents into categories with significantly different purchase rates. Neither measure eliminates the influence of the other; households who had a positive May to November change in their probability of buying were more likely to buy if they reported in November that their chances had increased than if they had reported no change or a decrease; households reporting no change or a decrease in November were more likely to buy if their computed change in probability was positive than if it were zero or negative. Tables 4 and 5 show that even when households are classified by their reported November 1968 probability, their likelihood of buying a car varies according to their responses to the November change question. For example, 13.8 percent of those households who reported in November 1968 that their chances of buying a car within 12 months was zero actually bought a car between November and May. When these households were classified by their response to the November change question, the purchase rates were 10.4 percent for those who said their chances had decreased, 10.8 percent for those who said their chances had not changed, and 31.0 percent for those who said their chances had increased.

Some regression results follow. The number of observations is the number of households for which reasonably complete data were obtained for the first four visits, about 3,500. The first result below shows the results of defining the dependent variable as the presence or absence of a reported car purchase during November 1968 - May 1969, and using the November 1968 six month probability, the May 1968 six month probability, the change between the two, and the November 1968 single change question as the independent variables:

(The latter variable was scaled as follows: one for increase, zero for no change or don't know, and minus one for decrease.)

$$\text{CARPUR (Nov. - May)} = .1267 + \underset{(.0020)}{.0087}\ \text{MAY PRB} + \underset{(.0023)}{.0423}\ \text{NOV PRB} + \underset{(.0121)}{.0399}\ \text{NOV CHANGE} \qquad R^2 = .14$$

As is usually the case, the percent of total variation explained is small. The most important variable by far is the November 1968 probability of buying a car, but both the May 1968 probability and the November 1968 change question have "t" values above 3. The computed May to November change in probabilities is not significant. When the dependent variable was defined as the change in actual purchases from the period May 1968 - November 1968 to the period November 1968 - May 1969, the following results was obtained:

$$\triangle\,\text{CARPUR} = -.0276 + \underset{(.0036)}{.0161}\ \text{MAY PRB} + \underset{(.0164)}{.2387}\ \text{NOV CHANGE} + \underset{(.0032)}{.0484}\ \triangle\,\text{PRB} \qquad R^2 = .21$$

In this test, both the computed May to November change in probabilities and the November 1968 change question are highly significant; each of the "t" values is approximately 15. The May 1968 probability variable is also significant with a "t" value of about 4. The performance of the November 1968 change variable is interesting. As we have mentioned earlier, the change in probability from May to November should be measured most accurately by subtracting one level from the other. The November 1968 change question should, in theory, be washed out by the more accurate computed change. The fact that the November change variable, scaled in a rather crude way, contributes as much as the computed change variable, is a measure of the unrealiability of many respondents' estimates of their probability of buying a car. For a substantial proportion of households, it suggests that computed changes in probabilities are not likely to be associated with changes in actual car buying behavior. It also suggests that the Census Bureau's program could be improved by the addition of a question on the order of the November 1968 change question.

Tables 6 and 7 show the relationship of the two methods for measuring changes in the holdings of assets. One thousand and eighty-five households reported in November 1968 that their holdings in savings accounts were higher than they had been in May 1968. When these are classified by the difference between the level category reported in May and the level category reported in November, we find that 577 reported a higher category in November, 278 reported the same category, and 320 reported a lower category. Because categories represent intervals, only the 320 can be considered to have given inconsistent replies. For the 1,284 households reporting that the amount they held in savings accounts had not changed, 316 reported a higher November category, and 554 reported the same category, and 414 reported a lower category. The 730 households reporting higher or lower categories gave inconsistent replies. Six hundred and seventeen households reported in

November that their holdings in savings accounts had decreased: 160 of these reported a higher November category, 133 reported the same, and 388 reported a lower category. The 160 reporting a higher category were inconsistent. For all households in table 5, 1,110 out of 2,986, or more than 35 percent, gave inconsistent replies. Table 6 shows the results for holdings in U.S. Government bonds, and the results are very similar. The percent of inconsistencies for all households in the table is again over 35 percent.

The differences in the two measures of asset change are striking. According to their response to the November 1968 change question, 617 households had a lesser amount in savings in November 1968 than in May 1968, but 1,122 households reported a lesser category in November than they had in May. Only 85 households reported in November 1968 that the amount they held in U.S. Government bonds had decreased since May, but 276 reported a lower category in November than they had in May.

Although it is not possible to measure the relative merits of the two measures of asset change, the degree of inconsistency serves to reinforce the evidence that survey data on the asset holdings of consumers will almost certainly contain very serious errors.

Table 1.--RESPONSE TO NOVEMBER 1968 QUESTION ON WHETHER THE CHANCES OF BUYING A CAR HAD INCREASED OR DECREASED BY COMPUTED MAY 1968 TO NOVEMBER 1968 CHANGE IN THE PROBABILITY OF BUYING A CAR WITHIN 6 MONTHS

Response to Nov. 1968 question on change and reported Nov. 1968 probability of buying a car within 6 months	All households	November 1968 6-month probability minus May 1968 6-month probability		
		Positive	Zero	Negative
Households reporting in Nov. 1968 that their chances had increased:				
Total	803	461	187	155
November 1968 6-month probability:				
0	206	-	118	88
10	19	15	1	3
20	102	68	17	17
30	14	10	-	4
40	6	5	-	1
50	93	68	6	19
60	12	9	-	3
70	24	18	-	6
80	123	94	22	7
90	32	25	-	7
100	172	149	23	-
Households reporting in Nov. 1968 that their chances had not changed:				
Total	2,142	381	1,365	396
0	1,581	-	1,297	284
10	93	62	8	23
20	210	134	29	47
30	16	12	-	4
40	12	7	-	5
50	68	48	7	13
60	5	4	-	1
70	12	9	-	3
80	74	54	10	10
90	12	6	-	6
100	59	45	14	-
Households reporting in Nov. 1968 that their chances had decreased:				
Total	544	78	226	280
0	457	-	216	241
10	20	5	1	14
20	35	16	3	16
30	1	-	-	1
40	3	3	-	-
50	8	3	1	4
60	-	-	-	-
70	3	3	-	-
80	9	4	2	3
90	1	-	-	1
100	7	4	3	-

Table 2.—RESPONSE TO NOVEMBER 1968 QUESTION ON WHETHER THE CHANCES OF BUYING A CAR HAD INCREASED OR DECREASED BY COMPUTED MAY 1968 TO NOVEMBER 1968 CHANGE IN THE PROBABILITY OF BUYING A CAR WITHIN 12 MONTHS

Response to Nov. 1968 question on change and reported Nov. 1968 probability of buying a car within 12 months	All households	November 1968 12-month probability minus May 1968 12-month probability		
		Positive	Zero	Negative
Households reporting in Nov. 1968 that their chances had increased:				
Total..............	803	469	172	162
November 1968 12-month probability:				
0....................	70	-	22	48
10....................	9	7	-	2
20....................	58	24	13	21
30....................	10	5	1	4
40....................	12	10	1	1
50....................	74	42	9	23
60....................	16	12	-	4
70....................	30	17	-	13
80....................	161	96	29	36
90....................	54	43	1	10
100....................	309	213	96	-
Households reporting in Nov. 1968 that their chances had not changed:				
Total..............	2,142	591	982	569
0....................	1,092	-	768	324
10....................	86	44	7	35
20....................	320	148	76	96
30....................	39	28	-	11
40....................	35	26	1	8
50....................	145	84	19	42
60....................	29	17	1	11
70....................	27	17	2	8
80....................	163	103	38	22
90....................	41	27	2	12
100....................	165	97	68	-
Households reporting in Nov. 1968 that their chances had decreased:				
Total..............	544	55	135	354
0....................	359	-	106	253
10....................	17	1	2	14
20....................	79	17	12	50
30....................	4	-	-	4
40....................	3	2	-	1
50....................	24	8	3	13
60....................	2	-	-	2
70....................	5	1	-	4
80....................	28	14	4	10
90....................	5	2	-	3
100....................	18	10	8	-

Table 3.—PERCENT OF HOUSEHOLDS BUYING A CAR BETWEEN NOVEMBER 1968 AND MAY 1969 BY NOVEMBER 1968 RESPONSE TO QUESTION ON WHETHER THE CHANCES OF BUYING A CAR HAD INCREASED OR DECREASED AND BY THE DIFFERENCE BETWEEN THE MAY 1968 AND NOVEMBER 1968 PROBABILITIES

Response to Nov. 1968 question on change	Total	Nov. 1968 Probability minus May 1968 Probability					
		Probability of buying a car within 6 months			Probability of buying a car within 12 months		
		Positive	Zero	Negative	Positive	Zero	Negative
All households......	21.1	38.0	14.7	20.4	29.3	16.4	18.3
Households reporting that their chances had increased.....	39.1	47.7	24.1	47.6	38.8	47.1	29.0
Households reporting that their chances had not changed...	16.2	28.1	11.9	19.9	22.0	11.6	18.3
Households reporting that their chances had decreased.....	14.3	26.3	12.8	13.9	23.6	12.6	13.6

Table 4.--PERCENT OF HOUSEHOLDS BUYING A CAR BETWEEN NOVEMBER 1968 AND MAY 1969 BY NOVEMBER 1968 RESPONSE TO QUESTION ON WHETHER THE CHANCES OF BUYING A CAR HAD INCREASED OR DECREASED AND BY MAY 1968 AND NOVEMBER 1968 PROBABILITIES OF BUYING A CAR WITHIN 6 MONTHS

Response to Nov. 1968 question on change and reported Nov. 1968 probability of buying a car within 6 months	Total	May 1968 six month probability (Number of chances in 100)				
		0	10-30	40-60	70-90	100
All households...	21.1	17.1	25.2	31.1	31.9	27.7
Nov. 1968 six month probability:						
0............	12.4	10.9	16.1	19.8	17.8	13.7
10-30.........	21.2	20.3	20.5	20.7	19.1	42.9
40-60.........	31.4	33.7	26.2	36.4	24.1	21.7
70-90.........	46.6	43.2	54.9	26.1	54.0	50.0
100...........	64.7	59.8	48.5	86.4	73.9	67.5
Households reporting in Nov. 1968 that their chances had increased:						
Total.........	39.1	34.7	36.4	41.3	52.6	46.4
Nov. 1968 six month probability:						
0............	17.0	11.9	26.9	6.3	29.6	26.3
10-30.........	29.6	32.9	24.1	33.3	11.1	33.3
40-60.........	32.4	28.6	25.0	45.5	53.8	28.6
70-90.........	49.7	47.6	55.6	23.1	56.3	57.1
100...........	66.3	65.2	42.9	88.2	77.1	65.2
Households reporting in Nov. 1968 that their chances had not changed:						
Total.........	16.2	13.3	21.6	24.0	25.8	31.4
Nov. 1968 six month probability:						
0............	11.8	10.7	14.4	16.7	20.4	18.4
10-30.........	18.5	9.9	20.5	23.5	23.1	45.5
40-60.........	31.8	42.5	29.4	30.0	-	50.0
70-90.........	39.8	34.1	57.1	25.0	43.8	43.8
100...........	61.0	46.2	80.0	75.0	70.0	71.4
Households reporting in Nov. 1968 that their chances had decreased:						
Total.........	14.3	13.0	13.6	28.2	14.0	13.0
Nov. 1968 six month probability:						
0............	12.3	11.6	14.3	31.0	11.8	6.9
10-30.........	16.1	15.8	13.3	-	16.7	50.0
40-60.........	18.2	20.0	-	-	-	50.0
70-90.........	53.8	50.0	-	50.0	100.0	50.0
100...........	57.1	50.0	-	100.0	-	66.7

Table 5.--PERCENT OF HOUSEHOLDS BUYING A CAR BETWEEN NOVEMBER 1968 AND MAY 1969 BY NOVEMBER 1968 RESPONSE TO QUESTION ON WHETHER THE CHANCES OF BUYING A CAR HAD INCREASED OR DECREASED AND BY MAY 1968 AND NOVEMBER 1968 PROBABILITIES OF BUYING A CAR WITHIN 12 MONTHS

Response to Nov. 1968 question on change and reported Nov. 1968 probability of buying a car within 12 months	Total	May 1968 12-month probability (Number of chances in 100)				
		0	10-30	40-60	70-90	100
All households...	21.1	13.8	21.3	23.6	29.9	32.7
Nov. 1968 twelve month probability:						
0............	10.2	7.4	12.8	14.9	16.4	13.3
10-30.........	15.0	11.8	16.0	15.0	23.3	14.5
40-60.........	20.3	24.8	17.6	9.8	21.3	26.3
70-90.........	31.7	23.8	31.5	28.4	33.6	43.0
100...........	54.7	54.3	47.8	58.0	49.5	59.9
Households reporting in Nov. 1968 that their chances had increased:						
Total.........	39.1	31.0	35.0	38.0	40.4	54.9
Nov. 1968 twelve month probability:						
0............	21.7	13.6	33.3	20.0	22.2	33.3
10-30.........	13.0	15.2	11.1	12.5	12.5	10.0
40-60.........	21.6	25.7	19.2	14.3	25.0	18.2
70-90.........	35.5	27.4	31.0	38.7	36.2	48.8
100...........	58.9	51.6	56.3	55.9	56.7	67.7
Households reporting in Nov. 1968 that their chances had not changed:						
Total.........	16.2	10.8	17.7	18.6	26.1	30.6
Nov. 1968 twelve month probability:						
0............	9.5	6.9	14.1	11.3	18.4	19.5
10-30.........	15.3	11.1	16.5	18.4	27.7	12.5
40-60.........	20.1	24.6	17.9	9.7	18.9	30.0
70-90.........	26.8	22.0	26.8	21.9	29.7	34.3
100...........	47.9	60.7	31.6	61.5	35.1	51.5
Households reporting in Nov. 1968 that their chances had decreased:						
Total.........	14.3	10.4	14.8	15.9	19.6	14.2
Nov. 1968 twelve month probability:						
0............	10.3	9.4	9.6	19.4	12.7	7.3
10-30.........	15.0	12.5	16.7	7.1	16.7	17.6
40-60.........	17.2	20.0	-	-	25.0	28.6
70-90.........	36.8	-	66.7	-	38.5	50.0
100...........	44.4	50.0	-	66.7	66.7	37.5

Table 6.--RESPONSE TO NOVEMBER 1968 QUESTION ON WHETHER THE AMOUNT HELD IN SAVINGS ACCOUNTS HAD INCREASED OR DECREASED BY COMPUTED MAY 1968 TO NOVEMBER 1968 CHANGE IN THE AMOUNT HELD IN SAVINGS ACCOUNTS

Response to Nov. 1968 question on change and reported Nov. 1968 amount held in savings		All households	November 1968 category minus May 1968 category		
			Positive	Zero	Negative
Households reporting in Nov. 1968 that their holdings had increased...		1,085	577	278	320
November 1968 amount:					
Interval	Category				
Under $250	1	75	27	31	17
$250 to $499	2	65	29	15	21
$500 to $749	3	72	39	14	19
$750 to $999	4	75	46	7	22
$1,000 to $1,999	5	161	56	59	36
$2,000 to $2,999	6	116	66	17	33
$3,000 to $3,999	7	90	57	18	15
$4,000 to $4,999	8	82	58	17	7
$5,000 to $7,499	9	93	49	29	15
$7,500 to $9,999	10	94	54	26	14
$10,000 to $14,999	11	66	37	19	10
$15,000 to $19,999	12	36	25	5	6
$20,000 to $24,999	13	22	13	6	3
$25,000 to $49,999	14	30	16	12	2
$50,000 or more	15	8	5	3	-
Households reporting in Nov. 1968 that their holdings had not changed		1,284	316	554	414
November 1968 amount:					
1.....................		177	16	98	63
2.....................		109	28	44	37
3.....................		106	27	36	43
4.....................		98	25	33	40
5.....................		191	42	94	55
6.....................		137	37	58	42
7.....................		92	25	34	33
8.....................		52	13	17	22
9.....................		101	34	40	27
10.....................		64	20	25	19
11.....................		71	20	34	17
12.....................		29	11	10	8
13.....................		24	5	14	5
14.....................		24	10	11	3
15.....................		9	3	6	-
Households reporting in Nov. 1968 that their holdings had decreased..		617	160	133	388
November 1968 amount:					
1.....................		98	5	34	59
2.....................		75	6	16	53
3.....................		68	15	8	46
4.....................		43	6	9	28
5.....................		86	15	15	56
6.....................		61	9	12	40
7.....................		42	9	7	26
8.....................		29	3	4	20
9.....................		49	7	16	26
10.....................		19	6	3	10
11.....................		23	9	3	11
12.....................		12	4	2	6
13.....................		4	1	-	3
14.....................		10	2	4	4
15.....................		-	-	-	-

Table 7.--RESPONSE TO NOVEMBER 1968 QUESTION ON WHETHER THE AMOUNT HELD IN U.S. GOVERNMENT BONDS HAD INCREASED OR DECREASED BY COMPUTED MAY 1968 TO NOVEMBER 1968 CHANGE IN THE AMOUNT HELD IN U.S. GOVERNMENT BONDS

Response to Nov. 1968 question on change and reported Nov. 1968 amount held in bonds	All households	November 1968 category minus May 1968 category		
		Positive	Zero	Negative
Households reporting in Nov. 1968 that their holdings had increased..	432	215	143	74
November 1968 amount:				
1.....................	116	37	63	16
2.....................	94	46	28	20
3.....................	67	35	18	14
4.....................	34	27	1	6
5.....................	58	29	19	10
6.....................	28	23	4	1
7.....................	9	3	3	3
8.....................	11	8	2	1
9.....................	2	1	1	-
10.....................	7	2	3	2
11.....................	5	3	1	1
12.....................	-	-	-	-
13.....................	-	-	-	-
14.....................	-	-	-	-
15.....................	1	-	-	-
Households reporting in Nov. 1968 that their holdings had not changed	976	286	533	157
November 1968 amount:				
1.....................	543	144	330	69
2.....................	165	58	75	32
3.....................	81	19	49	13
4.....................	45	15	16	14
5.....................	65	23	27	15
6.....................	26	11	8	7
7.....................	15	4	9	2
8.....................	5	2	3	-
9.....................	12	4	6	2
10.....................	3	1	1	1
11.....................	1	-	1	-
12.....................	8	3	4	1
13.....................	4	2	2	-
14.....................	3	-	2	1
15.....................	-	-	-	-
Households reporting in Nov. 1968 that their holdings had decreased..	85	15	24	45
November 1968 amount:				
1.....................	39	3	16	20
2.....................	18	6	5	7
3.....................	7	2	1	4
4.....................	8	-	-	8
5.....................	5	1	1	2
6.....................	4	1	1	2
7.....................	-	-	-	-
8.....................	2	1	-	1
9.....................	1	1	-	-
10.....................	1	-	-	1
11.....................	-	-	-	-
12.....................	-	-	-	-
13.....................	-	-	-	-
14.....................	-	-	-	-
15.....................	-	-	-	-

A STUDY OF THE EFFECT OF REMUNERATION UPON RESPONSE IN A HEALTH AND NUTRITION EXAMINATION SURVEY

Henry W. Miller, Jacqueline Kennedy, and E. Earl Bryant
*National Center for Health Statistics**

In April 1971, a major new national health survey was undertaken by the National Center for Health Statistics (NCHS). This survey, the Health and Nutrition Examination Survey (HANES), is the fourth health examination type survey to be conducted by NCHS in compliance with the National Health Survey Act of 1956.

The present program is similar to the past surveys in that the sample design is a highly stratified, multistage probability type so that national and regional estimates by various socioeconomic and demographic characteristics can be made of the findings. It also utilizes the same data collection mechanism of mobile examination centers employed in the previous surveys.

There are two primary differences which distinguish HANES from the other surveys. First, it is a dual purpose—to measure and monitor the nutritional status of the American people and to collect information on the health and health care needs of the adult population. Secondly, the target population includes all persons 1-74 years of age rather than a specific, narrower, age segment of the population. The previous three surveys were concerned with characterizing the health status of adults 18-79 years, 6-11 years and 12-17 years, respectively.

The design of HANES calls for a sample size of approximately 30,000 persons from 65 primary sampling units across the nation. The persons selected for the examination are chosen so as to provide a representative sample of the total population with oversampling of groups at high risk of malnutrition. In keeping with the dual purpose concept of HANES, a subset of persons aged 25-74 years receive, in addition to the nutrition examination, a more detailed examination designed to detect certain chronic diseases.

Data are collected by three teams of specialized personnel operating simultaneously in various parts of the United States. The examinations take place in specially built and equipped mobile examination centers which consist of three interconnecting trailers. The staffs of the mobile examination centers include physicians, dentists, nurses, laboratory and health technicians, and dietary interviewers.

As in past surveys, the first contact with a sample household is made by a Bureau of the Census interviewer. At that time, a household questionnaire is completed which identifies the household composition and obtains various socioeconomic and demographic items. The Census interviewer explains that if anyone in the household is selected in the sample, a representative of the U.S. Public Health Service will call again within a week or so to explain the survey.

After the household interviews are completed, and the sample of persons to be examined is selected, HANES interviewers visit the sample persons to administer a health questionnaire and to make appointments for them to be examined.

Even though a household questionnaire and medical history forms may have been completed, and an appointment made for the examination, a person is not considered a respondent unless he actually participates in the examination. Past HES samples and present HANES sample are defined as of the time of the first household contact. Consequently, in addition to the nonresponse due to refusal to participate, there is also a certain amount of built-in nonresponse since persons who move, go on vacation, become ill, or for other reasons are not physically available, cannot be examined. If the nonrespondents differ from respondents for a given measurement, the amount of nonresponse bias introduced into an estimate generally would be expected to vary with the amount of nonresponse. Therefore, response rates for a survey such as HANES are important indicators of possible nonresponse biases. Response rates for the first three surveys were 87, 96, and 90 percent, respectively.

As HANES progressed through 1971, it became increasingly apparent that the response rate would not nearly approach the rates of the previous surveys. Interviewer techniques and procedures were appraised, and retraining and observations were made of individual interviewers. Other measures were also undertaken such as seeking more publicity about the survey at individual locations, seeking assistance from community action groups, and using pamphlets to provide sample persons with more information about the survey. Although these measures may have improved the response rate to some extent, the rate remained below satisfactory levels. By the end of the first 15 stands or sample locations, only 64 percent of the total sample had been examined, the response rates for the stands ranging from 46 to 82 percent. If the survey was to succeed, it was obvious that some additional means had to be found for motivating people to respond.

A proposal was made that response might be increased if some remuneration was offered to those who would participate in the examination. In past surveys conducted by NCHS, the response rate was high enough so that payment for participation had not been considered necessary. It was felt that for HANES such payment would be reasonable in aiding response since the time involved in traveling to and from the mobile examination center and the examination itself requires more of the examinee's time than in the past surveys, resulting in many instances in loss of time from work with subsequent loss of pay, or requiring the housewife to hire a baby sitter. In addition, it was felt that there might be offsetting economies in terms of the reduced number of visits required by HANES interviewers to seek cooperation. Finally, if remuneration could increase response to a satisfactory level, the additional cost would be relatively small as compared to the importance of the total program.

Necessary clearances were submitted and plans and procedures developed in November 1971 to institute a study of the effect of remuneration upon response. The earliest possible date that the study could be started was January 1972, at which time operations would be starting at three sites - Tucson, Arizona; West Palm Beach, Florida; and San Antonio, Texas. The latter was selected for two primary reasons—a sample size of about 600 as compared with 350 and 500 at the other two; and the fact that the San Antonio population was expected to be more typical of future HANES stands, particularly with respect to income and age distributions, than either Tucson or West Palm Beach.

Experimental Design and Survey Procedures

The design for the study was superimposed on the HANES sample design for the San Antonio SMSA and the survey procedures that had been specified for the national survey. The sample was of fixed size and was selected in clusters or segments of an expected 6 households each. Segments located in Census enumeration districts (ED's) classified in the 1960 Census as having median family incomes of less than $3,000 per year were selected at a rate of 8 times that of segments located in ED's with higher median incomes. The expected result was that about a fourth of the sample persons would have family incomes of less than $3,000. The initial sample consisted of 651 households, of which 631 were interviewed. The 2,010 persons that composed the initial sample[1] were listed by age and sex, and a systematic sample of 603 persons was selected, oversampling preschool children, women 20-44 years of age, and persons over 65. These 603 persons came from 402 households and 138 segments.

The first step in the experimental design was to classify the segments by median family income according to the information that had been collected in the household interviews and by segment size. The segments were then sorted into 7 size-income classes as indicated by the following grid.

Number of occupied households in segment	Medium Annual Family Income of Segment	
	Under $4,000	$4,000+
1		
2		
3		
4		
5		

Segments in each cell were randomly paired with another segment in the same cell. One segment of each pair was then randomly selected to have all of the sample persons in that segment told about the remuneration (Procedure A). The sample persons in the other segment of that pair were not told about the remuneration (Procedure B). It should be noted that all persons who were examined received $10.00. The only difference was that some of them did not know that they would be paid until they arrived at the examination center.

The pairs of segments were randomly assigned to the interviewers so that each interviewer's assignment was composed of a representative subsample of the segments.

An attempt was made, therefore, to control on three variables thought to be related to response; namely income, segment size, and the interviewer. It seemed reasonable that an offer of $10.00 would influence persons with low income more than it would those with higher incomes. Also, it is well known that some interviewers are more successful than others in obtaining response in surveys. Thus it seemed necessary to try to balance any effect the interviewers might have in the experiment. Segment size was selected as a control because of the possible interaction within segments between the sample persons, and to provide the ability to regulate the size of assignments to interviewers. This type of control was important because some of the interviewers could work only two weeks before they had to report to another HANES stand. The assignment of too many sample persons would make it impossible for an interviewer to complete all of the segments in her subsample, which was a necessary condition for the study.

The design of the study was thoroughly explained to the interviewers before the interviewing began. They were told that they must conduct the survey precisely as was their practice in other HANES stands, except for occasions where the procedures had been changed to accommodate the experiment. The major difference between their usual routine and the procedures to be followed in this stand was that they must offer remuneration to all sample persons so designated, and under no circumstances were they to offer remuneration to those not designated. If a person in Procedure "B" or "not-to-be-told" group had heard about remuneration, then he would be told that payment of $10.00 would be made if he should be examined. In such cases, a record was made to indicate that he knew about the payment.

To assure that the interviewers used a standard approach in the offer of remuneration, a statement was prepared and made part of the introductory remarks that the interviewers normally make upon entrance to a household. The statement read: "______ the United States Public Health Service is conducting a study on the health of the American people. The people chosen for the study are part of a carefully selected scientific sample, representative of all people in the United States. For the study to accurately picture the health of the Nation, we need your help. Today, I will ask some questions about your health and related matters. Then I would like to make an appointment for you to receive a free health examination at our special examination center. As an expression of appreciation for your help in this important survey, and as compensation for your time and inconvenience, you will receive a fee of $10.00 after the examination. Also, we will send any significant findings of the examination to the physician and dentist that you may want to designate." (If more than one family member was in the sample the interviewer emphasized that each sample person would receive $10.00.) This statement was either read or paraphrased for each sample person assigned to Procedure A. For those assigned to Procedure B, the statement was altered to exclude the part about remuneration.

The interviewers were told to stress several times during the interview the importance of keeping appointments and to explain the complete examination to each person. This was done to reduce the possibility that they would decide later to cancel the appointment because of their lack of understanding and appreciation for the examination.

At the appointed time, a taxi cab picked up the sample persons at their home or other designated place and drove them to the examination center.

When the examination was completed, each person was paid $10.00 in cash and asked to complete a form designed primarily to determine whether the sample person knew about the remuneration before the examination. The principal question asked was: "before coming for the examination, did you know that you would receive payment or compensation for your time if you came?" Those answering "yes," were asked how they knew.

In any experiment of this kind, it is inevitable that the design will not be followed exactly. One problem encountered in this study resulted from the need to have interpreters accompany interviewers to households where no one could speak English. The number of times that interpreters were used is not known exactly, but it probably was required in as many as 10-15 percent of the households. Some training was given to all the interpreters. However, they were not randomly assigned to treatment groups, and, consequently, the experimental results are probably contaminated to some extent by interviewer effects.

Another problem arose because some interviewers were not able to complete all of their assignment before having to leave San Antonio to work at another HANES stand. The goal was to have the assigned interviewer complete at least a first contact with a sample person, make remuneration offers, and attempt to make examination appointments. At the end of the fourth week of the survey, 4 of the 6 interviewers had departed without having completed first contacts with 109 sample persons; fifty were in the Procedure "A" group and 59 in Procedure "B". These were randomly reassigned to the two remaining interviewers.

Also, there may have been some effect on final response rates for the two treatment groups due to the use of temporary interviewers who were hired near the end of the study to follow up on persons who had refused to be examined, broke appointments, etc. These people were well-trained, experienced interviewers, however, and their assignments included similar proportions from both treatment groups.

Findings of the Study

Remuneration had a positive effect on the response rate for the San Antonio stand; 82 percent of the sample assigned to Procedure "A" were examined as compared with 70 percent of those assigned to procedure "B". This difference of twelve percentage points is both statistically significant and large enough to have important implications for future HANES stands.

The differences observed between the two groups are probably conservative since some of the people were not told about remuneration when they should have been and a few were told when they should not have been. According to the records kept by interviewers there were 10 procedure "A" errors and 4 Procedure "B" errors. This may be an under count, however. According to answers given by the sample persons themselves in the Exit Interview, as many as 20 percent of the Procedure "A" group may not have known about remuneration, while 14 percent of those in the Procedure "B" group may have known. It is difficult to assess the accuracy of these figures, however, since there was some evidence that the questions were not thoroughly understood. This occurrence was apparent when a person specified how he knew about remuneration after answering "No" to the question: "Before coming for the examination, were you told that you would receive payment as compensation for your time if you came?"

The possible effects of this confounding should be kept in mind when interpreting the results, since the response rates were computed according to the original assignment to Procedure "A" and Procedure "B".

Although the study design did not control on age and sex nor on income at the household level, it is instructive to examine the response rates within these domains.

Table 1 shows that there was a general improvement in the response rates among the various subclasses of the population. Although a few of the differences shown are not statistically significant, most of them indicate a 10 percentage point or more increase in response rates when remuneration was offered.

One might hypothesize that the effect of remuneration upon response rates increases with decreasing income. This was not true for this study. Eighty-five percent of the sample in Procedure "A" with annual incomes of $4,000 or more were examined, a 13 percentage point increase over the rate for Procedure "B". For people with annual family incomes of less than $4,000 per year, the rates were 78 percent for Procedure "A" and 67 percent for Procedure "B".

The effect of remuneration was greater for males than females. In fact, for females in the age groups 1-19 and 45-74 there was essentially no difference between the response rates of the two treatment groups, the rates being the same (85 percent) for the younger age group and 56 as compared with 52 percent for the older age group. This finding is also consistent with the response rates observed in previous HANES stands where relatively high rates were obtained for those under 20 with a noticeable decrease in the rate with increasing age.

There are probably many reasons why the older women did not respond including possibly fear or reluctance of being examined by a strange physician, fear of having certain physical conditions diagnosed, and reluctance to ride alone in a taxi across town.

In an attempt to refine the analysis to determine whether remuneration was the reason for improved response rates rather than other factors such as follow-ups, interviewer effects, etc., tabulations were made of the proportion of the sample making appointments at first contact and the proportion of these who kept the appointments. These findings are presented in Tables 2 and 3.

Although the differences shown in Table 2 are not statistically significant, the pattern of response rates by income, sex and age provides some evidence that remuneration did influence people to make appointments. Overall, 66 percent of those assigned to Procedure "A" and 61 percent of those assigned to Procedure "B" made appointments the first time they were contacted.

Table 3 shows that those expecting remuneration kept their appointments more often than those who did not expect to be compensated. Again, the sample sizes were too small to detect significant differences. However, the proportion keeping appointments is consistently larger for Procedure "A" than for Procedure "B" for all of the variables studied.

The primary criterion for measuring the success of this experiment and for making a decision to begin paying respondents in the national survey was whether such payment would cause a substantial increase in the response rate. However, as in all surveys, cost was also a factor, and an additional cost of $10.00 per person would not be insignificant. The argument can be presented, however, that remuneration will not necessarily increase the survey cost in that some savings accrue because respondents are more cooperative and require fewer contacts to obtain response.

The data in Table 3 provide some support to this argument. Eighty percent of the Procedure "A" sample kept their appointments without requiring multiple contacts as compared with 72 percent of the Procedure "B" sample. This greater degree of cooperation was observed for each of the age, sex, and income classes shown in the table.

Another index of the amount of effort required in the attempt to get people to be examined is the rate of "disruptive" contacts made, that is, the number of contacts made following refusals, and broken appointments, per 100 sample persons. These rates are shown in Table 4 by Procedure according to age, sex, and family income. Again, it is apparent that less effort was made to complete the survey for those promised remuneration that for those not promised remuneration, the rate being 52 per 100 persons for Procedure "A" and 62 for Procedure "B". This difference did not prevail for all of the age, sex, income classes shown in the table, but it did for the majority of the classes.

Epilogue

The findings of this study were considered significant enough to include remuneration as a routine procedure in the national survey. Remuneration of $10.00 per person examined was initiated simultaneously at the twenty-first (Avoyelles, Louisiana) and twenty-second (San Francisco, California) stands in the sequence of operations to cover the 65 stands scheduled for the survey. At the present time, operations have been completed at a total of 10 stands at which remuneration was made, including San Antonio. Of the 4,284 sample persons at those stands, 78.5 percent have participated in the examination. This compares with 67.2 percent for 19 preceding stands where remuneration was not offered.

Although it is not possible to assess just how much of this rather substantial response difference was due to remuneration, it seems clear that remuneration was a major factor. One other factor which may explain part of the difference is that the group of no-pay stands included a number of places in the Northeastern United States and other large metropolitan areas where, on the basis of our experience in previous health examination surveys, response was expected to be low. If these stands are excluded from both groups (7 pre-remuneration and 1 post remuneration stand) the response rates are 71.9 and 81.5, respectively.

The recent improvement in the response rate, for whatever reasons, is a welcome development. It should be realized, however, that even if the rate should continue at the 78 percent level for the remaining stands, we would not be content. Every effort will continue to be made, consistent with available resources, to make further improvements in the HANES response rate.

FOOTNOTES

*We wish to acknowledge the assistance of Dr. Saul Rosenberg, Miss Jean Findlay, and Mr. Kenneth Harris of NCHS, who participated in the various phases of this study.

[1] Information for the 20 nonresponding households was obtained from neighbors and their household members were included in the sampling frame.

Table 1. Proportion of Sample Persons Examined by Experimental Procedure According to Family Income, Age, and Sex: HANES Remuneration Study

Variable	Experimental Procedure A			Experimental Procedure B			Standard error of difference
	Sample size[1]	Proportion examined	Sampling variance	Sample size[1]	Proportion examined	Sampling variance	
Total	303	.82	.00049	292	.70	.00072	.035
Total Income							
Under $4,000	115	.78	.00149	99	.67	.00223	.061
$4,000+	170	.85	.00075	173	.72	.00117	.044
Unknown	18	.72	.01073	20	.75	.00938	.142
Sex							
Male	128	.88	.00083	123	.74	.00156	.049
Female	175	.78	.00098	169	.67	.00131	.048
Age and Sex							
1-19 years	118	.90	.00076	110	.83	.00128	.045
Male	56	.95	.00085	55	.80	.00291	.061
Female	62	.85	.00206	55	.85	.00232	.066
20-44 years	86	.86	.00140	85	.67	.00260	.063
Male	27	.78	.00636	23	.74	.00837	.121
Female	59	.90	.00153	62	.65	.00367	.072
45-74 years	99	.69	.00216	97	.59	.00249	.068
Male	45	.84	.00299	45	.67	.00491	.089
Female	54	.56	.00456	52	.52	.00480	.097

[1]The initial sample contained a total of 603 sample persons of whom eight could not be contacted. Three of these were in experimental procedure A and five in B. Since these eight persons were not contacted, they are excluded from this analysis.

Table 2. Proportion of Sample Persons Making and Appointment at First Contact by Experimental Procedure According to Family Income, Age and Sex: HANES Remuneration Study

Variable	Experimental Procedure A			Experimental Procedure B			Standard error of difference
	Sample size[1]	Proportion appointed	Sampling variance	Sample size[1]	Proportion appointed	Sampling variance	
Total	303	.66	.00074	292	.61	.00081	.039
Family Income							
Under $4,000	115	.74	.00167	99	.70	.00212	.062
$4,000+	170	.64	.00136	173	.56	.00142	.053
Unknown	18	.33	.01338	20	.65	.01138	.157
Sex							
Male	128	.73	.00154	123	.70	.00171	.057
Female	175	.61	.00136	169	.55	.00146	.053
Age and Sex							
1-19 years	118	.68	.00184	110	.63	.00212	.063
Male	56	.71	.00368	55	.75	.00341	.084
Female	62	.65	.00367	55	.51	.00454	.091
20-24 years	86	.69	.00249	85	.66	.00264	.072
Male	27	.74	.00713	23	.61	.01034	.132
Female	59	.66	.00380	62	.68	.00351	.085
45-74 years	99	.62	.00238	97	.56	.00254	.070
Male	45	.73	.00438	45	.69	.00475	.096
Female	54	.52	.00462	52	.44	.00474	.097

[1]See footnote on Table 1.

Table 3. Proportion of Sample Persons Keeping Appointment Made at First Contact by Experimental Procedure According to Family Income, Age and Sex: HANES Remuneration Study

Variable	Experimental Procedure A			Experimental Procedure B			Standard error of difference
	Sample size	Proportion kept. appt.	Sampling variance	Sample size	Proportion kept appt.	Sampling variance	
Total	200	.80	.00080	179	.72	.00113	.044
Family Income							
Under $4,000	85	.74	.00226	69	.65	.00330	.075
$4,000+	108	.84	.00124	97	.75	.00193	.056
Unknown	7	.86	0	13	.77	.01362	.117
Sex							
Male	93	.82	.00159	86	.75	.00218	.061
Female	107	.79	.00155	93	.68	.00234	.062
Age and Sex							
1-19 years	80	.86	.00151	69	.75	.00272	.065
Male	40	.88	.00264	41	.78	.00419	.083
Female	40	.85	.00319	28	.71	.00735	.103
20-44 years	59	.71	.00349	56	.63	.00416	.087
Male	20	.65	.01138	14	.64	.01646	.167
Female	39	.74	.00493	42	.62	.00561	.103
45-74 years	61	.80	.00262	54	.76	.00338	.077
Male	33	.85	.00386	31	.77	.00571	.098
Female	28	.75	.00670	23	.74	.00837	.123

Table 4. Number and Rate Per 100 Sample Persons of Disruptive Contacts[1] by Experimental Procedure According to Family Income, Age and Sex: HANES Remuneration Study

Variable	Experimental Procedure A			Experimental Procedure B			Standard error of difference
	Number of disruptive contacts	Rate per 100 sample persons	Sampling variance	Number of disruptive contacts	Rate per 100 sample persons	Sampling variance	
Total	158	52.15	.00207	182	62.33	.00283	.070
Family Income							
Under $4,000	71	61.74	.00618	77	77.78	.01309	.139
$4,000+	76	44.71	.00341	95	54.91	.00339	.082
Unknown	10	58.82	.04444	10	50.00	.04444	.298
Sex							
Male	56	43.75	.00403	67	54.47	.00883	.113
Female	102	58.29	.00402	115	68.05	.00376	.088
Sex and Age							
1-19 years	45	38.14	.00404	44	40.00	.00437	.092
Male	22	39.29	.01018	21	38.18	.00841	.136
Female	23	37.10	.00647	23	41.82	.00922	.125
20-44 years	46	53.49	.00758	71	83.53	.01704	.157
Male	16	59.26	.01783	21	91.30	.15777	.419
Female	30	50.85	.01249	50	80.65	.01102	.153
45-74 years	67	67.68	.00759	67	69.07	.00609	.117
Male	18	40.00	.01051	25	55.56	.01167	.149
Female	49	90.74	.01626	42	80.77	.01210	.168

[1]Disruptive contacts are defined as contacts made following refusals and broken appointments.

PEARSONIAN TYPE I CURVE AND ITS FERTILITY PROJECTION POTENTIALS

S. Mitra, Emory University, and A. Romaniuk, Statistics Canada
and University of Montreal

1. THE MODEL FOR AGE-SPECIFIC FERTILITY RATES

The purpose of this paper is twofold: to reduce further the number of parameters used in constructing models of fertility graduation and simulation, utilizing primarily those readily available or easily estimated so as to expand the range of their application; and, secondly, to explore the models potentials for deriving the birth series for forecasting population by the component method. The paper may be regarded as a step forward in the development of a fully-fledged parametric model for fertility projections adaptable, primarily, to requisites of empirical material available in the developed countries. Fertility time series for Canada are used to test the validity of the procedures developed therein.

This model was developed on the basis of certain well known facts about child bearing patterns of women. First, children are mostly born to mothers in the age-group say, 15-50 years and, second, fertility rate continues to increase beyond age 15, reaches a maximum somewhere in the age interval 20-30 years and begins to decline thereafter.

Such a curve fits the description of type I of the Pearsonian family of curves, and this model was attempted earlier (Mitra, 1967) in order to describe the pattern of age-specific fertility rates of a number of countries subject to certain modifications. With origin at mode, the equation of the curve is

$$y = y_o \left(1 + \frac{x}{a_1}\right)^{m_1} \left(1 - \frac{x}{a_2}\right)^{m_2} \;......(1)$$

where y_o is the modal ordinate and $-a_1 \leq x \leq a_2$.

Also, $m_1/a_1 = m_2/a_2$ ......(2)

Ordinarily, the parameters of the type I distribution can be obtained by solving equations generated by equating the first four moments of the observed distribution with those (Elderton, 1930) that can be directly derived from (1). For the distribution of age-specific fertility rates it was assumed that the age interval of 15-50 years is generally adequate and could therefore be used as an additional restriction. Thus

$$a_1 + a_2 = 35 \;......(3)$$

together with (2) and the starting point at age 15, reduced the number of independent parameters to only two and the solutions were obtained from the following equations (Mitra, 1967).

$$m_1 + m_2 = \frac{u_1' (a_1 + a_2 - u_1')}{u_2} - 3 \;......(4)$$

$$m_1 = \frac{m_1 + m_2 + 2}{a_1 + a_2} u_1' - 1 \;......(5)$$

where u_1' is the mean and u_2 is the variance.

Equations (2) and (3) could then be used to determine a_1 and a_2 and for y_o, the following, namely,

$$y_o = \frac{N}{(a_1 + a_2)B(m_1 + 1,\, m_2 + 1)} \cdot \frac{m_1^{m_1} \; m_2^{m_2}}{(m_1 + m_2)^{m_1 + m_2}} \;.(6)$$

can be used where N is the sum of the age-specific fertility rates and B stands for the Beta function. In fact y_o may be regarded as a multiplier that equalizes the sum totals of the observed and the graduated distributions.

2. FURTHER VERIFICATION OF THE MODEL AND THEORETICAL CONSIDERATIONS

If data are available in greater details, the proposed model can be tested by carrying out the analysis at sufficient depth. With that in mind the authors of the present paper have examined these distributions for Canada which are available by single years of age beginning 1926. As such, these data are too detailed, since there are too many class intervals than are necessary to investigate the nature of the distribution function. According to Pearson, data condensed into twelve or so classes are detailed enough for curve fitting. However, such condensation has not been attempted because that would require the use of somewhat unconventional (two to three year) age-groups.

