I

POPULATION PROJECTIONS FOR SMALL AREAS

Chairman, MEYER ZITTER, U. S. Bureau of the Census

Page

METROPOLITAN AREA POPULATION PROJECTIONS: USE OF REGIONAL AND NATIONAL ECONOMIC FRAMEWORKS

Mannie Kupinsky[1]
National Planning Association

The topic of this session is small area population projections. As the title of this paper indicates, we believe there is an interrelatedness between what happens in a local area and what is going on demographically and economically in the area's region and in the nation. Were we to suggest a procedure very limited in scope we might rightly be charged with a narrowness comparable to that of the farmer in Robert Frost's poem, "Mending Wall", who kept repeating "Good fences make good neighbors".

We view the problem of small area population projections within the local area context and also a broader (regional and national) but still very relevant context. Small area population changes reflect many influences. On the basis of our judgment, it appeared necessary to use both contexts if we were to include the primary influences determining population change.

Despite the complex sounding title, the conceptual key to the projection procedure can be summarized very briefly. Population projection at the national level may most accurately be estimated through assumptions covering future fertility and mortality rates. At subnational levels, the smaller the area the greater the probability that population change can be significantly affected by nondemographic factors. It is the central assumption underlying the population projection described in this paper that the key nondemographic factor is the one of job opportunity.

Thus, if a reasonable estimate of future jobs in an area can be projected, it is assumed that a dependable estimate of population may then be derived. Demographic analysis (fertility and mortality rates) could then follow in order to define population characteristics such as age and sex. Therefore, the major effort, in developing the procedure described in this paper was directed toward determining how best to estimate future economic developments in an area. From such estimates, employment opportunities could be derived and these were to be used as a basis for estimating population.

There are special situations in some local areas where neither the approach suggested in this paper nor the strictly demographic procedure would be completely satisfactory in projecting the population for small areas. Such would be the case where an area attracted old and retired people or where an area served to provide domiciles (often called bedroom areas) for people who work in a neighboring area. In these cases, there might need to be some modifications of the procedure described in this paper.

In the next section of this paper, we outline some of the major assumptions and concepts underlying the projection procedure. That is followed by a description of the procedure, by the projections for several metropolitan areas and by some conclusions.

Major Assumptions

The major assumptions and concepts on which the projection procedure is based are, as follows:

1. The geographical unit to be used in developing the projections should be an integrated economic unit. For that reason, the standard metropolitan statistical area (SMSA) concept as defined by the Bureau of the Budget was used. The Bureau defines the SMSA as a county or group of contiguous counties containing at least one city (i.e., central city) of 50,000 inhabitants or more, or "twin cities" with a combined population of at least 50,000. In addition to the county, or counties, containing the central city or "twin cities", contiguous counties are included in the SMSA if they are essentially "metropolitan" in character and are socially and economically integrated with the central city.[2]

   The Technical Committee on Area Definitions, established by the Bureau of the Budget, reviews the geographical definition of each SMSA at regular intervals to determine whether counties should be added or substracted from existing SMSAs or whether an area not previously defined as an SMSA has become metropolitan in character - according to the criteria established by the Bureau of the Budget.

   The geographical boundaries of an SMSA, therefore, are not fixed. They change as the economy of the area changes in size. Though most Federal agencies maintain statistical series covering SMSAs on a fixed boundary basis (using the most recent set of boundary definitions and revising figures for past years accordingly), the procedure described in this paper does not follow that practice. The procedure rather deals consistently with metropolitan economies. It does this by using as the metropolitan area boundaries, those defined in each year of reference by the Bureau of the Budget. Therefore, historical data for metropolitan areas that are referred to in this paper reflect in each year the actual size of an area's economy (as defined by the Bureau of the Budget). As a result, geographical boundaries may differ from one year of reference to another. Projected data also reflect the actual size of an area's economy. We, however, have not defined the geographical boundaries for the projected years.

2. A set of national projections are needed if subnational projections of a realistic and reasonable character are to be developed. The national projections provide the framework which reflect explicit assumptions on patterns of consumption, technological and productivity trends, government and business expenditures, and assumptions on the kinds of government policies that may be forthcoming over the projected period.

3. The subnational projections need to be based on an assessment of the comparative advantages that each state or metropolitan area has over other states or areas. States are singled out here because, in the projection procedure to be described, single or combinations of states form the market or production region within which metropolitan area economic growth is gauged. The analytical regions are discussed in section (5) below. The state (regional) projections are consistent with and add up to the national projections (mentioned above in (2)). The projected variables within a state are consistent with one another, that is, they reflect a working out of probable comparative advantages. These state projections are primarily determined by industry employment trends which are influenced by specific industry changes in national industry employment and differential geographic shifts in industry employment. The state projections are, then, disaggregations of the national projections.

4. The metropolitan area projections are worked out within a market or production region. To accomplish this we use the state projections mentioned in (3) above. In moving from states to metropolitan areas we find that the metropolitan area projections are very sensitive to the geographical boundaries that establish the economic region to which the metropolitan area is related. The objective is to define a region in such a way that the major transactions of the metropolitan area being examined are included within the region's boundaries. This means that the region associated with the metropolitan area for which projections are being made may also include other metropolitan areas. Each of these may have a different region associated with it. This implies that the region defines the boundaries of the geographic area which is significant for the metropolitan area and not the relationship of the region to other regions or to the nation as a whole. The relevance of this is that the procedure for deriving analytical regions is not the result of a disaggregation of national totals but rather an aggregation into regions from core units (which for practical reasons are contiguous states).

In defining analytical regions we have associated a unique region with each metropolitan area. Ideally, we would have liked to identify separate regions for each industry in the metropolitan area.

Before describing our criteria for defining a region it is necessary to distinguish among three types of industries in the metropolitan area.

1) National industries. Industries which sell not only within a market defined by the boundaries of the region containing the metropolitan area, but also make substantial sales beyond these boundaries. These will usually be commodity industries.

2) Multi-state industries. Industries which sell primarily within a market defined by the boundaries of the region containing the metropolitan area with a more significant proportion of these sales being outside of the metropolitan area than within the area. These are also usually commodity industries.

3) Localized industries. Industries which sell primarily within a market defined by the boundaries of the region containing the metropolitan area with a more significant proportion of their sales being inside the metropolitan area than outside. These are usually noncommodity industries.

The criteria we used to define the region associated with each metropolitan area are:

(a) that it include the markets in the associated states comprising the region for the industries located within the metropolitan area.

(b) that it include the major sources of labor and material input for the industries located within the metropolitan area.

(c) that to the extent that there are industries with national markets in the metropolitan area, the region includes other metropolitan areas which have comparable national industries that compete for the national market with industries in the metropolitan area being considered. The region in this case is thought of as a production area.

5. The level of economic activity and of population growth in an area depends on the area's level of activity in certain "basic industries" and in the export component of localized industries.

Projection Procedure

The projection procedure used to estimate metropolitan area employment and population is described below. Before proceeding with that description let us summarize what we have said above.

We have stated that metropolitan area growth is primarily due to changes in levels of activity in basic industries and the export component of localized industries. That the changes in these levels reflect the working out of comparative advantages among metropolitan areas. That to estimate these changes requires working with an analytical region for each metropolitan area within a framework of national and regional projections.

Steps in the projection procedure were as follows:

1. Use of National Planning Association's National and Regional Economic Projection Series.3/

2. Delineation of market or production regions for each area.

3. Projection of change in levels of activity in basic industries and the export component of localized industries.

4. Projection of residentiary employment of localized industries.

5. Projection of population.

6. Review of initial estimates. Then, judgment used as basis for making changes and finalizing estimates.

2. Delineation of regions for metropolitan areas.

Analytical regions for each area consist of one or more states. These regions for the most part are an area's market region. They include the state or states that provide the market for a large share of the area's noncommodity and commodity exports. The regions were delineated through the use of numerous data sources. Some examples of these regions are presented in Exhibit A.

3. Estimation of basic industry employment and the export component of localized industries.

a. Basic industries employment.

Basic industries are defined to include agriculture, mining, construction, manufacturing, the Federal Government and state and local governments. This group includes industries that are primarily export oriented: agriculture, mining, a number of each area's manufacturing industries and Federal and State governments; industries that are growth generating because of their investment type or high wage characteristics: construction and nonexport oriented manufacturing; and industries that provide services that would facilitate area growth: local government industries.

The metropolitan area employment in each "basic" industry is projected by means of trend extrapolation of the metropolitan area's share in the analytical region's employment in each basic industry. This technique assumes that metropolitan area employment in basic industries is related to the exports from the metropolitan area to its analytical region. The major factor underlying the trend in the projected ratio is each metropolitan area's competitive position compared with other metropolitan areas in the same analytical region in respect to export of basic industry goods and services to the analytical region.

b. Export components of localized industries.

A metropolitan area's "localized" industries find their markets largely within the metropolitan area but a significant share of the sales of some of these industries may be made to residents of the rest of the region. To identify this export component, we develop localizations coefficients. We measure the per capita share of employment in each localized industry in the metropolitan area and in the analytical region. Positive differences of the metropolitan area's ratio over the region's ratio (localization coefficients) multiplied by metropolitian area population provide estimates of export employment in these industries.

The export component of each industry is projected by working up the ratio of the export component in a specific localized industry to the region's nonmentropolitan employment in that localized industry. The trend of this ratio is extrapolated. The projected export employment is obtained by applying the projected ratio to the nonmetropolitan employment in the analytical region's localized industry for the projected year in reference. (The latter figure, in the actual operation of deriving employment estimates, is not available until the next step in the projection procedure is completed). The sum of employment in basic industries and in the export component of localized industries are together termed growth generating employment.