For all these distributions, the Pearsonian index k has assumed values justifying the use of a type I curve for describing the distribution of age-specific fertility rates. This verifies the theoretical formulation of the pattern of the distribution as postulated above.

As mentioned before, the best fit of a given fertility distribution may perhaps be obtained by estimating all the parameters in the usual manner which in this instance will require the computation of the first four moments. While that may not be too restrictive for fertility distributions given by single year of age, there appear to arise a few operational as well as technical problems that seem to complicate matters during the process of searching for meaningful interpretations of the parameters of the model. For one thing, the parameters of the model should respond in a manner that is consistent with the trend in the fertility rates. Otherwise, variations in the distributions cannot be properly related to the variations in the parameters and therefore will have little practical value. With so many parameters, it is quite difficult to relate the effect of say a reduction in total fertility rate on modal age, on the fertile age span, etc.

For another, it is difficult to explain why the effective fertile interval should be subject to variation from time to time. Yet, this is what must happen when the type I distribution is obtained in a straightforward manner, i.e., without any restriction. Not only can that happen, but the start and end of the curve may be in direct conflict with reality (see Table 1).

The earlier study (Mitra, 1967) took care of these problems by introducing a restriction like the one specified in equation (3). Of course, there is no reason to believe that the definition of that equation is unique in any way. In fact, it seems only appropriate that some such equation should be arrived at in a realistic manner and be used consistently, unless there are valid reasons to do otherwise. The advantages of these restrictions are primarily, a reduction in the number of independent parameters from four to two, which, in turn, renders the study of the pattern of variations and interrelationships of these two parameters relatively simple and perhaps more meaningful.

3. OTHER VARIANTS OF THE ESTIMATING PROCEDURE

I. Of course, there are more ways to estimate the parameters of a type I curve once the initial restriction on the fertile age interval (like ages 15 to 50) is imposed. The method used before and described earlier in this section is based on equating the first two moments of the observed distribution with those of the theoretical distribution. Alternatively, the first moment or the mean, and another measure, say the mode, may be used in a similar manner. The latter is estimated by some conventional method, or simpler still, through the substitution of the midpoint of the single year age interval corresponding to which the age-specific fertility rate is the maximum. In this case, given the value of the mode or a_1, with the origin at the start of the curve, the remaining parameters can be obtained from the following:

$$a_2 = (a_1 + a_2) - a_1 \;......(7)$$

$$m_2 = \frac{a_2(a_1 + a_2 - 2u'_1)}{(a_1 + a_2)(u'_1 - a_1)} \;......(8)$$

and $m_1 = \frac{a_1}{a_2} m_2$ ......(9)

II. Another method based on the fertile age-range, modal age and modal fertility also seems reasonable and worthy of investigation. With orgin at start of the curve, and using the relationship

$$\frac{m_1}{a_1} = \frac{m_2}{a_2} = \frac{m_1 + m_2}{a_1 + a_2} = C \;......(10)$$

equation (6) for the modal fertility rate can be rewritten as

$$y_o = \frac{N}{(a_1+a_2)B(Ca_1+1, Ca_2+1)} \cdot \left[\left(\frac{a_1}{a_1+a_2}\right)^{a_1}\left(\frac{a_2}{a_1+a_2}\right)^{a_2}\right]^C \quad (11)$$

where N, for the single year distribution, is equivalent to the total fertility rate. The constant C, obtained through iterative procedure, can then be used in (10) to estimate the remaining parameters. Iteration techniques based on descent method of the type suggested by Keyfitz (1968) is still another approach that could be used to improve the fit. Such an approach proved to be useful in fitting the Gompertz curve to Canadian data on cumulative fertility (Murphy, Nagnur, 1972) and could be eventually, with due adaptations, applied to fertility data in the form of age-specific distribution.

4. THE METHOD AND THE DATA

The first three methods outlined above have been tested against the Canadian fertility series from 1926 to 1969. (Methods based on iterative techniques will be presented in another paper). For simplicity in future references, these three methods will henceforth be recalled by the codes shown in parentheses as in the following:

A. Method based on first four moments (4M)

B. Method based on first two moments (2M)

C. Method based on mean and mode (1M)

The numerals 4, 2, and 1 stand for the number of moments on which a particular method is based. It will be interesting to examine the estimates of the parameters obtained from these three models and these have been done next where, for simplicity in presentation, the years, 1926, 1931, 1941, 1951, 1961 and 1969 have been selected that represents the whole range for which such data are available. A priori, one may expect the goodness of the fit to improve as the number of moments used is increased. It follows that (4M) sets some standard against which estimates obtained by other methods can be gauged.

5. MODAL AGE, MODAL FERTILITY AND TOTAL FERTILITY

Of all the parameters of the type I distribution, the one that is most meaningful in this context seems to be the modal age. Since the method (2M) is restricted by a fixed fertile interval, estimate of the modal age will not be unique and will depend on the specific choice of that intercal. The results based on two alternative fertility intervals of 15-50 and 17-50 years by (2M) along with those calculated by (4M) are compared with observed modal ages in Table 1. It should be mentioned that observed values presented in the table have been adjusted for irregularities which were apparently due to random fluctuations and age misreporting. This adjustment has been made by tracing through the points of observed mode ages, a freehand curve. Also it should be remembered that method (1M) has no built-in mechanism for calculation of mode age and that the latter has to be given in order to enable us to make use of this method for derivation of other parameters.

Table 1 ESTIMATES OF MODAL AGE FOR CANADA FOR A FEW SELECTED YEARS

Year	(4M)	(2M)		(1M)	Total Fertility (per 1000)
		(15-50)	(17-50)	(observed)	
(1)	(2)	(3)	(4)	(5)	(6)
1926	28.1 (17.0-48.5)	28.9	27.9	28.0	3356
1931	27.6 (16.9-49.3)	28.5	27.6	27.7	3201
1941	26.3 (17.2-49.6)	27.6	26.4	27.1	2824
1951	25.5 (17.0-49.9)	26.6	25.4	25.3	3480
1961	24.2 (17.2-50.4)	25.7	24.3	24.0	3857
1969	24.0 (16.6-53.0)	25.1	23.7	23.8	2410

The figures in parenthesesin Col (2) are the estimates of fertile age range based on (4M), the lowest (16.6) and the highest (53.0) limits of which are both found in 1969. The lower limit is no larger than 17.2 years and the smallest value of the upper limit is 48.5 years. The latter appears to vary more than the former, and the variation, though not large, is of considerable magnitude, and the actual values are not quite in accord with the real data. Such inconsistencies, as mentioned earlier, are unavoidable when all the parameters are estimated from the moments of the distribution.

Table 2 ESTIMATES OF MODAL FERTILITY FOR CANADA FOR A FEW SELECTED YEARS

Year	(4M)	(2M)		(1M)	
		(15-50)	(17-50)	(15-50)	(17-50)
(1)	(2)	(3)	(4)	(5)	(6)
1926	178	181	181	152	184
1931	174	175	175	150	181
1941	159	159	160	147	173
1951	205	203	205	172	204
1961	239	231	236	195	231
1969	157	150	155	130	155

6. THE PARAMETERS m_1 AND m_2

Apart from the goodness of fit, the justification of a model also depends on logical and meaningful interpretations of the parameters. The parameter a, which also corresponds to the modal age is simple to understand while the relevance of m_1 and m_2 is not that apparent. However, it can be argued that among other things, the physiological ability to reproduce depends on age of mother which, in turn, can be expressed as the number of years elapsed since the beginning of the fertile interval (15-50 for example), as well as the number of years left to reach the end of that interval. Of these two, the former may be regarded as a positive force that has the effect of raising the fertility rate while the latter acts negatively to hold the rate down. The intensities of the pushing upward and pulling down effects of these two forces, thus defined, and measurable in this instance by m_1 and m_2, can then be regarded as largely responsible for generating a given form of the fertility curve. This is not to suggest, however, that the pattern of fertility is, in fact, controlled in this manner, but merely to indicate the probable consistency of such an explanation of the parameters in this particular context.

The location of the mode (Table 1, cols. 2-5) indicates that the curve has positive skewness (mode is closer to the start of the curve as $a_1 < a_2$) and therefore $m_1 < m_2$, because of equation (2). From Tables 3 and 4, it is quite apparent that m_1 and m_2 must be very much sensitive to the estimation procedure compared to the actual frequency distribution.

As could be expected, the values of the parameters are, and should be, independent of the total frequency, since their derivation are dependent primarily upon the distribution of the relative frequencies. Accordingly, changes in the parametric values are reflections of the changes in the pattern of the distributions themselves. Particularly sensitive to changing age pattern of fertility is m_2. Correlation between this latter and such measures of frequency distribution as mean age, variance, skewness and kurtosis exceeds 90.

Table 3 ESTIMATES OF m_1 FOR CANADA FOR A FEW SELECTED YEARS

Year	(4M)	(2M) (15-50)	(2M) (17-50)	(1M) (15-50)	(1M) (17-50)
(1)	(2)	(3)	(4)	(5)	(6)
1926	.90	1.55	.98	.83	1.05
1931	.95	1.55	.97	.89	1.09
1941	.83	1.49	.89	1.13	1.23
1951	.85	1.45	.82	.75	.81
1961	.70	1.37	.73	.64	.64
1969	.95	1.39	.72	.76	.74

Table 4 ESTIMATES OF m_2 FOR CANADA FOR A FEW SELECTED YEARS

Year	(4M)	(2M) (15-50)	(2M) (17-50)	(1M) (15-50)	(1M) (17-50)
(1)	(2)	(3)	(4)	(5)	(6)
1926	1.67	2.35	2.00	1.40	2.11
1931	1.93	2.46	2.08	1.57	2.28
1941	2.12	2.68	2.25	2.14	2.82
1951	2.43	2.92	2.43	1.81	2.40
1961	2.64	3.12	2.57	1.86	2.39
1969	3.71	3.42	2.80	2.25	2.84

7. GOODNESS OF FIT

The classical approach to measure the goodness of fit is to calculate X^2, the magnitude of which depends partly on the relative difference between the observed and the graduated values. This method was considered unsuitable in the situation characterized by a large number of intervals and, perhaps because of that (4M) estimates were zero in some cases, at the beginning and at the end of the observed fertile interval. Overall index values, comparing an observed with the expected distributions, can be obtained in a number of ways. One such index, namely the index of dissimilarity (Δ) is obtained by reducing the two distributions in percentage form and summing only the positive differences between corresponding percentages. A given value of this index indicates the percentage of observations that need be redistributed among intervals so that the two distributions become identical. It is easy to see that Δ can range from 0 to 100 and its magnitude depends also on the choice of class intervals.

Table 5 Δ VALUES FOR CANADA FOR A FEW SELECTED YEARS

Year	(4M)	(2M) (15-50)	(2M) (17-50)	(1M) (15-50)	(1M) (17-50)
(1)	(2)	(3)	(4)	(5)	(6)
1926	2.74	4.08	2.77	8.94	2.80
1931	2.66	3.78	2.59	7.15	2.27
1941	2.32	4.30	2.39	5.33	4.51
1951	2.28	4.03	2.41	9.42	2.41
1961	2.97	5.38	3.30	11.34	4.03
1969	3.48	5.23	3.99	10.39	4.13

According to this index, the discrepancy was at its highest level at both extremes, in 1961 when the total fertility rate was the maximum for the period, and also in 1969 when the rate was at its lowest. Of these two years, the index values for (4M) were higher for 1969 and those for (2M) were slightly higher for 1961. As could be expected, the index values were uniformly the smallest for the (4M) method, but considering the simplicity and logical consistency of the other methods, the values are not very high and in general, the model seems to be quite satisfactory, at least as a first approximation of a graduation formula for the fertility rates. The sudden increase of the index values in the latter years seem to indicate the possibility of systematic biases, rather than random fluctuations in the series generated by the differences between the observed and the graduated fertility rates.

8. POTENTIALS FOR FERTILITY PROJECTIONS

It seems that the period covered in this study is quite representative of the wide variation in fertility patterns of Canada in recent decades, and in all likelihood the observed ranges of all the principal characteristics will continue to represent their variations in the near future. An examination of the models, in terms of their parametric values and also in terms of their goodness of fit makes it quite clear that for Canada, the accuracy of fertility projections depends more on the accuracy of projecting the total fertility rate than on the relative distribution of these rates by age.

There are a number of approaches that could be followed in projecting total fertility rates. A cohort-orientated approach involving a three-step operation is being considered for use in the population projections in Canada and may be described briefly as follows.

The first step in this approach consists in projecting, for each successive generation of women, completed family size; that is, number of children a woman will achieve upon completion of her childbearing life. For women, who at the beginning of the projection, have already reached a sufficiently advanced stage of family formation, completed family size can be projected without a great risk of error by straightforward graphical extrapolation of the curve of cumulative fertility to date. Acceptable results can also be obtained through the fitting of a Gompertz curve (A. Romaniuk and S. Tanny, 1969). The real challenge for the forecaster is presented by the cohorts of women who have not yet entered childbearing but who only in a few years will form the majority of childbearers. Among various indicators, parity distribution, that is, completed fertility by birth order, offers the most effective basis for forecasting the movement of family size for future mothers. When plotted on the graph, series of completed fertility by birth order offers a clear-cut time perspective that is not apparent in the series of general fertility. The projection basis can further be enhanced by examining specific factors that determine the behaviour of fertility for each birth order. Variables such as age at marriage and proportion of never married as well as childlessness are clearly related to the fertility of the first birth order. The total number of children per woman is then obtained by adding completed fertility rates for each birth order.

As a second step, after future family size has been projected, changes in fertility age pattern should be considered in order to project mean age of fertility. It should be noticed in passing that, although shifts in timing of fertility appear on examination of past trends to be less violent than shifts in family size, their impact on total fertility rate is nonetheless appreciable and, therefore, should receive close scrutiny. It is well documented that variations in the total fertility rate (post-war baby boom and very likely also the recent dramatic fall in fertility) is the result of the combined effect of shifts both in the family size and in the timing of childbearing. Age at first marriage, child-spacing and parity distribution are of prime interest and anticipation of their future trends will help to formulate assumptions with regard to the mean age of fertility.

The final step in this procedure involves derivation of period total fertility rate for future years from the two cohort measures mentioned above -- family size and mean age of fertility. A cohort-to-period translation model of the Ryder type (1969) could be used to perform this operation. The latter model could be improved by a built-in factor that would make allowances for the fact that changes in the timing of fertility are not necessarily linear as implied in the Ryder translation model but might vary in magnitude from one cohort to another.

After the period total fertility rate was obtained all that remains to be done is to distribute it by age of mothers. It is at this stage that the model described in this paper is expected to render most valuable service to the forecaster. Indeed total fertility can be distributed by age by means of formula (1) from a few selected parameters, and the forecaster needs no longer to go through lengthy procedures of projecting fertility rates for individual ages.

In order to generate fertility distribution by formula (1), one has the alternative of projecting dependent parameters a_1, a_2, m_1, m_2, or of projecting independent parameters -- , moments and other measures associated with frequency distribution. Which one of these approaches shall be used is a matter of personal judgement. Maybe measures such as mean age or mode age are more meaningful concepts than the constants $\underline{a}$ and $\underline{m}$ and the rationales for their future movement can be more easily substantiated (Stone, 1970).

Selection of the independent parameters to be projected depends upon which particular procedure, among the three developed in this paper, will be used for calculating the constants implied in formula (1). The best results in terms of goodness of fit are obtained by the procedure involving four distribution moments (4M). But if only an approximate fit is sufficient for the purpose of projecting the births, as this is actually the case (see next paragraph), then the practical considerations, such as the easiness of projecting future movement, become overriding in the selection of specific independent parameters to be projected. In this respect procedure (1M) may be preferred. Instead of four moments, it involves only two relatively simple measures, mean and mode age of fertility. Once the mean age is projected mode age can easily be derived because there is a very high positive correlation between these two measures. Fertility age interval, which in this procedure is assumed to be a fixed one, can be set up in the light of some testing against historical series. It has been shown in this paper that in the case of Canada, the model performs better with 17 than with 15 as the starting age of childbearing.

In the making of population projections by age by means of the component method, the number of annual births for future years is an important input. In order to demonstrate how well our model performs in terms of generating annual number of births, Table 6 below has been prepared. It contains, for few selected years for Canada, ratios of estimated to actual number of births, whereby estimated numbers are obtained by multiplying the number of females in ages 15 to 49 by the age-specific fertility rates derived by means of three variants of calculating parameters in formula (1). One is impressed by the close agreement between actual and estimated figures. Deviation is negligible except in one case where it comes near to two per cent. It is true, this high agreement is achieved at aggregate level only, partly through compensation of errors at individual ages. Yet it should be borne in mind that what is required for population projections, as input, are not births at individual ages, but the total number of births, and in this respect our model performs almost to perfection.

Table 6 RATIO OF ESTIMATED TO ACTUAL NUMBERS OF ANNUAL BIRTHS FOR A FEW SELECTED YEARS FOR CANADA

Year	4M	2M (15-50)	2M (17-50)	1M (15-50)	1M (17-50)
1926	1.001	.986	.999	1.005	1.000
1931	1.001	1.001	1.000	1.010	.998
1941	.998	.997	.998	1.000	.996
1951	.998	.997	.997	.984	.996
1961	1.003	1.005	1.001	1.013	1.002
1969	1.008	1.006	1.011	1.018	1.009

CONCLUSION

Additional research might be needed before a parametric model of fertility projections of the type outlined here can be made fully operational. This paper contains, nevertheless, important ingredients for such a model. It is shown that fertility distribution by age can be derived by mathematical function from only a limited number of parameters that need to be projected. The paper offers alternative ways of calculating these parameters, it tests their results against fertility data for Canada and points to potentials for further developments in the areas of fertility projections. Reduction of age schedule of fertility to only a few meaningful parameters makes possible an in-depth analysis to a degree that cannot be achieved under conditions of conventional procedures used in fertility projection. Furthermore, the model is a powerful labour-saving device, for in defining fertility in terms of a mathematical function, computers can be utilized in performing many of the involved operations.

REFERENCES

Elderton, W.P., Frequency Curves and Correlation, Cambridge University Press, 1930.

Keyfitz, N., Introduction to the Mathematics of Population, Reading Massachusetts, Adison-Wesley, 1968.

Mitra, S., The Pattern of Age-Specific Fertility Rates, Demography, Vol. 4, No. 2, pp. 894-906 1967

Murphy, E.M. and Nagnur D.N. "A Gompertz That Fits: Applications to Canadian Fertility Patterns", Demography, Vol. 9, No. 1, 1972, pp. 35-50.

Romaniuk, A., and Tanny, S.M., 'Projection of Incomplete Cohort Fertility for Canada by means of the Gompertz Function', Analytical and Technical Memorandum, No. 1, Dominion Bureau of Statistics, Census Division, Canada, Ottawa, March, 1969.

Ryder, N.B., "The Emergency of a Modern Fertility Pattern; United States, 1917-1966" in Fertility and Family Planning, S.J. Behrman et al (ed.). Ann Arbor, Michigan: The University of Michigan Press, 1969.

Stone, L.O. "Parametric Approaches of the Age Distribution of Cohort Total Fertility for Projections", Unpublished Manuscript, Dominion Bureau of Statistics, Ottawa, July, 1970.

RANDOM-ROUNDING TO PREVENT STATISTICAL DISCLOSURES

M.S. Nargundkar, Statistics Canada
Walt Saveland, Indian and Northern Affairs

This is an extract from our "Random-rounding: A means of preventing disclosure of information about individual respondents in aggregate data". Anyone interested in using or critically examining the random-rounding procedure should request a copy of the full report from either author. The procedure is being used in the production of nearly all population and housing tabulations of the 1971 Census of Canada. Nevertheless, the authors are personally responsible for the assertions and arguments presented herein. Mr. Saveland was with Statistics Canada during most of the developmental work on random-rounding.

In Canada, our recently enacted Statistics Act requires (Queen's Printer for Canada, 1970, Section 16.1.b) that no sworn employee of Statistics Canada "shall disclose or knowingly cause to be disclosed, by any means, any information obtained under this Act in such a manner that it is possible from any such disclosure to relate the particulars obtained, from any individual return to any identifiable individual person, business or organization".

This requirement is very similar to its counterpart in the former Act. Its introduction, however, coincides with the increasing practical difficulty of preventing disclosures in statistical data. While the new Act was being drafted and legislated, planning and "tooling up" for the 1971 Census of Canada were also well under way. The Census tabulation programme is far more complex than that of any former Census. In addition to pre-planned tabulations, it includes efficient means of access to the data base in order to produce special-area and special-population tabulations as they are needed and requested. With outputs on paper, microfilm and magnetic tape and with the indeterminate accumulation of special requests in future years, the tabulation programme is much too large and interrelated to admit to manual checking for even the most rudimentary violations of individuals' confidences.

The rigorous wording of the new Act apparently requires prevention of possible disclosures in aggregate (tabulated) data, yet no concrete guides exist in legal precedent as to what would be and what would not be illegal disclosure. Therefore, it was the advantage of Statistics Canada to develop sufficient safeguards to avoid even the risk of violating the law or public trust. Because of the imminence and complexity of 1971 Census tabulations, an acceptable means of preventing disclosure had to be amenable to easy introduction into computer programmes which had already been developed. Also, an acceptable means of prevention must not distort the statistical meaning of data; it must be mathematically straightforward enough so that the statistical agency can give a general guarantee of its harmlessness.

1. Possibilities of disclosure

In order to have criteria with which to examine the effectiveness of random-rounding, we begin with a catalogue of the logical possibilities of disclosure in aggregate data. Adhering strictly to the Statistics Act, two assumptions underly the catalogue. First, a cell count or a measure derived from cell counts may be an illegal disclosure regardless of what additional information, from other counts and measures or from independent sources, must be mustered so that information about an individual person or organization can be deduced. Second, all cell counts and derived measures are treated as if exactly correct, as if completely free of error. Though the second assumption might be improbable in any given instance, the possibility of disclosure,not the likelihood, depends on the possibility of such exactitude. Hopefully, the caution which these assumptions prescribe will result in more than foiling the possibilities of disclosure. It should discourage most attempts to find disclosures and, perhaps, spare someone from being victimized by means of a seeming disclosure which is no less threatening just because it is inaccurate.

The simplest possibility is direct disclosure from a cell which either is empty or counts one individual (person, business or other organization). If a particular individual is known to be the only one who could be counted in that cell, the only one who could have the particular combination of attributes counted by that cell, then a count of 0 discloses that he does not have the particular combination of attributes, a count of 1 discloses that he does indeed have the particular combination of attributes.

Given the possibility of disclosure from a cell counts of 0 or 1, it is easy to imagine similar possibilities from counts of 1 or 2, 2 or 3, and so forth. If one individual is known to have a particular combination of attributes and a second individual is known to be the only other one who could have the same combination of attributes, then a cell count of 1 for the combination of attributes discloses that the second individual does not have them, a count of 2 discloses that he does have them. The same reasoning works if two, three, or any number of individuals have a particular combination of attributes and only one other individual could have the same combination. Because the individuals known to have the particular combination of attributes can be thought of as being eliminated from the cell count in order to see whether a 0 or 1 remains, we call this disclosure by elimination within a cell.

Given the possibility of elimination within a cell, it is again easy to imagine the next possibility - disclosure by elimination among cells. In the simplest instance, this would be to subtract one cell count from another in order to find a difference of 0 or 1. Of course, the two cells must be related such that one is the subset of the other - so that differencing their counts yields the count of a second subset, the subset which complements the first with respect to the entire set. If a particular individual is known to be the only one who could have the particular combination of attributes counted by the second subset, then a difference of 0 discloses that he does not have the particular combination, a difference of 1 discloses that he does with this simple instance in mind, it is obvious that the count of the set and/or of the first subset could be sums of cell counts. Also, the difference between counts of the set and the first subset might be greater than 1 and still subject to disclosure by elimination within a cell. The full report gives a detailed discussion of various eventualities.

The above three possibilities seem to us fundamental, and we only mention two derivative varieties. First, given the possibilities of disclosure about individual persons or organizations through cell counts which reveal whether or not they have particular combinations of attributes, it follows that disclosures about organizations might be derived from disclosures about the attributes of members and that disclosures about persons or sub-organizations might be derived from disclosures about attributes of organizations of which they are members. Because of the necessity to extend information from members to organizations or vice-versa, we call this disclosure by extended correspondence. The second variety is disclosure from derived measures. It is simply the derivation of disclosing cell counts from measures which themselves had originally been derived from cell counts - weighted sample counts, percentiles, means, rates, and percentages. The full report discusses these two varieties of disclosure in detail.

2. Random-rounding

Recognition of the danger of disclosure has often been focussed on the very small cell counts; replacing these small counts by an asterisk or other cipher has been a common remedy. Our catalogue of possible disclosures shows small cell counts to be only part of the danger: the general danger lies in the possibility of exactly correct, error-free cell counts - of any size. The strategy of random-rounding is to render improbable and indeterminable the correctness of each issued cell count by introducing an acceptably small and unbiased error into its original value - a value which itself might or might not be free of error. Each individual random-rounding error must be indeterminable, but the resultant loss of data reliability must be estimable.

Random-rounding is a variation on systematic (or conventional) rounding. Partly because the error introduced by systematic rounding need not be free of bias, partly because the rigid rules of systematic rounding can sometimes be used to deduce pre-rounded values, random-rounding must have two properties. First, any value to be rounded may be rounded either "up" or "down" so that, in the long run, positive and negative errors balance to yield zero bias. Second, every single decision to round "up" or "down" must be indeterminable so that there exist no circumstances in which a pre-rounded value can be deduced from its position in a distribution or cross-classification and from knowledge or rounding rules.

2.1 Probabilities of rounding up and down -

Imagine that all the counts in a table of, for example, m x n cells are being rounded to multiples of some integer value, called the rounding base. For each cell count, the remainder to be rounded up or down is calculated:

(2-1) $r_{ij} = c_{ij} \bmod b$

where r_{ij} is the remainder for the ij^{th} cell in the mxn table, c_{ij} is the underlying cell count, and b is the rounding base. The calculation of r_{ij} is clarified by an identity:

(2-2) $r_{ij} = c_{ij} - bk_{ij}$

where k_{ij} is the largest integer such that $bk_{ij} \le c_{ij}$. Clearly,

(2-3) $0 \le r_{ij} < b$ and r_{ij} must also be an integer value because all cell counts c_{ij}, are themselves integer values. Given r_{ij}, rounding "up" is defined as

(2-4) $\hat{c}_{ij} = c_{ij} + (b - r_{ij})$

where $\hat{c}_{ij}$ is the rounded cell count. The error, e_{ij}, introduced by rounding up is

(2-5) $e_{ij} = b - r_{ij}$ Substituting the value of r_{ij} from Eq. 2-2 in Eq. 2-4 shows that

(2-6) $\hat{c}_{ij} = b\,(k_{ij} + 1)$, that the underlying c_{ij} has indeed been rounded to a multiple of b.

Similarly, rounding "down" is defined as

(2-7) $\hat{c}_{ij} = c_{ij} - r_{ij}$ The corresponding error is

(2-8) $e_{ij} = -r_{ij}$, and the rounded value is again a multiple of b:

(2-9) $\hat{c}_{ij} = bk_{ij}$

Given the two alternatives of rounding up or down, imagine a probability being attached to each alternative, such that the sum of the two probabilities is unity. Expected rounding error for each r_{ij} can then be calculated as the sum of two products, each probability times the error which it introduces:

(2-10) $E(e_{ij}|r_{ij}) = P_{ij}\,(b - r_{ij}) + (1 - P_{ij})\,(-r_{ij})$ where $E(e_{ij}|r_{ij})$ is the expected rounding error for the r_{ij}, P_{ij} is the probability of rounding c_{ij} up and the error terms come from Eqs. 2-5 and 2-8.

One means of excluding bias from the rounding error is to require the expected error for each r_{ij} to be zero:

(2-11) $0 = P_{ij}\,(b - r_{ij}) + (1 - P_{ij})\,(-r_{ij})$.

Then, solving for P_{ij} reveals the probability to be attached to rounding up:

(2-12) $P_{ij} = r_{ij}/b$,

and the probability for rounding down is obvious:

(2-13) $(1 - P_{ij}) = 1 - (r_{ij}/b)$.

With the expected error for each $\hat{c}_{ij}$ equal to zero, the expected error for any sum of $\hat{c}_{ij}$'s will also equal zero. An appendix to the full report documents the superiority of random-rounding over systematic rounding in this respect.

2.2 Random decisions to round up and down -

Given probabilities to round up and down, such that positive and negative errors balance to yield zero bias, it remains to render indeterminable the direction of any particular instance of rounding. Then, no underlying count will be deducible from its position in a distribution or cross-classification and a knowledge of systematic-rounding rules.

An obvious means of making indeterminable each rounding decision (up? or down?) is to line the cell counts in a sequence and to match the sequence of cell counts to a sequence of random numbers: each random number in sequence will be unpredictable and will decide the rounding direction for the cell count with which it is matched. In order for the rounding decisions to conform with the rounding probabilities in Eqs. 2-12 and 2-13, the random numbers must be uniformly distributed, lying in real interval between zero and the rounding base:

(2-14) $0 \leq v_{ij} < b$

where v_{ij} is the random number matched with the ij^{th} cell count.

Given uniform distribution of the random numbers, the probability of v_{ij} being less than r_{ij} is as follows:

(2-15) $P(v_{ij} < r_{ij}) = \dfrac{r_{ij} - 0}{b - 0} = r_{ij}/b$

with r_{ij} as defined in Eqs. 1-1 and 1-2. By substitution from Eq. 2-12, this is clearly the probability of rounding "up".

(2-16) $P(v_{ij} < r_{ij}) = P_{ij}$.

Similarly, the probability of v_{ij} being greater than or equal to r_{ij} is as follows:

(2-17) $P(v_{ij} \geq r_{ij}) = \dfrac{b - r_{ij}}{b-0} = 1 - (r_{ij}/b)$.

By substitution from Eq. 2-13, this is the probability of rounding down:

(2-18) $P(v_{ij} \geq r_{ij}) = 1 - P_{ij}$.

Given Eqs. 2-16 and 2-18, it is a simple task to tie the rounding decisions for each cell count to the random number with which the count is matched: if v_{ij} is less than r_{ij}, then round up; otherwise round down. Rounding up and down being defined, respectively, in Eqs. 2-4 and 2-7, this is the random-rounding procedure for each cell count:

(2-19) IF $v_{ij} < r_{ij}$

THEN $\hat{c}_{ij} = c_{ij} + b - r_{ij}$;

ELSE $\hat{c}_{ij} = c_{ij} - r_{ij}$

where IF, THEN, and ELSE have their common meanings. Substituting from Eq. 2-1:

(2-20) IF $v_{ij} < c_{ij} \bmod b$

THEN $\hat{c}_{ij} = c_{ij} + b - c_{ij} \bmod b$;

ELSE $\hat{c}_{ij} = c_{ij} - c_{ij} \bmod b$.

2.3 Generating random numbers -

The random-rounding procedure stated in Eq. 2-19 and again in Eq. 2-20 yields the two necessary qualities--zero bias and indeterminacy. As shown in Eq. 2-20, three values - c_{ij}, b and v_{ij} must be at hand in order to random-round each ij^{th} cell. The cell count, c_{ij}, is given in each instance, and the rounding base, b, is predetermined. (The choice of b is discussed in the full report).

Given the practical need to generate sequences of random numbers, v_{ij}'s, a brief survey of literature led us to RANDU--the "power-residue" routine published by IBM (1968, p.77). Its only distinctly non-random feature is that idential "starting values" yield identical sequences of (random) numbers. In order to circumvent this weakness, each starting value is determined by the respective computer run's exact START time--a truly unpredictable value.

RANDU generates values, y_{ij}, in the open interval between zero and one:

(2-21) $0 < y_{ij} < 1$.

Multiplication of these values by b almost transforms them into the values required to satisfy Eq. 2-14:

(2-22) $0 < by_{ij} < b$.

The discrepancy between Eqs. 2-14 and 2-22 is that the former includes the exact value of zero, $\{0,b)$, while the latter excludes it, $(0,b)$. So trivial is this discrepancy that any effort to compensate for it would introduce additional error. For practical purposes:

(2-23) $v_{ij} = by_{ij}$.

It should be noted that computer execution of random-rounding, often as an addition to existing tabulation-producing programmes, is greatly eased by the inconsequence of exactly how sequences of random numbers are matched to cell counts. Any "route" may be taken through a table: our imaginary mxn table and ij^{th} cell have no other purpose than to indicate the independent and identical application of Eq. 2-20 to each cell count.

3. Final remarks

In addition to the topics covered in this extract, the full report discusses the acceptability of random-rounding error compared to other forms of error in census data, the effectiveness of random-rounding in preventing disclosures, and *a fortiori* estimation of random-rounding error in tabulations of cell counts.

Many people at Statistics Canada helped us to progress from a vague grasp of the confidentiality problem to a rigorous development and justification of random-rounding. We mention only a few - and them only for their most important contributions. E.M. Murphy brought the problem to our attention and guided initial work with bright ideas and practical sense. Ivan Fellegi, through earlier drafts of the paper referenced below, outlined both the problem and the general practical requirements of any solution. L.O. Stone and F.G. Boardman gave us notions which pointed our minds at random-rounding. R.J. Davy provided us with crucial managerial support and criticism. M.A. Mocken relentlessly chained our speculations to the practical qualities of the data-processing systems. Others, whom we do not even know, have taken much time and trouble to make random-rounding part of the computer programmes which produce 1971 Census tabulations.

REFERENCES

[1] Fellegi, I.P.: "On the Question of Statistical Confidentiality", Journal of the American Statistical Association, 67,337 (March 1972).

[2] IBM: "System/360 Scientific Subroutine Package, (360A-CM-03X) Version III, Programmer's Manual", IBM, White Plains (New York), 1968.

[3] Parliament of Canada: "Statistics Act". The Queen's Printer for Canada, Ottawa, 1970.

AN EXTENSION OF THE RESPONSE ERROR MODEL FOR APPLICATION TO COMPLEX ESTIMATORS BASED ON DIFFERENT SAMPLES

Gad Nathan, Hebrew University, Jerusalem and Central Bureau of Statistics, Israel*

0. Introduction and Summary: The basic response error model developed by the U.S. Bureau of the Census [5] has been generalized to the multivariate case by Koch [4]. The possibilities of application of the model to complex sample designs and to complex estimators are, however, limited by the requirement that all components of the vector random variable are measured for the same sample. This is not necessarily the case for many applications, such as difference, ratio or regression estimation, or panel-type sample designs, where component estimators may be based on different samples (e.g. sub-samples, partially matched samples, etc.). In the following, the basic multivariate response error model is extended to cover the case where each component of the vector variable is measured in a possibly different sample, all selected from the same finite population. The extended model is then applied to difference and ratio estimation in different response error structure situations and to the case of sampling on two occasions.

1. The Model: The formulation of the model follows that of Koch [4]. For each unit in the population, $i=1,\ldots,N$, let there be defined a p-component vector random variable, $\underset{\sim}{Y}_{it}$, by:

$\underset{\sim}{Y}'_{it}=(Y_{it}^{(1)},Y_{it}^{(2)},\ldots,Y_{it}^{(p)})$, where t indexes the sequence of repeated trials.

Any given sample is defined by the indicator random variable:

$$U_i^{(j)}=\begin{cases}1 : \text{if unit } i \text{ is in the sample for the } j\text{-th component}\\ 0 : \text{otherwise.}\end{cases}$$

So as to simplify the presentation, simple random sampling without replacement will be assumed for each of the component samples. Thus, if n_j is the sample size for component j, then:

$E\{U_i^{(j)}\}=n_j/N$ and: $E\{U_i^{(j)}U_{i'}^{(j)}\}=n_j(n_j-1)/[N(N-1)]$, $(i\neq i')$. The relationship between the samples is defined by:

$$P\{U_i^{(j)}=U_{i'}^{(j')}=1\}=E\{U_i^{(j)}U_{i'}^{(j')}\}=\begin{cases}v_{jj'} & : i=i'\\ w_{jj'} & : i\neq i'\end{cases},$$

where it is assumed that the expectation depends only on whether $i=i'$ and not on the specific values of i and i'. If we define: $n_{jj'}=n_jn_{j'}/(N\,v_{jj'})$, (where $n_{jj}=n_j$ and it is assumed for the time being that $v_{jj'}\neq 0$), then it is easy to see that:

$$w_{jj'}=\frac{n_jn_{j'}(n_{jj'}-1)}{n_{jj'}N(N-1)}\quad.$$

The statistic considered will be the sample mean: $\bar{\underset{\sim}{y}}'_t=(\bar{y}_t^{(1)},\bar{y}_t^{(2)},\ldots,\bar{y}_t^{(p)})$, where: $\bar{y}_t^{(j)}=\sum_i U_i^{(j)}Y_{it}^{(j)}/n_j$, as an estimate of the population mean: $\bar{\underset{\sim}{Y}}'=(1/N)\sum_{i=1}^{N}\underset{\sim}{Y}'_i$, where: $\underset{\sim}{Y}'_i=E_t\{\underset{\sim}{Y}'_{it}\}$ is the expected response for the i-th element over all trials. It will be assumed that there is no response bias, so that: $Y_i^{(j)}=E_t\{Y_{it}^{(j)}\}=E_t\{Y_{it}^{(j)}|U_i^{(j)}=1\}$. The variance-covariance matrix of $\bar{y}_t$ can be decomposed, as usual, as follows:

$$\underset{\sim}{V}=E_t\{(\bar{\underset{\sim}{y}}_t-\bar{\underset{\sim}{Y}})(\bar{\underset{\sim}{y}}_t-\bar{\underset{\sim}{Y}})'\}=\underset{\sim}{RV}+\underset{\sim}{SV}+2\overline{\underset{\sim}{IV}}\quad. \tag{1}$$

Letting $\bar{\underset{\sim}{y}}'=(\bar{y}^{(1)},\bar{y}^{(2)},\ldots,\bar{y}^{(p)})$ and $\bar{y}^{(j)}=\sum_{i=1}^{N}U_i^{(j)}Y_i^{(j)}/n_j$ be the sample mean of the expected values for the j-th component, this decomposition reflects the three sources of variation as follows:

Response Variance: $\underset{\sim}{RV}=E_t\{(\bar{\underset{\sim}{y}}_t-\bar{\underset{\sim}{y}})(\bar{\underset{\sim}{y}}_t-\bar{\underset{\sim}{y}})'\}$

Sampling Variance: $\underset{\sim}{SV}=E_t\{(\bar{\underset{\sim}{y}}-\bar{\underset{\sim}{Y}})(\bar{\underset{\sim}{y}}-\bar{\underset{\sim}{Y}})'\}$

Interaction: $2\overline{\underset{\sim}{IV}}=\underset{\sim}{IV}+\underset{\sim}{IV}'$, where: $\underset{\sim}{IV}=E_t\{(\bar{\underset{\sim}{y}}_t-\bar{\underset{\sim}{y}})(\bar{\underset{\sim}{y}}-\bar{\underset{\sim}{Y}})'\}$.

In the following, $V^{(j,j')}$, the (j,j') element of $\underset{\sim}{V}$, will be expressed by further decomposing each of its components:

$$V^{(j,j')}=RV^{(j,j')}+SV^{(j,j')}+2\,\overline{IV}^{(j,j')}\quad. \tag{2}$$

Define the simple response variance as:

$$SRV_i^{(jj')}=E_t\{(Y_{it}^{(j)}-Y_i^{(j)})(Y_{it}^{(j')}-Y_i^{(j')})|U_i^{(j)}=U_i^{(j')}=1\}$$

and:

$$SRV^{(jj')}=(1/N)\Sigma_i\ SRV_i^{(jj')} \tag{3}$$

and, similarly, the correlated response variance as:

$$CRV_{ii'}^{(jj')}=E_t\{(Y_{it}^{(j)}-Y_i^{(j)})(Y_{i't}^{(j')}-Y_{i'}^{(j')})|\ U_i^{(j)}=U_{i'}^{(j')}=1\}$$

and:

$$CRV^{(jj')}=\frac{1}{N(N-1)}\sum_{i\neq i'}CRV_{ii'}^{(jj')}\quad, \tag{4}$$

*This research was carried out while the author was Visiting Associate Professor at the University of North Carolina at Chapel Hill and at the Research Triangle Institute and was sponsored by U.S. Bureau of the Census Contract No. 1-35096.

(where the conditional expectations are well defined, since we assumed $v_{jj'}=P\{U_i^{(j)}=U_i^{(j')}=1\}\neq 0$ and it can easily be seen that $w_{jj'}=P\{U_i^{(j)}=U_{i'}^{(j')}=1\}\neq 0$, except in the trivial case $n_j=n_{j'}=1$).

The <u>response variance component</u> can be shown to be:

$$RV^{(jj')}=(1/n_{jj'})[SRV^{(jj')}+(n_{jj'}-1)CRV^{(jj')}]. \quad (5)$$

If we define the simple sampling variance as:

$$SSV^{(jj')}=\frac{1}{N-1}\Sigma_i(Y_i^{(j)}-\bar{Y}^{(j)})(Y_i^{(j')}-\bar{Y}^{(j')}), \quad (6)$$

the <u>sampling variance component</u> can be written as

$$SV^{(jj')}=(1/n_{jj'})(1-\frac{n_{jj'}}{N})SSV^{(jj')}. \quad (7)$$

The <u>interaction component</u>, $IV^{(jj')}$, which reflects the inter-relationship of sampling and response errors, is non-zero if:

$$Y_{ii'}^{(jj')}=E_t\{Y_{it}^{(j)}|U_i^{(j)}=U_i^{(j')}=1\}\neq Y_i^{(j)}.$$

If we define simple interaction between response and sampling deviations for the same units as:

$$SIV^{(jj')}=(1/N)\Sigma_i(Y_{ii}^{(jj')}-Y_i^{(j)})Y_i^{(j')}$$

and: $\overline{SIV}^{(jj')}=1/2(SIV^{(jj')}+SIV^{(j'j)})$; (8)

and define the simple correlated interaction between response and sampling deviations for different units as:

$$SCIV^{(jj')}=\frac{1}{N(N-1)}\sum_{i\neq i'}(Y_{ii'}^{(jj')}-Y_i^{(j)})Y_{i'}^{(j')}$$

and: $\overline{SCIV}^{(jj')}=1/2(SCIV^{(jj')}+SCIV^{(j'j)})$. (9)

Then:
$$\overline{IV}^{(jj')}=(1/n_{jj'})[\overline{SIV}^{(jj')}+(n_{jj'}-1)\overline{SCIV}^{(jj')}]. \quad (10)$$

Substituting (5), (7) and (10) in (2) we obtain:

$$V^{(jj')}=\frac{1}{n_{jj'}}\{SRV^{(jj')}+(n_{jj'}-1)CRV^{(jj')}$$
$$+(1-n_{jj'}/N)SSV^{(jj')}+2[\overline{SIV}^{(jj')}$$
$$+(n_{jj'}-1)\overline{SCIV}^{(jj')}]\}, \quad (11)$$

where the last two terms drop out if $Y_{ii'}^{(jj')}=Y_i^{(j)}$ for all (i,i'). It should be noted that this expression has the same form as that given by Koch [4] for the case where all components are measured on the same sample, with only the sample size n replaced by $n_{jj'}$. Thus, methods proposed by Chai [2] and Bailar and Dalenius [1] to estimate the population parameters on the basis of a single sample can be applied.