4. Projection of residentiary employment of localized industries.

Change in the residentiary employment of localized industries is judged, in the conceptual framework being described in this paper, to be determined by the change in the area's growth generating employment. The latter includes both export activity industries and selected other industries (construction and local government, for example). The rationale for this is that nationally and in metropolitan areas, there has been a shift toward noncommodity industries. This shift is projected to continue nationally and in our framework of projections it would therefore continue in metropolitan areas also.

The procedure for projecting the residentiary employment of localized industries is as follows: residentiary employment in a localized industry is computed as a ratio of growth generating employment. This ratio is projected and is applied to projected growth generating employment to give projected residentiary employment in localized industries. 4/

5. Projection of population

The population of a metropolitan area is defined as a function of the relationship between the population-employment ratio of a metropolitan area and the comparable ratio for the analytical region. The projection of this relationship provides projected metropolitan area population. All terms other than that of projected population for the SMSA are known from either the work described above or from the available state employment and population projections.

The procedure outlined above provides a set of projections. These are reviewed and changes are made whenever judgment dictates that they are needed. This may be considered a shortcoming by those who seek some completely mechanized technique which will provide reasonable projections without requiring the final touch-up. We feel it is an advantage to bring to bear on the initial set of projections the experienced mind containing much additional economic intelligence which has not been programmed into the projection techniques.

Selected Metropolitan Area Projections

In Table 1 we present some of the projections from the metropolitan area volumes cited in footnote 3. In respect to population, of the ten areas presented, Boston's population is projected to grow at the slowest rate (.6% annual rate 1962-75) and Phoenix's at the fastest (3.2% annual rate).

Despite the projected slow rate of populations growth (and the projected relatively slow rate of employment growth at 1.4 percent per annum), Boston's economy is expected to continue to provide a very substantial number of job opportunities for its population and a significant increase in per capita income. In 1975, employment participation of Boston's population (employment divided by population) is expected to be 25 percent above the national average. Per capita personal income is expected to increase at just about the nation's average annual rate of increase. (Boston 2.5 percent; U. S. 2.6 percent).

Boston's projected slow population growth can be explained by its physical location - bordered by an ocean and by other well developed "little" economies and by its already well developed economic size and maturity.

Phoenix's projected faster growth rate in population (and employment at a 4.5 percent annual rate) reflects a situation somewhat opposite to that cited for Boston. A much newer economic area, Phoenix is situated in a fast growing state. Favored by an excellent climate, Phoenix is attractive to industry and to those seeking a place for retirement. It is also developing finance and trade services for its immediate environs in the state.

Very briefly, the above analysis has attempted to illustrate how economic projections (in this case primarily those of industry employment) are pivotal in developing small area population estimates. The presentation was necessarily over-simplified and I refer you to the data in the three volume study on metropolitan area projections developed by NPA for a more adequate description of the economic bases for population projections.

Conclusions

The projection procedure described in this paper is based on the assumption that job opportunities determine population locations. It attempts to provide a demographic and economic framework which community leaders and others can use as part of the information base they need for making policy decisions. The great advantage of the procedure is that it forces area projections to remain within the bounds of what we consider reasonable regional control totals. Its chief shortcoming is that at present it provides limited detail about an area's economic and demographic circumstances. The National Planning Association plans to provide additional detail in the future as part of its Regional Economic Projection Series.

1/ Contributions to the methodology described in this paper were made by Sidney Sonenblum, Director of Research for NPA's Center for Economic Projections.

2/ For a detailed description of the criteria used in establishing the SMSA see "Standard Metropolitan Statistical Areas", U. S. Government Printing Office, Washington, D. C., 1964.

3/ National Economic Projection Series and Regional Economic Projection Series produced on a regular basis by NPA's Center for Economic Projections, Washington, D.C.

4/ The detailed mathematical formulation of the projection procedure is presented in Volume I of a three-volume report on metropolitan area projections. The report is titled: "Economic and Demographic Projections for Two Hundred and Twenty-Four Metropolitan Areas", Regional Economic Projection Series, Report No. 67-R-1, National Planning Association. The report was prepared by Dr. Joe Won Lee.

TABLE I. POPULATION AND EMPLOYMENT ESTIMATES FOR SELECTED METROPOLITAN AREAS

Metropolitan Area	Total Employment					Population				
	1950	1957	1960	1962	1975	1950	1957	1960	1962	1975
	(in thousands)									
Boston, Mass	983.1	1082.7	1145.9	1152.6	1382.5	2373.5	2473.2	2594.0	2638.5	2868.4
New York-N.E., N.J.,N.Y.,N.J.1/	5373.9	6103.1	6406.7	6439.0	7834.6	12955.5	14226.8	14809.1	15335.0	17344.6
Chicago-N.W. Indiana, Ill.-Ind.1/	2670.2	2913.9	2951.3	2969.8	3609.0	5521.7	6308.0	6828.1	6991.5	8749.2
Milwaukee, Wisconsin	406.3	493.2	506.0	514.1	697.4	874.9	1123.7	1201.0	1246.3	1604.9
Richmond, Virginia	158.7	179.0	190.8	200.9	267.3	329.9	384.6	410.8	458.2	565.8
Louisville, Ky.-Ind.	233.9	287.3	280.2	282.2	406.6	580.2	681.0	729.3	748.3	1017.7
Jacksonville, Florida	134.3	154.4	170.9	180.1	287.2	307.2	407.6	460.1	498.8	724.6
Phoenix, Arizona	118.1	188.4	233.8	252.2	447.8	337.7	548.6	675.4	775.8	1171.0
Los Angeles-Long Beach, Calif.	1764.6	2482.0	2647.2	2594.9	3929.3	4416.4	5984.9	6817.5	6377.0	8670.0
Seattle-Everett, Washington	296.4	378.1	424.9	467.3	598.6	737.5	874.2	1114.8	1149.3	1395.5

1/ Standard Consolidated Area

EXHIBIT A

Metropolitan Area	Analytical Region
Boston, Mass	New England, Middle Atlantic
New York - N.E., N.J., N.Y., N.J.3/	United States
Chicago - N.W. Indiana, Ill.-Ind. 3/	United States
Milwaukee, Wisconsin	Wisconsin, Minnesota, Illinois
Richmond, Virginia	Virginia
Louisville, Ky. - Ind.	Kentucky, Indiana
Jacksonville, Florida	Georgia, Florida
Phoenix, Arizona	Arizona
Los Angeles-Long Beach, California	Far West, Southwest
Seattle-Everett, Washington	Washington, Oregon, Idaho, Montana

POPULATION PROJECTIONS FOR COUNTIES AND METROPOLITAN STATISTICAL AREAS IN CALIFORNIA

Walter P. Hollmann, State of California, Department of Finance

Introduction

Perhaps when the history of population projecting or forecasting in the twentieth century is written, it will fall into three parts. The first, which began early in the nineteenth century, was characterized by concern for patterns of growth and only with the total population. A number of interesting attempts to fit curves to recorded changes have been chronicled culminating in the work of Pearl and Reed in 1920. Critics of this school claimed that the curves fit as well as they did only because they were used at a unique time in demographic history and that their success in describing what happened did not promise a corresponding success in describing what would or what might happen.

Starting in 1928 with the work of P. K. Whelpton, more analytic methods came into vogue. The cohort-survival or cohort/component technique initiated by Whelpton and developed by him with the aid of Warren S. Thompson and by a host of others is still widely used. I do not know to what extent its supremacy in the production of population projections for relatively short periods, 10 years to 20 years where the data are available would be challenged. For states and for the nation, detailed information is available, at least in the form of sophisticated estimates. For longer periods, for the look 25 and more years into the future that has become essential to major planning efforts, there is disagreement. This session is a demonstration of this.

The third period is that on the threshold of which we now stand. To an increasing extent, attempts will be made to forecast population or migration as a dependent variable, dependent upon some other and presumably independent variable, such as a quantified measure of economic activity.

We in California face several technical problems, which are not unique to us but nevertheless severely restrict the nature of our projection technique. One of these is our unusually high rate of migration and the inherent uncertainties in its composition; the other is the demand for small area detail resulting from our position in the state bureaucracy. The latter sometimes generates requests for more detail for smaller areas than prudent estimating would permit, so the work to be described represents one solution to the dilemma of insufficient data on the one hand and excessive demand for geographic detail on the other. Unfortunately we have been so busy mending holes in the data that we have not yet been able to set sail for the promised land of econometric models. It must be admitted that an area of relatively small population, high migration and insufficient demographic data might be amenable to an econometric model. I suspect that such an area might present problems whatever method were selected; certainly even an intercensal test of the method would be suspect.

The projection effort of the California Department of Finance has so far been restricted to the cohort-survival approach using the typical demographic inputs rather than attempting to depend upon non-demographic variables. There is a widespread practice in the literature to distinguish between forecasts and projections. The former term seems to refer to efforts which include judgment while the latter, presumably, are rigidly empirical. The distinction is more apparent than real. The present efforts may be regarded as projections since this is the word most commonly used in the demographic literature. No effort was made to gaze into the future beyond that required to select one of several possible patterns of change. For example, only two of many possible migration assumptions were selected and only two patterns of fertility and one of mortality were chosen. These selections were based upon a judgmental process, but no special claim of prescience is implied.

The preparation of projections is always a hazardous enterprise and anyone who claims to predict the future is, to quote Philip M. Hauser,

> "either a fool or a charlatan. Yet the projections of the demographers are more than exercises in arithmetic: they make it possible for us to see the implications of observed rates of growth."

The experienced user of population projections knows that they are neither a firm prediction of things to come nor a mere game played with the computer. Rather, they are a useful planning device to show us where we are going if our assumptions are correct. For this reason, projections are only as valid as their underlying assumptions and the user is urged to scrutinize such foundations with the greatest care.

The projections discussed in this paper, some of which are already completed and others of which are yet to be published, fall into three parts. First, a set of statewide civilian projections by age and sex was prepared, using the methods and assumptions described below. Next, a set of county total population projections was prepared using crude death rates, crude birth rates and numbers of net migrants with the addition of an assumed future military population for each county. In counties where State institutions were located, it was possible to remove the institutional population before establishment of the vital rates and to replace it in the final product. Better data on other special populations would permit the elaboration of this refinement. Each county's net migration was based upon its share of the State's net migrants within recent years with adjustments, by judgment, for local conditions. The totals of births, deaths, and net migrants for the counties were controlled to the totals for the state of these components, projected. A discussion of the final disaggregation of the age groups into geographical areas, the third and final effort, concludes the paper.

Projection Assumptions, the State

Assumptions used in projecting fall into two groups, the general and the specific. Almost invariably the general assumptions underlying population projections take the following form: it is assumed that

1. our democratic institutions and system of government will remain, and with them the right of every person to migrate where his whims or social or economic advantage dictate.
2. no major natural catastrophes will befall the State or the nation.
3. no major or world-wide war will break out.

The framing of specific assumptions requires, first of all, the selection of a basic orientation toward one or the other of two ideologies of projection. One assumes that employment or wages and salaries is the independent variable with respect to migration and that this variable is, in turn, dependent upon a host of other economic variables. The opposing view acknowledges the importance of economic parameters but argues that their bearing on migration is not yet sufficiently understood and, furthermore, that they cannot explain all migration. The prospect of an econometric model on which to base future population or migration is an attractive one. However, without ready access to the projected parameters needed as inputs to the model, no improvements in the validity of the projection are likely. An underlying article of faith is that whatever the level of migration assumed, the economic activity within the state will be adequate to support the population with tolerable levels of unemployment. None of the assumptions formulated envisions sustained migration as high as that of the average of the past decade.

From the standpoint of the policymaker who is to consider the various alternatives it would seem that the demographic projection would be preferable to the economic since it is easier to understand and the underlying assumptions are fewer and more clearly spelled out. The demographic projection has fewer hidden policy assumptions built into it. Furthermore, there seems to be a greater probability that the economic organization of society will differ from that which is now expected, than that the demographic events of births, deaths and migration will differ from those anticipated.

Similar considerations face the projector in examining the probable future courses of births and deaths. Changes in eating, smoking, or recreational habits or of major medical discoveries may cause changes in mortality, yet the actual level of mortality is more readily projected than changes in the determinants cited. Demographers seem to agree that, barring a major medical breakthrough, changes in mortality by age will not be critical.

This is not true of future additions to the population by births. The determinants of future fertility are many and complex and include such things as: availability of contraceptives and the religious and moral questions concerning their use; the types of housing to be produced in the future; the costs of education; future fundamental social values as they are reflected in styles of life, and a host of others.

To keep the projection problem to manageable proportions, the complexity of the determinants of future fertility, mortality and migration may be sidestepped in favor of summary assumptions concerning their level, formulated as specifically as possible. The user is urged to evaluate the plausibility of the assumptions and impose his own insights on them. The choice of assumptions is offered to enable him to do this.

Mortality There has been little change in mortality by age since 1955, therefore it is assumed that age-specific mortality rates will continue at their present level for all age groups.

Fertility The decline in fertility of American women since 1957 has dramatized the need for new projection assumptions. In the four series of population projections published by the United States Bureau of the Census in 1967[1], four different assumptions were used. The Bureau used a "cohort" measure of fertility, rather than the "period" or "calendar year" age-specific method. Briefly, this approach examines a birth cohort or group of women born in a specified period and projects what their future fertility will be in the light of age at marriage, children already born, age, and related factors. A group of age-specific birth rates for the projection period has been derived from the assumed cohort fertilities used in the Bureau of the Census' latest United States projections.

The differing United States fertility assumptions resulted in four projection series designated A, B, C and D. Series A assumes that future cohorts will experience, during their major childbearing years, the high fertility of the post World War II period, while series B assumes a somewhat diminished level. Series C and D assume that completed fertility will resemble that observed during the five decades preceding the postwar rise and it is to these two series that future California rates are tied. It is worthy of mention that of the four series offered as models of future fertility by the Census Bureau only a few years ago and first published in 1964, the two lowest are now considered as alternatives in the light of recent observed vital rates. The actual performance of the past several years for the United States suggests a recent level between C and D. Although this decline may merely reflect temporarily delayed births, a further diffusion of the small-family pattern, particularly among groups heretofore responsible for the larger families, will result in lower age-specific birth rates. The pyramid below compares the theoretical future age and sex structures of the State's population under fertility assumptions I-D and I-C.

(Figure 1)

Admittedly, a sociological judgment has been added to the projections, namely that the recent decline in fertility represents a return to a long-established historical trend. There is at least presumptive evidence that this is the case in the United States and in other countries with advanced technologies. The course of the birth rates for women of various ages under the assumptions of series C and D are presented graphically below.

(Figure 2)

For California, it was assumed that the difference between the State age-specific rates and the national rates observed in 1960 would tend to converge within 50 years.

Net Migration A satisfactory study of migration including the gross streams in and out of individual states has not been made since the Census of 1960, which asked, nationwide, a question on residence in 1955 of all persons five years of age and older within a sample population. The estimation of net migration, the excess of in-migrants over out-migrants, is a vital part of the population estimating process, especially in California. Since 1950 California's net civilian migration has varied from an annual estimated 268,000 in fiscal 1951 to 388,000 in fiscal 1957; since 1960, annual estimates of net civilian migration have varied between 369,000 in fiscal 1963 and a provisional estimate of 240,000 in fiscal 1966. International migrants to California, immigrants, are a significant element in population change and in the last ten years have varied between the 49,673 of 1959 and the 79,090 of 1963[2]. For the projections, annual immigration at a level of 65,000 has been assumed, and this element is included in the projections of net migration. It is as yet too soon to assess the effects of recent changes in the immigration laws, but this variable may increase in the years ahead.

Military migration is assumed to be zero unless a buildup or contraction of troop strength in California is underway. The dependents of military personnel are civilians, and for certain areas rapid changes in net civilian migration have been caused by military shifts. For projection purposes, and because no change in the military population of California after 1970 can be foreseen, net civilian migration and net total migration need not be distinguished. "Loss to military," that component of civilian population change measuring the net movement into and out of the armed forces, is projected at zero or it may be considered a part of the net migration to and from the State.

The determinants of net migration are certainly manifold, complex, interrelated and not completely understood. While attempts have been made to construct models which tie migration to other projected variables, these population projections make no attempt to isolate or separately to project the motivations underlying migration. In the familiar fashion they attempt to show the future population of the State, at various times and for various age and sex groups, implied by their underlying assumptions. The present "state of the art" imposes its limitations.

In the framing of assumptions concerning the future level of net migration one salient fact cannot be ignored. There are no concrete up-to-date data on the age and sex composition of the migrant population of all ages to California, neither in-, nor out-, nor net. However, using birth, death, school enrollment and social security data it is possible to estimate age groups annually. Any change in these age groups beyond the effects of births and deaths are assumed to be attributable to migration. Using this method, an average annual rate of migration by age from 1960 to 1965 was developed and applied throughout the projection period to obtain an age "mix" of migrants.

The decision to use this device to establish merely a migration mix and not a magnitude is attributable to a property inherent in the assumption. If a constant rate of net migration were applied to a growing population it would imply an ever-increasing number of migrants due only to the larger population. For this reason net migration is controlled to a previously assumed level.

In Series I, net migration is assumed to level at 300,000 per year. As an alternative, Series II is presented. In this series the level of immigration from foreign countries is assumed stable at 65,000 per year, while the domestic net migration of 235,000 is assumed to be declining at the rate of 4,700 per year. In effect, Series II assumes that although international migration will continue, and the high level of U. S. migration in and out of all states will also continue, in 50 years all interstate disequilibriums will have vanished; California (and all other states) will attract and repel migrants in equal measure. Series I represents the effects, conservatively stated, of the continuation of recent migration experience of the State. Series II measures the effects of a moderate decline in net migration. These two options are comparable to those used in population projections produced by the Bureau of the Census. The effect of these differing assumptions on the possible age structure of the State is shown by Figure 3.

Although Series I-D and I-C on the one hand and II-D and II-C on the other, differ only in their underlying fertility assumptions, some differences appear in age groups outside those affected by the differing births. This is attributable to the fact that net migration is held at a constant level as are net migration rates by age. The differing composition of the population under the differing fertility assumptions generate slightly differing "mixes" of the migrant population.

(Figure 3)

Summary, Statewide Projections

If the assumptions underlying Series I-D are realized, the civilian population of California will attain a level of about 26,100,000 in 1980 and 38,700,000 by the year 2000. Series II-D suggests a population of 25,600,000 in 1980 and 35,500,000 in 2000. By the year 2000 Series I-C and II-C suggest civilian populations of 41,600,000 and 38,200,000 respectively[3]. It is worth noting that the effects of varying assumptions became more pronounced with the passage of time. A short-term projection can be made with a single set of assumptions and yield information in which the user may have a fair degree of confidence, but a projection carried beyond a decade or two should be interpreted in terms of ranges. However, even such ranges should not be regarded as upper and lower limits to the possible future course of population.