For the special case, $v_{jj'}=0$ $(j\neq j')$, i.e. non-overlapping samples for the j-th and j'-th component, the resulting modification of the decomposition, (for $j\neq j'$), is:

$$V^{(jj')}=SCRV^{(jj')}+IRV^{(jj')}-\frac{1}{N}SSV^{(jj')}+2\overline{SCIV}^{(jj')}, \quad (12)$$

which is the limit of (11) as $1/n_{jj'}$ goes to zero. It should be noted that $V^{(jj')}$ is independent of sample size, in this case.

2. <u>Application to Complex Estimators:</u> The above model can easily be applied to a variety of sampling designs with respect to the relationship between samples for different components. In the following, the application of the extended model to complex estimators is considered. For the sake of algebraic simplicity, we shall assume no interaction between sampling and response deviations, in the sense that $Y_{ii'}^{(jj')}=Y_i^{(j)}$ for all $i,i'=1,\ldots,N$ and for all $j,j'=1,\ldots,p$. The results can easily be extended to the case of non-zero interactions. The variance of a linear combination of the sample means: $\bar{y}_{t\ell}=\Sigma_j\ell_j\bar{y}_t^{(j)}$, as an estimate of $\bar{Y}_\ell=\Sigma_j\ell_j\bar{Y}^{(j)}$, will be:

$$var(\bar{y}_{t\ell})=\Sigma_{j,j'}\ell_j\ell_{j'}V^{(jj')} \quad . \quad (13)$$

Similarly the variance of an analytical function of the sample means $g(\bar{y}_t^{(1)},\ldots,\bar{y}_t^{(p)})$ can be approximated by the appropriate Taylor expansion.

We shall consider two variables, X and Y, with observable values at the t-th trial for the i-th unit X_{it}, Y_{it}. Define $X_i=E_t(X_{it})$; $Y_i=E_t(Y_{it})$ and let $\bar{X}=(1/N)\Sigma_iX_i$, $\bar{Y}=(1/N)\Sigma_iY_i$, $\bar{X}_t=(1/N)\Sigma_iX_{it}$ and $\bar{Y}_t=(1/N)\Sigma_iY_{it}$, be the population means of the expected values and of the values for the t-th trial respectively. Let:

$V_{xy}=(1/N)\Sigma_iE_t(X_{it}-X_i)Y_{it}-Y_i)$; and similarly V_{xx} and V_{yy} ;

$C_{xy}=\frac{1}{N(N-1)}\sum_{i\neq i'}(X_{it}-X_i)(Y_{i't}-Y_{i'})$; and similarly C_{xx} and C_{yy} ;

$S_{xy}=\frac{1}{N-1}\Sigma_i(X_i-\bar{X})(Y_i-\bar{Y})$; and similarly S_{xx} and S_{yy} .

Consider a single simple random sample without replacement of size n, selected from the population of N elements and defined by the indicator random variables U_i $(i=1,2,\ldots,N)$. Let $\bar{x}_t=(1/n)\Sigma_iU_iX_{it}$; $\bar{y}_t=(1/n)\Sigma_iU_iY_{it}$; $\bar{x}=(1/n)\Sigma_iU_iX_i$; and $\bar{y}=(1/n)\Sigma_iU_iY_i$ be the sample means for the t-th trial and for the expected values, respectively.

In the following Y will be the variable for which an estimate of the population mean of expected values, $\bar{Y}$, is required. Measurements of Y are available only for the sample elements at a given trial, t , so that only $\bar{y}_t$ is known. X will be an auxiliary variable for which measurements are available for the whole population, in one of the following alternative ways:

(a) X is measured at the t-th trial for the population and for the sample. In this case $\bar{x}_t$ and $\bar{X}_t$ are known but not $\bar{x}$ or $\bar{X}$. This would be the most usual case in practice, giving rise to the difference estimate: $\bar{y}_{Dt1}=\bar{y}_t+k(\bar{x}_t-\bar{X}_t)$; or to the ratio estimate: $\bar{y}_{Rt1}=(\bar{y}_t/\bar{x}_t)\bar{X}_t$.

(b) X is measured at the t-th trial for the sample, so that $\bar{x}_t$ is known, and the population mean of the expected values, $\bar{X}$, is known (or alternatively $\bar{X}_t$ has no response error), but $\bar{x}$, the sample mean of expected values, is not known. This case would be rather unusual in practice but may arise if errorless measurements are available for the whole population (or only for its mean) but there is a practical difficulty in matching the sample back, to obtain the values of X_i for the sample elements. The difference estimate for this case is: $\bar{y}_{Dt2}=\bar{y}_t+k(\bar{x}_t-\bar{X})$; and the ratio estimate is: $\bar{y}_{Rt2}=(\bar{y}_y/\bar{x}_t)\bar{X}$.

(c) The expected values of X are known for the whole population and for the sample (or X is assumed to be measured without response error), so that $\bar{x}$ and $\bar{X}$ are known. In this case the difference estimate is $\bar{y}_{Dt3}=\bar{y}_t+k(\bar{x}-\bar{X})$; and the ratio estimate is: $\bar{y}_{Rt3}=(\bar{y}_t/\bar{x})\bar{X}$.

The variances of all these estimates can be obtained from (11) and (13), by considering the vector variables: $\underset{\sim}{Y}'_{it}=(Y_{it},X_{it},X_{it})$, $\underset{\sim}{Y}'_{it}=(Y_{it},X_{it},X_i)$ or $\underset{\sim}{Y}'_{it}=(Y_{it},X_i,X_i)$, for (a), (b) or (c), respectively, with sample sizes: $n_1=n_2=n$ and $n_3=N$ and values of $n_{jj'}$: $n_{11}=n_{12}=n_{22}=n$ and $n_{13}=n_{23}=n_{33}=N$.

The variance of the difference estimates is obtained by applying (13) with $\ell=(1,k,-k)$. The common component of the variances of the three estimates which is independent of k is the variance of the sample mean:

$$n\ \mathrm{var}(\bar{y}_t)=V_{yy}+(n-1)C_{yy}+(1-n/N)S_{yy} \ . \tag{14}$$

If we define the intra-trial correlation as:

$\delta^{(jj')}=\dfrac{CRV^{(jj')}}{SRV^{(jj')}}$ (if $SSRV^{(jj')}\neq 0$) ; so that: $\delta_{xx}=C_{xx}/V_{xx}$; $\delta_{xy}=C_{xy}/V_{xy}$; $\delta_{yy}=C_{yy}/V_{yy}$, then minimal variances of the difference estimates are:

$$V_1=\min_k[n\ \mathrm{var}(\bar{y}_{Dt1})]=n\ \mathrm{var}(\bar{y}_t) -(1-f)\frac{[V_{xy}(1-\delta_{xy})+S_{xy}]^2}{V_{xx}(1-\delta_{xx})+S_{xx}} \tag{15}$$

$$V_2=\min_k[n\ \mathrm{var}(\bar{y}_{Dt2})]=n\ \mathrm{var}(\bar{y}_t) -\frac{[V_{xy}\{1+(n-1)\delta_{xy}\}+(1-f)S_{xy}]^2}{V_{xx}\{1+(n-1)\delta_{xx}\}+(1-f)S_{xx}} \tag{16}$$

$$V_3=\min_k[n\ \mathrm{var}(\bar{y}_{Dt3})]=n\ \mathrm{var}(\bar{y}_t)-(1-f)S^2_{xy}/S_{xx} \ , \tag{17}$$

where $f=n/N$.

The last terms of the right hand sides of (15)-(17) represent the possible gains over the sample mean from the use of the three difference estimates. If we assume V_{xy}, S_{xy}, $\delta_{xy}>0$, then for n sufficiently large we will have: $V_2<V_1$ and $V_2<V_3$, so that the greatest gain could be made by using $\bar{y}_{Dt2}$, (i.e. using the trial sample mean , $\bar{x}_t$, and the population expected mean $\bar{X}$), even if errorless (i.e. expected) values are available both for the sample and for the population.

If, however, the response correlations, V_{xy} and C_{xy} , are small, relative to the response variances (V_{xx} and C_{xx}) , and n is not too large then $\bar{y}_{Dt3}$ may well have the smallest variance of the three estimates.

The variance of the ratio estimates, $\bar{y}_{Rt}$, will be approximated by applying (13) with $\ell=(1,-R,R)$. Thus (approximately):

$$n\ \mathrm{var}(\bar{y}_{Rt1})\doteq n\ \mathrm{var}(\bar{y}_t) +(1-f)\{R^2[V_{xx}(1-\delta_{xx})+S_{xx}] -2R[V_{xy}(1-\delta_{xy})+S_{xy}]\} \tag{18}$$

$$n\ \mathrm{var}(\bar{y}_{Rt2})\doteq n\ \mathrm{var}(\bar{y}_t) +R^2\{V_{xx}[1+(n-1)\delta_{xx}]+(1-f)S_{xx}\} -2R\{V_{xy}[1+(n-1)\delta_{xy}]+(1-f)S_{xy}\} \ ; \tag{19}$$

$$n\ \mathrm{var}(\bar{y}_{Rt3})\doteq n\ \mathrm{var}(\bar{y}_t)+(1-f)[R^2S_{xx}-2RS_{xy}] \ . \tag{20}$$

The conditions under which the ratio estimates are better than the sample mean, i.e. $\mathrm{var}(\bar{y}_{Rt})<\mathrm{var}(\bar{y}_t)$, are then:

$$R<2\,\frac{V_{xy}(1-\delta_{xy})+S_{xy}}{V_{xx}(1-\delta_{xx})+S_{xx}} \ , \text{ for } \bar{y}_{Rt1} \ ;$$

$$R<2\,\frac{V_{xy}[1+(n-1)\delta_{xy}]+(1-f)S_{xy}}{V_{xx}[1+(n-1)\delta_{xx}]+(1-f)S_{xx}} \ , \text{ for } \bar{y}_{Rt2} \ ;$$

$$R<2\,\frac{S_{xy}}{S_{xx}}\quad ,\ \text{for}\ \ \bar{y}_{Rt3} \tag{21}$$

If response correlations are small, relative to the correlations between expected values, the conditions (21) for $\bar{y}_{Rt1}$ and $\bar{y}_{Rt2}$ may be more stringent than the condition for $\bar{y}_{Rt3}$ and in particular if:

$$\frac{V_{xy}(1-\delta_{xy})}{V_{xx}(1-\delta_{xx})}<\frac{S_{xy}}{S_{xx}}\quad ,\ \text{for}\ \ \bar{y}_{Rt2}\ \ \text{and if:}$$

$$\frac{V_{xy}[1+(n-1)\delta_{xy}]}{V_{xx}[1+(n-1)\delta_{xx}]}<\frac{S_{xy}}{S_{xx}}\quad ,\ \text{for}\ \ \bar{y}_{Rt1}\ . \tag{22}$$

The last relationship would always hold if $\delta_{xy}<\delta_{xx}$ for large enough n .

Comparing $\bar{y}_{Rt1}$ and $\bar{y}_{Rt2}$, we have $\mathrm{var}(\bar{y}_{Rt2})>\mathrm{var}(\bar{y}_{Rt1})$ if:

$$R>2\,\frac{V_{xy}[1+(N-1)\delta_{xy}]}{V_{xx}[1+(N-1)\delta_{xx}]}=2\,\frac{E_t\{(\bar{Y}_t-\bar{Y})(\bar{X}_t-\bar{X})\}}{E_t\{(\bar{X}_t-\bar{X})^2\}}\ . \tag{23}$$

Thus, if the correlation between the trial means, $\bar{X}_t$ and $\bar{Y}_t$, is small, the ratio estimator $\bar{y}_{Rt2}$ would be preferred to $\bar{y}_{Rt1}$, i.e. the trial population, $\bar{X}_t$, should be used rather than the errorless expected population mean, $\bar{X}$, for blowing-up the trial sample ratio, $\bar{y}_t/\bar{x}_t$, even if $\bar{X}$ is known.

3. Application to Sampling on Two Occasions:

Let samples of the same sizes, n , be selected on each of two occasions such that $m(<n)$ elements are matched and measured on both occasions and $u(=n-m)$ are unmatched. Let Y be the variable for the second (current) occasion and X for the first and consider the unbiased estimate from the t-th trial for $\bar{Y}$:

$$\bar{y}'_t=a\bar{y}_{mt}+(1-a)\bar{y}_{ut}+b\bar{x}_{mt}-b\bar{x}_{ut}\quad , \tag{24}$$

where $\bar{y}_{mt}$, $\bar{x}_{mt}$ are the sample means for the matched part, $\bar{y}_{ut}$, $\bar{x}_{ut}$ are the sample means for the unmatched parts and a, b are any constants. Set: $\underset{\sim}{Y}'_{it}=(Y_{it},Y_{it},X_{it},X_{it})$

with: $n_1=n_3=m$; $n_2=n_4=u$;

$v_{12}=v_{14}=v_{23}=v_{24}=v_{34}=0$; $v_{13}=m/N$; $n_{13}=m$.

If we define

$$T_y^2=V_{yy}-C_{yy}+S_{yy}\ ;\ T_x^2=V_{xx}-C_{xx}+S_{xx}\ ;\ \text{and}$$

$$\rho=(V_{xy}-C_{xy}+S_{xy})/(T_xT_y)\ ,$$

then the minimal variance of (24) is:

$$\min_{a,b,U}\,[\mathrm{var}(\bar{y}'_t)]=C_{yy}-S_{yy}/N+\frac{T_y^2}{n}\cdot\frac{(1+\sqrt{1-\rho^2})}{2}\ , \tag{25}$$

where $U=u/n$.

The variance of the simple sample estimate for the second occasion, $\bar{y}_t=\frac{1}{n}(m\bar{y}_{mt}+u\bar{y}_{ut})$, is:

$$\mathrm{var}(\bar{y}_t)=C_{yy}-S_{yy}/N+T_y^2/n\quad . \tag{26}$$

The factor $\frac{1}{2}(1+\sqrt{1-\rho^2})$ in (25) represents the reduction in the third term of the variance of the simple estimate, (26), obtained by using matched samples and is the same in form as obtained in classical sampling theory without response errors (e.g. in Cochran [3]). However, if the correlated response error C_{yy} is large relative to T_y^2 , the matching will not significantly reduce the total variance.

4. References:

[1] Bailar, A. and Dalenius, T. "Estimating the Response Variance Components of the U.S. Bureau of Census' Survey Model." Sankhya B, 31 (1969), 341-60.

[2] Chai, J.J. "Correlated Measurement Errors and the Least Squares Estimator of the Regression Coefficient." Jour. Amer. Statist. Assoc. 66 (1971), 478-83.

[3] Cochran, W.G., Sampling Techniques, 2nd edition, Wiley, 1963, 342-4.

[4] Koch, G.G. "An Alternative Approach to the Multivariate Response Error Models, with Applications to Estimators Involving Subclass Means." (1972). Submitted for publication.

[5] Hansen, M.H., Hurwitz, W.N. and Bershad, M.A. "Measurement Errors in Censuses and Surveys." Bull. Inter. Statist. Inst. 38: 2, (1961), 359-74.

PRELIMINARY EVALUATION OF 1969 MONEY INCOME DATA COLLECTED IN THE 1970 CENSUS OF POPULATION AND HOUSING

by
Mitsuo Ono *
Social and Economic Statistics Administration, Bureau of the Census

Introduction

All persons 14 years old and over who were included in either the 15 or 5 percent sample of the 1970 Census of Population and Housing were asked a series of questions on their money income received in 1969. In order to reduce respondents' burdens of reporting income information, a skip pattern was used whereby respondents reporting not working at all in 1969 were not required to answer the first three earnings questions. Hence, depending on their responses to the work experience question, persons receiving the Census sample questionnaires were asked to complete either three or six income questions.

The income information requested in the 1970 Census covered money income regularly received before deduction for taxes, and excluded lump sum payments, such as net capital gains. See reference 1 for further details.

Since money income data compiled in the Decennial Censuses are used widely for many purposes, it was deemed advisable to present a summary evaluation on this subject using currently available data. This preliminary note is organized under five main headings: (1) analysis of income allocation rates, (2) comparison of income levels of families fully reporting their income information with others, (3) comparisons between income information compiled in the 1970 Census and in the March 1970 Current Population Survey (CPS), (4) comparisons between aggregate incomes reported in the 1970 Census with benchmark totals, and (5) a summary.

Analysis of Income Allocation Rates

In any large-scale statistical undertaking such as the Decennial Census, incomplete, illegible, or missing questionnaire items are bound to occur, even though every effort is made to minimize these problems. A computer editing operation corrects for inconsistencies. The allocation procedure assigns acceptable entries in place of nonentries on questionnaires because they were not reported, poorly marked, or were not read by the electronic scanning equipment. Income assignments (positive or negative dollar amounts or none) are based on fully reported income information of persons with similar social and economic characteristics. In the 1970 Census, this income "hot deck" allocation procedure covered four matrices, three for workers (covering earnings of all workers, income other than earnings for all workers other than farmers and farm managers, and income other than earnings for farmers and farm managers) and one for nonworkers relating to their incomes other than earnings. These matrices contained various combinations of persons' characteristics such as age, sex, relationship to household head, race, work experience, major occupation group, weeks worked, and class of worker through which the allocations take place. For example, there were 2,094 matrix cells for allocating earnings of persons reporting that they had worked in 1969 but who did not report their earnings on the questionnaire. The allocations took place in the order of the processing of individual records. When no reported income information was stored in the computer, initial missing income items were assigned based on a "cold deck" table of starting constant values.

In addition, because of enumerator or processing errors, it is sometimes necessary to assign all of the characteristics for a person or for a household through a substitution or replication procedure. The term "allocation" as used in the text of this paper also includes a small number of such substitution cases.

All income nonentries of family members were allocated prior to deriving family income. The allocation rates for family income shown in table 1 relate to all families for which any family member 14 years old and over had one or more income items allocated. In the 1970 Census, if there was either a positive or negative dollar amount reported on any of the three earnings questions, "none" assignments, in either the earnings or "other income" question, were not counted as income allocations. These allocation rates are unit counts and are not weighted by the income amounts contributed by family members, e.g., the family is counted as allocated although the family head, usually the major earner, had fully reported income but one of his teenage sons did not. The allocation rate for persons is derived by dividing the number of persons with one or more income nonentries which have been allocated by the number of persons 14 years old or older. The family allocation rate is obtained by dividing the number of families with one or more family members having an allocated income by the total number of families.

Preliminary Findings

The family income allocation rates presented in table 1 show that of approximately 51.2 million families enumerated in the 1970 Census, about 10.6 million or about 21 percent of all families had at least one member of the family with one or more income items allocated. This rate was about double the 10.6 percent computed for the 1960 Census. The definitions used to compute the allocation rates for the two censuses are about similar. It should be noted that there were three income questions in the 1960 Census and six income questions in the 1970 Census; undoubtedly this increase in the number of income items had a direct effect on the higher allocation rate in 1970.

It should be further noted that the allocations do not have as much impact on income reporting as these rates imply since, if ever one member of a family has one or more income items allocated, the family as a whole is treated as an allocated unit whether or not other family members had fully reported their income information. Hence, a more useful index is the proportion of the total income allocated. This analysis, planned for the future, is not yet available. An alternative measure to the family allocation rate is to compute the allocation rate on a persons basis. The persons allocation rate was 12.5 percent in the 1970 Census. Even though this is only about half the family rate, the persons allocation rates in the 1970 Census, were about double those from the 1960 Census. The breakdown of this rate shows that the rate for men (13.3 percent) was slightly higher than the rate for women (11.8 percent). The comparable rates from the 1960 Census were 6.2 percent overall, 6.4 percent for men, and 6.0 percent for women.

The family income allocation rate varied widely by family income intervals. As shown in table 1, of the 51.2 million families, the Census tabulated about 1.3 million families with incomes of less than $1,000 (including net losses). Approximately 28 percent of these families were families for which some or all of the income information had been allocated. At the other end of the distribution, there were 398,000 families with family income greater than $50,000. The allocation rate for this group was 27 percent. These two, 28 and 27 percent, were the highest allocation rates. The lowest allocation rate of 18.3 percent was recorded by families in the $10,000 to $11,999 family income class interval. Thus, the overall distribution of allocations follows a bimodal-type distribution with families at the two extremes of the distribution sharing the peak rates.

Table 2 shows the distribution of family income allocation rates by State areas. For total money income, it shows that the allocation rates ranged from a peak of 26 percent for the State of Nevada (30 percent for the District of Columbia) to a low of 17 percent for Montana. In the 1960 Census, these rates ranged from 15 percent for both of the States of Colorado and Florida (24 percent for the District of Columbia) to a low of 8 percent for Iowa. Also, a further analysis was made of the 1970 Census family allocation rates by states which were enumerated predominantly using the mailout-mailback system as compared with states using other enumeration procedures. The allocation rates between states using these alternative enumeration procedures showed no appreciable difference and hence, it appears that the type of enumeration procedure had no impact on varying the level of the allocation rates.

<u>Comparison of Income Levels of Families Fully Reporting Their Income Information With Others</u>

As shown in table 2, the difference in median family income levels "before and after" allocation was almost negligible for the country as a whole and by states. This difference for the United States as a whole was about $50 resulting from a "before allocation" value of $9,642 and an "after allocation" value of $9,590. The median income of family units with one or more members having some income information allocated was about 3 percent less than family units for which the income information was not allocated. These figures indicate that lower than median income families were more likely to have one or more members requiring some income information allocated.

<u>Comparisons Between Income Information Compiled in the 1970 Census and in the Current Population Survey (CPS)</u>

Since 1947, annual income information for the Nation as a whole has been compiled in the March supplement to the CPS. References 2 and 3 provide more information regarding the CPS. Since income information from the CPS, among others, is collected under more controlled conditions than under the Census operations, e.g., the interviewers are more experienced and trained to obtain information in depth, the CPS results provide a source by which the overall quality of Census results can be ascertained.

<u>Preliminary Findings</u>

As shown in table 3, the 1969 median income of families obtained from the March 1970 CPS was $9,433, approximately 1.7 percent less than comparable median income from the 1970 Census. Overall, data in this table show that the correspondence between the 1970 Census and the March 1970 CPS was closer than between the 1960 Census and the March 1960 CPS. The difference in 1959 median family incomes was about 5 percent as compared with the difference of about 2 percent in the 1969 median family incomes.

Table 4 shows aggregate money incomes by type of income as computed from the Census and the CPS. In the 1970 Census, aggregate income was computed at 635.5 billion dollars, about 5 percent higher than the 603.3 billion dollars computed from the 1970 CPS. The comparable rate in the 1960 Census was 9 percent. The implications of these figures regarding the probable improvements in collecting CPS income data relative to Census data are currently under investigation.

<u>Comparisons Between Aggregate Incomes Reported in the 1970 Census with Benchmark Totals</u>

Another indication of quality is the ratio of the money income amount collected in the Census to the money income amount that should be collected if there were no misreporting or underreporting of this information. Estimates of the latter, designated as BEA benchmark figures, were computed from data developed originally by the Bureau of Economic Analysis (BEA), which, in turn, derived their figures from administrative data sources. See references 3 and 4 for further details.

<u>Preliminary Findings</u>

As shown in table 4, total 1969 money income compiled in the 1970 Census was about 92 percent of the BEA benchmark total. These "benchmark" ratios were 100 percent for wage and salary, 91 percent for net self-employment income, (99 percent for nonfarm self-employment income, and 65 percent for farm self-employment income), and 60 percent for "other income" (82 percent for Social Security and railroad retirement benefits, 69 percent for public assistance payments, and 53 percent for the remaining "other income" types). The 1960 Census ratios between Census and benchmark totals were 94 percent for the total, 99 percent for wage and salary, 112 percent for self-employment income, and 62 percent for other income. The major difference in these ratios between the two censuses was the reduction in the 1970 Census benchmark ratio for net self-employment income. This is still under investigation. It may be possible that there was less reporting of gross self-employment income in the 1970 Census than in the 1960 Census.

Table 5 shows a preliminary tabulation of these "benchmark" ratios by states and by type of income. A number of states have Census wage and salary totals which are larger than BEA benchmark totals. This discrepancy is the result, among others, of problems in the benchmark figures and the misreporting of income in the wage and salary question. Thus, because of the phrase "from all jobs" in the wage and salary question, it may be possible that some had interpreted this question to cover all earnings or income instead of wage and salary income received during 1969. Another indication of reporting bias is the overreporting and underreporting (relative to benchmark totals) of the net nonfarm and farm self-employment income items. This problem involves not only getting better methods for reducing misreporting of these items in the field but also involves getting better comparable benchmark data, especially for farm household net self-employment income. These, and other problems, uncovered by table 5, are currently under investigation.

Summary

1. Although the family income allocation rate was about 21 percent, the difference in median income levels between "before and after" allocated family income was not significant.

2. The Census income data were more consistent with comparable income information from the March CPS in 1970 than in 1960.

3. Comparisons with aggregate income benchmark estimates indicate that the rate of total income reported in the 1970 Census was slightly lower than the rate reported in the 1960 Census. The reported rate differed primarily for the net self-employment income item.

References

1. U. S. Bureau of the Census, Census of Population: 1970, General Social and Economic Characteristics, Final Reports PC(1)-C1 through PC(1)-C52.

2. U. S. Bureau of the Census, Current Population Reports, Series P-60, No. 75, "Income in 1969 of Families and Persons in the United States," U. S. Government Printing Office, Washington, D. C., 1970.

3. U. S. Bureau of the Census, Income Distribution in the United States, by Herman P. Miller, U. S. Government Printing Office, Washington, D. C., 1966.

4. "Appraisal of Basic Data Available for Constructing Income Size Distributions": by Selma F. Goldsmith, in Studies in Income and Wealth, Volume 13, National Bureau of Economic Research, 1951, pp. 265-373.

* Statistical assistance and comments by members of the Family and Individual Income Statistics Branch are acknowledged.

Table 1.--Families by Total Money Income in 1969, Before and After Allocation or Substitution of Income Nonrespondents, for the United States

Total Money Income	Number			
	After Allocation or Substitution	Before Allocation or Substitution	Allocation or Substitution	Allocation or Substitution Rate
Total families	51,168,599	40,589,511	10,579,088	20.7
Less than $1,000........	1,277,006	918,679	358,327	28.1
$1,000 to $1,999........	1,733,205	1,324,402	408,803	23.6
$2,000 to $2,999........	2,260,578	1,749,836	510,742	22.6
$3,000 to $3,999........	2,499,946	1,938,347	561,599	22.5
$4,000 to $4,999........	2,601,863	2,021,902	579,961	22.3
$5,000 to $5,999........	2,934,453	2,307,430	627,023	21.4
$6,000 to $6,999........	3,146,245	2,497,712	648,533	20.6
$7,000 to $7,999........	3,451,531	2,776,852	674,679	19.5
$8,000 to $8,999........	3,640,466	2,952,407	688,059	18.9
$9,000 to $9,999........	3,457,835	2,815,029	642,806	18.6
$10,000 to $11,999......	6,585,510	5,377,909	1,207,601	18.3
$12,000 to $14,999......	7,031,917	5,709,257	1,322,660	18.8
$15,000 to $24,999......	8,176,995	6,441,997	1,734,998	21.2
$25,000 to $49,999......	1,972,996	1,467,407	505,589	25.6
$50,000 or more	398,053	290,345	107,708	27.1
Median income	$9,590	$9,642	$9,361	(X)

Source: Census of Population: 1970, General Social and Economic Characteristics, Final Report PC(1)-C2 to C52. United States total obtained by summing the states.

X - Not applicable.

Table 2.--PERCENT OF FAMILIES WITH ONE OR MORE INCOME ALLOCATIONS IN 1969 AND 1959 AND MEDIAN INCOME IN 1969 OF FAMILIES BEFORE AND AFTER ALLOCATIONS OF INCOME, FOR THE U. S., BY STATE

STATE	Allocation Rates: 1970 Census	Allocation Rates: 1960 Census 1/	Median Income - 1970 Census: After Allocation	Before Allocation	With Allocation	Ratio 2/
UNITED STATES, TOTAL	20.7	10.6 r	$9,590	$9,642	$9,361	0.995
NORTHEAST						
New England: Maine	19.1	11.2	8,205	8,183	8,314	1.003
New Hampshire	20.4	12.1	9,698	9,583	10,187	1.012
Vermont	18.8	9.6	8,929	8,899	9,087	1.003
Massachusetts	21.1	12.2	10,835	10,873	10,665	0.997
Rhode Island	23.6	14.2	9,736	9,794	9,521	0.994
Connecticut	23.0	11.6	11,811	11,886	11,509	0.994
Middle Atlantic: New York	20.9	13.4	10,617	10,669	10,391	0.995
New Jersey	21.8	12.9	11,407	11,491	11,051	0.993
Pennsylvania	19.6	10.8	9,558	9,598	9,365	0.996
NORTH CENTRAL						
East North Central: Ohio	19.8	10.7	10,313	10,356	10,102	0.996
Indiana	17.6	10.7	9,970	9,992	9,845	0.998
Illinois	21.1	13.6	10,959	10,985	10,845	0.998
Michigan	19.6	10.7	11,032	11,081	10,792	0.996
Wisconsin	21.3	9.5	10,068	10,091	9,968	0.998
West North Central: Minnesota	20.3	9.0	9,931	9,980	9,705	0.995
Iowa	18.4	8.2	9,018	8,997	9,130	1.002
Missouri	22.2	13.3	8,914	8,988	8,620	0.992
North Dakota	18.2	9.9	7,838	7,806	8,018	1.004
South Dakota	19.9	10.1	7,494	7,514	7,398	0.997
Nebraska	19.2	9.3	8,564	8,592	8,417	0.997
Kansas	19.2	9.2	8,693	8,717	8,572	0.997
SOUTH						
South Atlantic: Delaware	18.4	14.4	10,211	10,366	9,391	0.985
Maryland	20.7	11.9	11,063	11,137	10,735	0.993
District of Columbia	29.9	23.6	9,583	9,712	9,273	0.987
Virginia	20.3	10.3	9,049	9,073	8,949	0.997
West Virginia	18.6	10.0	7,415	7,450	7,235	0.995
North Carolina	21.5	10.3	7,774	7,854	7,452	0.990
South Carolina	21.8	9.7	7,621	7,645	7,527	0.997
Georgia	20.8	11.2	8,167	8,238	7,867	0.991
Florida	22.7	15.4	8,267	8,355	7,937	0.989
East South Central: Kentucky	19.6	11.3	7,441	7,486	7,236	0.994
Tennessee	22.3	10.4	7,447	7,492	7,277	0.994
Alabama	21.0	9.7	7,266	7,337	6,964	0.990
Mississippi	22.3	9.3	6,071	6,116	5,902	0.993
West South Central: Arkansas	21.0	11.2	6,273	6,301	6,161	0.996
Louisiana	24.7	11.6	7,530	7,638	7,165	0.986
Oklahoma	20.3	13.3	7,725	7,717	7,763	1.001
Texas	19.6	12.1	8,490	8,514	8,377	0.997
WEST						
Mountain: Montana	17.3	13.7	8,512	8,541	8,348	0.997
Idaho	20.2	11.2	8,381	8,325	8,633	1.007
Wyoming	19.8	10.7	8,943	8,911	9,119	1.004
Colorado	19.7	15.4	9,555	9,616	9,271	0.994
New Mexico	22.0	12.2	7,849	7,867	7,769	0.998
Arizona	22.5	12.1	9,187	9,220	6,062	0.996
Utah	20.8	11.3	9,320	9,297	9,413	1.002
Nevada	26.0	14.3	10,692	10,647	10,835	1.004
Pacific: Washington	20.3	9.5	10,407	10,457	10,182	0.995
Oregon	22.8	9.6	9,489	9,545	9,256	0.994
California	20.2	12.5	10,732	10,784	10,498	0.995
Alaska	21.4	14.0	12,443	12,502	12,200	0.995
Hawaii	21.3	10.7	11,554	11,450	11,993	1.009

1/ Nonresponse allocation rates for total family income are somewhat overstated. For a discussion on family allocation rates, see text, page LXXXVII in the U. S. Summary.
2/ Ratio of after allocation to before allocation   r = revised
Source: Census of Population: 1970, General Social and Economic Characteristics, Final Report PC(1)-C2 to C52, and Census of Population: 1960, General Social and Economic Characteristics, Final Report PC(1)-1C to 52C

Table 3.--INCOME IN 1969 AND 1959 OF FAMILIES BASED ON THE DECENNIAL CENSUSES AND THE MARCH CURRENT POPULATION SURVEYS, FOR THE UNITED STATES

Total Money Income	Family Income in 1969		Family Income in 1959	
	1970 Census	March 1970 CPS	1960 Census	March 1960 CPS
UNITED STATES				
Number............thousands......	51,169	51,237	45,128	45,062
Percent......................	100.0	100.0	100.0	100.0
Less than $1,000.................	2.5	1.6	5.6	5.1
$1,000 to $1,999.................	3.4	3.4	7.5	8.3
$2,000 to $2,999.................	4.4	4.6	8.3	9.3
$3,000 to $3,999.................	4.9	5.3	9.5	10.1
$4,000 to $4,999.................	5.1	5.4	11.0	11.7
$5,000 to $5,999.................	5.7	5.9	12.3	13.2
$6,000 to $6,999.................	6.1	6.4	10.7	11.0
$7,000 to $7,999.................	6.7	7.3	8.6	8.4
$8,000 to $8,999.................	7.1	7.4	6.6	10.6
$9,000 to $9,999.................	6.8	7.0	4.9	
$10,000 to $11,999...............	12.9	13.0	10.5	9.1
$12,000 to $14,999...............	13.7	13.7		
$15,000 to $24,999...............	16.0	15.6	3.3	2.4
$25,000 to $49,999...............	3.9	3.2	1.3	0.7
$50,000 or more..................	0.8	0.4		
Median income....................	$9,590	$9,433	$5,660	$5,417
Mean income......................	$10,999	$10,577	(NA)	(NA)

NA - Not available

Source: Census of Population: 1970, General Social and Economic Characteristics, Final Reports PC(1)-C2 to C52; Census of Population: 1960, General Social and Economic Characteristics, Final Report PC(1)-C1, U. S. Summary, and March 1970 and 1960 Current Population Surveys.

Table 4.--COMPARISON OF CENSUS, CURRENT POPULATION SURVEY, AND BUREAU OF ECONOMIC ANALYSIS ESTIMATES OF AGGREGATE INCOME IN 1969, 1959, AND 1949, BY TYPE OF INCOME, FOR THE UNITED STATES

Year and Type of Income	Estimates of Aggregate Income (In billions)				
	Census		Current Population Survey		Bureau of Economic Analysis (BEA) 1/
	Families and Unrel. Indiv.	Persons 14 years old and over	Families and Unrel. Indiv.	Persons 14 years old and over	
1969					
Total income	$635.5	$633.8	$603.3	$608.0	$692.4
Wage or salary income	499.4	(NA)	474.8	478.9	497.8
Self-employment income	56.7	(NA)	51.8	52.1	62.1
Nonfarm	47.9	(NA)	43.4	43.6	48.6
Farm	8.8	(NA)	8.4	8.5	13.5
Other income	79.4	(NA)	76.7	77.0	132.5
Social Security	22.2	(NA)	22.3	22.4	27.0
Public assistance	4.6	(NA)	4.9	5.0	6.6
Other	52.6	(NA)	49.5	49.6	98.9
1959					
Total income	332.3	331.7	304.5	306.7	353.1
Wage or salary income	(NA)	246.5	231.8	233.5	249.8
Self-employment income	(NA)	47.9	36.1	38.3	42.6
Other income	(NA)	37.3	32.6	32.7	60.6
1949					
Total income	155.2	173.2	160.2	159.8	191.0
Wage or salary income	(NA)	2/124.3	120.4	120.0	128.8
Self-employment income	(NA)	2/ 31.1	27.2	26.5	31.3
Other income	(NA)	2/ 16.6	13.2	13.3	30.9

Year and Type of Income	Ratio of --					
	Census to CPS		Census to BEA		CPS to BEA	
	Fam. and Unrel. Indiv.	Pers. 14 yrs. old and over	Fam. and Unrel. Indiv.	Pers. 14 yrs. old and over	Fam. and Unrel. Indiv.	Pers. 14 yrs. old and over
1969						
Total income	105	104	92	92	87	88
Wage or salary income	105	(X)	100	(X)	95	96
Self-employment income	109	(X)	91	(X)	83	84
Nonfarm	110	(X)	99	(X)	89	90
Farm	105	(X)	65	(X)	62	63
Other income	104	(X)	60	(X)	58	58
Social Security	100	(X)	82	(X)	83	83
Public assistance	94	(X)	69	(X)	74	76
Other	106	(X)	53	(X)	50	50
1959						
Total income	109	108	94	94	86	87
Wage or salary income	(X)	106	(X)	99	93	93
Self-employment income	(X)	125	(X)	112	85	90
Other income	(X)	114	(X)	62	54	54
1949						
Total income	97	108	81	91	84	84
Wage or salary income	(X)	104	(X)	97	93	93
Self-employment income	(X)	117	(X)	99	87	85
Other income	(X)	125	(X)	54	43	43

NA - Not available X - Not computable.

1/ - 1969 BEA estimates were prepared by the Bureau of the Census using the Bureau of Economic Analysis data.

2/ - These estimates are based on preliminary sample tabulations rather than on final results because the final data do not contain distributions of each type of income. The aggregate total income estimated from the preliminary sample is in close agreement with the comparable aggregate estimated from the final data.

Table 5.—TYPE OF INCOME IN 1969 AND 1959 OF FAMILIES AND UNRELATED INDIVIDUALS IN THE DECENNIAL CENSUSES AS A PERCENT OF COMPARABLE BUREAU OF ECONOMIC ANALYSIS ESTIMATES IN 1969 AND 1959, FOR THE UNITED STATES, BY STATE

STATE	1970 Census									1960 Census 1/			
	Total money income	Wage or salary income	Self-employment income			Other Income				Total money income	Wage or salary income	Self-employment income	Other income
			Total	Nonfarm	Farm	Total	Social Security income	Public assistance income	Other income				
UNITED STATES, TOTAL	92	100	91	99	65	60	82	69	53	94	99	112	62
NORTHEAST													
New England													
Maine	90	104	77	90	33	53	80	85	42	92	101	101	59
New Hampshire	96	104	109	104	347	60	77	69	54	100	106	126	62
Vermont	92	101	113	124	87	55	80	73	46	92	98	103	65
Massachusetts	90	100	93	93	87	57	79	69	50	92	98	134	55
Rhode Island	92	100	90	89	200	64	77	82	59	92	96	132	60
Connecticut	90	101	94	96	56	53	79	67	48	92	96	132	60
Middle Atlantic													
New York	88	98	93	93	99	55	79	61	50	90	94	130	51
New Jersey	92	101	96	96	99	58	80	84	51	96	100	121	60
Pennsylvania	91	99	90	88	120	61	81	71	53	92	97	125	56
NORTH CENTRAL													
East North Central													
Ohio	92	99	86	85	98	60	82	70	53	94	98	117	61
Indiana	92	99	83	90	66	63	83	101	54	96	99	115	69
Illinois	89	97	93	95	83	54	79	69	47	93	96	115	60
Michigan	92	100	91	89	123	58	83	72	51	95	99	117	58
Wisconsin	95	104	96	97	92	61	84	65	53	96	101	117	61
West North Central													
Minnesota	93	101	93	100	81	61	83	77	53	97	102	111	64
Iowa	88	105	76	96	64	53	83	55	43	94	103	102	58
Missouri	94	103	90	96	77	65	84	68	58	89	94	99	60
North Dakota	91	106	99	103	97	48	80	84	37	88	91	94	65
South Dakota	86	105	81	115	68	51	82	77	40	85	97	79	63
Nebraska	85	102	77	93	66	49	83	79	39	91	97	99	60
Kansas	88	99	91	100	80	54	83	64	45	98	109	99	61
SOUTH													
South Atlantic													
Delaware	88	101	73	95	28	46	75	59	40	80	92	111	37
Maryland	94	101	97	100	71	61	82	64	56	94	98	125	59
District of Columbia	83	92	135	133	(NA)	48	77	73	45	92	104	96	56
Virginia	96	103	109	108	119	61	82	90	56	99	103	119	64
West Virginia	93	102	96	87	-386	59	85	74	45	92	96	114	66
North Carolina	91	99	80	101	45	59	81	69	50	94	101	94	60
South Carolina	92	100	89	98	55	56	77	92	46	96	103	101	60
Georgia	92	99	83	103	38	60	80	77	52	97	101	103	73
Florida	99	107	91	112	23	83	97	84	79	98	104	102	79
East South Central													
Kentucky	93	101	85	90	73	66	86	71	57	98	102	121	65
Tennessee	94	102	95	95	95	58	82	82	48	96	100	117	65
Alabama	95	104	90	112	42	63	83	63	55	96	101	99	72
Mississippi	89	101	71	111	30	57	81	81	43	93	102	81	75
West South Central													
Arkansas	89	103	70	107	37	60	84	84	48	93	100	84	79
Louisiana	89	99	89	101	49	53	81	70	44	95	98	114	71
Oklahoma	97	107	106	112	88	62	86	74	53	100	105	110	72
Texas	93	102	97	113	56	56	82	77	49	95	100	106	63
WEST													
Mountain													
Montana	93	104	99	118	81	58	81	95	51	91	98	95	64
Idaho	96	112	77	101	56	63	82	78	54	98	100	112	69
Wyoming	95	105	111	109	114	55	81	76	47	91	97	102	57
Colorado	98	105	113	111	122	65	83	72	60	95	100	115	65
New Mexico	91	100	87	118	42	58	81	69	51	97	99	105	74
Arizona	96	103	84	108	23	78	90	87	74	102	104	107	89
Utah	99	105	110	111	108	65	85	73	59	99	101	122	68
Nevada	90	95	91	89	115	62	74	69	58	94	95	117	71
Pacific													
Washington	94	102	89	98	58	62	85	85	55	98	101	128	68
Oregon	97	105	97	100	82	69	86	87	63	95	101	103	67
California	92	100	95	106	36	65	84	64	61	95	100	107	65
Alaska	94	95	137	134	-325	57	60	65	56	101	103	137	65
Hawaii	91	99	92	97	58	50	78	63	45	103	106	134	68

1/ Based on persons 14 years old and over, not directly comparable with the 1970 Census.
NA - Not available.

SOME ASPECTS OF REDESIGN OF THE CANADIAN LABOUR FORCE SURVEY

R. Platek and M.P. Singh, Statistics Canada

1. INTRODUCTION

The Labour Force Survey, established in 1945, was completely redesigned in 1963 as a multi-purpose continuing field survey with the aim of providing monthly estimates of important labour force characteristics at the provincial and the national level. At present the survey covers all civilian, non-institutional population of 14 years of age and over from the ten provinces. All the population of the Yukon Territory and the Northwest Territories plus the Indian and Eskimo population living on reserves and crown land in the ten provinces, is excluded from the coverage of the Labour Force Survey (LFS).