The median age of the civilian population at the time of the 1960 Census was 30.3. This is expected to drop to a minimum of 27.8 in 1970, from which it will rise, attaining a level of 28.7 in 1980 and 31.0 in 2000 under Series I-D. Under Series II-D a slightly more steep rise to 28.8 in 1980 and 31.4 in 2000 may be anticipated. For Series I-C and II-C the median age remains lower throughout the period due to the assumed greater number of births under both C series. In 1970 I-C and II-C imply 27.7 years and by 2000 the median age is 28.4 for Series I-C and 28.7 for Series II-C.

Projections for Standard Metropolitan Statistical Areas[4]

California consists of 58 counties of which 36 are non-metropolitan. The remaining 22 are divided into 14 SMSA's. On the basis of tests which revealed tolerable success in the development of 1965 composition (as measured by a composite method of estimating) from the 1960 benchmark, it was decided to attempt projections for the SMSA's and for the non-metropolitan counties, the latter treated as a unit. The non-metropolitan area of California is more statistical convenience than administrative entity since it contains counties as dissimilar as hot, dry, Imperial with its lettuce crops, Alpine (population 397 in 1960) and cool, moist Del Norte and its coastal redwoods.

The projecting task falls into two sections. The first involves the development and testing of rates, using the first benchmark, 1960, and proceeding to the second benchmark, 1965. A set of SMSA estimates has been prepared by a composite method. The second step is the projection of the rates established in the first within the framework imposed by the statewide assumptions.

The inputs required are the state age-specific birth rates, death rates and rates of net civilian migration and loss to military for the years 1960-65. Furthermore, the number of births, deaths, net migrants and loss to military, specific by age, for the 15 areas for the same period are necessary. For the test, deaths were developed by application of rates from an especially prepared 1960 table of survivals for the state, by single year of age and by sex. The use of statewide rates is probably justified in the absence of evidence of critical regional differences. Completion of the test will indicate whether or not minor adjustments are warranted. Births were developed by a comparison of statewide age-specific birth rates for 1960 with those for each respective area by ratio. This ratio was applied to the derived state age-specific rates for each year between 1960 and the second benchmark year and the sum of births generated were controlled to the recorded births for the state by minor adjustments to the rates. The assumption selected implies that relative differences in age-specific birth rates among the areas will continue and that any change in the birth rates experienced by the state will be shared proportionately by its areas.

On the basis of composite estimates prepared for the areas for the years 1960-1965, unadjusted migrations were calculated by adding, age specifically and algebraically, population changes and deaths. Starting from 1960, migration rates based upon the unadjusted migrations were averaged for an annual value, specific for age, and applied, year by year, along with birth rates, death rates, and loss to military to estimate the 1965 population. Comparisons with the 1965 and earlier composite estimates were made. The net civilian migration rates were adjusted by an iterative method to produce net migrants sufficient to yield the 1965 estimates.

The projections of the SMSA's by age, not yet completed, will be produced from the benchmark year 1965 using adjusted migration rates, birth rates and death rates and assumed loss to military. In the cases of all three components, the sums of the numbers generated for the 15 respective areas are to be controlled to the previously projected state total by age.

---

1 Bureau of the Census, Current Population Reports Series P-25 No. 359.

2 United States Department of Justice, Annual Report of the Immigration and Naturalization Service, 1966.

3 The figures cited are by no means official since the final run has not been made. Those interested in the final figures and their breakdowns may request them from the Department of Finance, Sacramento 95814.

4 The contribution of Mrs. Isabel T. Hambright, who with programming help from Survey Research Center, Berkeley, made these projections possible, is acknowledged.

Figure 1

CALIFORNIA'S CIVILIAN POPULATION BY AGE AND SEX

1960-1980-2000

SERIES ID AND IC

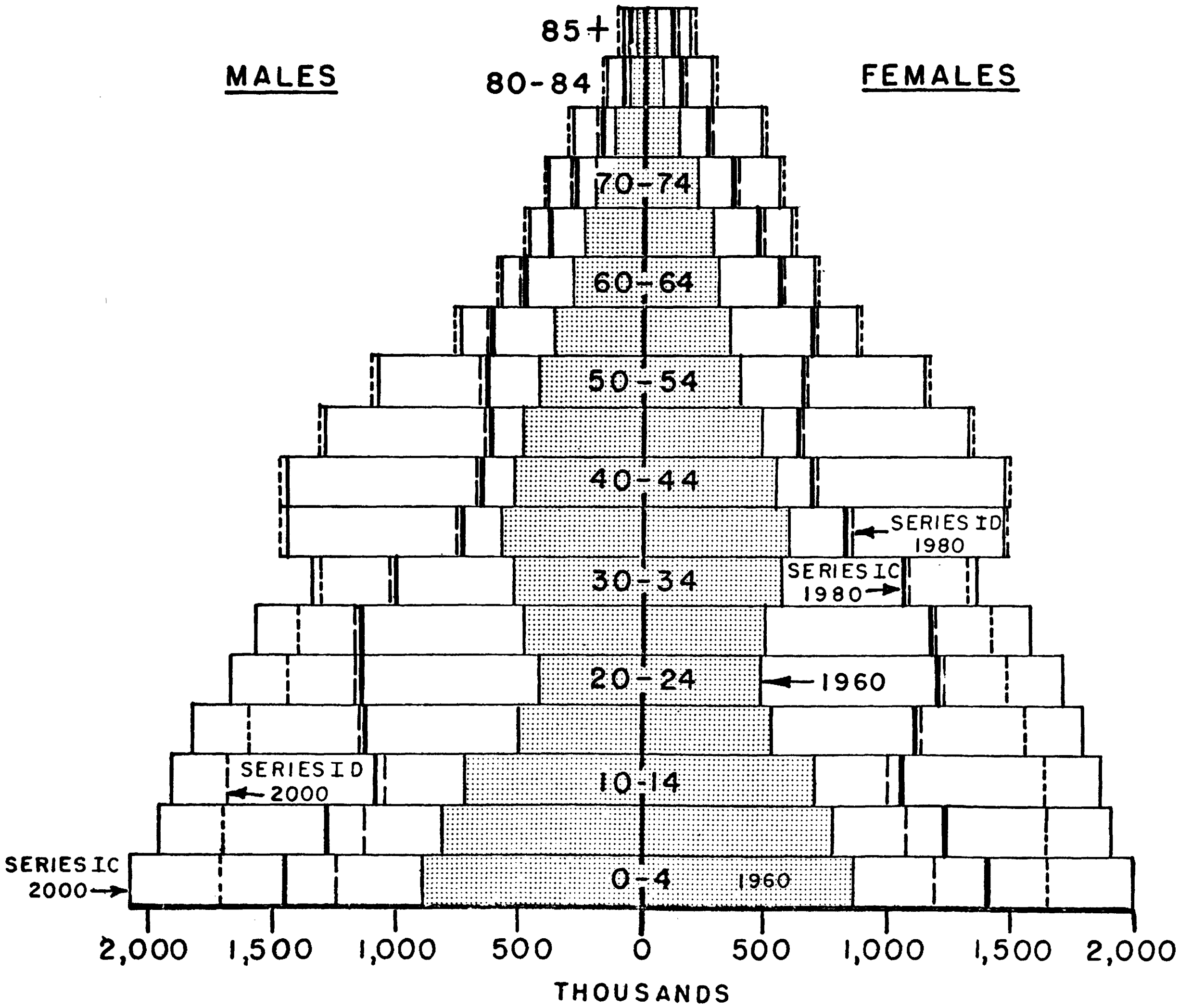


Figure 2

CALIFORNIA
AGE-SPECIFIC BIRTH RATES
1960 – 2000

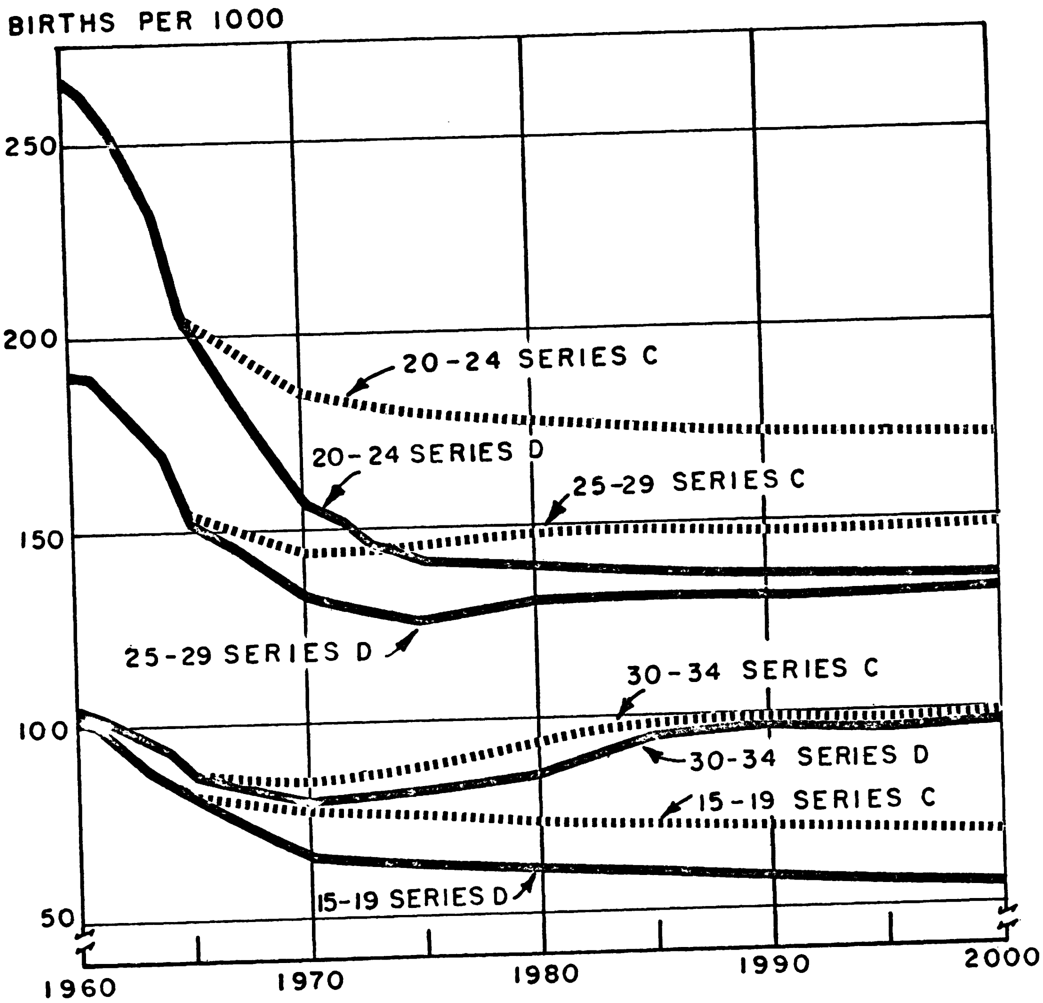
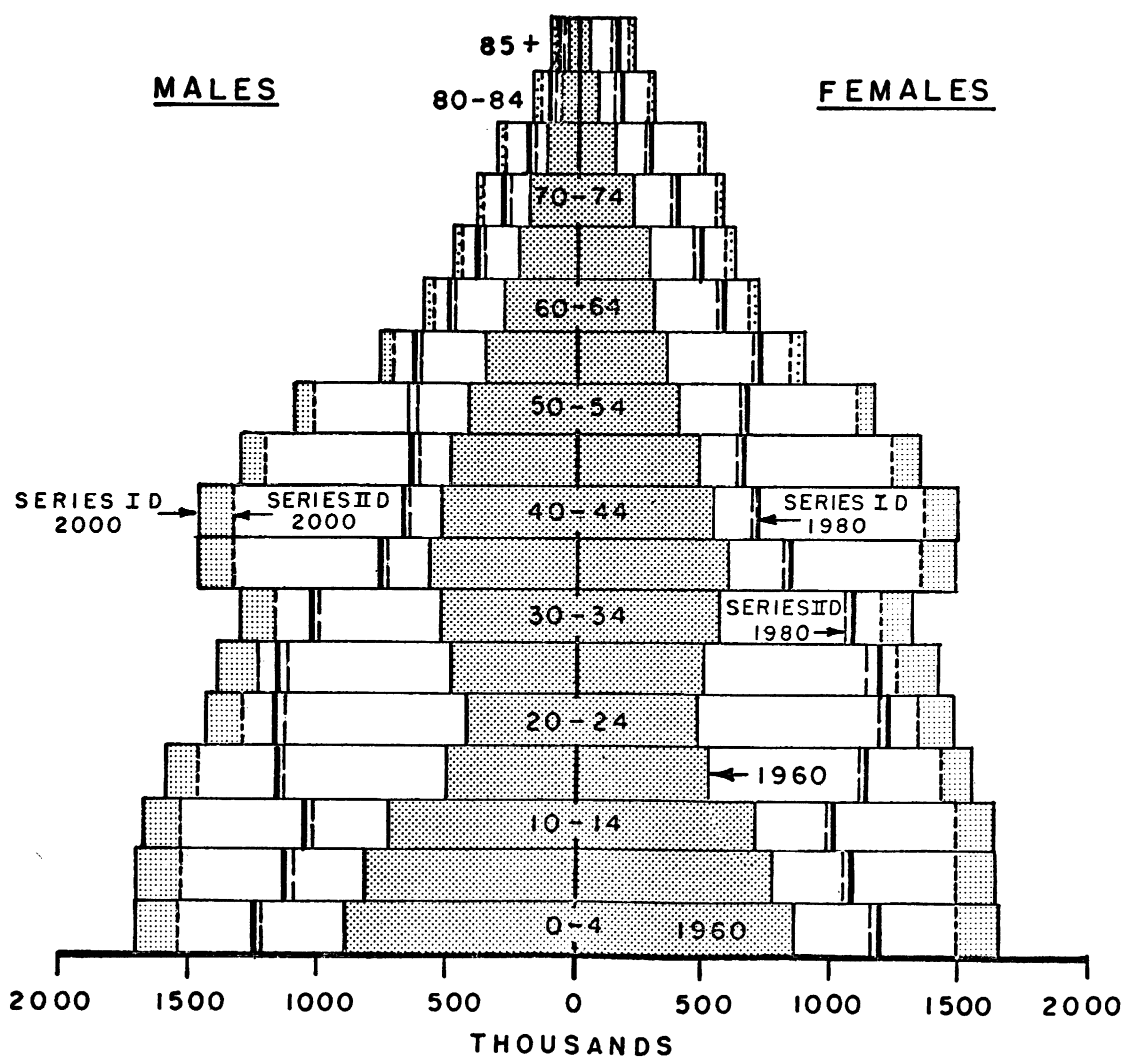


Figure 3

CALIFORNIA'S CIVILIAN POPULATION BY AGE AND SEX

1960-1980-2000

SERIES I D AND II D



ECONOMIC PROJECTIONS FOR LOCAL AREAS

Robert E. Graham, Jr., Daniel H. Garnick, and Allan L. Olson*
Office of Business Economics

The title of this paper has been changed from that listed in the program -- which emphasizes population projections -- to one that is more generalized. This change reflects the fact that our aim in the Regional Economics Division of the Office of Business Economics is to project three major aggregates -- income, employment, and population. Moreover, our main focus is on the separate and independent projections of income and employment, with population a derivative of the first two.

The decision to make population a derivative of income and employment stems from the assumption that the critical element in a population projection is regional migration and the major factor underlying migration is economic opportunity, or lack thereof. That is, population will move toward expanding economic opportunity and away from a shrinking or static regional economy. Consequently, it would seem that a better population projection can be made by concentrating more directly on the basic motivating factors and then deriving population from the results rather than via a population-to-employment approach.

Economic projections are often classified into 2 major groups: (1) Projections that represent mainly extrapolations of past trends which are usually termed simple or naive, and (2) projections made via an economic accounting model featuring income and product or input-output and which are termed sophisticated.

Without intent to set up a "straw man," it seems useful to point out that the foregoing characterizations of the two types of projections are not really valid.

Reliance on extension of past trends is not a distinguishing feature of the two methods, for both naive and sophisticated projections place equal reliance on past experience. To the extent that the sophisticated methods reflect no past experience, their results must generally be judged less reliable than those of simpler but historically-based series.

Use of past trends in the naive method is well-recognized. Not so obvious, however, is the equal reliance that sophisticated methods place on past developments. An input-output table, for example, to be really useful in making projections requires that at least two exogenous elements -- final demand and technical coefficients be projected, either explicitly or implicitly. To hold technical coefficients constant and project proportionately equal increases in all elements of final demand yield results of the most naive sort. Somewhat more valid is a regional projection based on input-output that relies on a simple projection of past trends to derive a population estimate which, in turn, undergirds a projection of final demand from which is derived a projected economic structure of the area.

Both naive and sophisticated methods, then, rely equally on past trends to the extent the basic data permit. If time series data are not available, and they tend to be scarce in direct proportion to the complexity of the projection framework being used, the technician must either hold all relationships constant or change them through deductive reasoning.

A pragmatic classification of projection methodologies into naive and sophisticated groups reflects more the complexity of the economic measure used in the projection process than the method used to extend that measure into the future. And, <u>ceteris paribus</u> the more complex or detailed the economic measure employed, the more useful will be the results to the extent that the systematic components dominate the random, nonsystematic components of the economic measure. However, in the real world of economic measures other things are seldom equal and in few areas are they more unequal than in regional economic measurement. Indeed, the regional field is characterized by a paucity of economic measures. Here, reference is not to the quantity of data available. What is lacking are time series for major constructs such as income and product tables, input-output accounts, employment and flow-of-funds series disaggregated both industrially and geographically.

In the final analysis the method used, a complex econometric approach versus a simple methodology with a good measure of judgment thrown in, really reflects the type and quality of the input data available. If input-output or income and product tables are available both historically and currently for the geographic areas under study, the so-called sophisticated method of projection would be the choice in nearly every instance. If data availability imposes its usual constraints, a simpler model tends to become the choice.

Our program of regional economic projections calls for the preparation of projections for 165 economic areas initially. Later the number may be increased to as many as 400. Given this very large number of geographic areas, data input becomes a crucial consideration.

---

* The authors are members of the Regional Economics Division.

To prepare input-output tables or income and product accounts that are something more than mirror images of their national counterparts for at least two years (in order to gauge trends) for 165 separate areas would be a task of near-impossible proportions. To measure personal income by local area for 5 selected years has required 2 years and an expenditure of close to $1 million. Preparation of the more detailed economic accounts, if indeed feasible, would require many times the resources needed for measuring personal income.