The LFS is comprised of two main parts: self-representing units (SRUs) and non-self-representing units (NSRUs). The SRUs are cities whose population exceeds 15,000 persons or whose unique characteristics demand their establishment as SR units. The NSRUs are the areas lying outside of the SRUs. In addition, apartment buildings in the larger SRUs and the special areas of census each form separate frames in the LFS. Special features of the present design for the different parts are briefly indicated in the corresponding sections and for a detailed description reference may be made to a report on Methodology-LFS (1965) and a paper by Fellegi, Gray and Platek (1966).

The present design of the LFS is based on the Census of 1961. The concepts, definitions, boundary demarcations, population counts, etc. of the 1961 Census were used at several stages in designing the survey. Several changes have taken place in the structure of the labour force since the Census of 1961, which affect the design to a varying degree. To a large extent the information on these changes is provided by the Census of 1971. A complete review of the LFS design is therefore necessary using the information provided by the 1971 Census and other sources.

Due to the difference in design and extent of updating used in different parts (viz. SRU, NSRU, apartment and special area) of the LFS, it was decided at an early stage of the development of the redesign program that each part would be reviewed individually and the decision on the extent of the redesign in one part would not directly influence the decision on another part.

In Section 2 the main objectives of the redesign are summarized while in Section 3 some discussion covering the use of theoretical studies in redesign is given. The effects of the changes between the 1961 and 1971 Censuses on the design of the LFS are discussed separately for the SR units and NSR units in Sections 4 and 5, respectively. Dependency of the design on censuses and the suitability of alternative schemes of selection are also discussed in these sections. Sections 6 and 7 deal with the possible improvements, on the basis of some studies undertaken, to the apartment frame and the special area frame. In Section 8, the possibility of extending the coverage of the LFS is discussed.

2. MAIN OBJECTIVES OF THE DESIGN

In a continuing large scale multi-purpose survey such as the Canadian LFS, the major objectives and priorities are decided at the inception of the survey, and the survey is designed accordingly with scope for incorporating changes depending on future requirements. Subsequent redesigns usually examine and incorporate changes with regard to (i) up-to-dateness of the sample, (ii) coverage, (iii) introduction of new methodology and (iv) level of data requirements. Changes in the design necessitated by change in the level of data requirement are not discussed in this paper. Other factors mentioned above are briefly discussed in relation to the objectives set for the present LFS redesign.

(i) Up-to-dateness: All aspects of the design which utilize 1961 Census or other obsolete data will be examined. Changes will be made using information available from the 1971 Census, other more recent sources and the results of theoretical studies. In some parts of the LFS, for example in SRUs and the apartment sample, attempts will be made to reduce the dependency of the design parameters on the census.

(ii) Coverage: As mentioned earlier, all the population of the Yukon Territory, the Northwest Territories plus the Indian and Eskimo population living on reserves and crown land in the ten provinces, is excluded from the coverage of the current LFS. Due to an increasing demand for LFS estimates for these populations the feasibility of conducting LFS in these areas will be studied.

(iii) New Methodology: Taking into consideration the results of continuing theoretical studies on LFS data and availability of vast computer resources, more refined methods of stratification, sample allocation, formation of sampling units, etc. will be examined. In addition, several new sampling methods suitable for the LFS will be compared with the existing methods for their efficiency, cost, variance estimation and operational convenience. The method (with necessary modifications) which is most suitable will be used.

3. THEORETICAL STUDIES

A regular feature of the LFS is to carry out theoretical studies using LFS data in order to assess the suitability of the ongoing design and implement changes when necessary. The studies undertaken on a continuous basis are (i) binomial factors, (ii) components of variance and (iii) cost components. Binomial factors measure the overall effects of stratification, allocation and the methodology. Past studies on binomial factors do not provide direct evidence of deterioration of the design. In the components of variance study (Gray, 1971) the total variance is split into different components and therefore the relative magnitude of these components would help to decide upon the sample allocation at different stages of sampling. The magnitude of these components, together with the corresponding cost components would determine the optimum size of the sampling units at different stages of sampling.

The LFS redesign procedure will not be based only on the results of detailed mathematical analysis of cost and variance. However, the results of these studies, together with some operational studies, will indicate the best of several alternatives that might be acceptable. In the finalization of the program an adjustment will be made to reconcile the theoretically desirable and operationally feasible.

4. SELF-REPRESENTING UNITS

4.1 Special Features: The self-representing units (SRUs) are cities whose population exceeds 15,000 persons or whose unique characteristics demand their establishment as SRUs. Larger SRUs are further sub-divided into compact subunits which in turn are sub-divided into segments (mostly city blocks with well defined boundaries). From each subunit six or a multiple of six segments are selected systematically with probability proportionate to size (PPS). The size measure is related to the number of households as of 1961 Census. The households within sampled segments are selected systematically. The selected households within each segment remain in the sample for six consecutive months when they are replaced by another group of households from the same segment. In addition, there is a similar scheme for the rotation of segments. On the basis of partial field counts (discussed in Section 4.3) the sample is regularly updated to take account of extra-ordinary growth in the subunits.

Due to regular updating of the present SRU sample and simplicity in the selection procedure, at an early stage of planning for redesign it was decided that the SRU sample should not be redesigned simply because a particular feature of the design requires revision. A decision was made to examine different features of the present design individually and then decide collectively the extent to which the SRU sample should be redesigned. Brief discussions on the choice of a selection scheme and problems of updating are given in the following sub-sections.

4.2 Selection Scheme: The present method of selecting segments, namely PPS systematic (referred to as Method 1) is quite simple to operate but suffers from certain disadvantages on theoretical considerations. From several 'unequal probability without replacement' selection methods available, two well-known methods were chosen, keeping in mind that six (or multiple of six) segments are to be selected. These methods are

Method 2: Random order--PPS systematic (Hartley and Rao, 1962).

Method 3: Random group--PPS (Rao, Hartley and Cochran, 1962).

Method 2 is essentially Method 1 applied to a random ordered list of segments within the subunit. In Method 3, the segments within a subunit are to be grouped into six (or multiple of six) random groups and then one segment selected with PPS from each group. These methods were examined for their suitability for the LFS and the following points noted with regard to the theoretical and operational considerations for each of three methods.

Theoretical Consideration:

(a) Comparison of Methods 2 and 3 under a super-population model indicates that in many situations met with in practice, Method 2 is likely to have smaller variance than Method 3. However, Method 3 possesses the exact theory for finite populations and hence exact variance and variance estimate can be obtained without any difficulty. Method 2 requires asymptotic approach and the variance and variance estimates are valid only for a large number of segments within each subunit.

(b) The efficiency of Method 1 currently in use is very much dependent upon the magnitude of serial correlation coefficient, which does not necessarily increase or decrease with sample size, as it is a function of variation within and between segments in addition to the size of the sample. Due to this the behaviour of the variance of the estimate is likely to be somewhat irregular for this method.

(c) Unbiased variance estimation is not possible in Method 1 whereas both Methods 2 and 3 provide unbiased variance estimators.

(d) Both Methods 2 and 3 are more efficient than PPS sampling with replacement.

(e) Method 3 provides simple and stable variance estimator as compared to many other unequal probability schemes.

(f) The joint inclusion probabilities are zero for most pairs of segments in Method 1 and in Method 2 their computation is quite involved when the number of segments is large.

Operational Consideration:

(a) In the context of LFS these three methods are equally simple to operate. Methods 2 and 3 would be expected to take approximately the same time as Method 1. Randomization needed in Methods 2 and 3 may not require extra time as some randomization is done also for Method 1 in the process of formation of rotation groups. In Method 3 the random groups themselves serve the purpose of rotation groups.

(b) Both Methods 2 and 3 are also applicable and simple for the rotation program used in the LFS.

(c) Method 3 possesses three special features, namely,

(i) Only one segment is selected with PPS from each random group.

(ii) After the formation of groups, sampling within the groups is done independently.

(iii) Each random group by itself is a subsample of segments from the subunit.

Use of these features can be well explored in updating the LFS sample and also in conducting special studies or surveys related to the labour force characteristics. Because of feature (iii) the LFS sample may serve as a suitable vehicle for special studies or surveys as they may be conducted in any one (or more) groups.

As mentioned previously, the LFS sample is continuously updated but due to the complexity of the updating method, among other reasons, it is limited only to areas of extraordinary growth (100% or more). Because of features (i) and (ii) above it would be possible to apply directly the Keyfitz method of revising selection probabilities to the individual random groups which have developed noticeable changes in the selection probabilities, without affecting the selections in the remaining groups of the subunit. This method does not affect the selections in the remaining groups of the subunit, also, because it is simple to execute even minor changes in selection probabilities could be revised. This aspect is discussed further in Section 4.3.

Remark 1 On the basis of the above discussions, both Methods 2 and 3 appear to be more suitable than Method 1 which is currently being used in SRUs. Comparing Method 2 and 3, from a theoretical standpoint one is equally preferable to the other. However, from operational viewpoints Method 3 seems to have very definite advantages over Method 2.

4.3 Uneven Growth and Updating Procedure: Population growth usually occurs at a faster rate in self-representing units than in non-self-representing units. Along with natural growth there is a marked tendency for persons living in rural or urban centers of NSRUs and also in smaller SRUs to migrate to larger SRUs. In addition, there is frequent movement of population within a particular SRU because of developments in certain parts of the SRU, location of area of work and other facilities. The result of all these is an uneven rate of growth in different parts of the SRU.

Uneven growth, in addition to creating administrative problems concerning enumerators' assignments, increases the sampling variance of the labour force estimates. This is because the size measures used in the present design were accurate when the LFS was redesigned in 1963 but they tend to become consistently less accurate as a result of uneven growth. Continued use of these original size measures does not affect the unbiasedness of the LFS estimates, but as most labour force characteristics are correlated with the size measure, the more out of date the measures become, the lower the correlation between them and hence the higher the resultant variance of the LFS estimates.

Partial Field Count: In order to check the deterioration of the design as a result of uneven growth, a scheme of partial field count is in operation in SRUs. In this scheme every subunit is checked annually, according to a fixed rotation program, and revised dwelling counts are then obtained only for those areas where growth is 100% (or more) of the original count. The sample in such subunits (or part thereof) is then revised using a prescribed method which does not affect the unbiased aspect of the estimates.

One of the major drawbacks of the present partial field count scheme is that revised counts are not available for segments for which growth is less than 100%. The updating procedure therefore allows for an increase in sampling variance of the LFS estimates to a magnitude which in some cases may be serious. This limitation of the updating procedure based on partial field counts necessitates the revision of the sample after every census in order to check the deterioration of the design.

Complete Field Count: This scheme is envisaged as a regular feature of the LFS replacing the present partial field counts. In this plan it is intended to obtain the dwelling counts, segment maps and other necessary information for all segments of the subunits irrespective of the rate of growth. According to a fixed rotation program each subunit will be completely checked in the field annually. The data collected from the first rotation will be used in the present redesign and the data from subsequent rotations will be used in the regular updating of the sample.

For the purpose of the present redesign, on an average the field counts so obtained would be only six months out of date, whereas the data obtained from the 1971 Census, if used, would be about three years out of date by the time the last province is redesigned. Field counts would thus have advantages over the census counts with regard to their up-to-dateness, and since the sample could be selected in stages as the data are received throughout the year, it would have a better stabilizing effect and uniform workload.

During subsequent years (rotations) the data collected would enable revision of the sample even when the growth is (for example) five to ten percent and thus the LFS estimates would retain their original efficiency.

Remark 2 Complete count scheme, together with random group method of selection and Keyfitz method of revising the sample (Section 4.2, item 3) should ensure that the sample in the self-representing units is up-to-date for a long period of time. As the design in SRUs will not depend on the Census of 1971, it would not require redesigning even after the next census, with the exception of the areas annexed to SRUs.

5. NON-SELF-REPRESENTING UNITS

5.1 Special Features: Non-self-representing units (NSRUs) are areas lying outside the SRUs. These are a combination of rural areas and small urban centers. Due to the relatively low density of population the sample in NSRUs is selected in four stages from each stratum of an economic region. The required number of primary sampling units (PSUs) are delineated and 2 PSUs are selected without replacement with unequal probabilities following Fellegi's method. Segments (rural and urban separately), clusters and households form the subsequent stages of selection.

The extent of redesign required in the NSRU sample, unlike the SRUs, would depend upon a single factor, namely the extent of restratification needed as a result of changes in boundaries of economic regions and census enumeration areas and also due to annexations. These factors were examined in detail and a brief resume is presented in Sections 5.2 to 5.5.PSU delineation and selection are discussed in subsequent sections.

5.2 Changes in the Economic Region Boundaries: Each province in Canada is divided into a number of economic regions (ERs). These ERs usually consist of a group of contiguous counties (or divisions) with a similar economic structure.

In the present design of LFS the DDP ERs of 1961 were used as primary strata. These ERs were further subdivided into strata following a particular principle ensuring minimum variation in the labour force characteristics within a stratum.

A review of DDP economic regions was carried out applying 'rigorous statistical techniques' leading to the delineation of regions. In November 1971, the Regional Statistical Policy and Coordinating Committee recommended a new set of regions. Main recommendation of the committee is that these 1971 ERs be 'utilized uniformly by the bureau divisions as an additional level of spatial reference for the purpose of data tabulation, development of new data and the survey design wherever its application is appropriate'.

For the purpose of LFS, use of these subprovincial regions has added advantages from both theoretical and operational considerations, such as:

(a) because the new set of ERs is formed on 'statistical principles' on the basis of more recent information, they are suitable as primary strata

(b) an ER as an area to be further stratified is more manageable than the province as a whole

(c) as the sample would be representative of each region of the province, estimates with a known degree of precision could be easily obtained at the subprovincial level. In addition it would allow more flexibility in the design since changes could be introduced, if needed, in any one (or more) regions without affecting the design of the other regions. Also it would be useful in conducting special studies and comparisons at the ER level.

In many cases the composition of the ERs in the new set is quite different from that of the old set of DDP regions. Where there are changes in the ER boundaries, the present strata will cut across the new ER boundaries. To avoid this and retain the efficiency of stratification it is essential that changes in the strata boundaries be incorporated to accommodate the 1971 ER boundaries. The extent to which restratification would be necessary can be obtained by comparing the composition of the two sets of ERs. The number of ERs with boundary changes is given in the following Section 5.3.

5.3 Annexation to SRU: The boundaries of many SRUs have changed due to growth or some other administrative and political considerations. As a result, parts of the NSR strata or in some cases the entire strata have been included in a neighbouring SRU. Consequently, the stratification in such ERs is adversely affected which could result in poor efficiency of the LFS estimates. The number of SRUs affected due to annexation is given in the following Table 1. Figures in column (5) refer to those ERs which are affected by annexation of 10% or more. Annexation in the province of Quebec has not been obtained as all ERs are affected by boundary changes and will require restratification.

Table 1: 1971 ERs Affected

Province	Total No.of ERs		No. of 1971 ERs affected due to		
	1961	1971	ER Boundary Change	Annexation to SRU	Joint consideration
(1)	(2)	(3)	(4)	(5)	(6)
Newfoundland .	6	6	-	1	1
P.E.I.	1	1	-	1	1
Nova Scotia ..	4	5	2	2	4
New Brunswick.	4	5	3	1	4
Quebec	10	10	9	n.a.	9
Ontario	10	10	-	4	4
Manitoba	7	7	-	-	-
Saskatchewan .	6	6	5	-	5
Alberta	7	8	5	1	5
B.C.	9	9	8	1	8
Total	64	67	32	11	41

5.4 Extent of Restratification: All the ERs which are affected by either boundary changes or annexation will require restratification. Table 1 shows that in 1971, 41 out of 67 ERs fall into this category. Of the remaining 26 ERs, 6 ERs consist of a single stratum and 6 ERs do not have any NSR area (these are distributed in different provinces). Thus, in 1971 there are only 14 out of a total of 67 ERs which do not require restratification on account of boundary changes or annexation.

Other operational and theoretical considerations must be taken into account when deciding upon restratification of an ER. Some of these considerations are boundary changes in enumeration areas, deterioration in the Index of Stratification, formation of new SRUs or extention of the coverage of existing SRUs.

Of the 14 ERs mentioned above, 8 were examined for the extent of changes in the boundaries of 1961 Census EAs: Ontario (4) Alberta (1) New Brunswick (1) and Manitoba (2). As the information on such changes between the Censuses of 1961 and 1971 is not yet compiled, the frequency of such changes was obtained between the Censuses of 1961 and 1966. If the relative frequency of changes between 1966 and 1971 is assumed to be the same as between 1961 and 1966, then the examination revealed that a percentage of the 1961 EAs which have changed their boundaries in the Census of 1971 varies from a low of 20% in some ERs to a high of 52% in others.

The above considerations show that for one reason or another almost all 1971 ERs would require restratification. As a result, it is contemplated that in NSR areas restratification would be carried out in all provinces with new strata formed in each of the 1971 economic regions. A project is undertaken to examine and extend the use of computer in stratification and PSU formation.

5.5 PSU Delineation and Selection: The primary sampling units (PSUs) in NSR areas consist of a group of contiguous rural enumeration areas and reasonably nearby urban areas within the stratum. While forming the PSUs in the present design attempts were made to make each PSU a replica of the stratum to which it belongs with respect to the 'discriminating characteristics' and the rural-urban population ratios. Restratification of NSR areas would imply redelineation of the PSUs. It is felt that the contiguity of the enumeration areas within PSU could be relaxed to some extent without affecting the cost of enumeration.

With complete restratification and redelineation of PSUs, the partial replacement of the sample (with maximum retention of the sampled PSUs) was found to be undesireable from both operational and theoretical considerations. A decision was made to reselect PSUs and units at subsequent stages.

As mentioned earlier, in the present design, two PSUs are selected from each stratum following Fellegi's method. This method has worked quite satisfactorily in NSRUs and is not difficult to operate for the selection of two PSUs. Unless, sufficient evidence is found in favour of some other method the PSUs in NSRUs will be selected following the present method. Consideration was given to random group method (Method 3, Section 4.2) for selection of PSUs. However, studies in different context have shown that Fellegi's method is likely to be more efficient than random group method for situations where (i) the coefficient of variation of the size variable is small and (ii) the sampling ratios (for the PSUs) is large. Both these conditions are likely to be satisfied for most of the strata because of the very principle of the PSU formation,and also the sampling ratios in terms of PSUs.

The segments of the sample PSUs will be placed in two groups (urban and rural) as in the present design. For reasons mentioned in Section 4.2, the required number of segments would be selected from each group, following Method 2 rather than Method 1, as at present. Use of Method 3 is not suitable here because of the small number of segments within each group. The clusters within sampled segments may be selected with stricter classification of possible configurations for two dimensional array of cluster selection and random start. The systematic sampling of households, the ultimate stage of sampling, will remain until a superior method is found.

6. APARTMENT SAMPLE

6.1 Definition and Coverage: In addition to the regular sample in SR areas a separate frame of apartment buildings, which have at least 30 units and 5 floors, is set up in 12 of the larger SRUs. The purpose of setting up this frame is to ensure the representation of apartment dwellers as they may be different in many characteristics from persons living in single dwellings or row housing.

6.2 Stratification and Selection Procedure: A study was conducted to examine the suitability of stratifying the apartment frame of each SRU on the basis of size of the apartment buildings, size being the number of suites in them. Frequency distribution for the SRU was obtained for 15 size classes and efficiency of stratification was calculated following two well known rules of stratification, namely (i) $\sqrt{f}$ rule and (ii) stratum total rule, for constructing two and three strata.

Both rules were found to be equally efficient for 6 out of 12 SRUs and the remaining 6 SRUs were equally distributed between the two rules. On forming two strata substantial gains over the unstratified sample were noticed in all cases. Division of the frame into three strata also showed some improvements over two strata. On practical considerations it was decided to form two strata in the seven medium sized SRUs and three strata in the two largest SRUs of Toronto and Montreal. In three smaller SRUs stratification was not considered necessary. A preliminary study conducted to examine the spatial representation of the sample revealed that before selection the two larger SRUs of Toronto and Montreal should be geographically stratified.

In the present design a two-stage probability sample is selected from the apartment frame. Apartment buildings (segments) are selected systematically with probability proportionate to size (number of suites) and households are selected systematically. The apartment frame is continuously updated according to a prescribed procedure. Consideration was given to the use of Method 3 (or 2) discussed in Section 4.2, for the selection of apartment segments from each stratum of a SRU. Due to open-endedness of the apartment frame within each strata, it is felt that PPS systematic sampling has an advantage over other methods. At the second stage, to ensure a representation of the different types of suites in the sample, a serpentine listing of the suites in the sampled apartment buildings is proposed.

7. SPECIAL AREA SAMPLE

The frame for this sample consists of special enumeration areas as defined in the Census of 1961 and the remote areas. Census special enumeration areas are grouped into three types (strata), namely, military establishments, hospitals and other institutions. The remote areas form the fourth stratum. The frame of special areas is a closed ended one. In most of the provinces, each stratum is divided into two substrata and samples are selected in two-stages from each substratum following the method used for the apartment sample. Although the labour force population in special areas is less than 2% of the total labour force, it was decided to treat them separately in order to have proper representation of the institutional population.

8. EXTENSION OF LFS TO NORTHWEST TERRITORIES AND YUKON TERRITORY

These areas differ significantly from other provinces in respect to population density, living conditions, transport and communication, seasonal variations, mobility of population, etc. These factors will have a significant bearing on the overall methodology of the LFS in these areas. As little information is known about such factors and their impact on the estimates, a pilot survey will be conducted for these areas with a view to (i) examine the operational feasibilities and to decide upon the enumeration methods, frequency of survey, rotation plan, etc., (ii) to determine the cost and other design parameters and (iii) to assess the general applicability of the LFS concepts and definitions.

Acknowledgements: We are grateful to Dr. I.P. Fellegi for his comments and encouragements during the completion of this paper. To our colleagues we are also grateful for enlighting discussions and active participation in the various projects undertaken.

REFERENCES

[1] Fellegi, I.P.: "Sampling with Varying Probabilities Without Replacement--Rotating and Non-Rotating Samples"; Journal of the American Statistical Association, Vol. 58, 183-201.

[2] Fellegi, I.P., Gray, G.B. and Platek, R.: "The New Design of the Canadian Labour Force Survey"; Journal of the American Statistical Association, Vol. 62, 421-453.

[3] Gray, G.B.: "Variance Components and Variance Functions"; talk presented at the 1st Canadian Conference on Applied Statistics, Sir George Williams University, 1971.

[4] Hartley, H.O. and Rao, J.N.K.: "Sampling with Unequal Probabilities Without Replacement"; Ann. Math. Stat., Vol. 33, 350-374.

[5] Keyfitz, N.: "Sampling with Probability Proportional to Size--Adjustment for Changes in Size"; Journal of the American Statistical Association, Vol. 46, 105-109.

[6] Rao, J.N.K., Hartley, H.O. and Cochran, W.G.: "On a Simple Procedure of Unequal Probability Sampling Without Replacement"; Journal of the Royal Statistical Society, Vol. 24 (2), 482-491.

Sampling and Survey Research Staff; Dominion Bureau of Statistics: Methodology - Canadian Labour Force Survey (1965).

SAMPLING ERRORS IN RATIO ESTIMATION

J. N. K. Rao and Richard A. Kuzik
University of Manitoba and Department of Indian Affairs and Northern Development, Canada

SUMMARY: The results of an extensive semi-empirical study on ratio estimators are reported. Assuming simple random sampling and a linear regression model $y = \alpha+\beta x+e$ with error variance $V(e|x)\propto x^g$, $g\geq 0$, the average mean square errors of seven ratio estimators of the population mean $\bar{Y}$ and the average biases and average mean square errors of two classical variance estimators and of the 'jack-knife' variance estimator, for a given x-population, are derived. Employing a wide variety of x-populations, the performances of these estimators and variance estimators are empirically investigated. The applicability of the results to other sample designs is discussed.

1. INTRODUCTION: In recent years considerable attention has been given to the investigation of the properties of ratio estimators. Suppose a bivariate simple random sample (y_i, x_i), $i=1,2,..,n$, is to be drawn from a population of N units with means $(\bar{Y},\bar{X})$. If $\bar{X}$ is known, the classical ratio estimator of $\bar{Y}$ is $\bar{y}_r = r\bar{X}$ where $r = \bar{y}/\bar{x}$ is the ratio of sample means. Beale's estimator

$$\bar{y}_B = \bar{y}_r[1 + (\frac{1}{n} - \frac{1}{N})\frac{s_{xy}}{\bar{x}\bar{y}}]/[1 + (\frac{1}{n} - \frac{1}{N})\frac{s_x^2}{\bar{x}^2}]$$

and Tin's estimator

$$\bar{y}_T = \bar{y}_r[1 + (\frac{1}{n} - \frac{1}{N})(\frac{s_{xy}}{\bar{x}\bar{y}} - \frac{s_x^2}{\bar{x}^2})]$$

are approximately unbiased in the sense that their asymptotic biases do not contain terms of order n^{-1} and order N^{-1}, where $s_{xy} = (n-1)^{-1}\Sigma(x_i-\bar{x})\cdot(y_i-\bar{y})$ and $s_x^2 = (n-1)^{-1}\Sigma(x_i-\bar{x})^2$. The well known unbiased ratio estimator of Hartley and Ross is given by

$$\bar{y}_H = \bar{r}\bar{X} + \frac{n(N-1)}{N(n-1)}(\bar{y} - \bar{r}\bar{x})$$

where $\bar{r} = n^{-1}\Sigma(y_i/x_i)$. If the sample is divided at random into $g(\geq 2)$ groups each of size p, $(n = pg)$, Jones's estimator of $\bar{Y}$ is

$$\bar{y}_J = w\bar{y}_r - (w-1)\bar{r}_g'\bar{X}$$

where $w = g\{1-(n-p)/N\}$, $\bar{r}_g' = r_j'/g$, $r_j' = (n\bar{y}-p\bar{y}_j)/(n\bar{x}-p\bar{x}_j)$, $\bar{y}_j$ and $\bar{x}_j$ being the sample means from the j^{th} group $(j=1,2,..,g)$. The asymptotic bias of $\bar{y}_J$ does not involve terms of order n^{-1} and order N^{-1}. Quenouille's estimator (originally proposed for infinite populations) is given by $\bar{y}_Q = g\bar{y}_r-(g-1)\bar{r}_g'\bar{X}$, but terms of order N^{-1} appear in its asymptotic bias. Finally, Mickey's unbiased estimator of $\bar{Y}$ is

$$\bar{y}_M = \bar{r}_g'\bar{X} + \frac{(N-n+p)g}{N}(\bar{y} - \bar{r}_g'\bar{x}).$$

$\bar{y}_M$ reduces to $\bar{y}_H$ when n=2. The approximately unbiased or wholly unbiased ratio estimators are useful in surveys with many strata and small samples within strata, especially if 'separate' ratio estimators are appropriate.

We consider three estimators for the mean square error (MSE) of $\bar{y}_r$:

$$v_1 = (n^{-1} - N^{-1})(s_y^2 - 2rs_{xy} + r^2s_x^2)$$

$$v_2 = (n^{-1} - N^{-1})(\bar{X}/\bar{x})^2(s_y^2 - 2rs_{xy} + r^2s_x^2)$$

and the 'jack-knife' estimator

$$v_3 = \bar{X}^2(1 - n/N)\{(g-1)/g\}\Sigma(r_j' - \bar{r}_g')^2$$

where $s_y^2 = (n-1)^{-1}\Sigma(y_i-\bar{y})^2$. The estimators $\bar{X}^{-2}v_2$ and $\bar{X}^{-2}v_3$ for MSE of r do not require the knowledge of $\bar{X}$.

The behaviour of estimators of $\bar{Y}$ and of estimators of MSE (or variance) of $\bar{y}_r$ (or r) may be investigated in a variety of ways, including the following (McCarthy [6]): (a) Exact analytic, in which the functional form of a distribution or a joint distribution is assumed; (b) approximate analytic, in which Taylor series approximations are used; (c) empirical studies, in which the data from actual surveys are used; and (d) Monte Carlo sampling from synthetic populations. We refer the reader to Hutchison [3], Rao and Rao [10], McCarthy [6] and Frankel [2] for details of the available results under the above categories. Some analytic results, exact for any sample size, have been obtained by assuming simple random sampling and the model

$$y_i = \alpha + \beta x_i + u_i$$

$$\varepsilon(u_i|x_i) = 0,\ \varepsilon(u_i^2|x_i) = \delta x_i^t,\ \delta > 0,\ t \geq 0 \quad (1)$$

$$\varepsilon(u_iu_j|x_i,x_j) = 0,\ i \neq j = 1, 2,...,N$$

with a gamma distribution for the variates x_i, where ε denotes the expectation operator (Rao and Rao [10]). However, it is difficult to obtain analogous exact analytic results for more complex sample designs. On the other hand, the empirical approach, employing actual survey data, permits the use of complex designs, and the properties of estimators of many parameters (including $\bar{Y}$ or $R = \bar{Y}/\bar{X}$) could be investigated with the help of a high speed computer. For instance, Frankel [2] investigated the properties of naive[1] estimators of ratios, regression coefficients, simple, partial and multiple correlation coefficients for three sample designs involving 6, 12 and 30 strata respectively and two clusters per stratum selected by simple random sampling. Employing the data collected by the U.S. Bureau of the Census in the March 1971 Current Population Survey, he generated 300 (or 200) independent samples for each design and then empirically investigated the behaviour of estimators and of variance estimators. Several y-characters and a single x-character (size of cluster) were considered in estimating R. On the other hand, Rao [9] considered a wide variety of (y,x)-populations, but confined himself to simple random sampling. An obvious limitation of the empirical approach is that the results are strictly applicable only to the particular population(s) considered. However, the empirical studies are extremely valuable in providing guidelines on the performances of various methods of estimation.

In this paper we employ a semi-empirical approach by using model (1) and a wide variety of x-populations. This combination of empirical and analytic approaches has obvious advantages as it throws further light on the performances of various methods of estimation. Moreover, it has not been possible to analytically investigate the stabilities of Beale's estimator $\bar{y}_B$ and the 'jack-knife' variance estimator v_3 (except for g=2). Although we confine ourselves to simple random sampling in this paper, it is possible to apply the present approach to more complex designs by employing suitable extensions of model (1). For instance, Konijn [5] has proposed an extension of model (1) suitable for two-stage sampling involving unequal probability sampling of primaries with or without replacement.

2. ESTIMATION OF $\bar{Y}$: The average MSE of $\bar{y}_r$, under model (1), is given by

$$\varepsilon\, \mathrm{MSE}(\bar{y}_r) = \varepsilon \underset{s}{E}(\bar{y}_{rs} - \bar{Y})^2$$
$$= \varepsilon \underset{s}{E} \{\frac{\bar{X}}{\bar{x}_s} (\alpha+\beta\bar{x}_s+\bar{u}_s) - (\alpha+\beta\bar{X}+\bar{U})\}^2$$

where $\bar{U} = \sum_1^N u_i$, $\underset{s}{E}$ denotes the average over all the $\binom{N}{n}$ possible samples s each of size n, $\bar{x}_s = \sum_{i\varepsilon s} x_i$ and $\bar{u}_s = \sum_{i\varepsilon s} u_i$. Noting that $n^2\varepsilon(\bar{u}_s^2) = \delta \sum_{i\varepsilon s} x_i^t$, $N^2\varepsilon(\bar{U}^2) = \delta\sum_1^N x_i^t$ and $nN\varepsilon(\bar{u}_s\bar{U}) = \delta \sum_{i\varepsilon s} x_i^t$, we get

$$\varepsilon\, \mathrm{MSE}(\bar{y}_r) = \frac{\delta}{N^2}\sum_1^N x_i^t + \alpha^2 \underset{s}{E}(A_{r\alpha s}) + \delta \underset{s}{E}(A_{r\delta s})$$

where the coefficients $A_{r\alpha s}$ and $A_{r\delta s}$ are functions of $\bar{X}$ and the x-values in the sample s. The expressions for $A_{r\alpha s}$ and $A_{r\delta s}$ are given in Appendix 1. Following the above lines, we derived the average MSE's of $\bar{y}_B$, $\bar{y}_T$, $\bar{y}_Q$, $\bar{y}_J$, $\bar{y}_H$ and $\bar{y}_M$ and the expressions for the coefficients of α^2 and δ ($A_{B\alpha s}$, $A_{B\delta s}$, etcetera) are presented in Appendix 1. It is possible to entertain more general models than (1) but the derivations become extremely tedious and the interpretation of the results will be difficult. For instance, if the model $y_i = \alpha + \beta x_i^{m+1} + \gamma x_i^{2(m+1)} + u_i$ with the same error structure as in (1) is used, we will have to consider the coefficients of α^2, β^2, γ^2, $\alpha\beta$, $\alpha\gamma$, $\beta\gamma$ for selected values of m and of δ for selected values of t, in order to compare the average MSE's of the estimators.

The computation of the coefficients $E(A_{r\alpha s})$, $E(A_{r\delta s})$, etcetera, for a given x-population, is done on a high speed computer. Table 1 describes the x-populations selected for the present study. The populations numbered 1-12 are natural populations with N ranging from 10 to 270 and the coefficient of variation (C.V.) of x from 0.17 to 1.03. Five artificial populations (nos. 13-17), generated from lognormal and gamma distributions, are also included. A computer program to draw all the $\binom{N}{n}$ possible samples of a given size n is available, but we adopted the following scheme to save computer time: If $\binom{N}{n} \le 2000$, draw all the $\binom{N}{n}$ samples from a given x-population and compute $A_{r\alpha} = \underset{s}{E}(A_{r\alpha s})$, $A_{r\delta} = \underset{s}{E}(A_{r\delta s}) + (\sum_1^N x_i^t/N^2)$, etc.; otherwise draw 2000 independent simple random samples each of size n when n<20, and approximate $A_{r\alpha}$ by $\sum_1^{2000} A_{r\alpha s}/2000$ and so on. We selected 1000 and 500 independent samples for N>100, n=20 and N>100, n=50 respectively. For the estimators involving g we consider the case g=n only, as the previous results indicate that g=n is a good choice. Moreover, three t-values in the range 0 to 2 (t=0, 1, 2) are selected as t is likely to lie between 0 and 2 in practice.

The values of the ratios $E_{rT\alpha} = A_{r\alpha}/A_{T\alpha}$, $E_{HM\delta} = A_{H\delta}/A_{M\delta}$ and so on for n=2, 4, 6, 8, 12 and 50 have been computed (Tables will be published elsewhere). The differences in the average MSE's of the estimators decrease as n increases and/or C.V.(x) decreases. Based on these results, we draw the following conclusions: (1) For n>2, $\bar{y}_M$ is preferable to $\bar{y}_H$ since $E_{HM\alpha}$ is substantially greater than 1, $E_{HM\delta}$>1 for t=0, 1 and $E_{HM\delta}$ for t=2 is $\ge$0.95 when C.V.(x) is not too large; the gain in efficiency of $\bar{y}_M$ over $\bar{y}_H$ is considerable when t=0. (2) $E_{MT\alpha}$ and $E_{MT\delta}$ are consistently greater than 1 so that $\bar{y}_T$ is more efficient than $\bar{y}_M$; the gain in efficiency generally increases with t when n>2. (3) $E_{JT\delta}$>1 for t=1, 2 and $E_{JT\alpha}$>0.97; for t=0, $E_{JT\delta}$ is close to 1 excepting for population 10 when n<4. The gain in efficiency of $\bar{y}_T$ over $\bar{y}_J$ is substantial for t=2, especially when n<4 and C.V.(x)>0.75. (4) $E_{rT\delta}$ when t=0 and $E_{rT\alpha}$ are significantly larger than 1; for t>1, $E_{rT\delta}$<1 but close to 1 when t=1. The gain in efficiency of $\bar{y}_r$ over $\bar{y}_T$ could be as high as 20% when t = 2 and $\alpha \doteq 0$. However, when α is significantly different from 0 and mnay strata employed, the approximately unbiased or wholly unbiased 'separate' ratio estimators are preferable to $\bar{y}_r$ (Hutchison [3]). (5) $\bar{y}_J$ could lead to small gains in efficiency over $\bar{y}_Q$. Moreover, the absolute bias of $\bar{y}_J$ is generally smaller than that of $\bar{y}_Q$, especially for small N. (6) $E_{BT\delta}$<1 for all t when n=2 or for t=1, 2 when n>2; $E_{BT\delta}$ for t=0 and $E_{BT\alpha}$ are close to 1, excepting the population 10. Moreover, as Beale has pointed out, $\bar{y}_T$ could occasionally take negative values when all sample pairs (y_i, x_i) are positive, whereas $\bar{y}_B$ is always positive. The conclusions (1)-(4) are in agreement with previous analytic results under a gamma distribution for x.

3. ESTIMATION OF MSE($\bar{y}_r$): We turn now to the performances of v_1, v_2 and v_3 as estimators of MSE($\bar{y}_r$). The results obtained here are equally applicable to the estimators of MSE(r) but only $\bar{X}^{-2}v_2$ and $\bar{X}^{-2}v_3$ are relevant as $\bar{X}$ is usually unknown when estimating ratios.

The average bias of v_i as an estimator of MSE($\bar{y}_r$) is given by

$$B(v_i) = \varepsilon E\{v_i - \mathrm{MSE}(\bar{y}_r)\} = B_{i\alpha}\alpha^2 + B_{i\delta}\delta^2 \text{ (say)},$$

i = 1, 2, 3, where $B_{i\alpha}$ and $B_{i\delta}$ are given in Appendix 2. The values of $B_{i\alpha}$ and $B_{i\delta}$ for n=2, 4, 6, 8, 12, 20 and 50 have been computed. Based on these results, a major result is that both v_1 and v_2 (t>0) underestimate MSE($\bar{y}_r$) whereas v_3 (with g=n) overestimates MSE($\bar{y}_r$) for all t; v_2 leads to overestimation when t=0 and $\alpha \doteq 0$ (this conclusion is in agreement with the analytic result under a gamma distribution for x). Other conclusions are: (1) $|B_{2\delta}|$ is smaller than $|B_{3\delta}|$ for t<1 but the difference is small for t=1, especially when n>4; the comparison between $|B_{2\alpha}|$ and $|B_{3\alpha}|$ is not clear cut. $|B_{3\delta}|$ is substantially smaller than $|B_{2\delta}|$ for t=2. (2) v_3 or v_2 is preferable to v_1 with respect to the absolute bias.

The stabilities of the estimators v_i may be judged by comparing their average MSE's. To simplify the algebra, we assume that the errors u_i are independently and normally distributed, and use the measure $\varepsilon E[v_i - \varepsilon\mathrm{MSE}(\bar{y}_r)]^2 = \varepsilon Ev_i^2 - 2(\varepsilon Ev_i)(\varepsilon\mathrm{MSE}\ \bar{y}_r) + (\varepsilon\mathrm{MSE}\ \bar{y}_r)^2$ rather than $\varepsilon E[v_i - \mathrm{MSE}(\bar{y}_r)]^2$, but the conclusions are not likely to be different from the latter measure. Let $\varepsilon E[v_i - \varepsilon\mathrm{MSE}(\bar{y}_r)]^2 = \alpha^4 M_{i\alpha} + \alpha^2\delta M_{i\alpha\delta} + \delta^2 M_{i\delta}$ and $\varepsilon Ev_i^2 = \alpha^4 \underset{s}{E}(M'_{i\alpha s}) + \alpha^2\delta \underset{s}{E}(M'_{i\alpha\delta s}) + \delta^2 \underset{s}{E}(M'_{i\delta s})$. The formulae for $M'_{i\alpha s}$, $M'_{i\alpha\delta s}$ and $M'_{i\delta s}$ will be published elsewhere. The values of $M_{i\alpha}$, $M_{i\alpha\delta}$ and $M_{i\delta}$ for n=2, 4, 6, 8, 12, 20 and 50 have been computed. On the basis of these results we draw the following conclusions: (1) $M_{1\alpha}$ is smaller than $M_{2\alpha}$ which in turn is much smaller than $M_{3\alpha}$ and a similar pattern for $M_{1\alpha\delta}$, $M_{2\alpha\delta}$ and $M_{3\alpha\delta}$ for all t, especially for smaller n and for populations with large C.V.(x). (2) $M_{2\delta}$ is smaller than $M_{3\delta}$ for all t but the difference is small for t=0 and 1, when n>20. (3) $M_{1\delta}$ is smaller than $M_{2\delta}$ for t=0 and 1, but much larger for t=2; $M_{1\delta}$ is slightly smaller than $M_{3\delta}$ for t=2. The results on v_1 and v_2 are in agreement with the analytic results under a gamma distribution for x (no analytic result on the average MSE of v_3 with g=n is available).

4. CONCLUDING REMARKS: Our results are immediately applicable to single-stage cluster sampling within strata, provided simple random sampling and 'separate' ratio estimators are used. We simply replace the variates y_i and x_i by the cluster total Y_i and the cluster size M_i respectively ($i=1,\ldots,N$). Cochran [1, p.256] proposed a model of the form (1) for cluster sampling: $Y_i=\alpha+\beta M_i+u_i$ with $E(u_i|M_i)=0$ and $E(u_i^2|M_i)=M_i^{2-g}$, g>0, where g is likely to lie between 0 and 1 which corresponds to our t between 1 and 2. Our results on estimation of $\bar{Y}$ are also applicable to sub-sampling of clusters. If $\bar{y}_i$ denotes an unbiased estimator of the cluster mean $\bar{Y}_i$, the ratio estimators are as before provided y_i and x_i are replaced by $M_i\bar{y}_i$ and M_i respectively. The between-cluster component of the MSE is the same as the MSE in single-stage cluster sampling. Consequently, the comparisons in Section 2 provide guidelines for the choice of a ratio estimator even when the clusters are sub-sampled. The variance estimators v_1, v_2 and v_3 are applicable to sub-sampling provided the clusters are selected with replacement or n/N is negligible, but their properties remain to be investigated.

If one uses the naive estimators ignoring the design effect and the model (1) with a gamma distribution for x is valid, the average MSE's would be independent of the sample design and, consequently, the previous analytic results for simple random sampling remain valid. However, the model (1) may not be realistic for some of the sample designs. For instance, with stratification it may be more reasonable to assume different intercepts α_h and/or different regression coefficients β_h between strata. Similarly, a model of the form (1) with correlated errors u_i is more realistic when systematic sampling is employed.

This work has been supported by a reasearch grant from the National Research Council of Canada.

REFERENCES

[1] Cochran, W. G. (1963). Sampling Techniques (2nd edition), New York: Wiley.

[2] Frankel, M. R. (1971). An empirical investigation of some properties of multivariate statistical estimates from complex samples. Ph.D. Thesis, University of Michigan.

[3] Hutchison, M. C. (1971). "A Monte Carlo comparison of some ratio estimators", Biometrika, 58, 313-21.

[4] Kish, L. (1965). Survey Sampling, New York: Wiley.

[5] Konijn, H. S. (1962). "Regression analysis in sample surveys", J. Amer. Statist. Assoc., 57, 590-607.

[6] McCarthy, P. (1969). "Pseudoreplication: Further evaluation and application of the balanced half-sample technique", National Center for Health Statistics, Washington, D. C., Series 2, no. 31.

[7] Murthy, M. N. (1967). Sampling Theory and Methods, Calcutta Statistical Publishing Society.

[8] Quenouille, M. H. (1959). "Tables of random observations from standard distributions", Biometrika, 46, 178-202.

[9] Rao, J. N. K. (1969). "Ratio and Regression Estimators", in New Developments in Survey Sampling, eds. N. L. Johnson and Harry Smith, Jr., pp.213-34, New York: Wiley.