Given data requirements and data availability, we have chosen to project an economic aggregate that is moderately comprehensive; that can be constructed to show adequate geographic and industrial detail; and for which a time series can be prepared. Personal income meets those requirements more adequately than any alternative.

Specifically, historical estimates of personal income by local area have been prepared for 5 selected years of the span 1929 to 1962. The years include 1929, 1940, 1950, 1959, and 1962. In general, each of these except 1940, represents a roughly comparable point on the business cycle, thereby eliminating the potentially distorting influence of the cycle on economic change. Since area employment estimates statistically comparable to the personal income series have not been completed, we are using employment from the decennial censuses of population. From this a series showing about 36 separate industries by local areas has been assembled.

Because it was desirable to project both employment and income in as much industrial detail as possible, a set of geographic areas in which the various industrial components of income and employment would bear reasonably stable relationships to one another was constructed. These economic areas are based on the nodal-functional area concept. That is, to each urban center are attached the surrounding county units in which economic activity is focused directly or indirectly on the center. Each economic area combines the place of residence and place of work of employees as nearly as possible so that there is a minimum of commuting across economic area boundaries.

Each economic area specializes in the production of certain types of transportable commodities and of nontransportable special services such as education at Cambridge, recreation at Miami, and finance in New York. The production locus of such goods and services is determined not so much by transportation costs as it is by the costs associated with special resources and by the economic benefits derived from economies of scale. Different commodities are associated with production processes requiring different input relationships and the comparative advantage of a region for the production of a commodity is determined by the region's relative endowment of the factors of production. In addition, in many industries the effort to maximize returns to the factors of production leads to expanded production as a means of exploiting the economies of scale. This process, which can be implemented only if trade can be carried on with other areas, further reinforces regional comparative advantage and specialization.

In contrast, each economic area approaches self-sufficiency in its residentiary industry sector; that is, while each area specializes in producing goods and/or services for "export" to other economic areas (and abroad) most of the services (and some goods) required by local residents and businesses are provided within the area.

Thus, the economic areas correspond to the closed trade areas of central place theory in which the number and type of establishments and their size and trade areas are bounded by the relative transportation costs from the hinterland to competing centers. Each area approaches closure with respect to residentiary industries which include general and convenience retail and wholesale trade activities and those other services which are difficult or impossible to transport and are most efficiently consumed in the vicinity of their production.

Application of the foregoing criteria to the U. S. economy yielded 165 areas each of which formed a complete and integrated economic unit characterized by comparative stability in interindustry relationships. Having delineated the 165 economic areas, we then considered alternative projection methodologies.

The first method examined was a naive model, characterized by a complete absence of theoretical underpinnings in its formulation. It was devoid of systematic or interacting components and all projected elements were exogenously determined. It was essentially a "no change" model.

The exogenous determinant or predictor in this naive regional model was the national change in employment or income in a given industry. That is, the base period ratio of regional employment or income to national employment or income in each industry was applied to the projected national level of employment or income for the corresponding industry.

$$(1) \quad E^t_{ij} = (E^o_{ij}/E^o_{io})\, E^t_{io}$$

Where the subscripts $\underline{i}$, $\underline{j}$ refer to the $\underline{i}$th industry and the $\underline{j}$th region, the subscript $\underline{o}$ refers to a summation: when in the right hand position, it is the summation of regions (= the Nation), when in the left hand position, it is the summation of industries (= total employment or income); superscripts $\underline{t}$, $\underline{o}$ refer to the projected period and the base period, respectively.

The naive model, though reflecting no more than the national industrial growth rates in each individual industry in each region, does, nonetheless, reflect an aggregate growth rate that differs from that for the Nation when the region's industrial composition differs from the national in the base period. Such a model, however, fails to take account of regional differences in rates of growth among individual industries. To take account of this, we turned to shift-share analysis.

Shift-share analysis is designed to discern regional departures from national industrial growth rates, and while its history goes back to 1943, most of the work using, clarifying and elaborating on the technique appeared only in the late 1950's and in the 1960's. [1] In its simplest form, the shift-share technique distinguishes a proportional growth element and a differential growth element between a region and the Nation in each industry.

$$(2)\quad E^t_{ij} = (E^t_{io}/E^o_{io})\, E^o_{ij} + C^t_{ij}$$

Where C^t_{ij} equals the difference between the level attributable to the national growth rate of the industry and the regional growth rate actually attained in the industry.

It is the attention paid to the differences between regional and national growth rates in each industry that distinguishes the naive share model of equation (1) from the shift-share model of equation (2). Thus, the first term on the right hand side of equation (2) is equal to the entire right hand side of equation (1). The second term on the right hand side of equation (2) is called the share effect (C_{ij}) in shift-share analysis. It is, in fact, the difference between the "hypothetical growth," accounted for by the first term, and the attained level of the left hand side. In basic or export industries the share effect is presumed to be connected with some regional competitive advantage (or disadvantage if the term is negative) in the industry. That is, the region presumably grows faster or slower than the rest of the Nation with respect to the industry in question because of a difference in the marginal productivity of capital in the region relative to all other regions. Thus, the shift-share projection model departs from the naive-share extrapolation model, at least implicitly, insofar as it treats regions as relatively open economies among which capital and labor may flow. In contrast, the naive-share extrapolation model treats each region as a miniature reproduction of the national economy with all national developments occurring proportionally in each region's economy.

The causal economic factors associated with C_{ij} are the essence of industrial location theory. But, over the last 20 years, there has been very little correspondence between developments in industrial location theory and the empirical studies undertaken with respect to locational patterns. [2]

Since industrial location theory has produced so little empirical evidence of the causal factors that determine industrial location patterns, projecting the C_{ij} term is still in an experimental stage. Two approaches have been tested. An econometric model which uses multiple regression to "explain" and project the C_{ij} effect for each of 50 industries has been developed in the Regional Economics Division. In it, the share effect is projected for each industry by a multiple regression analysis. That is, the C_{ij} effect in the most recent period for which data are available is regressed against a number of independent variables that relate to the preceding period or to a preceding point in time. This use of lagged variables obviates the necessity of making separate projections for each independent variable. The most significant of these variables is the C_{ij} effect in the preceding period. Additional independent variables include measures such as the size and rate of growth of the industry, total population, level of income, and the C_{ij} effect in related industries. Inclusion of this last variable makes it possible to establish appropriate interindustry linkages in the regression equations. Regression coefficients are calculated by "cross-sectional" analysis in which the value of the variable in each area forms an observation.

As empirical evidence is gathered, and as regional economic measurement is refined, it will be possible to select independent variables that have a closer and more stable relationship to the C_{ij} effect. At that time the foregoing method would seem to offer the most potential for development. However, in view of the paucity of data with which to measure past changes in the geographic location of industries and the comparative lack of information on factors underlying these changes, the foregoing approach to projecting the C_{ij} element of industrial change with its considerable emphasis on mathematical precision seemed unsatisfactory.

Accordingly, the second approach to projecting the C_{ij} term was a simple one that was less demanding of data and that could make maximum use of available information. For each industry, a simple curve was fitted to each region's share of the national total of income and employment (separately). This curve was then extended into the future and the values of the region's future share read.

This last approach is actually a variation of "shift-share" analysis with regional share effects (C_{ij}) calculated implicitly rather than explicitly. That is, from equation (2) the following relationship between changes in the regional share of the national industry (E_{ij}/E_{io}) and the regional-share effect (C_{ij}) of the shift analysis holds:

$$(2')\quad E^t_{ij}/E^t_{io} = E^o_{ij}/E^o_{io} + C^t_{ij}/E^t_{io}$$

$$(2'')\quad C^t_{ij} = E^t_{io}\left[E^t_{ij}/E^t_{io} - E^o_{io}/E^o_{io}\right]$$

$$= E^t_{io}\, \Delta\, (E_{ij}/E_{io}).$$

Statistical tests were applied to the several models as well as to others not described here. Although results at this stage are inconclusive, indications are that model 3 gave best results. Accordingly, it was chosen for further development. Model 3, it will be recalled, was the curvilinear extension by simple regression of a region's percentage share of the national total of income and employment in each industry. This mechanistic approach was modified in two ways.

First, substantial judgment was used in extending the curves. Such judgment reflected analysis of the numerous erratic observations in the historical time series; the timing of basic developments in a series; the status of the supply of the natural resource on which a particular industry depended; and the shape of the curve fitted to the measured observations. This approach permitted the full utilization of all information that could be assembled on any given industry in any region.

The projections made of the basic industries as outlined above (and specified in equation 2) were considered final. However, analysis of the interindustry relationships that prevailed in both income and employment in the 165 functional economic areas led to the modification of the shift-share projection model (as in model 3) to incorporate some features of an older basic-service model in projecting residentiary industries. This comprised the second of the two modifications referred to above.

Studies of the relationships of local-service or typically residentiary activities to export or basic industries in a region have given rise to an often used basic-service model. The interactions of the exogenous and the local-service industries result in a multiplier effect very similar to a Keynesian consumption multiplier. In the case of the basic-service model, the endogenous or internally determined sector is comprised of local-service activities such as trade, local transportation and other service activities. Since the function of these local-service or residentiary activities is to supply the local businesses and households with commodities and services which do not enter into interregional trade in substantial amounts, the magnitude of these residentiary activities is determined by the size of the population and income of the region. Thus, regional residentiary employment or income is functionally determined by regional total employment or income and hence must be solved simultaneously with the latter two aggregates.

The total employment and total income necessary to solve the residentiary industry equation were obtained by summing the final projections for basic industries and the preliminary projections for the residentiary group. Final projections of residentiary industries were then endogenously determined by functional relationships estimated in cross-section studies and projected forward by means of these relationships together with projected changes in the exogenous sector.

$$(3) \quad E^t_{oj} = \sum_{i=1}^{k} E^t_{ij} + \sum_{i=k+1}^{n} E^t_{ij}$$

$$(3') \quad \sum_{i=1}^{k} E^t_{ij} = \sum_{i=1}^{k} \left[(E^t_{io}/E^o_{io})E^o_{ij} + C^t_{ij} \right] = B^t_{oj}$$

$$(3'') \quad \sum_{i=k+1}^{n} E^t_{ij} = f(E^t_{oj}) = a_{oj} + b_{oj}E^t_{oj} = S^t_{oj}$$

Where E^t_{oj} = the sum of regional employment in industries $(i = 1, \ldots, n)$, of which the exogenous industries $(i = 1, \ldots k)$ are projected by means of the shift-share model as in equation (3) and the endogenous local-service industries $(i = k+1, \ldots, n)$ are jointly determined with total regional employment (here shown as a simple linear relationship where the a + b parameters are estimated by cross-section analysis).

Substituting (3') and (3'') into (3) and simplifying gives the multiplier value similar in structure to the Keynesian multiplier (1/1 - marginal propensity to consume locally produced goods and services).

$$(3''') \quad E^t_{oj} = B^t_{oj} + a_{oj} + b_{oj}E^t_{oj} = {}^1/_{1-b_{oj}} \, (B^t_{oj} + a_{oj})$$

Indeed, cross-section analysis undertaken by the Regional Economics Division has permitted estimates of industrially disaggregated regional residentiary sector multipliers.

$$(3'''') \quad E^t_{oj} = B^t_{oj} + \sum_i (a_{io} + b_{io})E^t_{oj} = ({}^1/_{1} - \sum_i b_{io}) \, (B^t_{oj} + \sum_i a_{io})$$ [3]

Clearly, since the a_{io} and b_{io} parameters represent a national central tendency, they do not necessarily fit the current case for individual regions. The Regional Economics Division is adjusting them for regional use however, by trending the current residentiary mix with respect to total regional employment toward the "national" parameters over the projection period. Thus, the working assumption is that regional local consumption patterns will trend toward national uniformity.

It is a deficiency of basic-service models of the type represented in the (3) series equations as well as of regional input-output models, that regional growth is caused entirely by external stimulation through the growth of the exogenous sector. While this deficiency is not altogether redressed in this methodology, it is diminished to the extent that the relationship between the basic and the local-service sectors is stable. Such stability is, of course, greater, the more successful we are in delineating nodal regional configurations.

Thus, if r_e and r_b are the rates of growth of regional total employment, and regional basic or exogenous employment, respectively, it can be shown that in this model the rate of growth of basic employment determines jointly the rates of growth of total regional and local-service employment.

$$(3''''') \quad \frac{E^t_{oj}}{E^o_{oj}} = (1 + r_e)t = \frac{1}{E^o_{oj}} \; \frac{1}{1-b_{oj}} \; (a_{oj} + B^t_{oj})$$

$$= \frac{1}{\frac{a_{oj}+B^o_{oj}}{1-b_{oj}}} \; \frac{1}{1-b_{oj}} \; [a_{oj} + B^o_{oj}(1+r_b)^t]$$

$$= \frac{a_{oj} + B^o_{oj}\;(1 + r_b)^t}{a_{oj} + B^o_{oj}}$$

This deficiency is all the greater, the greater the variability in the relationship between the exogenous and the endogenous sectors. Conceptually, the relationship between the exogenous and the endogenous sectors, here the basic and the local-service industries respectively, is most stable in a nodal regional delineation scheme and least stable in an arbitrary or administrative delineation scheme. Thus, the relatively closed trade area incorporated in the functional economic area concept would hypothetically permit less variance between exogenous and endogenous sectors than would regions which were identical with county boundaries or delineation based on administrative or homogeneous groupings of county units. Empirical studies performed by the Regional Economics Division with respect to indexes of industrial centralization and of relative regional specialization support the hypothesis that less variation in the basic-service relationship occurs in such nodal regions as OBE Economic Areas than in non-nodal regions comprised of single counties or homogeneous or arbitrary groupings of counties. [4] Hence, the validity for projecting such cross-section relationships as basic-service interactions forward in time decreases as regional delineations depart from the nodal regional concept.

Footnotes

1. For a detailed explanation of this type of analysis, see Growth Patterns in Employment by County, 1940-1950 and 1950-1960, Lowell D. Ashby, Office of Business Economics, U. S. Department of Commerce, 1965.

2. B. H. Stevens and C. A. Brackett argue that this lack of correspondence is in part attributable to the inability of existing theory to generate testable hypotheses. Cf. Industrial Location, A Review and Annotated Bibliography of Theoretical, Empirical and Case Studies, Regional Science Research Institute, Philadelphia, 1967. This dearth of hypothesis testing has also been noted by J. Meyer, "Regional Economics: A Survey," American Economic Review, LIII, No. 1, March 1963.

3. Insofar as the parameters are central tendencies over all the regions of the Nation, they represent, as it were, national coefficients, hence, the index notation a_{io} and b_{io} where the right hand notation position indicates their national character. This deficiency of relying on national or adjusted national coefficients is shared with many regional input-output projection models. Indeed, it can be shown that if the sectors were the same, the parameters in equation (3'''') are algebraically the same as those in the employment multiplier model estimated in the input-output projection study for the New York SMSA. Cf. B. Berman, B. Chinitz and E. Hoover, Projections of a Metropolis, Harvard University Press, Cambridge, 1961, pp. 8-9.

4. It must be borne in mind that for any given industrial sectoralization, increasing the size of the region has the tendency to decrease the variation among regions. Since increasing size actually means aggregating contiguous counties each with their own industrial mix, each county added to a regional configuration implies a discrete and not necessarily monotonic change. Nonetheless, counties lying in the hinterlands of urban centers, on the average, do exhibit the tendency toward reducing regional specialization when added to the urban centers. That is, when location quotients are the means for distributing portions of the industrial sectors among the basic and residentiary sectors we have:

$$\frac{E_{ij}/E_{oj}}{E_{io}/E_{oo}} = L_{ij}$$

$$B_{oj} = \sum_{i, L_{ij} > 1} E_{ij}\left(1 - \frac{1}{L_{ij}}\right) = \sum_{i, L_{ij} > 1} [E_{ij} - E_{io}(E_{oj}/E_{oo})]$$

Therefore, regional specialization, B_{oj}/E_{oj}, tends to zero as regions are summed, $\sum_j (E_{oj}/E_{oo}) = 1$, and $E_{ij} - E_{io}$ tends toward zero. Thus, we have to distinguish between decreases in the regional specialization index resulting purely from the size effects implicit in the mechanical aggregation of counties from those decreases resulting from nodal regional delineation.

DISCUSSION

Harry M. Rosenberg, Battelle Memorial Institute

Introduction

At a meeting of the St. Louis Chapter of the American Statistical Association, Jacob Siegel made the following comments about the present state of the art:

> "It is now generally held by the more prominent and learned members of the fraternity of professional demographers that, with the present knowledge and techniques, it is not possible to forecast the population of small geographic areas accurately."(1)

Against the background of these remarks made 15 years ago I wish to discuss the papers by Mr. Kupinsky and Mr. Hollman, since they are, in my opinion, excellent examples of the most recent developments in making population projections.

Mr. Hollman's paper describes the methods by which a service organization, the California Department of Finance, prepares population projections, using time-tested demographic techniques, whose underlying assumptions can be simply stated, whose results can be readily reproduced, and whose reliability is to some extent known on the basis of past applications and tests. In contrast, Mr. Kupinsky's paper describes methods by which a research organization, the National Planning Association, has prepared industry-employment, personal income, and population projections using an "eclectic" and "pragmatic" approach, which projects and imposes consistency on several exogenously determined elements of a very complex system. Achieving this consistency is attained by a number of mechanical iterations and some educated judgements about the "consistency" and "reasonableness" of the results. The methods used by the National Planning Association are in a sense experimental; they have not been extensively tested and evaluated. The underlying assumptions cannot be simply stated, and the results cannot be readily reproduced. Moreover, the extent to which errors associated with the various parameters, such as the "critical ratios" used in the method, can ramify throughout the system is not well-understood.

The California Department of Finance projections represent the application of methods developed during what Hollman would characterize as the second era of population projection history. These are methods that are extensively used by public agencies because their behavior is well-understood, because their data requirements are fairly modest, and because their application is mechanically simple. The National Planning Association projections, in comparison, represent a bold excursion into the frontier of the third era; here, population growth is considered within the broader context of economic change. To my knowledge, no other organization has developed projections of employment, income, and population in such rich geographic detail. I shall discuss how the NPA population projections, based on economic considerations, compare methodologically with the more traditional demographic methods; I shall mention some of the conceptual and practical problems associated with the use of the NPA projection procedures; and I shall speculate on the likelihood that the resulting regional and metropolitan projections are more reliable than those, like Mr. Hollman's, that are generated using standard cohort-survival methods.

NPA Projections

As indicated by the title of Mr. Kupinsky's paper, the NPA employment and population projections were developed within a national and regional projection framework. First, state industry employment projections were developed by apportioning projected national growth among the states; then metropolitan projections were derived by relating metropolitan growth to projected regional employment change. Regions, according to NPA usage, are those areas defined by state boundaries in which the major economic transactions of the metropolitan area under consideration take place.