[10] Rao, Poduri S. R. S. and Rao, J. N. K. (1971). "Small sample results for ratio estimators", Biometrika, 58, 625-30.

[11] Sampford, M. R. (1962). An Introduction to Sampling Theory, Edinburgh and London: Oliver and Boyd.

[12] Sukhatme, P. V. and Sukhatme, B. V. (1970). Sampling Theory of Surveys with Applications, Rome: F. A. O. of the U. N.

[13] Yates, F. (1960). Sampling Methods for Census and Surveys, London, Griffin.

TABLE 1: Description of x-population

Pop. no.	Source	N	x	C.V.(x)
1	[1], p.204	10	eye est. wt. of peaches	0.17
2	[7], p.131	176	length of timber	0.42
3	[13], p.159	43	no. persons in a kraal	0.45
4	[7], p.127	128	no. persons	0.60
5	[12], p.256	89	no. villages in a circle	0.61
6	[7], p.178	108	geographical area	0.69
7	[11], p.61	35	area under crops and grasses	0.71
8	[7], p.228	80	capital and output	0.75
9	[7], p.228	80	no. workers	0.95
10	[4], p.625	270	no. dwellings	0.99
11	[1], p.156	49	size of cities	1.01
12	[1], p.183	34	no. 'placebo' children	1.03
13	[8]	50	lognormal	0.73
14	[8]	100	gamma	0.79
15	[8]	200	gamma	0.82
16	[8]	100	lognormal	0.82
17	[8]	200	lognormal	0.85

APPENDIX 1: Formulae for the coefficients in the average MSE's of ratio estimators.

The average MSE of $\bar{y}_r$, under the model (1), is given by

$$\varepsilon[\mathrm{MSE}(\bar{y}_r)] = \frac{\delta(x^t)_p}{N^2} + \alpha^2 \underset{s}{E}(A_{r\alpha s}) + \delta \underset{s}{E}(A_{r\delta s})$$

where $(x^t)_p = \sum_1^N x_i^t$, $\underset{s}{E}$ denotes the expectation over all possible samples of a given size, and the coefficients $A_{r\alpha s}$, $A_{r\delta s}$ are given below. Similar notation is used for the other estimators $\bar{y}_H$, $\bar{y}_M$, $\bar{y}_Q$, $\bar{y}_J$, $\bar{y}_B$ and $\bar{y}_T$.

Classical ratio:

$$N^2A_{r\alpha s} = (na_{0s} - N)^2$$

$$N^2A_{r\delta s} = a_{0s}(x^t)_s(a_{0s}-2)$$

where

$$(x^t)_s = \sum_{i\varepsilon s} x_i^t, \quad a_{0s} = (x)_p/(x)_s.$$

FERTILITY AND INCOME IN CANADA

N. Baskara Rao *
Department of Sociology, University of Alberta

The main objectives of this study are to examine: i) the relationship between fertility and income using both time series and cross section data, ii) the extent of this relationship in Quebec and Ontario which have high proportions of Catholic and Protestant population respectively and iii) certain problems in the comparability of the time series and cross section as well as the micro and macro relationships.

Data and methodology:

The variables used in the time relationship are TFR (total fertility rate) and per capita personal disposable real income. Total fertility rates are from the Vital Statistics Reports and the income figures are from the National Accounts Reports. Personal income refers to the gross national income that accrues to the personal sector of the economy which is different from the government and the investment sectors. Time series analysis is restricted to the period 1926-64. Cross section data are from the census and they consist of the number of unmarried children less than age 25 and staying at home, and the earnings (wages & salaries) of families.

For the time series relationship fertility rates, lagged by one year, are regressed on income. The estimating equation is of the following form:

$$Y_i = a + bX_i + U_i \qquad \text{----- ----- -----} \quad (1)$$

where, Y_i = Total fertility rate for i^{th} year

X_i = Personal per capita real income for i^{th} year

U_i = error term

i = 1,2 . . . . T

and a & b are the estimated parameters.

Presence of autocorrelation in the residuals (U_i) was evaluated through the Durbin-Watson test and in all cases it was necessary to transform the data in order to reduce the extent of autocorrelation. Transformation of data was done as follows:

$$Y'_t = Y_t - r\,Y_{t-1} \qquad \text{----- ----- -----} \quad (2)$$

$$X'_t = X_t - r\,X_{t-1}$$

$$\text{where, } r = \sum_{t=2}^{N} U_t\,U_{t-1} \Big/ \sum_{t=2}^{N} U^2_{t-1}$$

The estimated parameters used in our discussion are from the regression equation fitted to the new series of values, X'_t & Y'_t and in all cases autocorrelation is not significant either at 5% or 1% level. We have tried to evaluate the presence of heteroscedasticity by using the approximate method of examining whether there is any systematic variation in the range of the residuals ordered according to income or time. We find that the range of the residuals does not show any systematic variation with time or the level of income. Due to fewer number of observations regression analysis could not be done for the cross section data.

Time series relationship:

In Canada, income started declining from 1928, reaching the lowest level in 1933. The period 1934 to 1964 is generally one of rising income. 1927-37 and 1959-65 are periods of declining fertility. During 1937 and 1959, the period of baby boom, fertility rose by about 49%.

Because of the nature of the trend in fertility and the possible effects of war, the period 1926-64 is divided into four sub-periods: 1926-39, 1940-45, 1946-57 and 1958-64. Since TFR is allowed to lag by one year, the corresponding subperiods are 1927-40, 1941-46, 1947-58 and 1959-65. We have fitted a single equation for the whole period using dummy variables in such a way as to estimate the parameters separately for each of the sub-periods.

Table 1. Estimated relationship between fertility and income, Canada

Period	Slope Coefficient	Correlation
1926-39	+1.342* (0.773)	+0.517
1940-45	+0.510 (1.153)	+0.138
1946-57	+2.318** (0.942)	+0.783
1958-64	-4.984** (1.526)	-0.759

* = Significant at .05
** = Significant at .01
Standard errors are given within brackets

One would normally expect that births would increase during favourable economic conditions and decrease during unfavourable economic conditions, thereby, resulting in a positive correlation between income and fertility. However, the estimated relationship (Table 1) is positive during 1926-57 and negative during 1958-64. We will examine this inconsistency in the direction of the relationship with the help of available data.

During the late 50's and the early 60's there was a moderation in the growth of Canadian economy (Economic Council of Canada, 1964: 7-53). This is indicated by the extent to which the actual growth rate falls short of the potential growth rate. The potential growth rate refers to 'a calculated measure of the total volume of production consistent with reasonably full and efficient use of the economic resources available to a nation' (Economic Council of Canada, 1964: 31). During 1957-63 actualgrowth in output consistently fell short of the potential growth.

Statistics on unemployment show similar slackening in economic activity. A growing economy requires that at least 97% of the labour force be employed. During 1946-54 unemployment rate

* The author gratefuly acknowledges the help received from Professors Karol J. Krotki, P. Krishnan, Frank Roseman, John Forster & Vic Matthews.

went beyond the 3% level only twice, while during 1954-64 it was always higher than 3%. During 1958-64 the average unemployment rate was about 6%.

Table 2. Annual average unemployment rates, Males, Selected age groups, Canada

Age groups	1950-53	1954-57	1958-61	1962-65
14-19	6.7	9.8	15.9	12.7
20-24	4.8	7.2	11.8	8.3
25-34	2.8	4.4	7.5	4.9
35-44	2.4	3.6	6.0	4.2
14 +	3.2	4.8	7.9	5.7

Source: Averages calculated from the annual unemployment rates in Sylvia Ostry, Unemployment in Canada

The extent of unemployment was quite high among the younger age groups. The duration of unemployment also increased. In 1958-65 an average of 33% of the total unemployed were seeking work for a duration of four months and above, as against only 24% during 1950-57.

Table 3. Proportion of owned dwellings, Canada

Year	Dwellings	
	owned	total
1941	56.6	100
1951	65.6	100
1961	66.0	100
1966	63.1	100

Source: Census of Canada, 1941, 51, 61 & 66

Table 4. Per cent changes in median income (constant dollars) of non-farm families by age of head, Canada

Period	Age of head		
	<25	25-34	35-44
1951-54	17.4	17.5	13.7
1954-57	5.1	10.0	7.9
1957-59	3.2	2.5	6.9
1959-61	-0.3	5.9	8.8

Source: Per cent changes calculated from data in Podoluk, J.R., Incomes of Canadians, 1968

Table 5. Average holding of 'total selected assets' by income groups among non-farm families and unattached individuals, Canada

Income ($)	Average holding of assets in dollars 1955	1958	1963
<1000	1187	3277	3081
1-1999	1332	4459	4465
2-2999	1277	5394	4657
3-3999	1093	5947	5145
4-4999	1910	8461	7171
5-6999	2281	10529	8906 10763
7-9999	4598	14796	13240
10,000 +	12468	30373	25043

Source: DBS, Income, Liquid Assets and Indebtedness of Non-farm Families in Canada, for the years 1955, 1958 and 1963.

Note: i) For 1963 and for the income group 5000-6999, average holdings refer to income groups 5000-5999 and 6000-6999.

ii) Total selected assets are: a) current savings deposits with chartered banks, savings deposits in post office savings banks, insurance companies etc. b) government of Canada bonds, public utilities, municipal and provincial bonds, industrial and corporate bonds; c) mortgages and agreements of sale on residential and other types of property; d) loans to other persons and e) estimated market value of owner occupied homes (1955 figures do not include this item).

Data on family income, ownership of dwellings and asset holdings also indicate that the economic conditions as experienced by the families during the later post war years were relatively unfavourable. Though family income increased between 1951 and 1961, most of this increase was confined to the period 1951-57. If the rise in income is sufficient enough to keep pace with the rising aspirations among the people, one would normally expect an increase in the holding of assets like houses, government bonds etc. Ownership of dwellings increased substantially during the 40's, it was more or less stable in the 50's and decreased slightly during 1961-66. Similarly, the average holding of other types of assets declined sharply during 1958-63. This is so in most of the income classes.

Women who were at their prime reproductive ages in the late 50's and the early 60's generally spent the early part of their life during a period of economic prosperity. They might have had high expectations regarding their standard of living quite in keeping with their childhood experience. High expectations on the one hand and a relatively unfavourable economic climate on the other might have contributed at least partly to the decline in TFR since 1959. However, the evidence presented here in support of a positive relationship between fertility and income in Canada cannot be considered as conclusive. It would have been better if we had comparable data on the socio-economic characteristics of individual families for the pre war period as well. These data are hard to obtain. Secondly, better indicators will be needed to measure the economic burden of the families at various points in time. It may be mentioned here that the behaviour of personal income during the post war recessions was different from that during the depression of the 30's. In recent times the share of transfer payments (e.g. unemployment compensation payments) in personal income increased considerably, thereby making personal income less sensitive to economic fluctuations. Perhaps, a single indicator or variable may not explain the variations in fertility for both the pre war and the post war periods. Further, the variations in TFR are due to the changes in the proportion married, completed family size and in the age pattern of childbearing and it is important to know how these components are related to the changes in the economic conditions.

Cross section relationship:

Cross section data available in the census consist of the number of wage earning families cross classified by the number of children (less than age 25, unmarried and staying at home at the time of enumeration), income (wages and salaries) of the family and the age of head. Wage earning families constitute approximately two-thirds of the total number of families and most of them are husband and wife families. Since data on the number of children living are available only for 1941 (not for 1951 and 1961) we had to use, for the sake of comparability, the number of unmarried children staying at home which will differ from the actual family size. However, when we take into consideration only those families where the age of head is less than 45, the difference between these two indicators of family size is negligible as can be seen from the 1941 census data. Secondly, from the 1941 data and from other evidence (Henripin, 1968: 282-83) we find that the direction of the income-fertility relationship is the same whether we use the number of children living or the number of unmarried children staying at home.

Table 6 shows the relationship between family size and income at specified age groups. The reduction in family size for all income groups in the 35-44 age group in 1951 as compared to 1941 can be attributed to the postponement of births also resulting in a smaller completed family size. That is, the 1907-11 and the 1912-16 birth cohorts who were in the 40-44 and 35-39 age groups in 1951 had to face economic difficulties at the time of their prime childbearing periods and consequently they had less time available for a possible recovery of births.

Using the differentials given in the table we can make a crude comparison of the extent and the direction of the relationship. During 1941-61 there seems to be a weakening of the negative relationship as well as a reversal to a positive relationship. For example, in 1961, the group less than age 25 shows a more consistent positive relationship. However, it is difficult to conclude that these cohorts will show the same positive relationship when their family size is completed. It may be that among the recent cohorts, earlier childbearing is more common in the higher income groups as compared to the lower income groups, thereby showing a positive relationship at a point in time. In the absence of data on the differences in the timing of births among the various income groups it is difficult to comment on a possible reversal from a negative to a positive relationship between completed family size and income.

Table 6. Average number of children (per 1000 families) according to income and age of head, wage earning families, Canada

1941 Income ($)	1941 Age < 35	1941 Age 35-44	1951 Income ($)	1951 Age < 35	1951 Age 35-44
< 450	1506	2938	500	1494	2671
450-949	1305	2661	500-999	1598	2819
950-1449	1182	2417	1000-1449	1541	2695
1450-1949	1169	2190	1500-1999	1398	2464
1950-2949	1134	1969	2000-2499	1400	2305
2950-3949	1115	1810	2500-2999	1422	2225
3950-4949	1204	1850	3000-3999	1427	2098
4950 +	1369	1938	4000-5999	1455	1949
			6000 +	1614	2005
All income groups	1259	2413		1437	2306
Differentials	-9%	-34%		+8%	-25%

1961 Income ($)	Age < 25	Age 25-34	Age 35-44
< 2000	947	2109	3111
2000-2999	882	1982	2854
3000-3999	896	1932	2747
4000-4999	943	1948	2654
5000-5999	954	1972	2606
6000-6999	957	1982	2554
7000-9999	1037	2042	2543
10,000 +	1114	2230	2679
All income groups	917	1971	2706
Differentials	+18%	+6%	-14%

Source: Census of Canada 1941, 1951 & 1961.

Differentials: Change in the average number of children in the highest income groups as compared to that in the lowest income group, expressed as per cent of the latter.

Quebec and Ontario:

The general trends in income in Ontario and Quebec are similar to those for the country as a whole. Throughout the period of study, the income level was lower in Quebec. In the case of fertility there are significant differences between these two provinces. During 1937-59, the period of baby boom, TFR in Ontario rose from 2161 to 3773 or by 75%. In Quebec on the other hand the increase was only 20%, i.e., from 3268 to 3928. During 1926-37 and 1959-65 the decline in fertility was more rapid in Quebec. As a result of this differential rate of change, TFR in Quebec was 4% less than that of Ontario in 1965 though it was 58% higher in 1926.

Table 7. Estimated relationship between fertility and income, Ontario & Quebec

Period	Slope	Correlation
Ontario		
1926-39	1.559** (0.796)	+0.614
1940-45	2.501* (1.517)	+0.440
1946-57	2.289** (0.915)	+0.639
1958-64	-4.525** (1.396)	-0.873
Quebec		
1926-39	2.200** (0.795)	+0.543
1940-45	1.322 (1.962)	+0.384
1946-57	1.142 (1.030)	+0.509
1958-64	-4.948** (1.209)	-0.867

* = Significant at .05
** = Significant at .01
Standard errors are given within brackets

The negative relationship in both Ontario and Quebec during 1958-64 has to be interpreted in the light of the observations made earlier regarding the slackening of economic activity since the middle of 1950's. Unlike in Ontario, during 1940-57 the relationship is notsignificant in Quebec. It may be noted that at the start of the baby boom TFR was relatively high in Quebec. In 1937 the year of very low fertility, TFR in Ontario was 2161 while in Quebec it was 3268 or about one and a half times larger than the former. In other words, the historic transition from high to low fertility was achieved much earlier in Ontario than in Quebec. During 1940's along with the rise in income, Quebec experienced certain social changes, like the rapid increase in urbanization and enrollment in schools, the weakening of the hold of religion on education (Whyte, D.R., 1968) and the rise of the middle class more conscious of its right to economic security (Guindon, 1968). These social changes combined with the relatively high fertility might have reduced the chances of any substantial increase in fertility. Most of the increase in TFR which Quebec experienced can be attributed to the shift toward an younger age at childbearing. In the 30-34 age group the family size of the 1922-26 birth cohort was 83% complete as against 74% in the 1907-11 cohort. Similar changes can be observed in Ontario. However, unlike in Ontario, completed family size in Quebec did not show any marked increase. For example, in Quebec, compared to the 1907-11 cohorts, the 1917-21 and the 1922-26 cohorts enlarged their completed family size by 4% and 4.6% respectively. For Ontario, on the other hand, the corresponding figures are 19% and 30%.

The Cross section data show that in Ontario during 1941-61 there was a slight reversal from a negative to a positive relationship between income and family size at specified age groups. In Quebec on the other hand even though the negative relationship had weakened, there was no such reversal toward a positive relationship. Further research on fertility differentials by education, occupation, rural-urban residence etc. in the provinces may throw some light on the relationship between the stage reached by certain population in the demographic transition and the nature of the relationship between fertility and socio-economic factors.

Time series and cross section, micro and macro relationships:

There seems to be some "discrepancy" between the positive time series relationship on the one hand and the shift from a negative to a positive cross section relationship on the other. Previous studies on the relationship between fertility and income have, with a few exceptions, shown positive time series and negative cross section relationships. For our discussion here, cross section data will mean the distribution of families by income and the number of children, and the time series data refers to the mean income and family size of a population for a series of time points. In the former there is individual correspondence between income and family size while in the latter we do not know whether there is such a correspondence (i.e., whether those who increased their family size were the same who experienced an increase in income as well). If there is no individual correspondence in the time series (aggregate) relationship how do we explain the reproductive behavior of the individuals on an average? In a way the cross section-time series "discrepancy", interpretation of a macro relationship (or the consistency between the micro and the macro relationships) and the identification of the true relationship seem to be related to one another (Chipman, 1957). In the following paragraphs we intend to highlight some of these problems.

Specification bias:

Time series and cross section findings can differ with respect to the direction and magnitude of the relationship as a result of excluding a relevant independent variable. Let us assume that the true relationship is of the following form (using deviations from the mean):

$$Y = \beta_1 X_1 + \beta_2 X_2 \qquad \text{----- ----- -----} \quad (3)$$

where, Y = fertility (total fertility rate or family size),
X_1 = income
X_2 = education (number of years of schooling)

In the estimating equation let education be ommitted, then

$$Y = bX_1 \qquad \text{----- ----- -----} \quad (4)$$

Since Y is influenced by both X_1 and X_2,

$b=\beta_1 + \beta_2 \cdot \alpha$

where,

$$\alpha = \Sigma X_1 X_2 / \Sigma X_1^2$$

The bias in 'b' as a result of excluding X_2 is given by the terms other than $\beta_1 \cdot \beta_2$ is a measure of the effect of the left out variable on Y, and 'α' or the auxiliary coefficient refers to the covariation between the included and the excluded variables. In other words, in (4) income captures the influence of education (on fertility) and the extent of this influence is determined by the auxiliary coefficient. In (4) we have the unconditional effect (direct effect of income as well as the indirect effect of education through income) of income on fertility, while in (3), β_1 measures the conditional effect, i.e., the effect of income on fertility given that education is held constant.

The sign of 'b' depends on the algebraic size of the quantity $\beta_2 \cdot \alpha$. Let us suppose that fertility and education are negatively related, while fertility and income, income and education are positively related. And we fit equation (4) to both time series and cross section data. In this case the positive covariation between education (number of years of schooling) and income will be larger in the cross section than in the time series data. Because of this, the auxiliary coefficient will be larger in the cross section data. Given positive income effect ($+\beta_1$) and negative education effect ($-\beta_2$) a larger value of 'α' (in cross section) can result in 'b' having a negative sign in the cross section relationship and positive sign in the time series relationship. In general, the way in which certain variables are operationalized and the extent of variability in the variables (or indicators) can be such that the cross section and time series findings may not be consistent with respect to the magnitude and the direction of relationship.

There are certain situations where it is difficult to specify the theoretical relationship in the time series data. Let us take for example, the interrelationship among the opportunity cost of bearing and rearing children, husband's income, wife's ability and willingness to earn income and family size. It is difficult to specify this interrelationship in a time series equation. One may argue that these variables represent certain characteristics of the individuals which are not strictly comparable to the macro variables generally used in the time series equation. Theoretically, the variables that influence the individual family size are the same whether we analyse the data cross sectionally or over time. The problem here is one of identifying the theoretical relationship given certain type of data. Consequently there are some difficulties when we try to interpret the statistical relationship in order to explain and predict the reproductive response. Instances are not rare when cross section equations are used to predict changes in fertility over time. Similarly, from time series relationship inferences are drawn about differential fertility among couples with varying levels of income. Let us briefly examine these aspects.

It is reasonable to assume that the relationship between, say, income and family size is determined by the individual decision making units. Through cross section and time series equations we seek to understand what this individual relationship (on an average) is like. Each couple or family has a tendency to change (increase) their family size given certain change (increase) in their income. This relative change in family size and income can be called individual elasticity which is equal to or greater than zero. Now, the distribution of family size (according to income) at a point in time can be considered as a function of the initial distribution of income among the couples, individual elasticity of demand for children and the rate of increase in income in the past. If the poor have higher elasticity than the rich, an unit increase in income will result, at a point in time, in a larger family size for the poor as compared to the rich. The continuation of this negative cross sectional relationship at successive points in time will depend on the extent of the difference in elasticity (between the rich and the poor) and in the rate of increase in income. However, mean income and family size may both increase over time. This positive time series relationship per se does not indicate whether the elasticity is higher or lower among the poor as compared to the rich. This is because of the lack of individual correspondence between changes in income and family size. For the same reason, given the positive time series relationship, it is not necessary that the poor and the rich should have small and large family size respectively.

<u>Aggregation bias</u>:

The lack of individual correspondence in the time series relationship poses some problems in interpreting the relationship. The question is whether the degree of the relationship estimated from the time series data can be interpreted as the response of the couples (on an average) to an increase in income. That is, whether the macro (time series) slope is equal to the mean of the micro slopes or individual elasticities. Aggregation of individual or micro functions is different from a function fitted to aggregate variables. In time series equation only aggregate variables are used. It has been shown elsewhere (Allen, 1956, 694-722; Brown, 1965: 145-161) that given unequal micro slopes (elasticities differ among the couples) the macro slope 'b' can be equal to the mean of the micro slopes ($b_i \cdot 1/N$) only when the income distribution is constant during the period. In the case of systematic changes in income (i.e., those couples with atypical elasticity experience violent changes in income) the macro slope is equal to the mean of the micro slopes plus some bias which is called aggregation bias. As a result of this bias the macro slope can underestimate or overestimate the mean of the micro slopes. If a section of the population has a high rate of growth in their income, the macro slope is more likely to reflect the fertility experience of this particular group. So far we have assumed that the differences in the elasticities (between the rich and the poor) are constant over time. If they change, which is quite likely, there is the additional problem of identifying the 'changing structure'.

We do not argue that the cross section studies are superior to the time series or vice versa. The point is that the type of data used and the way in which the variables are measured and classified have certain implications for explaining and predicting the reproductive behavior. And our discussion of income-fertility relationship has to be viewed in the light of these problems.

REFERENCES

1. Allen, R.G.D., Mathematical Economics, 1956, 694-722.
2. Basavarajappa, K.G., 'The Influence of Fluctuations in Economic Conditions on Fertility and Marriage Rates, Australia, 1920-21 to 1937-38 and 1946-47 to 1966-67', Population Studies, 1971, 25(1).
3. Becker, Gary S., "An Economic Analysis of Fertility" in National Bureau of Economic Research, Demographic and Economic Change in Developed Countries, 1960.
4. Bird, R.C. et al., " 'Kuznets Cycle' in Growth Rates: The Meaning', International Economic Review, 1965:6.
5. Blake, Judith, 'Are Babies Consumer Durables? A Critique of the Economic Theory of Reproductive Motivation', Population Studies, 1968: 22(1).
6. Brown, M.T., Specification and Uses of Econometric Models, 1965: 145-161
7. Charles, Enid., The Changing Size of the Family in Canada, 1941.
8. Chipman, J.S., The Review of Economics and Statistics, 1957: 39(2), 233-235.
9. Cho, Lee-Jay, et al., Differential Fertility in the United States, 1970.
10. Easterlin, R.A., Population, Labour Force and Long Swings in Economic Growth, 1968.
11. Economic Council of Canada, First Annual Review, Economic Goals for Canada to 1970, 1964
12. Guindon, Hubert, 'Social Unrest, Social Class and Quebec's Bureaucratic Revolution', in B.R. Blishen et al. (ed.), Canadian Society, 1968.
13. Gujarati, Damodar, 'Cyclical Behaviour of Help Wanted Index and the Unemployment Rate', The Review of Economics and Statistics, 1969: 51.
14. Haavelmo, T., 'Family Expenditure and the Marginal Propensity to Consume', Econometrica, 1947: 15(4), 335-341.
15. Henripin, J., Tendances et Facteurs de la Fecondite au Canada, 1968, 282-83.
16. Kirk, Dudley, 'The Influence of Business Cycles on Marriages and Birth Rates', in National Bureau of Economic Research, Demographic and Economic Change in Developed Countries, 1960.
17. Krotki, K.J., 'Consequences of the Demographic Wave in Western Canada', in Card, B.Y., (ed.) Perspectives on Regions and Regionalism and other papers, 1968
18. Long, L.H., 'Fertility Patterns Among Religious Groups in Canada', Demography, 1970: 7(2).
19. Peitchinis, S.G., The Economics of Labour, 1965.
20. Simon, J.L., 'The Effect of Income on Fertility', Population Studies, 1969: 23(3).
21. Sweezy, A., 'The Economic Explanation of Fertility Changes in the United States', Population Studies, 1971: 25(2).
22. Whyte, D.R., 'Religion and the Rural Church', in B.R. Blishen et al., (ed.), Canadian Society, 1968.

ROOT ESTIMATORS FOR THE MEAN OF SKEWED DISTRIBUTIONS

L. J. Ringer, O. C. Jenkins, and H. O. Hartley
Texas A&M University

For a random sample, y_i, $i = 1, 2, \ldots, n$, from a finite population of size N, the most commonly used estimator of $\bar{Y}$ is the sample mean $\bar{y}$. In fact, Hartley and Rao [1968] have shown that $\bar{y}$ is the best estimator in any competition in which only unbiased estimators are allowed.

For certain distributions, it may be that there is a biased estimator which in some sense is preferable to $\bar{y}$. In this paper we examine the class of square root estimators when the distributions have positive skewness. The mean square error (MSE) is used as the basis for determining the best estimator.

Square root estimator.

Let y_i, $i = 1, 2, \ldots, n$ denote a random sample without replacement from a finite population of size N, with all $y > 0$. Let $u = \sqrt{y}$. The two competitors for estimating the population mean are the usual sample mean

$$\bar{y} = \sum_{i=1}^{n} y_i/n$$

and the square root estimator

$$\hat{y} = C_1\bar{y} + C_2\bar{u}^2$$

where C_1 and C_2 are predetermined constants.

The comparison of the expectations and mean square errors of these two estimators will be facilitated by the use of the finite population cumulants κ_r and the sample k-statistics (e.g. Wishart [1952]) of u, where

$$k_1 = \Sigma u_i/n = \bar{u}$$

$$k_2 = \Sigma(u_i-\bar{u})/(n-1) \quad .$$

In terms of the κ parameters of the distribution of u

$$V(\bar{y}) = (\tfrac{1}{n} - \tfrac{1}{N})[\kappa_4 + 4\kappa_{31} + 4\kappa_{211}] + 2\left\{(\tfrac{n-1}{n})^2(\tfrac{1}{n-1} - \tfrac{1}{N-1}) + (\tfrac{n-1}{n})[\tfrac{1}{n(n-1)} - \tfrac{1}{N(N-1)}] - (\tfrac{2n-1}{n(N-1)})(\tfrac{1}{n} - \tfrac{1}{N})\right\}\kappa_{22} \quad .$$

which, since $E(\bar{y}) = \bar{Y}$, is also $MSE(\bar{y})$.

Let us first consider the special case, $C_1 + C_2 = 1$, so that the estimator becomes

$$\hat{y} = (1-C)\bar{y} + C\bar{u}^2$$

$$= (1-C)\,\frac{n-1}{n}\,k_2 + k_1^2 \quad .$$

Then, the bias and mean square error of the estimator, $\hat{y}$, are

$$B(\hat{y}) = -\,C\kappa_2(1-1/n)$$

$$MSE(\hat{y}) = C(1-\tfrac{1}{n})\{C(1-\tfrac{1}{n})\kappa_2^2 + (\tfrac{1}{n}-\tfrac{1}{N})[C(1-\tfrac{1}{n}) - 2]\kappa_4 + [2(C-2)(1-\tfrac{1}{n})(\tfrac{1}{n-1}-\tfrac{1}{N-1}) + \tfrac{4}{N-1}(\tfrac{1}{n}-\tfrac{1}{N})]\kappa_{22} - 4(\tfrac{1}{n} - \tfrac{1}{N})\kappa_{31}\} + V(\bar{y}) \quad .$$

Note that the bracketed term must be negative for $MSE(\hat{y})$ to be less than $V(\bar{y})$. The $MSE(\hat{y})$ is minimized by choosing $C = C_o$ where

$$C_o = \frac{(\frac{1}{n-1})(\frac{N-n}{N})\,[\kappa_4 + 2\kappa_{22} + 2\kappa_{31}]}{\kappa_2^2 + (\frac{1}{n}-\frac{1}{N})\kappa_4 + 2(\frac{1}{n-1}-\frac{1}{N-1})\kappa_{22}} \;; \quad n > 1 \quad .$$

For N moderately large, there is little loss in looking at limits as $N\to\infty$. The corresponding quantities then become

$$\lim_{N\to\infty} MSE\ (\hat{y}) = \frac{C}{n}(1 - \tfrac{1}{n})\{[C(1 - \tfrac{1}{n}) - 2]\kappa_4 + [C(n+1) - 4]\kappa_2^2 - 4\kappa_3\kappa_1\} + V(\bar{y})$$

$$\lim_{N\to\infty} V(\bar{y}) = \frac{1}{n}\{\kappa_4 + 4\kappa_3\kappa_1 + 4\kappa_2\kappa_1^2 + 2\kappa_2^2\}$$

$$\lim_{N\to\infty} C_o = \frac{(\frac{1}{n-1})[\kappa_4 + 2\kappa_2^2 + 2\kappa_3\kappa_1]}{\frac{1}{n}\kappa_4 + \frac{n+1}{n-1}\kappa_2^2} \quad .$$

It can be shown that for any distribution for which the third central moment is positive the square root estimator will be more efficient than $\bar{y}$ if and only if $0 < C < 2C_o$.

Example distributions for which the square root estimator is more efficient.

It should be apparent that both C_o and $R = MSE(\hat{y})/V(\bar{y})$ are not affected by changes in the scale of the observations. Therefore, in the discussion to follow the scale parameters of the distributions were set equal to one for simplicity.

To illustrate the usefulness of the square root estimator three families of distributions were examined, ranging from a fairly symmetric gamma distribution to a highly skewed Pareto distribution. The specific distributions are:

Family	Form	Curve Number
	$\frac{1}{6} y^3 e^{-y}$	1
Gamma	$\frac{1}{2} y^2 e^{-y}$	2
	$y e^{-y}$	3
	e^{-y}	4
	$\frac{1}{12} y e^{-\sqrt{y}}$	5
Wishart	$\frac{1}{4} \sqrt{y} e^{-\sqrt{y}}$	6
	$\frac{1}{2} e^{-\sqrt{y}}$	7
	$\frac{1}{2} y^{-\frac{1}{2}} e^{-\sqrt{y}}$	8
	y^{-2}	9
	$2y^{-3}$	10
Pareto	$3y^{-4}$	11
	$4y^{-5}$	12
	$5y^{-6}$	13

For sample sizes n = 2(1)20. the maximum value of $C(=2C_o)$ and the value of R at $C = C_o$ are given in Table 1.

Two points are obvious from the table. First, the efficiency of the square root estimator decreases with increasing sample size. Hence, the square root estimator would be appropriate only for small samples. Secondly, for a given value of n, a particular value of C will give near optimum results for a wide class of distributions. Consequently, it is not as important to know the exact distribution of $y_i(u_i)$ but knowledge of the general shape of the frequency distribution should be sufficient.

General square root estimators.

If we allow C_1 and C_2 to be arbitrary constants we find

$$B(\hat{y}) = - [(1 - C_1 - \frac{C_2}{n})\kappa_2 + (1 - C_1 - C_2)\kappa_{11}] \quad ,$$

$$V(\hat{y}) = C_1^2(\frac{n-1}{n})^2 V(k_2) + (C_1 + C_2)^2 V(k_1^2)$$

$$+ 2C_1(C_1 + C_2)(\frac{n-1}{n}) \mathrm{Cov}(k_2, k_1^2) \quad ,$$

where

$$V(k_2) = [\frac{1}{n} - \frac{1}{N}]\kappa_4 + 2[\frac{1}{n-1} - \frac{1}{N-1}]\kappa_{22}$$

$$V(k_1^2) = [\frac{1}{n} - \frac{1}{N}][\frac{1}{n^2}\kappa_4 + \frac{4}{n}\kappa_{31} + 4\kappa_{211}]$$

$$+ 2\{(\frac{n-1}{n})[\frac{1}{n(n-1)} - \frac{1}{N(N-1)}]$$

$$- \frac{1}{n(N-1)}[\frac{1}{n} - \frac{1}{N}]\}\kappa_{22}$$

$$\mathrm{Cov}(k_1^2, k_2) = [\frac{1}{n} - \frac{1}{N}]\ [\frac{1}{n}\kappa_4 - \frac{2}{N-1}\kappa_{22} + 2\kappa_3\kappa_1] \quad ,$$

so that

$$\mathrm{MSE}(\hat{y}) = V(\hat{y}) + [B(\hat{y})]^2 \quad ,$$

and the optimal values of the coefficients are

$$C_1 = \frac{DE_1 - BE_2}{AD - B^2}$$

$$C_2 = \frac{AE_2 - BE_1}{AD - B^2}$$

where

$$E_1 = (\kappa_2 + \kappa_{11})^2$$

$$E_2 = (\kappa_2 + \kappa_{11})\ (\frac{\kappa_2}{n} + \kappa_{11})$$

$$A = (\frac{n-1}{n})^2 V(k_2) + V(k_1^2) + 2(\frac{n-1}{n})\mathrm{Cov}(k_2, k_1^2) + E_1$$

$$B = V(k_1^2) + (\frac{n-1}{n})\ \mathrm{Cov}(k_2, k_1^2) + E_2$$

$$D = V(k_1^2) + (\frac{\kappa_2}{n} + \kappa_{11})^2 \quad .$$

The values of the optimal coefficients for the distributions considered earlier are given in Table 2. In this case, the values are dependent on the form of the distribution. The efficiency ratio of the square root estimator for optimal values of C_1 and C_2 will always be better (smaller) than the efficiency ratio for optimal C under the restricted situation. However, for general C_1 and C_2, the knowledge of the form of the distribution is much more critical for near optimal results.

Conclusions.

A marked gain in efficiency over the sample mean may be obtained, for small sample sizes, by using the square root estimator $\hat{y} = (1-C)\bar{y} + C\bar{u}^2$. While the value of C depends on the sample size and distribution type, for a given n a particular value of C can be determined which will give near optimal results over a wide class of distribution types.

REFERENCES

[1] Hartley, H. O. and Rao, J. N. K., "A New Estimation Theory for Sample Surveys," Biometrika, 55 (1968), 547-557.

[2] Wishart, J., "Moment Coefficients of the k-Statistics in Samples from a Finite Population," Biometrika, 39 (1952), 1-13.

TABLE 1. VALUES OF C FOR WHICH $MSE(\hat{y}) < V(\bar{y})$ FOR THIRTEEN DISTRIBUTIONS; n = 2, 20

Distribution	n= 2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
(1) $\frac{1}{6} y^3 e^{-y}$	**.96	.96	.96	.97	.97	.97	.98	.98	.98	.98	.98	.98	.98	.99	.99	.99	.99	.99	.99
	*2.74	2.06	1.66	1.38	1.18	1.04	.92	.82	.76	.70	.64	.60	.56	.52	.49	.46	.43	.41	.39
(2) $\frac{1}{2} y^2 e^{-y}$	.94	.94	.95	.96	.96	.96	.97	.97	.97	.97	.98	.98	.98	.98	.98	.98	.98	.98	.98
	2.77	2.08	1.67	1.39	1.19	1.04	.93	.83	.76	.70	.64	.60	.56	.52	.50	.46	.44	.42	.40
(3) $y\ e^{-y}$	.92	.92	.93	.93	.94	.95	.95	.96	.96	.96	.96	.97	.97	.97	.97	.97	.98	.99	.99
	2.83	2.13	1.70	1.42	1.22	1.07	.95	.85	.78	.71	.65	.61	.57	.53	.50	.47	.45	.43	.41
(4) e^{-y}	.84	.84	.86	.87	.88	.90	.91	.91	.92	.93	.93	.94	.94	.94	.95	.95	.95	.95	.96
	2.98	2.23	1.79	1.50	1.29	1.13	1.01	.91	.83	.76	.70	.65	.61	.57	.54	.51	.48	.46	.43
(5) $\frac{1}{12} ye^{-\sqrt{y}}$	.73	.73	.75	.77	.79	.81	.83	.84	.85	.86	.87	.88	.88	.89	.90	.90	.91	.91	.91
	4.00	3.00	2.45	2.08	1.81	1.61	1.45	1.32	1.21	1.12	1.04	.97	.91	.86	.81	.77	.73	.70	.67
(6) $\frac{1}{4} \sqrt{y}e^{-\sqrt{y}}$	.67	.67	.69	.72	.74	.76	.78	.80	.81	.82	.84	.84	.85	.86	.87	.87	.88	.89	.89
	4.00	3.00	2.46	2.10	1.85	1.65	1.49	1.35	1.25	1.16	1.08	1.01	.95	.90	.86	.82	.77	.74	.71
(7) $\frac{1}{2} e^{-\sqrt{y}}$	.57	.57	.60	.63	.66	.68	.71	.73	.75	.76	.78	.79	.80	.81	.82	.83	.83	.84	.85
	4.00	3.00	2.48	2.41	1.89	1.70	1.55	1.42	1.31	1.22	1.14	1.07	1.01	.95	.91	.86	.82	.79	.75
(8) $\frac{1}{2} y^{-\frac{1}{2}} e^{-\sqrt{y}}$	.40	.40	.43	.47	.50	.53	.56	.58	.60	.62	.64	.66	.68	.69	.70	.71	.72	.73	.74
	4.00	3.00	2.53	2.22	2.00	1.82	1.68	1.57	1.46	1.37	1.29	1.23	1.17	1.11	1.06	1.01	.97	.93	.89
(9) y^{-2} $(y \geq 1)$	.28	.28	.33	.36	.39	.42	.44	.46	.48	.50	.52	.54	.54	.55	.57	.58	.59	.60	.62
	5.26	3.85	3.38	3.03	2.79	2.59	2.43	2.29	2.17	2.07	1.97	1.89	1.81	1.74	1.67	1.62	1.56	1.51	1.46
(10) $2y^{-3}(y \geq 1)$	.22	.22	.23	.24	.26	.27	.29	.30	.32	.33	.34	.36	.37	.38	.39	.40	.41	.42	.43
	6.08	4.56	3.99	3.67	3.45	3.29	3.15	3.04	2.95	2.85	2.77	2.71	2.64	2.57	2.52	2.47	2.41	2.36	2.31
(11) $3y^{-4}(y \geq 1)$	.35	.35	.36	.37	.39	.40	.41	.42	.44	.45	.46	.47	.48	.49	.50	.51	.52	.52	.53
	7.20	5.40	4.72	4.34	4.08	3.89	3.73	3.59	3.48	3.37	3.28	3.19	3.12	3.04	2.97	2.91	2.85	2.79	2.73
(12) $4y^{-5}(y \geq 1)$	.45	.45	.47	.48	.50	.52	.53	.55	.56	.57	.59	.60	.61	.62	.63	.64	.65	.66	.66
	10.09	7.57	6.56	5.96	5.54	5.21	4.94	4.76	4.51	4.33	4.17	4.03	3.89	3.77	3.65	3.55	3.45	3.35	3.27
(13) $5y^{-6}(y \geq 1)$	.51	.51	.52	.54	.56	.57	.59	.61	.62	.63	.65	.66	.67	.68	.69	.70	.71	.71	.72
	12.89	9.66	8.32	7.51	6.93	6.48	3.11	5.79	5.51	5.27	5.05	4.84	4.66	4.49	4.34	4.20	4.07	3.94	3.82

*Maximum values of C for which $MSE(\hat{y}) < V(\bar{y})$.
Optimum value is ½ the listed value.

**Value of $R = \frac{MSE(\hat{y})}{V(\bar{y})}$ at C_0.
Minimum value of C for which $EMS(y) < V(\bar{y})$ is zero in each case.

TABLE 2. VALUES OF C_1 AND C_2 FOR OPTIMUM EFFICIENCY FOR THIRTEEN DISTRIBUTIONS; n = 2, 20

Distribution	n = 2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
(1) $\frac{1}{6} y^3 e^{-y}$	.89	.92	.94	.95	.96	.96	.97	.97	.97	.97	.98	.98	.98	.98	.98	.98	.98	.98	.98	C_1
	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	C_2
(2) $\frac{1}{2} y^2 e^{-y}$	.86	.90	.92	.94	.95	.95	.96	.96	.97	.97	.97	.98	.98	.98	.98	.98	.98	.98	.98	C_1
	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	C_2
(3) $y\ e^{-y}$	.80	.86	.89	.91	.92	.93	.94	.95	.95	.96	.96	.96	.96	.97	.97	.97	.98	.98	.98	C_1
	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	C_2
(4) e^{-4}	.67	.75	.80	.83	.86	.88	.89	.90	.91	.92	.92	.93	.93	.94	.94	.94	.95	.95	.95	C_1
	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	C_2
(5) $\frac{1}{12} y e^{-\sqrt{y}}$	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	C_1
	.73	.85	.94	.99	1.03	1.05	1.08	1.09	1.11	1.12	1.13	1.14	1.15	1.15	1.16	1.16	1.17	1.17	1.18	C_2
(6) $\frac{1}{4} \sqrt{y} e^{-\sqrt{y}}$	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	C_1
	.67	.82	.91	.98	1.03	1.06	1.09	1.12	1.14	1.15	1.17	1.18	1.19	1.20	1.20	1.21	1.22	1.22	1.23	C_2
(7) $\frac{1}{2} e^{-\sqrt{y}}$	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	C_1
	.57	.75	.87	.96	1.03	1.08	1.12	1.16	1.19	1.21	1.23	1.25	1.26	1.28	1.29	1.30	1.31	1.32	1.33	C_2
(8) $\frac{1}{2} y^{-\frac{1}{2}} e^{-\sqrt{y}}$	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	C_1
	.40	.60	.80	.90	1.00	1.09	1.16	1.23	1.28	1.33	1.37	1.41	1.44	1.47	1.50	1.52	1.54	1.56	1.58	C_2
(9) y^{-2} $(y \geq 1)$	-1.17	-1.17	-1.18	-1.16	-1.15	-1.14	-1.13	-1.12	-1.11	-1.10	-1.10	-1.09	-1.09	-1.09	-1.08	-1.08	-1.07	-1.07	-1.07	C_1
	1.97	2.29	2.49	2.54	2.60	2.64	2.67	2.69	2.71	2.73	2.74	2.75	2.76	2.76	2.77	2.78	2.78	2.79	2.80	C_2
(10) $2y^{-3}$ $(y \geq 1)$	-2.09	-1.48	-1.25	-1.13	-1.06	-1.01	- .98	- .95	- .93	- .91	- .90	- .88	- .87	- .87	- .86	- .85	- .85	- .84	- .84	C_1
	3.10	2.56	2.37	2.27	2.21	2.17	2.14	2.11	2.10	2.08	2.07	2.06	2.05	2.05	2.04	2.04	2.03	2.03	2.02	C_2
(11) $3y^{-4}$ $(y \geq 1)$	-2.68	-1.88	-1.59	-1.45	-1.36	-1.30	-1.26	-1.23	-1.20	-1.18	-1.17	-1.15	-1.14	-1.13	-1.12	-1.11	-1.11	-1.10	-1.10	C_1
	3.69	2.92	2.64	2.51	2.47	2.37	2.33	2.30	2.27	2.25	2.24	2.23	2.22	2.21	2.20	2.19	2.18	2.18	2.17	C_2
(12) $4y^{-4}$ $(y \geq 1)$	-4.37	-3.25	-2.85	-2.64	-2.52	-2.43	-2.37	-2.33	-2.29	-2.20	-2.24	-2.22	-2.20	-2.19	-2.18	-2.16	-2.15	-2.15	-2.14	C_1
	5.39	4.29	3.90	3.69	3.59	3.49	3.43	3.38	3.35	3.32	3.30	3.28	3.26	3.25	3.23	3.22	3.21	3.21	3.20	C_2
(13) $5y^{-6}$ $(y \geq 1)$	-6.01	-4.61	-4.09	-3.83	-3.67	-3.56	-3.48	-3.42	-3.38	-3.34	-3.31	-3.28	-3.26	-3.24	-3.23	-3.21	-3.20	-3.19	-3.18	C_1
	7.04	5.64	5.13	4.87	4.71	4.60	4.53	4.47	4.42	4.39	4.36	4.33	4.31	4.29	4.27	4.26	4.25	4.24	4.23	C_2

PROPOSAL FOR CONTINUOUS, ALPHABETICALLY STAGGERED CENSUS

David Rothman

Technical Abstract

Each inhabitant is given, based on initial letters of surname, one of 3650 reporting dates for every decade. (If surname is changed, and new report scheduled $x \leq 10+\epsilon$ years after last report, person is counted as $x/10$ people. If $x > 10+\epsilon$, next report is made 10 years later instead.) Ethnic stratification, rather than strict alphabetical ordering, is utilized to minimize effect of immigration and internal growth differentials on population estimates. Temporary exactness (twenty months late), the quaint pseudo-virtue of one-day head-counts, is sacrificed in favor of long-term precision and speedier results at less cost, since special hires are eliminated and peak postal load is reduced. Legislative reapportionment will be more timely, rather than up to twelve years late from present technique. New follow-up methods must evolve to supplement address register, but reporting percentage should increase from improved public consciousness of daily census activities.

History

Phase I: Generation

Late 1967 or early 1968: Author conceives apparently original idea of staggering census according to letters of last name. Idea utilized in a draft of a proposed new constitution which achieves no more than a limited private circulation. (At this stage the problems of name changes and ethnic differentials are unappreciated. To see that the former is a problem, imagine a Perpetually Remarrying Woman. This PRW remarries every 3651 days in such a way as to report only once in her lifetime, unless special rules are devised to ensure reporting at least once every $10+\epsilon$ years. As for ethnicity, if the names were simply taken in alphabetic order, then there would be long stretches from a single ethnic group, such as "Mc" names, leading to bias in population estimates. Thus groups of this nature have to be spread throughout each decade.) Author can't believe that idea which seems so simple and so obviously right could possibly be original, so no particular effort is made to push what is thought of as a minor procedural revision.

1969: Major New York publishers reject manuscript. Author retreats imperceptibly from megalomania.

Phase II: Appeal to Ad Hoc Power

14 November 1970: Having read of the appointment of the President's Commission on Federal Statistics, the author finds that one of the members is located in New York City. Author writes letter to Solomon Fabricant to broach subject of alphabetically staggered census.

16 November 1970: Dr. Fabricant responds to author that letter is being forwarded to Daniel B. Rathbun, Executive Director of Commission.

19 November 1970: Dr. Rathbun responds to author, saying that copy of letter is being sent to Joseph F. Daly, Associate Director for Research and Development, Bureau of the Census.

24 November 1970: Dr. Daly responds to author, rejecting procedure.

Date unknown: Author writes to Dr. Daly and/or Dr. Rathbun to answer objections. (No copy kept by author.)

13 April 1971: Author writes reminder to Dr. Rathbun.

16 April 1971: Dr. Rathbun responds to author, restating objections raised by Dr. Daly.

22 April 1971: Author sends Dr. Rathbun a two-page preliminary feasibility study. (Ethnicity now appreciated.)

4 May 1971: Dr. Rathbun responds to author again rejecting proposal, quoting new rejection from Dr. Daly. Only now does author begin to comprehend that crux of negative opinion is lack of foolproof follow-up analogous to address register for present one-day head-count method.

8 May 1971: Author responds to Dr. Rathbun (copy to Dr. Daly), giving alternate follow-up procedures for proposed method, such as telephone directories, etc.

21 May 1971: Dr. Rathbun writes author that proposal will be "made available to all Commission members...along with other methods of improving the census".

18 June 1971: Dr. Rathbun writes author to suggest "a detailed proposal aimed at, say, Congress".

1 July 1971: Author writes Dr. Rathbun to inquire about Commission support for such a proposal.

12 July 1971: Dr. Rathbun writes author that "all available resources have been committed", but no reply yet on "Census resources".

Phase III: Appeal to Permanent Power in Government

1 September 1971: Author writes to the Chairmen of Senate Commerce and House Government Operations Committees to inquire as to how to approach Congress.

3 September 1971: M. Q. Romney, Associate General Counsel, House Government

Operations Committee, responds to author, giving correct committees in House and Senate with "legislative jurisdiction over the Census".

24 September 1971: Author writes to the Honorable Thaddeus J. Dulski, Chairman of House Committee on Post Office and Civil Service, inquiring how to make proposal. Also to Honorable Gale W. McGee, Chairman of Senate Committee on Post Office and Civil Service.

29 September 1971: Sen. McGee responds to author, and "would be happy to consider an outline" of proposal.

30 September 1971: Rep. Dulski responds to author, enclosing Congressional Record for 14 June 1971 ("Administration Withdraws Its Support for a Mid-decade Census") and for 22 September 1971 ("Mid-decade Census Proposal Is Still Under Active Consideration"). He explains how his "staff has done some research on" the proposal, is "informed" about previous replies from the government, and is somewhat negative. He mentions "hearings on a mid-decade census program earlier this year" held by the Honorable Charles H. Wilson, Chairman, Subcommittee on Census and Statistics, and mentions subcommittee commitment to "this legislation a matter of top priority". He does "not expect to schedule...any further hearings during" Ninety-second Congress which "would encompass" author's proposal.

6 October 1971: Author writes to Rep. Dulski, answering objections, opposing mid-decade census as "a halfway measure to meet the goal of updated population estimates", and offering to appear before Committee.

6 October 1971: Author writes to Sen. McGee outlining proposal.

12 October 1971: Rep. Dulski responds to author, indicating that letter of 6 Oct. has been referred to Rep. Wilson.

14 October 1971: Sen. McGee acknowledges receipt of outline from author, saying it will be studied.

14 October 1971: Sen. McGee sends above correspondence to Dr. George Hay Brown, Director, Bureau of the Census.

21 October 1971: Dr. Brown writes Sen. McGee explaining why Bureau "cannot endorse" proposal. Alternative to address register now is clearly the central problem.

26 October 1971: Sen. McGee, enclosing letter from Dr. Brown, writes author to reject proposal.

1 November 1971: Author writes slightly snide letter to Sen. McGee (copy to Dr. Brown) answering objections.

9 November 1971: Author writes to Prof. Charles B. Nam, asking to present paper on proposal to Montreal meeting of ASA in August 1972.

9 December 1971: Author writes reminder to Sen. McGee.

16 March 1972: Author reads that Dr. Brown will be in New York City on 4 April and 13 April, and writes him asking for a meeting.

23 March 1972: Author writes strong letter to Rep. Wilson asking that certain research monies be withheld from the Bureau of the Census "until Dr. Brown digs a bit deeper into the idea of a continuous census".

27 March 1972: Dr. Brown responds to author, arranging telephone conversation on 4 April.

4 April 1972: Dr. Brown speaks with the author for about 45 minutes. Author agrees to inquire about constitutionality of proposal, and Dr. Brown agrees to send results of informal preliminary studies indicating that, without an address register follow-up, as much as 7% of population may be missed.

4 April 1972: Author writes to Attorney General asking for opinion on constitutionality (copy to Dr. Brown).

Phase IV: Appeal to Permanent Power in Profession

21 April 1972: Author writes Dr. Nancy R. Mann, a member of ASA Census Advisory Committee (and former associate), to ask for opinion.

15 May 1972: Dr. Mann responds to author offering to "inquire further at the next advisory comm. meeting in September", and mentioning conversation with Dr. Brown in which he indicated "a problem about redefining census tracts".

20 May 1972: Author writes Prof. Leslie Kish, Chairman of ASA Census Advisory Committee, asking if committee members would consider Montreal paper so that an "informed discussion" could take place at committee meeting in September.

22 May 1972: Prof. Kish's secretary replies to author, stating that he will be in Yugoslavia until 11 June.

5 July 1972: An Assistant Attorney General responds to author, stating that legal opinions can only be furnished to "the President and heads of the executive departments".

11 July 1972: Author writes Dr. Brown

about above response and asks him (or a representative) to attend presentation in Montreal.

16 August 1972: Author presents paper at Montreal meeting, conjectures that continual reapportionment entailed in this system would tend to take redistricting procedure out of political arena and into professional hands. Session is poorly attended (Tukey is major competition). In informal discussion after session, birthdays and Social Security Numbers are rejected as basis for continuous census, since only last name is held in common by entire family units.

5 September 1972: Prof. Kish mails form to author indicating he will be in Europe until about 30 June 1973.

Status of Project (22 September 1972)

None of the four individuals (Sen. McGee, Rep. Wilson, Dr. Brown, or Prof. Kish) with political or professional power to push such a proposal is interested in doing so.

The author has seen no report to substantiate the claim of a significant percentage of omissions due to follow-up difficulty.

DEMOGRAPHIC AND SOCIO-ECONOMIC CHARACTERISTICS OF MEDICAL LABORATORY MANPOWER

Adolph Scolnick, U.S. Bureau of Health Manpower Education

Introduction

Information on the demographic characteristics of medical laboratory manpower has never been gathered in detail and on a regular basis even though such an evaluation of the current work force has long been needed by the medical profession, government agencies and other organizations in the health field.

Prior to this survey, made in late 1969 and early 1970, information on registered laboratory workers was collected by the National Committee for Careers in Medical Technology (now The National Committee for Careers in the Medical Laboratory) through a salary survey in 1959 and 1963, and a more detailed study of salaries and employment status in 1966. An earlier study of salaries was made in 1954 by Dr. J. L. Arbogast. Analyses of the 1959, 1963 and 1966 surveys were published in the National Committee for Careers in the Medical Laboratory (NCCML) publication GIST, issues 7, 22 and 36.

Late in 1969 an opportunity came to NCCML to include a questionnaire in an annual mailing concerning re-registration sent by the Board of Registry of the American Society of Clinical Pathologists to all currently certified laboratory personnel. Since that was the last year in which annual registration was compulsory, NCCML deemed it an opportunity to collect a variety of basic data from this group of individuals.

After extensive consultation and revision, a questionnaire was devised covering type of registration, age, highest education attained, recent participation in continuing education courses and workshops, place of employment, level of principal employment, duties of principal employment, salary and work history. From the medical technologist no longer working in the laboratory, data was sought on present status and reason for it, work history, intentions about returning to work and need for refresher training. Certain questions covered persons in the armed forces and those whose work was in the field of environmental health.

The original questionnaires were sent to all American Society of Clinical Pathologists (ASCP)-registered medical laboratory personnel by the ASCP in Chicago in December 1969 and were returned directly to NCCML in the early months of 1970. Those listed in "Register 1969, Certified Medical Laboratory Personnel," published by ASCP, total about 60,100. For this analysis, some 29,000 questionnaires, based on the response to the item, State of residence, were considered useable, reflecting a 48.3 percent response. (Table 1).

In addition, a similar questionnaire altered only to reflect titles of laboratory personnel in California, was mailed by the Department of Public Health of the State of California early in 1970 with returns coming directly to NCCML. The total number of licensed laboratory personnel in California at that time was 14,500 with about 4,400 questionnaires, based on the response to the item, State of residence, used for this analysis. This response was 30.5 percent. Responses were not used if the respondent stated they had already answered the earlier questionnaire.

Thus the useable replies totaled about 33,500 for a total response rate of 44.8 percent. Intersociety controversy in the medical laboratory field at that time undoubtedly kept a number of individuals from returning their questionnaires.

The Universe Versus the Respondents

The universe of registered and/or licensed medical laboratory manpower and the number of respondents to the questionnaire items, State of residence and/or place of principal employment, are compared in Table 2. The latter group, those who were currently working full-time, part-time, or irregularly in the medical laboratory field, totaled about 24,000, or 71.7 percent, of the former group responding to the item State of residence.

Sex and Age

The descriptive term "manpower" in the phrase "medical laboratory manpower" is a misnomer since females constituted 87 percent of those responding to the questionnaire item. Almost half of all the respondents to the item on age were in the age category 25-34 years. The high concentration of medical laboratory manpower in the age category 25-34 years contrasts sharply with projections by the Bureau of Labor Statistics of the age distribution of the female labor force for 1970. The Bureau of Labor Statistics projects fewer female workers in the ages 25-34 years than in the older years 35-44 years of age, or even of those 45-54 years. The comparison may be invalid when 29 million female persons in the total labor force is compared with some 30 thousand registered/licensed medical laboratory manpower. Further research of the characteristics of the nonrespondents on the registries is surely needed to affirm conclusively the sex-age attachment to the labor force of medical laboratory manpower.

Possibly due to educational or other restrictions for ASCP registration or California licensure, hardly any respondent was under 20 years of age, while 240 of the older respondents were age 65 years and over, and some still employed at the age of 80 (Table 3).

Age and Activity of Respondents

Among the approximately 34,200 respondents, both active and inactive, about 9,900, almost 29 percent, were inactive -- i.e., almost one out of three respondents were not currently working either full-time, part-time, or irregularly in the medical laboratory field (Table 4). Considering that irregular employment may be considered as "frictional" or "marginal" employment, the rate of inactivity reported is probably a minimum. Since the Registry/Licensure population described earlier totals about 74,600, the almost 9,900 inactive respondents, or 13 percent, is relatively significant. If however, the inactive respondents are for the most part not "actively seeking work" in the medical laboratory field, then they possibly should not be considered in the "labor force" of medical laboratory manpower.

It is then possible that the rate of inactivity, in the sense of measuring "unemployment", may be overstated.

Active employment of the respondents who stated their age was 71 percent overall. Three age categories were below the overall average: age 25-34 years (68 percent); age 35-44 years (66 percent); and age 65 and over (52 percent). One may speculate that the smaller percentages of the first two ages are due to family responsibilities normally associated with female workers in those age groups. The younger age category, 20-24 years, had the highest activity rate (86 percent), which is what one might expect from our knowledge of labor force participation rates.

Age, Activity and Type of Registration

The inactive medical laboratory manpower were concentrated most heavily among medical technologists (ASCP registered and California licensed) and least heavily among the Cytotechnologists (ASCP registered) (Table 5). Based only on the number of respondents, the medical technologists rate of inactivity was 43 percent and the rate for cytotechnologists was 17 percent. The rates among the total population could very well be different from those given here, since the respondents constituted less than 45 percent of the entire population of registered or licensed medical laboratory manpower. Inactivity was highest in the age groups, 25-34 years and 35-44 years, and probably due to the high female to male ratio of medical laboratory manpower. No doubt, family commitments serve to keep many females from working full-time, part-time, or even irregularly in the medical laboratory field.

Highest Education Attained

The bulk of the medical laboratory manpower-- some 70% of the respondents-- hold a baccalaureate degree, which is not surprising since the minimum educational requirement for medical technologist (ASCP) is 4 years of college. Medical technologists held 96% of all the baccalaureate degrees among those who responded to the question on highest education attained (Table 6).

Continuing Education

Medical laboratory personnel are participating in educational opportunities in continuing education. Attendance of respondents at continuing education workshops or courses, of at least one day's duration in the year prior to the survey was reported by almost 40% of the respondents (Table 7). Attendance by a participant at more than one course during the year is, of course, not reflected here. Attendance did vary by type of registration-- of the 20,300 medical technologists (ASCP registered and California licensed) who were actively employed, almost 11,000, or 54 percent, attended continuing education workshops or courses in the medical laboratory field. Attendance by cytotechnologists was equally high (54 percent). Attendance by clinical laboratory assistants and histologic technicians dropped off to 45 percent and 35 percent, respectively. These percentages should not be construed to apply to the universe; they apply only to the sub-set of respondents.

The primary duties of the actively employed medical laboratory worker is also associated with participation in continuing education. General laboratory work was reported as the primary duties of 5,000 persons--presumably the normal, but certainly not the only, duties of the medical technologist. Some 4,300 medical laboratory workers described their primary duties as specialists in the laboratory. In contrast to those engaged in general laboratory work, laboratory specialists are usually limited to work in one area of the laboratory. One such specialist is the cytotechnologist engaged in screening slides in the search for abnormalities that are warning signs of cancer. Another is the histologic technician who specializes in cutting and staining body tissue for microscopic examination.

Space is not available to report on other major findings: e.g. current employment in hospitals; medical laboratory manpower in environmental health; salaries of full-time workers; reasons given for stopping work or current unemployment status; length of inactivity in and expectations for returning to the medical laboratory field.

Conclusions and Recommendations

The present survey, which is one of a series originating with the National Committee for Careers in the Medical Laboratory's need for information on salaries of laboratory workers, has been used primarily as a recruitment tool to entice persons into the medical laboratory field. In some respects, the current survey is deficient because response was low, less than 45 percent, and this raises questions about the characteristics of the non-respondents. Procedures to reduce non-response, if used at all, was ineffective. Questionnaires were accepted for data processing, even in those cases where only a minimal number of items had been completed, and without regard to the importance of these items. Since no classifying or screening variables were identified, no legible returned questionnaire was rejected. Turning to the contractor's final report, wherein the analysis for the most part, was done item by item, there is no respondent analysis reported on data sets. The data on tape presented some problems since the Bureau of Health Manpower Education used the tape in two ways: (1) to compare the data on tape with the contractor's final report; and, (2) to produce additional tabulations for the Bureau's own use. Without going into the specifics, we often were unable to reconcile the data items on tape with the data in the final report. Continuing effort is being made to reconcile these differences.

The introduction of probability sampling holds the most promise for improving the data collected from the medical laboratory Registry/Licensure population. With the use of probability sampling, more of the financial resources can be allocated to increase follow-up procedures, and hopefully to reduce non-response to a manageable proportion consistent with dollar costs and sampling variances of the more important characteristics. The complexity of this kind of survey, involving difficult computer programming for handling cross-tabulations, suggests that professional assistance of survey statisticians be initiated early in the planning stage.

Table 1 Comparison of ASCP Registered Population and Respondents, Medical Laboratory Manpower, By States: 1969-70

State of Residence	REGISTRY POPULATION Number Registered	RESPONDENTS Number Responding	RESPONDENTS Percent of Registry Responding
TOTAL	60,113	29,040	48.3
Alabama	891	384	43.1
Alaska	93	50	53.8
Arizona	548	283	51.6
Arkansas	453	214	47.2
California	5,650	2,835	50.2
Colorado	1,161	599	51.6
Connecticut	917	441	48.1
Delaware	195	101	51.8
District of Columbia	216	86	39.8
Florida	1,874	922	49.2
Georgia	1,201	578	48.1
Hawaii	319	140	43.9
Idaho	223	131	58.7
Illinois	3,290	1,572	47.8
Indiana	1,444	728	50.4
Iowa	917	480	52.3
Kansas	1,062	516	48.6
Kentucky	1,108	503	45.4
Louisiana	1,286	572	44.5
Maine	172	96	55.8
Maryland	1,118	541	48.4
Massachusetts	1,370	555	40.5
Michigan	2,887	1,384	47.9
Minnesota	1,937	849	43.8
Mississippi	503	210	41.7
Missouri	1,328	639	48.1
Montana	308	171	55.5
Nebraska	657	321	48.9
Nevada	124	71	57.3
New Hampshire	223	127	57.0
New Jersey	1,480	685	46.3
New Mexico	301	156	51.8
New York	2,813	1,421	50.5
North Carolina	1,200	563	46.9
North Dakota	254	145	57.1
Ohio	3,128	1,539	49.2
Oklahoma	874	422	48.3
Oregon	827	398	48.1
Pennsylvania	3,222	1,541	47.8
Rhode Island	281	141	50.2
South Carolina	520	256	49.2
South Dakota	243	124	51.0
Tennessee	1,235	544	44.0
Texas	3,657	1,747	47.8
Utah	330	188	57.0
Vermont	185	90	48.6
Virginia	1,400	718	51.3
Washington	1,424	728	51.1
West Virginia	502	254	50.6
Wisconsin	2,283	1,091	47.8
Wyoming	113	76	67.3
Guam	5	5	100.0
Puerto Rico	355	106	29.9
Virgin Islands	6	3	50.0

Table 2 Comparison of Registry/Licensure Population and Number of Tape Records, by State of Residence and Place of Principal Employment: Medical Laboratory 1969-70

	ASCP Registered	California Licensed
Registry/Licensure Population	60,113	14,500
Number of Tape Records	30,263	4,522
Number of Codes Records	30,234	4,518
State of Residence:		
Number of Coded Records with State Codes	29,436	4,482
United States, including Guam, Puerto Rico, and the Virgin Islands, but excluding Foreign, illegal codes, and non-reporters	29,040	4,428
Place of Principal Employment:		
Number of Coded Records with State Codes	21,149	3,256
United States, including Guam Puerto Rico, and the Virgin Islands, but excluding Foreign, illegal codes, and non-reporters	20,796	3,214

Table 3 Sex and Age of Respondents: Medical Laboratory Manpower, 1969-70

	ASCP Registered (number)	California Licensed (number)
SEX:		
Female	26,715	3,386
Male	3,273	1,112
Total	29,988	4,498
AGE:		
Total	26,667	4,293
Under 20 years of age	6	2
20-24 years	4,150	198
25-34	13,579	1,683
35-44	5,470	1,183
45-54	2,455	816
55-64	867	311
65 and over	140	100

Table 4 Age Distribution of Active and Inactive Respondents: Medical Laboratory Manpower (ASCP Registered and California Licensed), 1969-70 1/

AGE:		Active	Inactive
Total	34,239	24,386	9,853
Age not reported	3,735	2,745	990
Age reported	30,504	21,641	8,863
Under 20 years of age	107	90	17
20-24 years	4,227	3,618	609
25-34	15,097	10,250	4,847
35-44	6,547	4,325	2,222
45-54	3,166	2,419	747
55-64	1,136	823	313
65 and over	224	116	108

1/ Total of Medical Technologists (ASCP Registered and California licensed), Histologic Technicians (ASCP), Cytotechnologists (ASCP), Certified Laboratory Assistants (ASCP), Other Not Specified.

Table 5 Age of Active and Inactive Respondents, by Type of Registration: Medical Laboratory Manpower, 1969-70

	Medical Technologist (ASCP Reg. & Calif. Lic.)	Histologic Technician (ASCP)	Cytotech-nologist (ASCP)	Clinical Laboratory Assistant (ASCP)	All Others
Total--All Ages					
Active	20,305	1,320	979	1,499	283
Inactive	8,842	428	173	349	61
Under 20 years of age					
Active	3	11	0	75	1
Inactive	2	4	0	11	0
20-24					
Active	2,547	239	137	682	13
Inactive	340	66	20	183	0
25-34					
Active	8,754	528	458	419	91
Inactive	4,395	204	112	112	24
35-44					
Active	3,693	239	158	148	87
Inactive	2,120	53	17	17	15
45-54					
Active	2,056	162	79	74	48
Inactive	715	17	3	5	7
55-64					
Active	747	5	34	16	21
Inactive	253	51	2	2	5
65 years or more					
Active	105	5	0	1	5
Inactive	94	1	3	1	9
Age unknown					
Active	2,400	131	113	84	17
Inactive	923	32	16	18	1

Table 6 Highest Education Attained by Respondents: Medical Laboratory Manpower, 1969-70

	ASCP Registered	California Licensed
Total	30,101	4,508
High School	640	10
Vocational or technical training-- post high school	1,526	66
Associate degree-- Junior College	1,403	135
Some college-- no degree	4,965	436
Baccalaureate (B.A. or B.S.)	20,505	3,462
Master's Degree	998	250
Ph.D., D.V.M., or Dr. in Education	53	73
M.D., D.D.S., or DMD	11	76

Table 7 Attendance of Respondents at Continuing Education Workshops or Courses in the Medical Laboratory Field, by Type of Registration and by Primary Duties: Medical Laboratory Manpower, 1969-70 1/

Number of Respondents--Total	34,223
Not Attending Workshops/Courses	20,739
Attending Workshops/Courses	13,484
Type of Registration (according to State of Residence):	
Number of Respondents--Total	12,760
Medical Technologists	10,953
Histologic Technicians	460
Cytotechnologists	524
Clinical Laboratory Assistants	674
Others	149
Primary Duties (according to State Where Employed):	
Number of Respondents--Total	11,608
General Laboratory Work	5,013
Laboratory Specialist	4,269
Administration (Laboratory or Education)	1,394
Teaching	389
Research	444
Others	99

1/ Attendance of at least one day's duration in the prior year.

ACKNOWLEDGMENT

This paper was presented at the Annual Meeting of the American Statistical Association, Montreal, Quebec, Canada, August 14- 17, 1972. The research upon which this paper is based was supported in part by Contract No. NIH 71- 4032 with the U.S. Public Health Service, Department of Health, Education and Welfare.

REFERENCES

[1] National Committee for Careers in the Medical Laboratory. Gist. Issues 7, 22 and 36.
[2] op cit Issue 50.
[3] U.S. National Center for Health Statistics. Health Resources Statistics, 1971. DHEW Publication No. (HSM) 72-1509.
[4] Virginia Coordinated Health Survey Committee. Health Manpower 1971.
[5] New York State Department of Health. Hospital Manpower Survey, 1969 (Feb. 1970).
[6] Texas Health Careers Program and Governor's Office of Comprehensive Health Planning. Allied Health Manpower in Texas, 1970.
[7] U.S. Department of Labor, Bureau of Labor Statistics. Projections of the Labor Force, 1970-80; Labor Force Projections, by Color, 1970-80, and Special Labor Force Report, Nos. 49 and 73.
[8] Kish, Leslie. Survey Sampling. John Wiley & Sons, Inc. 1965
[9] Cochran, William G. Sampling Techniques, 2nd edition. John Wiley & Sons, Inc. 1963

FEEDBACK IN PATH ANALYSIS: SOME SOCIOLOGICAL EXAMPLES

B. Krishna Singh and J. Sherwood Williams, Virginia Commonwealth University

I. INTRODUCTION

There has been a considerable amount of work incorporating path analysis in the sociological methodology since Boudon's (1965) and Duncan's (1966) introductions of dependence or path analysis to social science research, which is based on Wright's (1921, 1934) pioneering works. Sociologists and other social scientists have found this type of analysis to be a useful tool in attempting to bridge the gap between sociological theory and statistical analysis. The linear causal models have gained a greater prominence due to path analysis.

With few exceptions path analysis has been used only with unidirectional causal models. Although statistical principles for the consideration of reciprocal interaction (feedback) in path analysis are available (e.g., Turner and Stevens, 1959; Tukey, 1954; Wright, 1960b) they have been largely ignored by sociologists who have used path analysis. Duncan (1966) noted that feedback in path analysis was a neglected aspect of the technique. With only one exception (Duncan, Haller and Portes, 1968; Duncan, 1970) our literature review of sociological research which has used path analysis indicates that feedback remains a neglected topic.

This benign neglect can be attributed to several factors: (1) the system of reciprocal interaction is not a topic in vogue yet although it is talked about in sociological research, (2) from a methodological perspective, reciprocity or feedback cannot be subjected to the simple interpretation as is the case for unidirectional causal models (3) since sociological research tends to deal more with cross-sectional data rather than longitudinal data, the need for grappling with the concept of feedback may not yet have been seriously felt, (4) sociologists may not yet recognize reciprocal interaction or feedback as a useful theoretical and methodological device. Other factors that might help explain why feedback has been neglected could be presented, but that is not the purpose of this paper. Regardless of the reasons, feedback largely has been ignored by sociologists who have used the path analytic technique.

The major concern of this paper is to discuss the issue of feedback in path analysis from a sociological research perspective. It should be noted that such a model will require somewhat different theoretical, statistical, causal, measurement and sampling assumptions (discussed in Section III) than with the unidirectional path analysis. We also will provide theoretical and empirical examples of feedback models in sociological research, and we will examine how feedback can be handled from the path analysis perspective.

II. THE CONCEPT OF FEEDBACK AS A METHODOLOGICAL DEVICE

In its simplest form, feedback or reciprocal interaction may be defined as a continuous "cause and effect" sequence in a system of relationships. To illustrate, variable A affects variable B with some defined function, and variable B also affects variable A with some defined function. A pictorial representation of this type of feedback can be shown as:

FIGURE 1

A B A' B' A" B"

A B

What is important here is that the variables are in some way interacting. There are two basic types of feedback that may be considered: (1) instantaneous feed (feedback with time lag) and (2) delayed feedback (feedback with time lag). These distinctions are made on the basis of whether time is a real parameter in the reciprocal interaction of two or more variables. For example, the reciprocal interaction relationship A B can be thought of as A_t B A_t is a measurement of variable A at time one and A_t is a measurement of the same variable at time two. In other words, there is a time lag between the two measurements of A.

An advantage of the feedback is that it does not require all of the assumptions of causal models. The utility of feedback model is maximized if we keep in mind that usually the research questions raised are not always subject to unidirectional causal models. Frequently, we are unable to meet some of the causal assumptions or establish the causal priorities necessary for unidirectional causal models, and our research efforts become fraught with serious theoretical and methodological problems.

Much of the sociological research concerned with temporal dimensions involves the reciprocal interaction or feedback situations. For example, two interrelated propositions from demographic research state that (1) as the fertility rates increase, the infant mortality rates also increase, (2) as the infant mortality rates decrease, the fertility rates also decrease (Bogue, 1969: 831). Another example can be drawn from the research on modernization and urbanization. The two propositions can be stated as: (1) the greater the modernization of a society, the greater the urbanization, and (2) the greater the urbanization of a society, the greater the modernization (Lerner, 1958; Black, 1967). At first glance these examples might appear to be circular. A closer examination will indicate that there are reciprocal relationships between (1) fertility rates and infant mortality rates, and (2) modernization and urbanization of societies. Similar examples can be drawn from many social/psychological and sociological research.

In our efforts to make sociology consistent with the goals of science, we attempt to arrive at (1) estimation, and (2) structural models (Heise, 1969). While the current perspective of path analysis in sociology is sufficient to deal with estimation models, it is not adequate for analyzing structural models. It is our contention that sociological theories will be more accurately tested through the use of structural models than through the use of estimation models.

III. ASSUMPTIONS[1]

The reciprocal interaction (feedback) in path analysis may either include or ignore any time lags. While the assumptions discussed below usually apply to both of these situations.

1. <u>There is a reciprocal relationship between the variables under consideration.</u>

This assumption suggest that if A and B are reciprocally related with each other, either A or B can be used as a starting point for analysis. While it may be feasible (or even sometime pragmatic) to establish a causal order, it is essential that both A and B act in a "cause and effect" sequence. The notion of recursive systems (especially causal order) may be helpful in determining a starting point for the analysis (Blalock, 1964, 1967).

2. <u>The variables have linear relationships.</u>

This assumption is frequently made but usually not tested in sociological research. If known nonlinear relationships do exist, they can be reduced to linear form by transformations. If transformation is necessary, the user should be aware of the difficulty of specifying parameters using this method.

3. <u>The reciprocal interaction relationships of the system under study are sufficiently established from a theoretical point of view.</u>

This assumption is often made in sociological literature but is rarely empirically established. Much of the research in sociology has been concerned with single variable situations, and, consequently, they did not establish interactive relationships. It should be pointed out that there are many "grand" and "middle range" social theories that are best tested with feedback models. At the same time, it can be readily argued that some of the "micro" theories may be more clearly subject to interactive systems of relationships than unidirectional causations. While it must be admitted that sociological theories are not developed enough to effectively utilize the path analytic and causal inference approaches, feedback models of path analysis may circumvent some of these problems. In addition, we tend to view human behavior, as well as human societies, with sets of interrelated (oftentimes circular) propositions. While we cannot debate, at this point, whether a unidirectional model or a feedback model better fits such theorizing, the writers submit that the latter might be a better approach.

4. <u>Observations (measurement) of each variable are made to allow the intervals of maximum effects.</u>

This assumption can be problematic in sociological research. The reciprocity of interactions can be thought of as a standardized normal distribution and, consequently, will reach a peak. The usual path analytic approach applies (in a strict sense) to point variables only. To achieve the intervals of maximum effects, Wright (1960: 424) suggested, "The intervals between observations or the lenghts of the averaged periods should be equal to or be an aliquot part of the lag." While such controls may be applied in studies of feedback with time lag, the problem may be somewhat less cumbersome in the instantaneous feedback situation. In the latter case, one can simply assume that these adjustments are being made both ways, and, consequently, the observations remain essentially linear point variables.

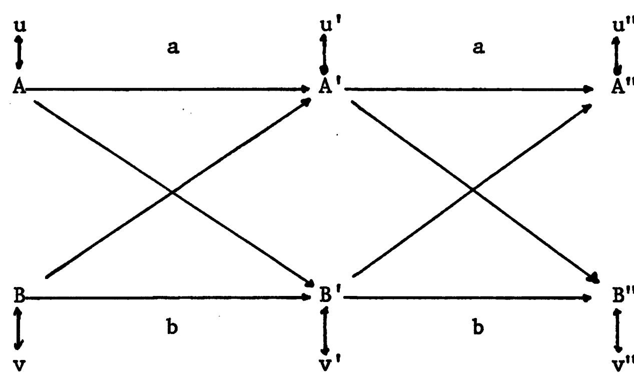
IV. FEEDBACK MODELS IN PATH ANALYSIS

As mentioned earlier, the mathematical formulation of feedback already is available (Wright, 1960; Duncan, Haller and Portes, 1968; Duncan, 1970). Other works in the area of econometrics, which have considered simultaneous structural equations include Johnston (1963), Goldberger (1964), Ezekiel and Fox (1959). Our focus in this section will be the application of feedback models of path analysis in sociological research. As previously noted, there are very few studies which have utilized such a perspective. Consequently, most of our examples will be theoretical in nature with an attempt to demonstrate how they might be used from reciprocal interaction perspectives. The examples to be discussed are bivariate, multivariate; and multivariate-multiphasic feedback models.

<u>Bivariate Feedback Models</u>

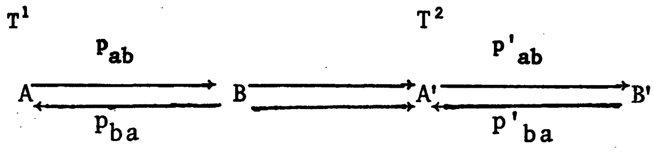
The simplest form of feedback model is the bivariate model as illustrated in Figure 2,

FIGURE 2



where A and B are the known variables and u and v are unknown residual variables. The above figure is an example of a path diagram illustrating delayed feedbacks between variables A and B and can be extended through as many cycles as are necessary (Wright, 1960b). Another way of looking at this type of feedback model is from a multiphasic perspective (with known time parameters), which can be represented as shown in Figure 3.

FIGURE 3



where T = time and p = path coefficient.

The models of instantaneous feedback are perhaps more frequent in cross sectional research or these situations which imply known exogenous variables, and known interacting variables. An example of such a feedback model is given by Wright (1960b) for studying quantity marketed and price under the classical economic theory of supply and demand. Some sociological and social/psychological theories that can fit this type of model are exchange, interpersonal and human communication theories.

An example: A widely accepted argument proposed by Lerner (1958) is that urbanization and modernization are reciprocally related. In this proposition there are two unknown residual variables (A and B) that leads to the following:

Urbanization (U) = C_{um} Modernization (M) + $C_{ua}A$ (2)

Modernization (M) = C_{mu} Urbanization (U) + $C_{ub}B$ (3)

These equations may be stated in a standardized form as:

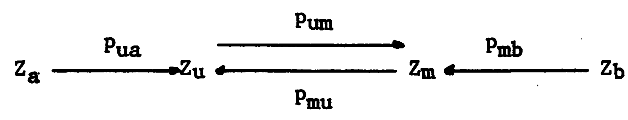
$$U = P_{um} z_m + P_{ua} z_a \quad (4)$$

$$M = P_{mu} z_u + P_{ub} z_b \quad (5)$$

Although we would rarely attempt to explain a bivariate model of reciprocal interactions, it will be of some value to understand its fundamentals. First, in this bivariate feedback model we find simple ways of looking at the reciprocal interaction without time lag. The equations may be extended with known time lags if there is need for it. The bivariate models can be easily extended into bivariate multiphasic models. It should be kept in mind that successive determinations will make such a model partially recursive. Second, as is true for the unidirectional path analysis, the reciprocal path coefficients will have simpler interpretation in the bivariate case than either the bivariate multiphasic or the multivariate multiphasic case (Land, 1969: 10-12).

The presentation of a bivariable model of this type is illustrated in Figure 4.

FIGURE 4

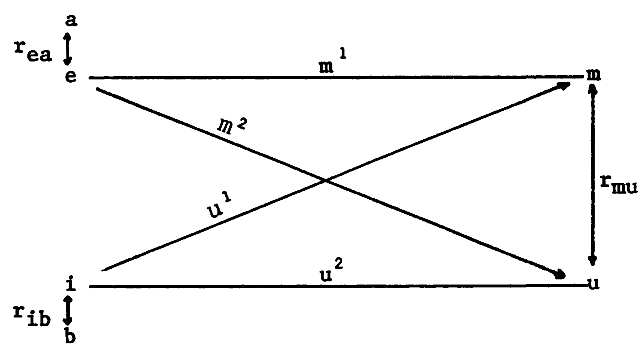


The computation of bivariate feedback path models will be quite similar to those for the simple bivariate path models with the exception that both variables will change ordinates as the case may be. There also will be two path coefficients and two coefficients of alienation.

Multivariate Feedback Path Models

In order to save space, readers are referred to Duncan, Haller and Portes (1968) and Duncan (1970) for specific computational details. Here we will attempt to expand the example suggested for the bivariate analysis. In addition to reciprocal interactions between urbanization and modernization, two known exogenous variables to this system are education and industrialization, as suggested by Lerner and others (e.g., Moore, 1963; Roger, 1969; Black, 1967). It may be assumed that the education and the industrialization of a society vary independently. Two unknown variables that are related with education and industrialization are A and B respectively. This model is represented in Figure 5.

FIGURE 5



where e = education, i = industrialization, m = modernization, u = urbanization, A and B = residual variables and r = correlation coefficient.

Considering the residual variable (A) we may write the following equations:

$$r_{ma} = m_1 r_{ea} \quad (4)$$

$$r_{ua} = u_1 r_{ea} \quad (5)$$

$$r_{mu} = m_1 u_1 + m_2 u_2 \quad (6)$$

$$r_{mm} = m_1^2 + u_2^2 = 1 \quad (7)$$

$$r_{uu} = u_1^2 + u_2^2 = 1 \quad (8)$$

$$m_2^2 u_2^2 = (1 - m_1^2)\ (1 - u_1^2) \quad (9)$$

$$(1 - m_2^2)_2 (1 - u_1^2) = 1 - m_1^2 - u_1^2 - m_1^2 u_1^2 \quad (10)$$

$$m_2^2 u_2^2 = (r_{mu} - m_1 u_1)^{2} \qquad (11)$$

$$(r_{mu} - m_1 u_1) = r_{mu} - 2m_1 u_1 r_{mu} + m_1 u_1 \qquad (12)$$

$$m_1 = \frac{r_{ea}}{r_{ua}} u_1 \qquad (13)$$

$$m = \sqrt{(1 - r_{mu}^2)/(1 - 2\frac{r_{ma}}{r_{mu}} + \frac{r_{ma}^2}{r_{ua}^2})} \qquad (14)$$

$$u_2 = \sqrt{1 - u_1^2} \qquad (15)$$

$$m_2 = \sqrt{1 - m_1^2} \qquad (16)$$

Similar equations can be written for B if it is a known variable. The regression equations for m and u can be written as:

$$m - \bar{m} = C_{me} (e - \bar{u}) + C_{ui} (i - \bar{u}) \qquad (17)$$

$$u - \bar{u} = C_{ue} (e - \bar{u}) + C_{mi} (i - \bar{u}) \qquad (18)$$

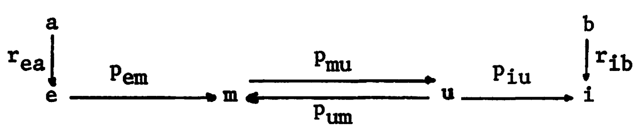
These coefficients can be calculated as ratios in which the unknown standard deviations cancel:

$$\varepsilon = \frac{u_1 \sigma_u}{m_1 \sigma_m} = C_{um} \qquad (19)$$

$$\eta = \frac{u_2 \sigma_u}{m\ m} = \frac{1}{C_{um}}$$

Thus to illustrate the feedback between urbanization and modernization, our two endogenous variables, we can look at Figure 6.

FIGURE 6



The major consideration should be given to overidentification of structural equations. Such a problem can be approached from the "two stage least squares" (2SLS) perspective. Some suggestions have been made that reciprocal interaction equations can be expressed with "unilateral causal dependence" with positioned dependent variables and solved separately through least squares (Wold and Jureen, 1953). Such a set also can be thought of as a "recursive system" of relationship when delayed feedback models are used (Ezekiel and Fox, 1959: 927-28).

The reader is reminded of the unique problems of selections and measurement of variables in the preceding situations. Although some of the inclusions of "unnecessary" variables can be controlled through 2SLS, measurement may still be a problem.

Multivariate, Multiphasic Feedback Models

Such feedback models also may be considered as partially recursive in the sense that the endogenous variables are successively determined with a known causal time order and with a continuance "cause-effect" sequence. Such a model is discussed in Duncan, Haller and Portes' (1968) article.

This model can be delineated in two, three or four phases, as the case may be. One major problem, however, is the usual overdetermination of the model. Consequently, we will have more structural coefficients. This problem can be handled, of course, through multistage least square method, which we are suggesting as an extension of the two stage least square method. Such a method is suggested by Turner and Stevens (1959) and elaborated by Duncan, Haller and Portes (1968). Essentially, it implies a combination of causal modeling and factor analysis based on heuristic considerations.

V. METHODOLOGY OF FEEDBACK AND SOCIOLOGICAL THEORIES

In this paper we have alluded to the fact that our current sociological perspectives and feedback models and their interrelationships.

It should be obvious from the preceding discussion that feedback methods have considerable promise in theory construction and verification. While much focus in sociology has been upon construction of unidirectional causal theories, we need to keep in mind that we may be ignoring the understanding of the structure of sociological theories. While we do not deny the useful contribution of the unidirectional causal approach, in our opinions, the gap between sociological theories and research will be filled by utilizing reciprocal interaction models.

We can assume that in a crude way the major sociological traditions can be categorized into (1) evolutionary, (2) structural-functional, (3) conflict, and (4) cyclical perspectives. We can further assume that all of these traditions attempt (even though they may not succeed) to offer multilinear models of society. Thus, it becomes imperative that these theoretical traditions and derivations thereafter involve interdependence and often-times reciprocal relationships with or without time lag. In addition, explanations and dynamics of social change (which seem to be one of our major concerns without much directed effort) are embedded with sets of reciprocal interactions.

While we accept the fact that unidirectional causal models are first steps in constructing causal models, we need to extend the understanding of reciprocal relationships as the necessary second step for theory construction.

VI. CONCLUSIONS AND IMPLICATIONS

We are usually at a loss to utilize our most frequently discussed social theories because we have not had methodological tools to handle some theories. While the systems analysis approach has been available to us for some time, it has been largely unused due to needed assumptions of total interdependence. We know that the variables we deal with in sociology are rarely independent of one another. The reciprocal interaction models stand in the middle in that they can treat interdependence on a smaller but understandable level.

Certainly the value of the notion of feedback is well known to sociological researchers. Many areas of social research, such as communications, decision making, social control, cannot be meaningfully analyzed without accounting for feedback in some way. Our purpose in this paper has been to suggest a technique for dealing with feedback that has been relatively ignored by sociologists. Currently researchers usually deal with feedback by means of ANOVA, simulations and various other types of controlled experimentation.

Since path analysis has been shown to be a powerful technique, particularly in helping us to determine various cause and effect relationships, it seems essential that its methods of dealing with feedback should become more fully explicated and developed. Many social situations cannot be characterized in simple cause and effect terms unless the situation is viewed as being static. Many, if not most, social situations require that we understand the reciprocal causal effects of the variables under consideration. The notion of feedback in path analysis is one method by which we may assess reciprocal causation. We recognize that this technique will not eliminate all of the confusion that results from "causal analysis." However, it should assist us in a variety of situations encountered by sociologists. It is unrealistic to assume that social systems are closed and, thus, we must, in some manner, seriously deal with the notion of feedback if our discipline is to continue to develop.

FOOTNOTES

1. We will be concerned with only those assumptions that are unique to feedback models. Assumptions about sampling and measurement are almost the same as those necessary for the unidirectional path analysis. See Heise (1969) for these assumptions.

REFERENCES

Black, C. E.
1967 The Dynamics of Modernization. New York: Harper and Row.

Blalock, H. M., Jr.
1964 Causal Inferences in Nonexperimental Research. Chapel Hill: The University of North Carolina Press.

1967 "Path coefficient versus regression coefficient." American Journal of Sociology. 71: 37-47.

Bogue, D. J.
1969 Principles of Demography. New York: John Wiley.

Boudon, R.
1965 "A method of linear causal analysis: dependence analysis." American Sociological Review 30: 365-72.

Duncan, O. D.
1966 "Path analysis: sociological examples." American Journal of Sociology 72: 1-16.

1970 "Duncan's correction of published text of 'peer influence on aspirations: a reinterpretation." American Journal of Sociology 75: 1042-46.

Duncan, O. D., Haller, A. O. and Portes, A.
1968 "Peer influence on aspirations: a reinterpretation." American Journal of Sociology 75: 119-37.

Ezekiel, M. and Fox, K. A.
1959 Methods of Correlation and Regression Analysis. New York: Wiley, 2nd ed.

Goldberger, A.
1964 Econometric Theory. New York: Wiley.

Heise, D. R.
1969 "Problems in path analysis and causal inferences." Pp. 38-73 in E. F. Borgatta (ed.) Sociological Methodology 1969. San Francisco: Jossey-Bass.

Johnston, J.
1963 Econometric Methods. New York: McGraw-Hill.

Land, K. C.
1969 "Principles of path analysis." Pp. 1-37 in E. F. Borgatta (ed.) Sociological Methodology 1969. San Francisco: Jossey-Bass.

Lerner, D.
1958 The Passing of Traditional Society: Modernizing the Middle East. New York: The Free Press.

Moore, W.
1963 Social Change. Englewood Cliffs: Prentice-Hall.

Roger, E. M.
1969 Modernization Among Peasants: The Impact of Communication. New York: Holt, Rinehart and Winston.

Tukey, J. W.
1954 "Causation, regression and path analysis." Pp. 35-66 in O. Kempthorne and others (eds.) Statistics and Mathematics in Biology. Ames, Iowa: Iowa State College Press.

Turner, M. E. and Stevens, C. D.
1959 "The regression analysis of causal paths." _Biometrics_, 15: 236-258.

Wold, H. and Jureen, L.
1953 _Demand Analysis: A Study in Econometrics_. New York: Wiley.

Wright, S.
1921 "Correlation and causation." Journal of Agricultural Research 20: 557-85.

1934 "The method of path coefficient." _Annals of Mathematical Statistics_ 5: 161-215.

1960a "Path coefficient and path regression: alternative or complimentary concepts?" _Biometrics_ 16: 189-202.

1960b "The treatment of reciprocal interaction with or without lag, in path analysis." _Biometrics_ 16: 423-45.

A SAMPLE DESIGN FOR INTERVIEWING A HARD TO LOCATE POPULATION

Bernard R. Siskin and Richard D. Leone, Temple University

In selecting an appropriate sample design, the problem was twofold: (1) to complete a sufficient number of interviews permitting responsible inferences not only about the overall sample, but also about each stratum, given budgetary constraints, and (2) to make adjustments for the bias concomitant with nonresponse in a study where it is safe to assume that a sizeable percentage of the sample cannot be located and the probability of response is correlated with the dependent variables.

The low-response rate gives rise to a problem, because, realistically it can be assumed that the ability to locate a person is a function of the person's mobility, and the latter is a function of demographic characteristics. In turn, the dependent variable of the investigation might well be correlated with these demographic characteristics or with mobility itself. Thus the response rate may differ greatly from one stratum to another, and a significant nonresponse bias is highly probable.

In assessing the problem alluded to above, an assumption was posited that the population could fall into one of three possible categories: (1) easy to interview, (2) difficult to interview, (3) impossible to interview.

What differentiates the first category from the second is the amount of time it takes to trace a client and ultimately to interview him. The easy-to-get represent people who were at the same address or left a clear path to their new residency. It is assumed, moreover, that the three categories represent a continuum with respect to the dependent measures. Implicit in this assumption is that the impossible-to-interview represents people who cannot be located in the interview time period. Moreover, it is assumed that, if they could have been located, their mobility and other characteristics would be similar to those in the second category as compared to those in the first who remained at their address of record or who left tracks that were easy to follow. This, of course, is a simplification. The impossible-to-interview can in fact be subdivided into categories and various subcategories.

1. Located but not interviewed.
   (a) Explicitly refuse.
   (b) Implicitly refuse; located, but consistently avoid being interviewed.
   (c) Institutionalized.[1]
2. Cannot be located.[2]

We hypothesized, however, that the bulk of the impossible to interview would fall into the second subdivision, and subsequent experience verified this.

Common Procedures

Our awareness of the biases associated with the nonresponse problem led us to consider three rather common sampling procedures: (1) oversampling, (2) substitution, and (3) oversampling with subsampling of the nonrespondents.

Oversampling. In oversampling, the initial sample size is increased to adjust for an anticipated low response rate. This procedure encourages interviewing only the "easy-to-interview," which reduces the time spent tracing. Two major shortcomings are tied to this procedure: (1) the fundamental bias of the nonrespondent class is ignored, and (2) there is no guarantee that a reasonable response rate for each stratum will be forthcoming. We felt extremely uncomfortable with positing an a priori personal judgment as to the different expected response rates in each stratum. Hence, to double merely each stratum, or to allocate "incorrectly" the increases in each stratum by some other arbitrary formula, might result in an underrepresentation of those strata where there is a low response rate. Consequently it might be impossible to make reasonable inferences about the subpopulations.

Substitution. This procedure recommends that as interviewing in the field unfolds, a person possessing similar demographic characteristics is substituted for a sample member who cannot be located. The substitution method answers the problem of disproportionate responses for each stratum because the matching process should dynamically adjust the amount of oversampling in each stratum to compensate for the varying strata response rates. There exists, however, a real doubt as to whether this procedure reduces the nonrespondent bias [Kish, 1965]. To the extent that only the easy-to-locate are interviewed, this would in no way compensate for the inherent nonrespondent bias.

Oversampling and Subsampling. The third and final common sampling procedure we considered was to oversample and to subsample for all nonrespondents. Similar to simple oversampling, this method does not assure adequate representation in each stratum. It does, however, provide some clue as to the nonrespondent bias. Indeed, the results of the nonrespondent sample could be laid out on a time continuum and fit into a mathematical model to estimate the nonrespondent bias [Housman, 1953].

Additional costs are often insignificant for the follow-up of nonrespondents, but the cost factor often militates against a follow-up procedure in those studies where interviewers are in the field, and it is difficult to locate the interviewee. The response rate for the nonresponse segment of the sample might well be extremely low, and, if a reasonable sample of "nonrespondents" is desired, the "nonresponse" sample would have to be very large, perhaps even the census of nonrespondents. The principal barrier to interviewing this type of population is in locating the person to be interviewed, and not usually in calling back after location is determined. The amount of interviewer time spent in tracking and subsequently actually conducting an interview can be quite long. The costs associated just with tracing are quite high, especially when one considers that an interviewer may have to make numerous separate trips to an area before he can even complete one interview. Also, the morale of interviewers will tend to wane, even if they are offered a bonus, as they spend excessive amounts of time without completing an interview.

Recommended Procedure

Having weighed the pros and cons of the sampling procedures usually employed, we developed a design which combined the substitution method with the nonrespondent subsampling technique. This combined design calls for the initial selection of a stratified sample of size N_i in stratum i where N_i=desired number of completed interviews in stratum i.[3] For a fixed time period, (t_1) this initial sample is tracked and (if possible) interviewed. At the end of this predetermined period of time, all clients not interviewed will be classified as the "difficult-to-interview" and will constitute a special subsample. For each client included in this "difficult-to-interview" subsample, a substitute client possessing matched demographic characteristics will be selected and included in an alternate sample. During another specified time period (t_2) tracing and interviewing continues for both the "difficult-to-interview" group and the newly created alternate sample. At the end of this period (t_2), the process will be repeated: creating additional alternate samples at the end of each time period reflecting the interview need for each stratum, the process is terminated when the desired number of interviews are completed.

Let us consider another hypothetical case in which we shall compare oversampling with census subsampling of nonrespondents with the recommended procedure. Assume that there are two distinct strata of equal research interest. Table 1 shows the true unknown response rate for each of the strata, where response rate is a function of time spent tracing and interviewing.

The desired final completed interviewed sample is 500, preferably 250 from each of the strata. The costs for these results will vary depending upon the design chosen. Let us reasonably assume, moreover, that the costs associated with tracing and interviewing is an increasing function of the time a sample element remains in the field and is sought. Consider the two following cost situations:

1. Where the costs per time period expended in the field in quest of any sample element is the same for both design and is (a) C_t=10 + 10t t=1, ...4 where t=number of time periods the element is in the field, and
2. Where the cost per time period expended in the field on any sample element is as above where oversampling transpires, but due to averaging out travel costs over a number of interviews is (b) C_t= 15+5t t=1, ...4 for the combined recommended design.

Table 2 shows basically two alternatives: (1) the expected results from oversampling with census subsampling of the nonrespondents in time period 1, and (2) the anticipated results from the recommended design discussed above.

As this hypothetical example illustrates, the recommended procedure has certain advantages. First, since it used substitution, it results in adequate representation both overall and in each stratum. While oversampling with subsampling also reaches the target for total interviews, the representation for stratum B is inadequate. This, of course, could have been compensated for if we knew the relative response rates a priori. Second, inasmuch as the recommended process includes specifically a subsample of the "difficult-to-interview" category, the nonresponse bias can be estimated more readily, since the number of "difficult-to-interview" completion is higher. Third, this recommended sample procedure costs less per completed interview and considerably less in total cost under the second than does oversampling with subsampling in this case.

Results of Recommended Procedure

Now let us turn our attention to our experience in actively implementing this design. As noted above, we were interested in studying the effect of a program change which was implemented as the Concentrated Employment Program in Philadelphia. Although this change in the program was initiated in January 1970, this research did not commence until 1971. We found existing program data inadequate and thus we deemed it necessary to interview 250 clients from 1969 and 250 clients from 1970, since we were interested in finding whether the program change improved the employability of 1970 clients over 1969 enrollees. We were also concerned about the impact of the program modification upon different types of clients. Therefore, we selected matched samples from each year stratified on the basis of demographic characteristics, so that we could evaluate each stratum separately as well as each year as a whole.

A pilot field study, which lasted one month, demonstrated that within the time allocated an interview completion rate of approximately 50 percent was attainable.[4] It indicated, moreover, that low response was a function of the difficulty in locating the clients and was not caused by implicit or explicit refusal to be interviewed. In light of these test findings, we decided to use the design outlined above and to use a time schedule with four time frames of six weeks each. The interviewers would track sample members over each time period, and after which decisions could be made concerning substituting.

The six-week period was chosen to allow adequate time for call backs in those cases where the client's residence was known but where the actual interviewing required more than one call. The results of the sampling procedure are presented in Tables 3 through 6.

As might be expected, the dynamics of the substitution process altered the sample allocations in each stratum in order that the distribution of completed interviews was similar to the initial sample distribution. Table 7 shows the adjustment of the sampling distribution as a result of the substitution process. When the sampling results are analyzed on the basis of specific demographic subgroups, the results are as we anticipated except for two subcategories, race and age. Spanish speaking people were easier to locate than Negroes; this result was not expected because the Spanish tended to be at the lower end of the socioeconomic scale in Philadelphia. More significantly, we were able to locate younger clients easier than older ones.[5] These findings illustrate the risks involved if a priori judgments are made about where to oversample to compensate for differing stratum response rates.

Conclusion

The research design and tracing techniques discussed herein have resulted in a completion rate which is high enough to make inference about the population and each subpopulation of interest within our budgetary constraints. Moreover, we were able to estimate the nonrespondent bias. It should be noted that from an administrative point of view, however, the sampling procedure requires a high degree of coordination and control. Not only are numerous forms needed to monitor the interviewing as it unfolds, but this particular design also necessitates: (1) continually updating activities in the field, (2) expeditiously replacing those clients who are designated as members of the nonrespondent sample, and (3) staying abreast at all times with the status of each individual within the sample so that the appropriate decision concerning replacing can be made.

In light of this experience it is suggested that if this design is to be used, that a period somewhat longer than six weeks for each wave is necessary because of the coordinating problems. It is also necessary to provide a longer period of time between the start up of the initial wave and the subsequent wave so as to be able to identify the nonrespondent sample cohort and to execute the preliminary mailing procedure for the replacement sample. On the other hand, experience has demonstrated that the delay period cannot be too long; otherwise it is impossible to retain experienced interviewers and the cost of interviewing drastically increases. This increase in cost arises when there is no overlap between waves. When a reasonable overlap exists, interviewers are able to search for clients considered hard to find while they are working in those areas actually interviewing clients who previously indicated that they were willing to cooperate in the study. Hence, travel costs are shared with the cooperating group and the hard-to-locate one.

Acknowledgement

The authors are most grateful to the Department of Labor for its support. This paper was prepared as part of a report for the Manpower Administration, U.S. Department of Labor, under Research Contract No. 41-1-002-42 authorized by the Economic Opportunity Act. Since contractors performing research under government sponsorship are encouraged to express their own judgment freely, the report does not necessarily represent the Department's official opinion or policy. Moreover, the authors are solely responsible for the factual accuracy of all material in the report.

Footnotes

1. In this study the decision was made not to interview those in this category, but possibly they could have been interviewed.
2. Included a small proportion (10 percent) who had moved out of the Philadelphia Standard Metropolitan Statistical Area, to which interviewing was restricted.
3. This means the total of the "easy-to-interview" and "hard-to-interview" completions. If only the number of "easy-to-interview" completions are specified, then we must set $N_i >$ desired N_i, since N_i will approach the desired N_i as the number of time periods goes to infinity.
4. Budget constraints restricted the size of the pilot sample. Consequently reliable estimates of strata response rates were not available.
5. Other characteristics such as sex, race and education were distributed equally among the age groups.
6. This response rate should have been larger, but a delay in getting into the field resulted in some persons becoming institutionalized and some moving out of the area.
7. For Wave one (1) we received assistance from one of Bell Telephone's local offices. This source was willing to search not only current telephone directories, but also other files. This proved to be quite time consuming for the local office personnel. Thus when we requested assistance in Wave two (2), the local office referred us to Bell Telephone's Community Service Office located at their corporate headquarters. Since this latter division was willing to search only telephone directories (something we could do ourselves) we did not request their assistance in subsequent waves.

References

Cochran, W. G. Sampling Techniques. 2nd ed. New York: John Wiley and Sons, 1963.
Hansen, M.H., Hurwitz, W.N. and Madow, W.G. Sample Survey Methods and Theory. Vol. I.
Hendrichs, W.H. "Adjustments for Bias by Non-Response in Mailed Surveys." Agricultural Economics Research, 1 (1949), 52.
Houseman, E.E. "Statistical Treatment of Non-Response Problem." Agricultural Economics Research, 5 (1953), 12-18.

TABLE 1. RESPONSE RATES AS A FUNCTION OF TIME SPENT TRACING AND INTERVIEWING

Time Period	Strata A	Strata B
1	.700	.350
2	.790	.385
3	.811	.405
4	.815	.408

TABLE 2. HYPOTHETICAL RESULTS: OVERSAMPLING WITH SUBSAMPLING AND RECOMMENDED DESIGN

Sample Size Results and Costs	Oversampling and Subsampling Stratum A	Stratum B	Stratum A + B	Recommended Sample Design Stratum A	Stratum B	Stratum A + B
Total sample size	385	385	770	354.2	586.7	940.9
Easy-to-interview Completed	269	134	404	248	205	453
Difficult-to-interview Completed	44	50	94	49	63	112
Total response	313	184	498	297	268	565
Total response rates	82	48	64	83	45	60
Total costs[a]						
(a)		$51,851.00			$52,173.00	
(b)		51,851.00			45,372.00	
Average cost per interview[a]						
(a)		103.97			92.21	
(b)		103.97			80.20	

[a]Under different cost assumptions A and B.

TABLE 3. INTERVIEWING RESULTS

Sample, Size, Status and Costs	Results
Initial sample size	500
Total sample size	996
Easy-to-interview Completed	469
Difficult-to-interview Completed	68
Total response	537
Total response rate	53.92
Total costs	$25,776.
Average cost (per interview)	$48.

TABLE 4. RESPONSE RATE AS A FUNCTION OF TIME SPENT TRACING AND INTERVIEWING

Time Period (Six Weeks Each)	Cumulative Response Rate
1	43.27
2	50.21
3	53.93
4	56.13[a]

[a]Note 56.13 is larger than the actual response rate as shown in Table 3 since not all elements are in the field four periods. This figure would represent the response rate for oversampling with census subsampling.

TABLE 5. COMPOSITION OF "IMPOSSIBLE-TO-INTERVIEW" CATEGORY

Subdivision		Number	Percent of Total
Located but not interviewed		113	24.6
Explicit refusal	(15)		
Implicit refusal	(65)		
Institutionalized	(33)		
Cannot be located		346	75.4
Total		459	100.0

TABLE 6. RESULTS OF SAMPLING BY DEMOGRAPHIC CHARACTERISTICS

Enrollee Characteristics	Percent Initial Sampling			Percent Final Sampling			Distribution of Sample Results	Completion Rates	
	1969	1970	1969 & 1970	1969	1970	1969 & 1970	1969 & 1970	1969	1970
Sex									
Male	64.8	65.6	65.2	69.6	70.2	69.9	61.4	39.0	43.9
Female	35.2	34.4	34.8	30.4	29.8	30.1	38.6	56.0	65.0
Race									
Negro	79.2	79.6	79.4	76.6	80.2	78.3	82.1	48.4	50.4
Spanish	20.8	20.4	20.6	23.5	19.8	21.7	17.9	30.6	49.5
Age[a]									
18 and under	14.4	14.8	14.6	12.6	13.5	13.1	14.9	55.4	14.9
19-21	30.8	31.2	31.0	31.0	29.0	30.0	31.6	45.0	54.7
22-24	22.4	21.2	21.8	22.3	22.7	22.5	21.8	37.4	54.1
25-34	27.2	28.0	27.6	27.5	29.4	28.4	26.6	45.8	42.5
35-54	5.2	4.8	5.0	6.6	5.4	6.0	5.1	35.3	46.2
Grades Completed									
4 and under	2.0	2.4	2.2	2.1	2.3	2.2	1.0	18.2	27.3
5-7	8.0	6.8	7.4	9.5	8.5	9.2	6.2	26.5	39.0
8	3.2	4.4	3.8	3.1	4.4	3.7	4.5	62.5	52.9
9-11	58.0	58.0	58.0	58.5	58.8	58.6	59.3	46.7	48.6
12	28.8	28.4	28.6	26.7	26.0	26.7	29.0	44.9	59.2
Program Status									
Completor	57.6	58.0	57.8	58.1	56.5	57.3	59.1	47.0	50.2
Dropout	42.4	42.0	42.2	41.9	43.5	42.7	40.9	40.3	50.2

[a]Age is determined at entrance of program and is age at last birthday.

TABLE 7. SUBSTITUTION EFFECT ON OVERSAMPLING BY DEMOGRAPHIC CHARACTERISTICS

Demographic Characteristics	Increase or Decrease in Percentage of Sample (final-initial)
Sex	
Male	+4.7
Female	-4.7
Race	
Negro	+0.8
Spanish	-0.8
Age[a]	
18 and under	-1.5
19-21	-1.0
22-24	+0.7
25-34	+0.8
35-54	+1.0
Grades Completed	
4 and under	0
5-7	+1.7
8	-0.1
9-11	+0.6
12	-2.2
Program Status	
Completor	-0.5
Dropout	+0.5
Year	
1969	+1.8
1970	-1.8

[a]Age is determined at entrance of program and is age at the last birthday.

CONSUMER ANTICIPATIONS AND HOUSEHOLD CHARACTERISTICS

Thomas Stoterau, Bureau of the Census

I. Introduction

The intent of this paper is to examine the probability approach to ex-ante consumer behavior for different types of households. The general hypothesis to be explored is whether income or education, for example, affect the general suitability of purchase probability questions. In the past, any such concern has been parried by the assurance that the concept of chance is commonplace in a variety of American rituals. Horse racing, World Series' pools, Weather Bureau reports, and the "numbers" have always been cited as examples. It, admittedly, seems reasonable that nearly everyone should have encountered the idea of chance in at least one of these situations. Regardless, this argument provides no empirical evidence about the ability of all types of respondents to assess the likelihood of future personal events. Thus it is that we will look at auto purchase probabilties and subsequent purchase behavior in conjunction with income, education, type of family, and respondent.

In 1966 the Census Bureau's program to measure ex-ante purchase behavior pioneered a subjective probability methodology. The survey abandoned its intentions to buy questions for what seemed to be a rather logical extension of the intention's rationale. The new approach consisted of asking respondents to numerically describe their chances of purchasing various durables over several time horizons. This technique utilized an eleven point "flash card" (0,10,20,...,100). It was believed that this would allow respondents to rationally speculate about their purchase probabilities. Whereas responses had previously been restricted to yes-no (0,1), the likelihood of the purchase event could now be described with greater discrimination. This approach has continued as the central feature of the Survey of Consumer Buying Expectations ever since.

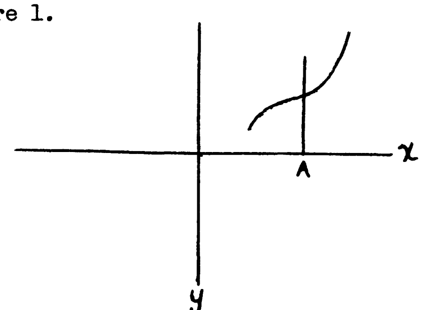
II. Description of Data and Analysis

The data used in this paper were gathered by the Census Bureau from a special panel known as the Consumer Anticipations Survey (CAS). The CAS study was developed as an experimental effort which would be used to revise and strengthen the Bureau's quarterly survey of consumer anticipations. Both the Census Bureau and the National Bureau of Economic Research participated in the design of the survey and questionnaires. In the study, about 3,500 households were asked detailed questions about their economic and demographic characteristics. These households were visited five times at roughly six-month intervals between May 1968 and November 1970. The sample selection was non-random and purposive. All of the selected census tracts were suburban and had moderately high to high median incomes. Such areas were purposely selected because these households were likely to yield a high frequency of positive responses to questions about economic activity.

The CAS survey provided information on a wide range of items concerning household decision making, and was the basis of an invited session at the 1971 ASA meetings in Fort Collins.1/

The analytical process which supports this paper is cross-sectional. That is, each household is considered an observation, and the dependent variable is the actual purchase of a new car (0,1) in a given six-month period. The independent variables, of course, vary with each equation. This is the standard, cross-section regression analysis. The work presented below differs from previous analytical efforts in that here the regression equations are run on partitioned independent variables.2/ In this case, the reasons for partitioning an independent variable are: (1) to examine the fit of the regression equation and (2) to observe the impact of the partitioned variable, over two or more bounded areas, upon the other variables in the equation. The first hopes that by segmenting a curvilinear relationship, we can find more nearly linear situations. Figure 1 presents a graphical illustration of a hypothetical case.

Figure 1.



As can be seen, regression equations fit over the two segments are more nearly appropriate than a single equation fit over the entire range of values. If, however, the relationship between the dependent and the independent variable is linear, the equations fit over the bounded segments will approximate the one fit over the entire x-axis. But even if a given partition does not improve the fit of the regression equation, it may still be useful analytically as a means of telling us about the impact of the partitioned variable upon the other variables. In this study, households were partitioned into those with incomes above $10,000 and those $10,000 and below. Regression equations with the same variables were then run for each of them. Other partitioned variables were treated similarly. All of the variables for these equations are defined in table 1, below.

III. Income and Anticipations

The selection of a partition is arbitrary. In this case, $10,000 was selected because it is about the median U.S. family income. The regression results for the $10,000 income partition can be seen in tables 2 and 3 at the end of the paper. In general, the R-squares are about the same for both groups. The six-month and twelve-month purchase probabilties ($_6PP^*$ and $_{12}PP^*$) are much more significant for the upper income segment. That is, $_6PP^*$ and $_{12}PP^*$ explain much more of the variance in new car purchases for households with incomes above $10,000 than those under. The t-values are almost twice as large for the upper group.

Total family income ($I_{A,B}$) and other income ($OI_{A,B}$) are also more significant for the upper income segment. In contrast, the type of family ($TF_{A,B}$) is not significant for households above $10,000, but it is significant for the lower group. Since most of the households were husband - wife families, the type of family variable is almost a scaled variable for the number of children in the family (see Table 1). The respondent variable ($RS_{A,B}$) is also more significant for the lower group. Somewhat surprisingly, none of the other variables were significant for either group. These included the probability of the head being unemployed, education of head, and whether the household was currently making payments on a car. These results would suggest that the purchase probability method may be more suitable for households with incomes above $10,000. It does not, however, suggest that the approach is necessarily inappropriate for other households.

IV. Education and Anticipation

The first attempts to partition the education variable were not fruitful. These efforts first used four years of high school and then one to three years of college. Later, a partition at four or more years of college yielded some worthwhile results. The regression equations and their related coefficients can be seen in tables 4 and 5 at the end of this paper. One of the most interesting findings is that a dummy variable for respondent is not significant for households with a head with four or more years of college, but it is significant with a positive sign for those with less education. The respondent variable assumed a value of 1 if the head was involved in the interview and a zero if the head was not. This implies that for households with college educated heads it does not matter who the respondent is, but that for less well educated households it does. There may be implications in this finding which affect the on-going Survey of Consumer Buying Expectations (CBE). Our regular CBE survey accepts any person living in the household who is eighteen or older. It is generally believed that it is too difficult and expensive to restrict the respondent to one specified member of the households in a fairly large scale, current program. And despite the findings above reported, it is very unlikely that the benefit would warrant the cost of restricting the respondent.

As can be seen in table 4, below, the purchase probabilties ($_6PP^*$ and $_{12}PP^*$) are more significant for the upper education segment. In contrast, type of family ($TF_{A,B}$) and total family income ($I_{A,B}$) are generally more significant for the households whose head did not have four years of college. None of the other variables explained a significant portion of the variance in any of the equations.

Table 1.--CAS VARIABLES INCLUDED IN REGRESSION ANALYSIS

$RS_{A,B}$ - Respondent in interview (subscript refers to whether it is an above or below partition household)

1 = Head involved in interview
0 = otherwise

$EH_{A,B}$ - Education of head

1 = no education
2 = 1 to 8 years of elementary school
3 = 1 to 3 years of high school
4 = 4 years of high school
5 = 1 to 3 years of college
6 = 4 years of college
7 = 5 or more years of college

$U^*_{A,B}$ - Probability of head being unemployed in the next twelve months (0, 10, 20,..., 100)

$OI_{A,B}$ - Other income for the household (interest and dividends, capital gains, rent, own business, and pensions)

Actual dollar amount (5 digits)

$I_{A,B}$ - Total annual income from all sources

Actual dollar amount (5 digits)

$TF_{A,B}$ - Type of family

1 = single male head, no children
2 = single male head, 1 or more children
3 = single female head, no children
4 = single female head, 1 or more children
5 = husband and wife, no children
6 = husband and wife, 1 child
7 = husband and wife, 2 children
8 = husband and wife, 3 children
9 = husband and wife, 4 or more children

$MP_{A,B}$ - Presently making payments on a car

0 = No
1 = Yes

$PG_{A,B}$ - Present grade in school of oldest child (if under 21)

00 to 18

$_6PP^*_{A,B}$ - Probability of purchasing a new car in the next six-months (0, 10, 20,..., 100)

$_{12}PP^*_{A,B}$ - Probability of purchasing a new car in the next twelve-months

V. Conclusion

Contrary to my a priori suspicions, the regression equations for the various segments of the partitioned variable did not differ greatly. As is often the case with cross-section studies, all of the R-squares are disappointing. And due to the nature of the data, comparisons among the standard errors of the estimate (Sy.x) are not very meaningful. We did find that the purchase probabilties have larger t-values for households with incomes above $10,000 and for households whose head has four years of college. It also seemed possible that we would find that for some types of households, purchase probabilties are not significant. This did not happen. The t-values of $_6PP^*$ and $_{12}PP^*$ are fairly healthy in all of the equations.

Footnotes

1/ See Proceedings of the Social Statistics Section, American Statistical Association, Washington, D.C., pp. 126-170

Table 2.--REGRESSION EQUATIONS FOR HOUSEHOLDS WITH INCOME ABOVE $10,000 WITH OBJECTIVE AND ANTICIPATORY VARIABLES - DEPENDENT VARIABLE IS THE PURCHASE OF A NEW CAR

(t-values are shown in parenthesis)

Equation number	R^2_A	REGRESSION COEFFICIENTS										r	$S_{y.x}$
		RS_A	EH_A	U^*_A	OI_A	I_A	TF_A	$_6PP^*_A$	$_{12}PP^*_A$	MP_A	PG_A		
I.....	.112	.012 (1.96)	-.008 (1.79)	-.004 (1.02)	-	-	-	.035 (18.98)	-	-	-	.104	.295
II....	.099	-	-.008 (1.75)	-	.0000 (3.03)	-	-	-	.0254 (17.05)	-.008 (.67)	.0052 (.389)	.087	.297
III...	.153	.004 (.746)	-002 (.39)	-034 (.79)	-	-	-	.044 (22.86)	-	-	-	.066	.295
IV....	.012	-	-.0035 (.704)	-	-	.0000 (6.00)	-.0004 (5.76)	-	-	-	-	.081	.319
V.....	.100	.0089 (1.49)	-	-.025 (.588)	-	.0000 (3.53)	.0006 (.090)	-	.0249 (16.62)	-	-	.007	.296

NOTE: The A subscript indicates that the variable is for households above the income partition.

Table 3.--REGRESSION EQUATIONS FOR HOUSEHOLDS WITH AN INCOME OF $10,000 OR LESS WITH OBJECTIVE AND ANTICIPATORY VARIABLES - DEPENDENT VARIABLE IS THE PURCHASE OF A NEW CAR

(t-values are shown in parenthesis)

Equation number	R^2_B	REGRESSION COEFFICIENTS										r	$S_{y.x}$
		RS_B	EH_B	U^*_B	OI_B	I_B	TF_B	$_6PP^*_B$	$_{12}PP^*_B$	MP_B	PG_B		
I.....	.124	.0236 (2.19)	.0011 (1.42)	.0018 (.297)	-	-	-	.0439 (8.58)	-	-	-	.058	.255
II....	.126	-	-.003 (.381)	-	.000011 (.378)	-	-	-	.0331 (9.19)	-.025 (1.13)	.005 (.389)	.052	.255
III...	.049	.0097 (.934)	.0025 (.391)	-.0079 (1.38)	-	-	-	.0248 (5.17)	-	-	-	.036	.247
IV....	.063	-	-	-	-	-	-	-	.023 (6.41)	.0069 (3.24)	-	.0378	.245
V.....	.138	.0235 (2.21)	-	.0013 (.230)	-	.000009 (.764)	-.022 (2.13)	-	.0328 (9.24)	-	-	.065	.253

NOTE: The B subscript indicates that the variables is for households below the income partition.

Table 4.—REGRESSION EQUATIONS FOR HOUSEHOLDS WITH FOUR OR MORE YEARS OF COLLEGE - DEPENDENT VARIABLE IS THE PURCHASE OF A NEW CAR

(t-values are shown in parenthesis)

Equation numbers	R^2_A	REGRESSION COEFFICIENTS										λ	$S_{y.x}$
		RS_A	EH_A	U^*_A	OI_A	I_A	TF_A	$_{6}PP^*_A$	$_{12}PP^*_A$	MP_A	PG_A		
I.....	.133	.0056 (.77)	-	-.0039 (.62)	-	-	-	.0375 (16.71)	-	-	.0040 (.308)	.088	.284
II....	.111	.0042 (.57)	-	-.0042 (.659)	-	.000009 (1.604)	.0011 (1.37)	-	.0265 (14.47)	-	-	-.0134	.287
III...	.086	-	.0014 (1.02)	-	-	-	.0081 (.997)	-	.0223 (13.13)	-	-	-.0151	.291
IV....	.0221	-	-.0134 (.83)	-	.000014 (1.51)	-	-	.0149 (6.00)	-	-	-	.077	.311
V.....	.058	.0054 (.72)	-	-.0034 (.53)	-	.000010 (2.33)	-	-	.0156 (9.75)	-	-	.037	.300
VI....	.0096	-	-	-	-	.000018 (3.875)	-.0090 (1.047)	-	-	.0199 (1.39)	-	-.042	.306

NOTE: The A subscript indicates that the variable is for households above the education partition.

Table 5.—REGRESSION EQUATIONS FOR HOUSEHOLDS WHOSE HEAD HAS LESS THAN FOUR YEARS OF COLLEGE - DEPENDENT VARIABLE IS THE PURCHASE OF A NEW CAR

(t-values are shown in parenthesis)

Equation numbers	R^2_B	REGRESSION COEFFICIENTS										λ	$S_{y.x}$
		RS_B	EH_B	U^*_B	OI_B	I_B	TF_B	$_{6}PP^*_B$	$_{12}PP^*_B$	MP_B	PG_B		
I.....	.096	.0226 (2.996)	-	-.0002 (.512)	-	-	-	.0334 (12.77)	-	-	-.0185 (1.42)	.093	.293
II....	.104	.0194 (2.58)	-	-.0003 (.076)	-	.0000 (3.73)	-.0135 (1.88)	-	.0245 (11.82)	-	-	.037	.292
III...	.059	-	-.002 (.213)	-	-	-	-.0197 (2.67)	-	.0183 (10.02)	-	-	.105	.299
IV....	.007	-	.186 (1.15)	-	.0000 (1.34)	-	-	.008 (2.86)	-	-	-	.078	.311
V.....	.062	.021 (2.79)	-	-.0000 (.004)	-	.0000 (4.30)	-	-	.0136 (7.88)	-	-	.035	.299
VI....	.008	-	-	-	-	.0000 (3.56)	.0008 (.11)	-	-	.005 (5.64)	-	.032	.293

NOTE: The B subscript indicates that the variable is for households below the education partition.

ALLOCATION IN STRATIFIED SAMPLING BASED ON PRELIMINARY TEST OF SIGNIFICANCE II

Victor K. T. Tang and B. V. Sukhatme
California State University, Humboldt and Iowa State University, Ames

I. Introduction

Let a population consisting of N units be classified into k strata, the i-th stratum having N_i units $i = 1,2,\ldots,k$ so that $\sum_{i=1}^{k} N_i = N$. Let Y be the characteristic under study and consider the problem of estimating the population mean $\bar{y}_N = \frac{1}{N}\sum_{i=1}^{N} y_i$ from a stratified random sample of size $n = \sum_{i=1}^{k} n_i$ where n_i units are drawn by simple random sampling without replacement from the i-th stratum $i = 1,2,\ldots,k$. An unbiased estimate of the mean $\bar{y}_N$ is given by

$$\bar{y}_{st} = \sum_{i=1}^{k} W_i \bar{y}_{n_i} \tag{1.1}$$

where W_i is the proportion of units in the i-th stratum and $\bar{y}_{n_i}$ is the simple mean unbiased estimate of $\bar{y}_{N_i}$, the mean for the i-th stratum. If N_i is so large that $\frac{N_i}{N_i-1} \cong 1$, $V(\bar{y}_{st})$ can be written as

$$V(\bar{y}_{st}) = \sum_{i=1}^{k} \frac{W_i^2\sigma_i^2}{n_i} - \frac{1}{N}\sum_{i=1}^{k} W_i\sigma_i^2 . \tag{1.2}$$

If the total sample size n is fixed in advance, the classical problem of allocation of sample sizes in stratified sampling is to determine a vector $(n_1,n_2,\ldots,n_k)$ of k non-negative integers such that $\sum_{i=1}^{k} n_i = n$ and for which $V(\bar{y}_{st})$ is minimum. The allocation so determined, commonly known as Neyman allocation (Neyman, 1934) is given by

$$n_i = nW_i\sigma_i / \sum_{i=1}^{k} W_i\sigma_i . \tag{1.3}$$

Neyman allocation however depends on strata variances σ_i^2 which are generally not known. One way out of this difficulty (Sukhatme, P. V., 1935) is to draw an initial sample of fixed size m from each stratum to estimate σ_i^2 which in turn are used to estimate n_i from (1.3). In this case, n_i is estimated by

$$n_i = nW_is_i / \sum_{i=1}^{k} W_is_i \tag{1.4}$$

where s_i^2 is an unbiased estimate of σ_i^2. The allocation (1.4) will be called Modified Neyman allocation.

Another allocation which is frequently used in practice and does not require knowledge of strata variances σ_i^2 is proportional allocation in which n_i is given by

$$n_i = nW_i . \tag{1.5}$$

If the strata variances σ_i^2 do not differ significantly among themselves, modified Neyman allocation may turn out to be less efficient than proportional allocation (Evans, 1951).

Before deciding on the method of allocation, it is therefore proposed to carry out a preliminary test of significance concerning the homogeneity of strata variances. If on the basis of the test of significance the strata variances are found to be homogeneous, the sample sizes to be drawn from the different strata will be determined according to proportional allocation. This allocation based on preliminary test of significance will be called 'sometimes proportional allocation'. In an earlier paper (Sukhatme, B.V. and Tang, 1970), some results concerning the efficiency of sometimes proportional allocation with respect to proportional allocation and modified Neyman allocation were presented for the relatively simple case of two strata when $\sigma_1^2 \le \sigma_2^2$. In this paper, some further results are presented concerning the efficiency of sometimes proportional allocation and optimum choice of level of significance for the case of two strata when $\sigma_1^2 \ne \sigma_2^2$. The results for three or more strata will be presented at a later date.

2. Variance of $\bar{y}_{st}$ under sometimes proportional allocation

The sometimes proportional allocation is defined as

$$\begin{aligned} n_i &= nW_i \qquad \text{if } \frac{s_j^2}{s_i^2} < \lambda \quad \text{for } i,j = 1,2 \text{ and } i \ne j \\ &= n\frac{W_is_i}{\sum_{i=1}^{2} W_is_i} \qquad \text{otherwise,} \end{aligned} \tag{2.1}$$

where λ is a known constant. Let the event A_0' be defined by $\{\frac{s_j^2}{s_i^2} < \lambda \text{ for } i,j = 1,2 \text{ and } i \ne j\}$ and A_1' be the complementary event of A_0'. The variance of $\bar{y}_{st}$ is given by

$$V(\bar{y}_{st})_S = \sum_{i=0}^{1} E_i\{V(\bar{y}_{st}|A_i')\}P(A_i'), \tag{2.2}$$

where E_i denotes that the expectations are taken with reference to the set A_i' and S stands for sometimes proportional allocation. To evaluate the variance, it will be assumed that $\frac{(m-1)s_i^2}{\sigma_i^2}$ is approximately distributed as chi-square with $t = m-1$ degrees of freedom. It can be seen that

$$V(\bar{y}_{st})_S = \frac{\sigma_1^2}{n}(W_1^2 + W_2^2\theta_{21}) - \frac{\sigma_1^2}{N}(W_1 + W_2\theta_{21})$$
$$+ \frac{W_1W_2}{n}\sigma_1^2[(1+\theta_{21})\{I_{p_{21}^{(1)}}(\tfrac{t}{2}, \tfrac{t}{2})$$
$$- I_{p_{21}}(\tfrac{t}{2}, \tfrac{t}{2})\} + G\theta_{21}^{1/2}\{I_{q_{21}^{(1)}}(\tfrac{t-1}{2}, \tfrac{t-1}{2})$$
$$+ I_{p_{21}}(\tfrac{t-1}{2}, \tfrac{t-1}{2})\}] , \qquad (2.3)$$

where $\theta_{21} = \frac{\sigma_2^2}{\sigma_1^2}$, $p_{21} = \frac{\theta_{21}}{\lambda + \theta_{21}}$,

$$q_{21}^{(1)} = 1 - p_{21}^{(1)} = \frac{1}{1 + \lambda\theta_{21}} ,$$

$G = \frac{2\Gamma(\frac{t-1}{2})}{\{\Gamma(\frac{t}{2})\}^2}$ and $I_{\cdot}(.,.)$ is the incomplete beta distribution.

If we let $\lambda \to \infty$, we obtain the variance of $\bar{y}_{st}$ under proportional allocation, namely,

$$V(\bar{y}_{st})_P = (\tfrac{1}{n} - \tfrac{1}{N})\sigma_1^2(W_1 + W_2\theta_{21}) , \qquad (2.4)$$

where P stands for proportional allocation.

If we put $\lambda = 1$, we get the variance of $\bar{y}_{st}$ under modified Neyman allocation, namely,

$$V(\bar{y}_{st})_{\underline{N}} = \frac{\sigma_1^2}{n}(W_1^2 + W_2^2\theta_{21}) - \frac{\sigma_1^2}{N}(W_1 + W_2\theta_{21})$$
$$+ \frac{W_1W_2}{n}\sigma_1^2 G\theta_{21}^{1/2} , \qquad (2.5)$$

where $\underline{N}$ stands for modified Neyman allocation.

3. Efficiency of sometimes proportional allocation

We shall first discuss the relative efficiency of sometimes proportional allocation with respect to proportional allocation. If $e_1^*(\lambda, \theta_{21})$ denotes the relative efficiency of sometimes proportional allocation with respect to proportional allocation, it is easy to see that

$$e^*(\lambda, \theta_{21}) = \frac{V(\bar{y}_W)_P}{V(\bar{y}_{st})_S}$$

$$= 1/\Big\{1 - \frac{W_1W_2}{W_1 + W_2\theta_{21}}[(1+\theta_{21})\{I_0(p_{21}) \qquad (3.1)$$
$$+ I_0(q_{21}^{(1)})\} - G\theta_{21}^{1/2}\{I_{-\frac{1}{2}}(p_{21}) + I_{-\frac{1}{2}}(q_{21}^{(1)})\}]\Big\} ,$$

where $I_i(\alpha) = I_\alpha(\frac{t}{2} + i, \frac{t}{2} + i)$. Clearly, if $e_1^*(\lambda, \theta_{21}) \geq 1$, sometimes proportional allocation is at least as efficient as proportional allocation. We shall now present some results concerning the behavior of the efficiency function $e_1^*(\lambda, \theta_{21})$.

We shall first consider the case when λ is an arbitrary but fixed number. Then it can be seen that for any given $\lambda \geq 1$,

i) $\lim_{\theta_{21} \to 0} e_1^*(\lambda, \theta_{21}) > 1$,

ii) $\lim_{\theta_{21} \to 1} e_1^*(\lambda, \theta_{21}) \leq 1$,

iii) $\frac{\delta}{\delta\theta_{21}} e_1^*(\lambda, \theta_{21}) < 0$ for $0 < \theta_{21} < 1$

iv) $\exists\ \theta' > 1$ such that $\frac{\delta}{\delta\theta_{21}} e_1^*(\lambda, \theta_{21}) > 0$ for every $\theta_{21} > \theta'$ provided

$$1 - \frac{G}{2} + \frac{(\lambda^{t-\frac{1}{2}} + 1)(\lambda^{\frac{1}{2}} - 1)}{\lambda^{\frac{t}{2}} B(\frac{t}{2}, \frac{t}{2})} > 0 ,$$

and

v) $\lim_{\theta_{21} \to \infty} e_1^*(\lambda, \theta_{21}) > 1$.

As a consequence of the above, we obtain the following result.

<u>Theorem 3.1</u> Let $\lambda \geq 1$ be an arbitrary but fixed number such that

$$1 - \frac{G}{2} + \frac{(\lambda^{t-\frac{1}{2}} + 1)(\lambda^{\frac{1}{2}} - 1)}{\lambda^{\frac{t}{2}} B(\frac{t}{2}, \frac{t}{2})} > 0 .$$

Then $\exists\ \theta_0^{(1)}$ in $(0, 1)$ and $\theta_0^{(2)} > 1$ such that

$$e_1^*(\lambda, \theta_0^{(1)}) = e_1^*(\lambda, \theta_0^{(2)}) = 1$$

and

$$e_1^*(\lambda, \theta_{21}) > 1 \quad \forall \quad \theta_{21} < \theta_0^{(1)} \text{ or } \theta_{21} > \theta_0^{(2)} .$$

Theorem 3.1 assures us that there exist $\theta_0^{(1)}$ between 0 and 1 and $\theta_0^{(2)}$ larger than 1 such that for each $\theta_{21} < \theta_0^{(1)}$ or $\theta_{21} > \theta_0^{(2)}$, sometimes proportional allocation is always more efficient than proportional allocation.

We shall now consider the case when Θ_{21} is an arbitrary but fixed number less than $\frac{1}{2}(G^2 - 2 - G\sqrt{G^2 - 4})$ or larger than $\frac{1}{2}(G^2 - 2 + G\sqrt{G^2 - 4})$. Then it is easy to see that $e_1^*(0, \Theta_{21}) > 1$. Further, $e_1^*(\lambda, \Theta_{21})$ tends to 1 as its horizontal asymptote from below. It is clear that there exists λ_0 such that $e_1^*(\lambda, \Theta_{21}) > 1$ for every $\lambda < \lambda_0$. We have thus proved the following result.

Theorem 3.2 Let Θ_{21} be an arbitrary but fixed number less than $\frac{1}{2}(G^2 - 2 - G\sqrt{G^2 - 4})$ or larger than $\frac{1}{2}(G^2 - 2 + G\sqrt{G^2 - 4})$. Then $\exists\ \lambda_0$ such that

$$e_1^*(\lambda_0, \Theta_{21}) = 1$$

and

$$e_1^*(\lambda, \Theta_{21}) > 1 \quad \text{for every} \quad \lambda < \lambda_0 .$$

After having obtained the above results, it is now possible to prove the existence of a pair of numbers $(\lambda_1^*, \lambda_2^*)$ with $\lambda_1^* \leq \lambda_2^*$ such that for each λ outside the interval $(\lambda_1^*, \lambda_2^*)$ the relative efficiency of sometimes proportional allocation with respect to proportional allocation is never less than a preassigned value $e_0 < 1$. The result is stated in Theorem 3.3 without proof.

Theorem 3.3 Let e_0 be a real number such that $0 < e_0 < 1$. Then $\exists\ \lambda_1^* \leq \lambda_2^*$ such that $e_1^*(\lambda, \Theta_{21}) \geq e_0$ for every λ outside the interval $(\lambda_1^*, \lambda_2^*)$.

We shall now discuss the relative efficiency of sometimes proportional allocation with respect to modified Neyman allocation which is given by

$$e_2^*(\lambda, \Theta_{21}) = \frac{V(\bar{y}_{st})_{\underline{N}}}{V(\bar{y}_{st})_S}$$

$$\doteq 1/[1 - \frac{W_1W_2D^*}{W_1^2 + W_2^2\Theta_{21} + W_1W_2G\Theta_{21}^{1/2}}] ,$$

where

$$D^* = G\Theta_{21}^{1/2}\{I_{-\frac{1}{2}}(p_{21}^{(1)}) - I_{-\frac{1}{2}}(p_{21})\}$$

$$- (1 + \Theta_{21})\{I_0(p_{21}^{(1)}) - I_{-\frac{1}{2}}(p_{21})\} .$$

The results concerning the behavior of $e_2^*(\lambda, \Theta_{21})$ can be obtained in a similar manner and are stated below.

Theorem 3.4 Let $\lambda > 1$ be an arbitrary but fixed number. $\exists\ \Theta_0^{(1)}$ in (0, 1) and $\Theta_0^{(2)} > 1$ such that

$$e_2^*(\lambda, \Theta_0^{(1)}) = e_2^*(\lambda, \Theta_0^{(2)}) = 1$$

and

$$e_2^*(\lambda, \Theta_{21}) \geq 1 \quad \text{for every} \quad \Theta_0^{(1)} \leq \Theta_{21} \leq \Theta_0^{(2)} .$$

Theorem 3.5 Let Θ_{21} be an arbitrary but fixed number such that $\frac{1}{2}(G^2 - 2 - G\sqrt{G^2 - 4}) \leq \Theta_{21} \leq \frac{1}{2}(G^2 - 2 + G\sqrt{G^2 - 4})$. Then $\exists\ \lambda_0$ such that

$$e_2^*(\lambda_0, \Theta_{21}) = 1$$

and

$$e_2^*(\lambda, \Theta_{21}) \geq 1 \quad \text{for every} \quad \lambda \geq \lambda_0 .$$

Theorem 3.6 Let e_0 be a real number such that $0 < e_0 < 1$. Then $\exists\ \lambda_1^* \leq \lambda_2^*$ such that $e_2^*(\lambda, \Theta_{21}) \geq e_0$ for every λ outside the interval $(\lambda_1^*, \lambda_2^*)$.

4. Optimum choice of the level of significance of the preliminary test

As we have seen in Section 3, the relative efficiency of sometimes proportional allocation with respect to proportional allocation as also modified Neyman allocation depends on W_1, Θ_{21} and λ. Generally W_1 is known while Θ_{21} is not known. The question naturally arises concerning the choice of the level of significance as determined by λ. We would like to choose that value of λ for which the relative efficiency of sometimes proportional allocation with respect to either of the other two allocations is as high as possible. For example, if Θ_{21} is likely to be very much different from 1, it would seem desirable to choose λ as small as possible. If on the other hand, Θ_{21} is likely to be closer to 1, it would seem desirable to choose λ as large as possible. If however, nothing is known concerning the likely range of values of Θ_{21}, difficulty arises concerning the choice of λ. Theorems 3.3 and 3.6 provide useful results from this point of view. Let e_0 be a real number such that $0 < e_0 < 1$. Then we shall restrict our choice to those values of λ for which the relative efficiency of sometimes proportional allocation with respect to proportional allocation or modified Neyman allocation is at least e_0. Using this criterion, the following sets of values of λ are obtained for different values of m.

Within a particular set of values of λ, we shall choose that value of λ for which gain in efficiency of sometimes proportional allocation with respect to either of the other two allocations is maximum. Using this criterion, certain values for λ have been suggested in the last col-

Table 1

Sets of values of λ for which the relative efficiency of sometimes proportional allocation with respect to either of the other two allocations is at least e_0

m	e_0	Set of values of λ	Suggested value of λ
6	0.96	$3.3 \leq \lambda \leq 4.0$	3.3
7	0.97	$3.2 \leq \lambda \leq 3.7$	3.2
8	0.98	3.1	3.1
9	0.98	$2.6 \leq \lambda \leq 3.0$	2.6

Table 2

Relative efficiency of sometimes proportional allocation

		θ_{21}								
λ	with respect to	0.4	0.7	1.0	3.0	5.0	7.0	9.0	11.0	13.0
2.6	Proportional allocation	1.003	0.984	0.982	1.017	1.076	1.131	1.178	1.218	1.254
	Modified Neyman allocation	0.992	1.004	1.009	0.990	0.990	0.993	0.995	0.997	0.998
3.1	Proportional allocation	0.998	0.983	0.981	1.009	1.065	1.119	1.166	1.208	1.245
	Modified Neyman allocation	0.991	1.008	1.012	0.986	0.983	0.986	0.990	0.992	0.994
3.2	Proportional allocation	0.993	0.978	0.976	1.004	1.058	1.111	1.158	1.198	1.235
	Modified Neyman allocation	0.992	1.008	1.012	0.989	0.983	0.985	0.980	0.990	0.992
3.3	Proportional allocation	0.986	0.969	0.967	0.998	1.049	1.100	1.145	1.186	1.222
	Modified Neyman allocation	0.995	1.009	1.013	0.990	0.983	0.984	0.986	0.988	0.990

umn of Table 1. Table 2 gives the relative efficiency averaged over W_1 of sometimes proportional allocation with respect to proportional allocation as also with respect to modified Neyman allocation for suggested values of λ over a wide range of values of θ_{21}.

It is seen that sometimes proportional allocation is almost as efficient as modified Neyman allocation. It is also seen that sometimes proportional allocation is almost as efficient as proportional allocation for values of θ_{21} close to 1 while it is considerably more efficient than proportional allocation for values of θ_{21} away from 1.

Acknowledgement

This work was supported in part under OEC Contract No. 0-70-3655 by the U. S. Office of Education, Department of Health, Education and Welfare.

References

[1] Evans, W. D. (1934). "On stratification and optimum allocation," Journal of the American Statistical Association, 46, 95-104.

[2] Neyman, J. (1934). "On the two different aspects of the representative method: The method of stratified sampling and the method of purposive selection," Journal of the Royal Statistical Society, 97, 558-606.

[3] Sukhatme, B. V. and Tang, Victor K. T. (1970). "Allocation in stratified sampling based on preliminary test of significance," American Statistical Association Proceedings of the Social Statistics Section, 368-371.

[4] Sukhatme, P. V. (1935). "Contribution to the theory of the representative method," Journal of the Royal Statistical Society, Supplement 2, 253-268.

SOME COMPLIANCE TESTS FOR A UNIVERSITY SEGMENT

Henry Thomassen, University of Nebraska-Lincoln

If an individual is to improve himself, it is usually supposed that his environment must first be adjusted. Yet, when looking at the American university, people call for self-reform within the context of existing conditions. Is the university then free of environmental shackles and capable of rebirth at will? The findings of the study highlighted here suggest it is not. Instead, the university, or at least its school of business, seems to respond to the institutional and economic environment in regular fashion. So regular is the response, in fact, that the "compliance" of a single university segment with the field it faces may be tested against the behavioral background of a national sample.

Of course, the university is a complex organization with many behavioral facets. The one chosen for emphasis here is the offering of instructional topics in various subject fields. This university output is thought separable from other outputs like library services, academic atmosphere, and hypothesis formulation and testing, and is judged distinct from the community's utilization of instructional offerings in the productive process called higher education. In addition, this particular output has the advantage of being quantifiable in dimensions of diversity, dominance, and depth. Diversity here registers the number of topics in a series of subject fields which a school stands ready to "teach" to qualified members of society. Included are units like principles of economics, theory of matrices, and regression analysis. Dominance, on the other hand, measures the diversity of offerings in one subject field (like marketing, mathematics, or accounting) relative to those of the school as a whole. Lastly, depth registers the relative frequency of topic sequences of two or three-or-more units linked through a strict system of prerequisites.

To locate the variables in the environment upon which a school's diversity, dominance, and depth depend, one must fish in the environment with a theory. The theory adopted holds that the university's output is a social good. Offerings of instruction satisfy many individual preferences simultaneously and one person's enjoyment (as opposed to his utilization) of the output in no way reduces the amount available for another. So seen, no individual is apt to reveal his true preferences for the university's output nor to be willing to pay as much for the offer of instruction as he would if its availability to him depended upon his sacrifice of resources alone. By the same token, a university can count on no special rewards for carefully adjusting its output to suit the implicit preferences of its consumers. Thus, as the consumer is content to pay less than his maximum for the presence of a university's output which is not ideal, so the university is satisfied and permitted to produce an output that caters partially but not fully to the populace which supports it. A wedge is then driven into the chain of accountability between user and producer such that the university wins some slack for pursuing objectives of its own choosing. Admittedly, supporting taxpayers and donors may require certain output characteristics before an exchange between them and the university becomes possible at all. Perhaps the university must equip itself with fringes of research or athletics, must offer instruction in American Free Enterprise or in Black Capitalism, or must exclude Agricultural Price supports from discussions of farm policy. Nevertheless, with such basics satisfied, the responsiveness of the university will tend to be loose rather than rigid and categorical rather than specific.

What this theory says about the university's reaction to environment is easily summarized. Output will be sensitive to forces within the school as well as without it and the long list of conventional influences bearing on preferences of individuals and groups and on public goods exchanges become relevant for consideration.

Having settled on a representative of output for the university and a direction for locating its controls in the environment, the theme of response must next be confronted by reality. To do this, a random sample of thirty-nine schools has been drawn from the 1970 membership of The American Association of Collegiate Schools of Business. Of the thirty-nine, twenty-four prove to be public, nine are graduate, thirteen are graduate-undergraduate, and seventeen are undergraduate. Furthermore, twenty-three are linked to universities with academic rankings in the nation's upper ninety-eight as measured on the Cartter scale.

For each of the schools of the sample, catalogs give information sufficient to measure diversity, dominance, and depth of offerings. In this quantification of output, departmental lines are ignored and course descriptions are used to put offerings within one of ten conventional subject fields. These are accounting, marketing, management, finance, real estate and insurance, general business, mathematics, computer science, statistics, and economics. Of note, perhaps, is the fact that general business here treats the interaction of business with forces of demography, culture, politics, technology, organization, communication, and natural environment in social feedback systems; it clearly does not refer to the catch-all property of some misfit departments and courses. As might be expected, the three measures of output possess high coefficients of variation. Indeed, these are the coefficients which the theme of this study would explain by differences in setting.

It would take too long and be too tedious to relive the drama of selecting proxies and testing them for service in the theme. Suffice it to say that the method of selection has been that of step-wise regression applied to those forces inside and outside the university which possess properties of "mental fit". Using the F-test, only relationships significant at the 0.05 level (or better) have been saved and, using the t-test, only variables significant at the same level have been included. While superficial investigation of correlation matrices among inputs has lead to redefinition or rejection of all but one of any subset for which intercorrelation was obvious, no careful analysis of multicollinearity has been judged worthy of attempt.

On the whole, the application of the noted procedure gives results which uphold the hypothesis of university sensitivity. The significant relationships do draw from both inside and outside forces as supposed and numerous (nineteen) forces are required to "explain" the behavior of output. Indeed, the coefficients of determination for the response patterns, mainly in the 0.50 to 0.70 range, give room for additional as well as more appropriate inputs.

Coming to specifics, diversity and dominance prove to be much more sensitive to conditions than does depth. Diversity rises with the size of the total university faculty, the population of the city wherein the school is located, the public nature of the institution, and undergraduate emphasis of the college. It falls with the percent of the population in the school's state with a high-school education, a suggestion that an informed and interested group of supporters increases pressures for university accountability and curbs proliferation.

Among the ten subject fields, dominance responses are significant in all except economics and statistics. For finance and computer science, only external forces matter. Finance offerings grow in relative frequency with total personal income while those in computer science rise with employment in defense industries and with the percent of employment in goods-producing industries in the school's home state. Offerings in finance are at the same time restricted by the home state's expenditures for education per one-thousand dollars of personal income. In so far as the positive forces are concerned, concentrations of particular interest groups are evidently swamping the effects of field proliferation that wider and more affluent consumer audiences bring. Simultaneously, negative influences are reflecting the pressure for accountability that comes with public awareness. In the fields of general business and mathematics, internal forces alone rule. Of major account are positive effects tying to a school's academic rank (Cartter scale) and to its faculty quality (represented by salary levels). The negative influences relate to field diversification that comes with growth in the size of the school in terms of both students and faculty. In the four remaining subject fields, dominance responds both to inside and outside forces in generally interpretable fashion. Of special concern, nevertheless, is the negative influence which faculty quality has on offerings in real estate and insurance. Are such offerings gratifications which faculties take for themselves when monetary rewards are inadequate? If so, the faculty quality that has been associated with compensation levels must include activation of a strict perspective on higher education in business.

Among the measures of depth, only three-or-more sequences are "explained" by environment. In this case, the relative size of the business school within the university acts positively while the quality of the faculty acts negatively.

To test a single school's compliance with its environment, the relationships developed from the sample can be applied to the relevant measure of the school's own internal and external circumstance. This has been done for the business school of the University of Nebraska-Lincoln. Though Nebraska's environmentally-conditioned response is mainly in tune with that of the sample, its offerings in both finance and general business appear to be "under-dominant". In the case of finance, this might suggest that Nebraskans have an insufficient concentration of interest in financial aspects of business and government to swamp the diversification of fields that their personal income brings or that their University is made especially sensitive to public wishes by the quality of expenditures on education per thousand dollars of income. Furthermore, in the case of general business, one could say that the University's academic rank is really not as high as the Cartter ratings suggest (where it stands eighty-second) or that the salaries paid overstate the quality of the Nebraska faculty. However, recalling the speaker's affiliation, the proper interpretation of Nebraska's non-compliance in these two instances is left as an exercise for the listener, and hopefully also, for the reader of the complete study. [1]

REFERENCES

[1] Thomassen, H. (1972). Some compliance tests for a university segment. Complete paper of which the foregoing is a synopsis.

EXPECTED, DESIRED AND CUMULATIVE FAMILY SIZE

Robert H. Weller, Florida State University

Measuring a woman's fertility performance is made much more complicated than a mere counting operation by the multidimensional nature of this performance. The present paper represents an examination of three dimensions of fertility performance and the manner in which they are interrelated as well as the extent to which they are associated with such demographic variables as current age, duration of marriage, and age at marriage. These three aspects of fertility behavior--which certainly do not exhaust the universe of its dimensions--are cumulative family size, expected completed family size, and desired family size.

The first represents the number of children a woman has borne by a specific point in time. The customary measure of cumulative family size is the number of children ever born. Although the determination of this number is a relatively straightforward matter, the use of cumulative family size presents a problem when the investigator wishes to analyze groups of women who have not completed their childbearing. It reflects differences in current age and may also contain the effects of differences in age at marriage, the duration of marriage, and differences in timing and spacing of births. Because of the desirability of having at least rough controls for these factors, the researcher who analyzes the differences in cumulative family size frequently finds the cell frequences declining at an alarming rate as additional controls are introduced. Often the investigator must choose between controlling for extrademographic variables believed to be crucial in their independent explanatory power and controlling for demographic variables whose intervening influence is certain to be felt in the cumulative family size measure, e.g., duration of marriage, current age, and age at marriage.

One method of avoiding this particular pitfall when studying fertility behavior is to use expected completed family size. This involves several assumptions. One is that expected completed family size is a realistic estimate of the number of children a woman will have borne by the time she completes childbearing. Although this is often not the case with individual women, it is a fairly realistic assumption for aggregates of women (Westoff, Mishler and Kelly, 1957; Freedman, Whelpton and Campbell, 1959:218-219; Goldberg, Sharp and Freedman, 1959:376; and Whelpton, Campbell and Patterson, 1966:29).

Expected completed family size may be determined in several ways. One way is to determine the number of live births a woman already has experienced and add to this figure the number of additional children she expects to have. In the present study expected completed family size was determined through use of the question: "Altogether, how many children do you expect to have (from this present marriage)? Please count those you already have plus any others you expect to have in the future." Those respondents who hesitated or who did not respond in numerical terms were given the additional instruction: "Well, just give me your best guess as to how many children you will have, in all."

The third dimension of fertility behavior considered in the present paper is desired family size. This probably is the most difficult dimension of fertility behavior to measure unambiguously, for in large part the response reflects the wording of the question used as well as any conditions implied or explicitly stated by the question. The question utilized in the present study is quasi-retrospective as well as hypothetical. The respondent was asked what her desired family size would be if she were just getting married. The question is retrospective in that presumably a woman's response reflects her own experiences (and those of her spouse) since she married (cf. Freedman, Coombs and Bumpass, 1965, for a demonstration of the extent to which family size desires change over the course of marriage in response to economic and social contingencies as well as in response to fertility experience and fecundity problems). The question is hypothetical in that it asks the respondent to imagine she were just getting married. Moreover, there is no way of determining what, if any, conditions and subsequent developments the respondent associates with this imagined union.

Source of Data

Data are taken from two representative samples of the population of the state of Rhode Island. The sampling procedure employed was a multi-stage, area probability sample of households. The data utilized in this study were obtained in the falls of 1968 and 1969 under the auspices of the Population Research Laboratory at Brown University, with a wide variety of information being collected with respect to general medical behavior, demographic behavior, and socioeconomic characteristics. A detailed discussion of the sample's design and the objectives and scope of the project may be found in Bouvier (1971) and Organic and Goldstein (1970).

The following analyses are limited to currently married females ages 20 through 49 who have been married only once, who have not been sterilized, and who know of no physiological reason they cannot have children. Even though they were obtained separately, the 1968 and 1969 samples have been combined and treated as one sample in order to increase the number of cases in the various cells. Organic and Goldstein (1970) have reported that the sampling technique utilized makes such a cumulation valid.

Results

The three measures of family size are presented in Table 1. With the exception of a difference in location between cumulative family size and the other two aspects of fertility behavior, the overall picture is that of similarity between the three distributions.

In Table 2 it can be seen that only cumulative family size is even weakly related to the current age of the woman. This relationship is not linear, but reverses its slope somewhat after age 35-39. This suggests that cohort changes have been occurring in the cumulative family size of Rhode Island married females that are not fully evident in the cross-sectional data at our disposal.

A similar situation is present when duration of marriage is considered. As duration of marriage increases, cumulative family size also increases. However, neither expected nor desired family size is related to the length of time the female has been married. A third demographic variable that can affect family size is age at marriage. It can be seen in Table 2 that the older the age at marriage the lower the number of children ever born. However, expected family size does not decrease with age at marriage, nor does the number of children desired. An excepto this occurs among those females who were age 26 or older when they married. These women desire, and expect to have, substantially less children than the other females.

The relationship between expected completed family size and the number of children ever born is shown in Table 3. There is some misreporting, with about 2 percent of the respondents expecting a smaller family size than they already have. About 55 percent of the women report identical scores on the two measures, and about 43 percent of the women expect to have more children than they already have. Except in the case of childless women, expected and cumulative family size increase together.

When expected and cumulative family size are compared, the outcome is strongly associated with the female's current age (Table 4). The older the women the greater the probability that the expected and cumulative family size measures coincide. This is also true of duration of marriage. This gives additional evidence to the notion of expected completed family size as a valid substitute for the number of children a female will have borne before completing childbearing. Age at marriage appears unrelated to the outcome of the comparison between expected completed and cumulative family size. Although the probability of agreement between expected and cumulative family size increases with parity, the outcome of the comparison is not clearly related to either expected or desired family size.

It can be seen in Table 5 that desired and cumulative family size tend to increase as the other increases. How much of this tendency is due to the influence of past fertility behavior upon family size desires and/or vice versa are not known. Only a longitudinal research design would be appropriate to deal adequately with issues like that. However, over one half of the women report a desired family size larger than their number of children ever born. About one seventh of the women already have a cumulative family size that is larger than the desired family size measure. This suggests that desired and cumulative family size may be relatively independent. When desired and cumulative family size are compared, the outcome of the comparison is not consistently related to any of the six demographic variables utilized in this study (Table 6). However, the older the female the greater the probability she will have a cumulative family size larger than desired family size and the smaller the probability of desiring more than the number already born. The same types of relationship exist with respect to duration of marriage and the number of children ever born.

When the desired and expected family size measures are compared, about 58 percent of the women expect to have the same number of children they would desire to have if they were just getting married. Slightly more than one sixth expect a greater number, and the remaining 23 percent expect less. These last two categories of fertility behavior are of great heuristic value because they are indicators of the extent to which females do not have the number of children they desire.

The first category (those expecting to have more children than desired) has been classified elsewhere as "excess fertility" (Ryder and Westoff, 1971; and Weller and Chi, 1972). The category of women expecting less children than they would desire has been termed "deficit fertility" (Weller and Chi, 1972). The proportions observed in Table 7 are fairly close to those obtained in the second GAF study (Whelpton, Campbell and Patterson, 1966:52-53) and the 1965 National Fertility Study (Ryder and Westoff, 1971: 74).

The demographic variable most related to the outcome of the comparison between expected and desired family size is cumulative family size. As the number of children ever born increases, the probability of excess fertility increases (Table 8). The probability of deficit fertility decreases with parity, although not as regularly as excess fertility increases. The probability of excess fertility is negatively related to age at marriage, but bears little relationship to duration of marriage. An exception to the latter statement is that women married less than 5 years are clearly differentiated from the other women. Given the cross-sectional design of the study, it is not known to what extent these women's behavior will gradually resemble the behavior of the other women who have had greater exposure to the marital experience. The number of children desired and the total expected completed family size seemingly are not associated with the outcome of the comparison between desired and expected completed family size.

Conclusions

The extent to which different scores on cumulative, desired, and expected completed family size are obtained for the same women suggests that these are relatively independent dimensions of fertility behavior--at least at the time of the interview. Of course this is less true of expected completed and cumulative family size than of the two other possible relationships. With several notable exceptions, the associations between the measure of each of these dimensions of fertility behavior and the demographic variables of current age, age at marriage, and duration of marriage are not as strong as one would expect on a *priori* basis. This is particularly the case with the expected completed family size and desired family size. This suggests that perhaps it is not really necessary to control for all of these factors when analyzing the effects of extra-demographic variables upon fertility behavior.

At the same time, this investigator publicly wonders if the same results (and non-results) would have been obtained if the data had been collected and studied longitudinally rather than cross-sectionally. To understand the real relationship between these factors, it may be necessary to select a representative sample of women at age n and periodically collect measures of the relevant variables from these women as they progress through the life cycle. This has been tried on a small scale, but not with an entire cohort of women who are representative of the na- population of women. Of course it would also be necessary to design the study in such a way that an interaction between the process of obtaining repeated scores and the fertility measures themselves would be minimized and measurable.

A second conclusion that may be drawn is that fertility behavior is multi-dimensional rather than unidimensional. Hence it may be over-simplistic to write or speak about the relationship between variable X and fertility. Rather, it may prove more accurate to specify which specific aspects of fertility behavior are under consideration.

Table 1
Moments of Measures of Fertility Dimensions

	Children Ever Born	Expected Family Size	Desired Family Size
Mean [a]	2.5	3.3	3.3
Median	2.0	3.0	3.0
Standard Deviation	1.6	1.4	1.4
Skewness [b]	.9	.6	.6
Kurtosis [c]	2.9	3.0	3.1
Number of Cases	396[d]	396[d]	392

a. For computational ease, all cases in the category 6+ have been assigned the value 6.5. There were 11 or less for each of the variables.
b. The measures of skewness utilized is Pearson's (modified) coefficient of skewness. For discussion of this measure, see Yeomans (1968): 114-118.
c. The amount of kurtosis in a series is measured by the fourth moment around its mean divided by the fourth power of the standard deviation, i.e., $\frac{\Sigma (Xi - \bar{X})^4}{NS^4}$
Any value greater than 3 indicates a leptokurtic distribution, while one of less than 3 shows platykurtosis. (Yeomans, 1968: 118-119).
d. Includes 4 women for whom desired family size is not known.

Table 2. Family Size and Selected Demographic Characteristics

	Children Ever Born	Expected Family Size	Desired Family Size
Age of Wife			
20-24	1.0	3.3	3.3
25-29	2.0	3.1	3.0
30-34	2.4	3.3	3.4
35-39	3.5	3.5	3.5
40-44	3.1	3.1	3.5
45-49	2.9	2.9	2.8
No. of Cases*	393	396	387
Duration of Marriage			
0-2	0.4	3.3	3.4
3-4	1.4	3.0	3.2
5-9	2.2	3.2	3.0
10-14	3.3	3.4	3.3
15+	3.3	3.3	3.3
No. of Cases	392	391	386
Age at Marriage (Wife)			
≤17	3.3	3.4	3.2
18-19	2.5	3.2	3.1
20-21	2.4	3.4	3.3
22-23	2.3	3.2	3.3
24-25	2.6	3.3	3.3
≥26	1.5	2.0	2.8
No. of Cases	391	391	385

*In this and subsequent tables, those women who are not classified on one or more of the variables have been excluded. The number of such women can be obtained by subtracting the number of cases from 396, the base number of women in the study group.

Table 3
Expected Family Size by Children Ever Born

Children Ever Born	Expected Family Size 0	1	2	3	4	5	6-9	No. of Cases	$\bar{X}$
0	0	1	15	12	13	5	1	47	3.0
1	0	15	19	15	8	0	1	58	2.3
2	4	0	60	22	17	2	1	106	2.6
3	1	0	3	62	17	3	0	86	3.2
4	1	0	0	1	38	5	3	48	4.2
5	0	0	0	0	2	20	1	23	5.0
6-9	0	0	0	0	0	1	15	16	6.4
No. of Cases*	6	16	97	112	95	36	22	384	
$\bar{X}$	2.5	0.9	1.5	2.2	2.7	3.9	5.3		

%Expected < CEB = 1.8
%Expected = CEB = 55.6
%Expected > CEB = 42.6
*For ease of presentation, those women reporting a number greater than 9 on one of the fertility measures have been excluded from Tables 3,5,7.

Table 5
Desired Family Size and Children Ever Born

Children Ever Born	Desired Family Size 0	1	2	3	4	5	6-9	No. of Cases	$\bar{X}$
0	0	0	17	14	15	4	2	52	3.2
1	4	1	18	22	9	1	2	57	2.8
2	1	1	43	14	36	2	3	100	3.0
3	1	2	13	37	26	2	3	84	3.2
4	2	0	10	1	28	0	4	45	3.6
5	0	0	3	5	3	8	3	22	4.2
6-9	0	1	3	0	4	2	6	16	4.5
No. of Cases	8	5	107	93	121	19	23	376	
$\bar{X}$	2.1	3.1	2.0	2.0	2.6	3.4	3.8		

%Desired < CEB = 14.9
%Desired = CEB = 32.7
%Desired > CEB = 52.4

Table 6
Comparison Between Desired Family Size and Children Ever Born by Selected Demographic Characteristics.

	Desired % < CEB	Desired % = CEB	Desired % > CEB	Total	No. of Cases *
A. Age of Wife					
20-24	2.5	8.9	88.6	100.0	79
25-29	11.4	30.7	57.9	100.0	88
30-34	14.8	45.9	39.3	100.0	61
35-39	34.0	32.0	34.0	100.0	53
40-44	18.3	42.3	39.4	100.0	71
45-49	28.1	43.8	28.1	100.0	32
No. of Cases					
B. Duration of Marriage					
0-2	0.0	3.8	96.2	100.0	53
3-4	2.0	10.0	88.0	100.0	50
5-9	10.3	38.5	51.2	100.0	78
10-14	26.8	40.8	32.4	100.0	71
15+	24.5	43.8	31.7	100.0	139
No. of Cases					
C. Wife's Age at Marriage					
≤ 17	26.5	41.1	32.4	100.0	34
18-19	15.9	37.8	46.3	100.0	82
20-21	17.4	28.8	53.8	100.0	132
22-23	10.4	28.6	61.0	100.0	77
24-25	14.3	45.2	40.5	100.0	42
≥ 26	8.7	13.0	78.3	100.0	23
No. of Cases					390

Table 6 (Cont.)

	D < CEB	D = CEB	D > CEB	Total	No. of Cases
D. Children Ever Born					
0	0.0	0.0	100.0	100.0	52
1	6.9	1.7	91.4	100.0	58
2	2.0	43.1	54.9	100.0	102
3	18.8	43.5	37.6	100.0	85
4	28.3	60.8	10.9	100.0	46
5	54.5	36.4	9.1	100.0	22
6-9	68.7	31.3	0.0	100.0	16
No. of Cases					381
E. Expected Family Size					
0					6
1	25.0	6.3	68.7	100.0	16
2	3.1	36.5	60.4	100.0	96
3	9.8	36.6	53.6	100.0	112
4	16.0	31.9	52.1	100.0	94
5	42.8	22.9	34.3	100.0	35
6-9	45.4	36.4	18.2		22
No. of Cases					381
F. Desired Family Size					
0				100.0	8
1				100.0	5
2	27.1	40.2	32.7	100.0	107
3	6.5	40.9	52.7	100.0	93
4	6.6	23.0	70.5	100.0	122
5	10.5	42.1	47.4	100.0	19
6-9	16.0	20.0	64.0	100.0	25
No. of Cases					379

Table 7
Expected Family Size and Desired Family Size

Number of Children Expected	Desired Family Size 0	1	2	3	4	5	6-9	No. of Cases	$\bar{X}$
0	0	0	4	1	1	0	0	6	2.5
1	4	1	6	3	1	1	0	16	1.9
2	1	0	64	9	24	0	1	99	2.6
3	1	2	15	70	22	1	4	115	3.1
4	2	1	13	6	66	1	4	93	3.0
5	0	0	6	5	5	13	5	34	4.3
6-9	0	1	3	0	5	2	10	21	4.7
No. of Cases	8	5	111	94	124	18	24	384	
$\bar{X}$	2.1	3.5	2.3	3.0	3.6	4.8	5.1		
% Desired < Expected					18.7				
% Desired = Expected					58.4				
% Desired > Expected					22.9				

Table 8
Comparison Between Expected Family Size and Desired Family Size
By Selected Demographic Variables

A. Age of Wife	% Desired < Expected	%Desired =Expected	%Desired >Expected	Total	No. of Cases
20-24	6.2	83.9	9.9	100.0	81
25-29	20.2	58.5	21.3	100.0	89
30-34	22.6	59.7	17.7	100.0	62
35-39	37.0	31.5	31.5	100.0	54
40-44	16.7	45.8	37.5	100.0	72
45-49	26.5	44.1	29.4	100.0	34
No. of Cases					392
B. Duration of Marriage					
0-2 years	5.7	84.9	9.4	100.0	53
3-4	2.0	72.0	26.0	100.0	50
5-9	24.4	60.2	15.4	100.0	78
10-14	29.6	46.5	23.9	100.0	71
15+	24.5	43.8	31.7	100.0	139
No. of Cases					391

C. Wife's Age at Marriage					
≤ 17	32.4	47.0	20.6	100.0	34
18-19	17.1	68.3	14.6	100.0	82
20-21	23.5	55.3	21.2	100.0	132
22-23	18.2	50.6	31.2	100.0	77
24-25	16.7	61.9	21.4	100.0	42
≥ 26	4.4	47.8	47.8	100.0	23
No. of Cases					390
D. Children Ever Born					
0	1.9	76.9	21.2	100.0	52
1	8.5	55.9	35.6	100.0	59
2	9.5	61.9	28.6	100.0	105
3	21.6	54.5	23.9	100.0	88
4	36.2	51.0	12.8	100.0	47
5	50.0	36.4	13.6	100.0	23
6-9	75.0	25.0	0.0	100.0	16
Total Number of Cases					389
E. Expected Family Size					
0					6
1	25.0	6.3	68.8	100.0	16
2	1.0	64.0	35.0	100.0	100
3	15.7	60.8	23.5	100.0	115
4	23.2	69.4	7.4	100.0	95
5	45.7	40.0	14.3	100.0	35
6-9	63.7	31.8	4.5	100.0	22
Total number of cases					389
F. Desired Family Size					
0				100.0	8
1				100.0	5
2	33.3	57.7	9.0	100.0	111
3	11.7	74.5	13.8	100.0	94
4	8.8	52.8	8.4	100.0	125
5	15.8	68.4	15.8	100.0	19
6-9	16.0	32.0	52.0	100.0	25
Total Number of Cases					387

ACKNOWLEDGEMENTS

Portions of this research have been supported by United States Public Health Service Grant HS-00246 from the National Center for Health Services Research and Development awarded to Brown University and by National Science Foundation Grant GS-30975. The author thanks the Population Research Laboratory at Brown University for permission to use the data and George Cannon who performed the feat of converting them for use on the Florida State University computer.

REFERENCES

[1] Bouvier, Leon F. (1971), "The effect of Catholicism on the fertility of Rhode Island women: 1968-1969," Unpublished Ph.D. dissertation in sociology, Brown University.

[2] Freedman, Ronald, Coombs, L. C., and Bumpass Larry (1965), "Stability and change in expected family size: a longitudinal study," Demography 6, 250-275.

[3] Freedman, Ronald, Whelpton, Pascal K., and Campbell, Arthur A. (1959), Family Planning Sterility and Population Growth, McGraw-Hill, New York.

[4] Goldberg, David, Sharp, Harry, and Freedman, Ronald (1959), "The stability and reliability of expected family size data," Milbank Memorial Fund Quarterly 37, 369-385.

[5] Organic, Harold N. and Goldstein, Sidney (1970), "The Brown University Population Research Laboratory: its purposes and initial progress," The Community as an Epidemiologic Laboratory: A Casebook of Community Studies, Kessler, Irving I. and Levin, Morton L., ed., pp. 212-232. John Hopkins Press, Baltimore Maryland.

[6] Ryder, Norman B. and Westoff, Charles F. (1971), Reproduction in the United States 1965, Princeton University Press, Princeton, New Jersey.

[7] Weller, Robert H. and Chi, Peter S. K. (1972), "Excess and deficit fertility and population policy," Paper presented at annual meetings of the Population Association of America.

[8] Westoff, Charles F., Mishler, Elliott G., and Kelly, E. Lowell (1957), "Preferences in size of family and eventual fertility twenty years after," American Journal of Sociology 42, 491-497.

[9] Whelpton, Pascal K., Campbell, Arthur A., and Patterson, John E. (1966), Fertility and Family Planning in the United States, Princeton University Press, Princeton, New Jersey.

[10] Yeomans, K. A. (1968), Applied Statistics: Statistics for the Social Scientist, Penguin Press, Baltimore, Maryland.

Due to limitations of space imposed by the publisher, Table 4 has been omitted. Copies may be obtained by writing to the author, Department of Sociology, Florida State University, Tallahassee, Florida 32306.

MINUTES OF THE ANNUAL MEETING OF THE SOCIAL STATISTICS SECTION

Montreal, Canada, August 15, 1972

The meeting was opened by Eva Mueller, the Chairman, at 5:30 p.m. As a result of the recent election, the list of officers for 1973 is:

Chairman	Theodore D. Woolsey
Chairman-Elect	Charles B. Nam
Vice-Chairman	Daniel G. Horvitz
Vice-Chairman	Denis F. Johnston
Secretary	Regina Loewenstein
Section Representative on the Board of Directors	Evelyn Kitagawa
Section Representative on the Council	Eli S. Marks
Publications Liaison Officer	Monroe Lerner

Edwin D. Goldfield was reappointed Editor of the Proceedings.

Suggested topics for the Annual Meeting in 1973 included: welfare reform and income maintenance, costs and effectiveness of social statistics, social indicators, value of indices of cost-of-living and of unemployment, objective methods of evaluating budgets, income distribution, political polls, heredity and intelligence (Jensen report), surveys of victims of crimes, implications of undercount in 1970 Census, improvements of survey methodology, and joint meeting with Caucus for Women in Statistics about data needed to study problem of equal opportunity.

The Social Statistics Section has been asked to recommend persons to chair and serve on the Committee on Social Indicators of the American Statistical Association. Ten persons were suggested by persons attending the meeting. The work of the Committee will first emphasize conceptualization, then techniques to develop indicators, and last evaluation of proposed social indicators.

The informal sessions in the Annual Meeting that were tried for the first time this year were very poorly attended. Publicity in The American Statistician and improvement of mechanics were suggested for future years.

The contributed papers sessions have had decreased attendance and little discussion of papers. Suggestions for improvement were:

(1) More sessions with fewer papers to allow time for discussions of each paper. But attendance at each session would decrease.

(2) Enforce a regulation of ten minutes of presentation and five minutes of discussion of each paper.

(3) Speakers to be asked to prepare a lengthy paper for publication and a brief paper for presentation that will stimulate discussion.

(4) Revise existing instructions sent to speakers. For example, stress that not more than four or five mathematical expressions should be given in any one presentation.

1972 Officers of the Social Statistics Section

Chairman	Eva Mueller
Chairman-Elect	Theodore D. Woolsey
Vice-Chairman	Charles B. Nam
Vice-Chairman	Sidney Goldstein
Secretary	Regina Loewenstein
Section Representative on the Board of Directors	John D. Durand
Section Representative on the Council	Eli S. Marks
Publications Liaison Officer	Karl E. Taeuber
Editor of Proceedings	Edwin D. Goldfield