Both the state and the metropolitan employment projections are based on techniques that reflect the "export base" theory of economic change. Stated simply, this maintains that economic activities in large regions and smaller geographic areas may be divided into two classes that are different--both with respect to the forces that activate them and with respect to the contributions they make to the particular subnational economy. The first class is composed of "base" or "regional building" activities, which, according to theory, are industries that export goods beyond the boundaries of the area; the second class is composed largely of "service" or "region filling" activities, which are mainly local to the region. The service activities complement the base and react to changes in it. The forces of change, however, according to the theory, develop in the base industries, and, therefore, for purposes of analysis, the base industries are more significant.

There are close parallels in the procedures followed to project industry employment for states and for the metropolitan areas; at both levels of geographic detail, employment in basic industries, predominantly commodity-producing, was projected first relative to the larger geographic area; then employment in predominantly nonexport industries was projected using an export base multiplier. However, it is of great interest that the methodological similarities end when the population projections are developed. To project the population of states (and analytical regions) consistent with the industry employment projections, NPA developed essentially two independent sets of population projections, the first by expanding the employment projections to population projections. This was done by a straight-forward application of projected labor-force participation rates and employment results to the employment projections. The second set resulted from developing cohort-survival projections for each state. The two series were reconciled mainly by adjusting the net migration component of the cohort-survival projections. The importance of this procedure lies in its "constraining" function to the extent that in this way NPA could determine if the employment projections implied net migration patterns that were reasonably consistent with the past experience of the states under consideration. Net migration, the interface between population and employment at the subnational level, is useful here as a constraint on the system.

For the metropolitan areas, no independent demographic projections were developed. Rather, NPA made population projections by using the extrapolated relationship between employment/population for the metropolitan area relative to employment/population for the analytical region. These coefficients were applied to employment projections for the metropolitan areas to derive population projections through 1975. Whether the net-migration patterns implicit in the resulting population projections are "reasonable", that is consistent in sign and magnitude with historic experience, is impossible to say. Moreover, for reasons to be discussed later, it is extremely difficult to develop reasonable net-migration estimates that would correspond in a meaningful way with the NPA population projections.

Conceptual and Practical Problems

The NPA metropolitan-area projections are unquestionably a monumental and important work. They present for the first time a consistent set of industry-employment, personal income, and population projections for 224 metropolitan areas through 1975. However, I think that their pedagogical value may be at least as important as their value to policy planners who need "hard" projections as an element in the decision-making process. By pedagogical I mean that they can serve as a valuable teaching device to demographers, who for so many years have been secure with component-projection methods but who also have recognized the need for making and understanding the social and economic assumptions that underlie purely demographic techniques. The NPA projection reports demonstrate that the economic-demographic linkage can be effected at the national and subnational levels; Mr. Kupinsky's paper clearly describes how these projections were made. However, I am disappointed that neither the NPA reports in their well-documented methodological sections nor the presentation this morning discussed some of the serious conceptual and practical problems associated with this undertaking. I say this because demographers and planners for some time to come are likely to view the NPA projections as the "authoritative" source of internally consistent demographic and economic projections at the subnational level.

I will discuss some of these problems as they relate specifically to NPA's work and more generally to the problems of developing regional projections. These problems include (1) the definition of analytical regions, (2) the validity of export base theory for projection purposes, (3) the relationship between migration and employment, and (4) geographic-boundary problems in small area projections.

I want also to discuss very briefly the promise I see for more complex economic models, such as social accounts and input-output analysis, in relation to population projections; the importance of evaluative studies; and the direction that I hope future work in this area will take.

Defining Analytical Regions

Because the NPA projections are carried out within a national and regional framework, the research has had to face squarely the difficult problem of determining the relationship between the metropolis and its hinterland. This was done on the basis of several very general criteria relating mainly to the economic transactions between the metropolitan area and the region from which it draws its resources and which consumes its products. Kupinsky states that the projections are very sensitive to the geographic boundaries that establish the economic region to which the metropolitan area is related. In reviewing the text of the NPA projections publications, I see no mention of this problem. I feel that it would be well to develop in detail how the boundaries were actually established, how the economic functions played by the metropolitan centers relate to the analytical region, and--if the projections are highly sensitive to the definition of the relevant analytical region--, the extent to which confidence can be placed on the boundaries of the analytical regions.

Validity of Export Theory

The NPA projection methodology relies heavily on economic-base theory, as stated by Kupinsky, "The level of economic activity and of population growth in an area depends on the area's level of activity in certain 'basic industries' and the export component of localized industries". During the 1950's the validity of the economic-base theory was frequently, and I might add, effectively challenged. One study, for example, found no significant relationship between basic activity and population growth for selected areas. In fact, tests suggested that in the dual classification (basic and service industries) of economic activity, the service component might be a more important indicator of growth potential than the basic component.(2) In another study it was contended that the applicability of the basic-nonbasic concept tends to decrease with increasing size of a metropolitan area, and that large metropolitan areas exist, survive, and grow because their highly developed business and consumer services enable them to substitute new export industries for those that decline; it was argued that nonbasic industries are the permanent and constant element, in fact, the truly basic element of the metropolitan area economy, while the export activities were the more variable element, subject to continual change and replacement.(3) So far, Wilbur Thompson has had the last word on the problem on export-base theory; he maintains that in the short run, the primacy of the demand export products in effecting economic change in an urban area is uncontestable, but that, over the long term, the service sector becomes increasingly important.(4)

Two other points worth mentioning in this regard relate to classifying industries as export or residentiary and to the export multiplier. Studies have shown that the manner in which the classification of industry employment into the basic and nonbasic categories can have a significant effect on the employment projections, particularly if the "misclassified" industry has a predicted growth that is different from the predicted growth rate for the total primary, or basic, sector.(5) In addition, the base multiplier, which expresses the relationship between export- and nonexport-industry employment has been found to be unstable for some large metropolitan areas; this instability again can have a telling impact on the magnitude of employment projections that developed using this approach.(6)

I raise these questions about export-base theory not because I challenge its usefulness as an analytical tool, but because I feel that the validity of this method for making employment and population projections--despite its widespread acceptance and increasing use in regional studies--is still seriously open to question; these considerations should be brought to the attention of the users of the NPA projections.

Migration and Employment

It is generally recognized that there is a strong relationship between employment and population size on the subnational level and that the most important equating variable between the demand for labor and its supply is migration to and from the region. Other important variables are the labor-force participation rate, which expresses the net relationship between the labor force and population size; the unemployment rate which expresses the net relationship between the number of unemployed and the labor force; and the natural increase of the resident labor force. Studies on the magnitude and the timing of these linked relationships--while of considerable importance for making regional employment and population projections--have been extremely few in number. To my knowledge their findings have not been incorporated into any projection methodology in current use. Rather, a one-way nexus, expressed as a simple ratio, has usually been assumed between employment and population. Implicit migration effects are assumed to be treated as a residual following from the posited employment/population relationship.

I am in no position to evaluate the NPA projections in this regard--nor am I aware that evaluations of alternative methods, including the simple ratio assumption, have ever been made. The relevant points are that a model of population change can be developed in which migration is jointly and explicitly determined with employment change and that the resulting demographic projections may be more realistic than those developed from existing models.

Ira Lowry's recent study is extremely important in this respect.(7) Lowry examined the relationship between net migration and several other variables for 52 SMSA's between 1950 and 1960. He found that net migration was related: (1) negatively to natural increase in the resident labor force, (2) positively to changes in the number of resident military personnel, (3) positively to changes in the number of school enrollees 14-29 years, (4) negatively to changes in the median income of families, and (5) positively to changes in employment. The most impressive statistic was that 98 percent of the variance in migration could be explained by changes in employment, and that the model accounted for almost all the variation in net migration during the period. His model suggests that, on the average, an increment of 100 jobs is associated, *ceteris paribus*, with a net in-migration of 143 persons of labor-force age; while an increment of 100 residents of labor-force age would reduce the influx of migrations by 65 persons, rather than displacing it altogether. The importance of Lowry's monograph is in demonstrating, for metropolitan areas, the quantitative relationship between net-migration changes in the labor market expressed in terms of employment. Lowry shows how this approach can be adapted to projecting population, and presents a substantial argument for using this strategy rather than a ratio method, which lumps net migration, labor-force participation, employment status, and population into one coefficient whose magnitude has no structural meaning.

Geographic Boundaries

The considerations above are not specifically related to the NPA projections but bear more generally on the problems of making regional economic and population projections. The following comments relate directly to the NPA metropolitan projections. Kupinsky has stated that the NPA projections do not follow the practice of using constant SMSA boundaries:

> "Although metropolitan areas' statistical series for historical years readjusted to current boundary definitions may have many uses, we believe that such a series is inappropriate for a study of factors underlying

metropolitan area economic growth, since an extension of geographic boundaries is itself a means for accommodating such growth".

As a consequence of this procedure, the projected employment and population figures relate to an amorphous geographic area. This area presumably includes those counties defined as metropolitan at the benchmark date, but it may include additional counties by 1975. Given the NPA practice, there is no way to determine if the boundaries have changed during the projection period. I cannot see how this practice is consistent with defining economic regions (one or more states) by administrative boundaries, since surely the spatial characteristics of regional economic growth are the characteristics of metropolitan areas writ large. This is not to say that one can argue on this basis for flexible state boundaries but, rather, that fixed geographic boundaries are the constant about which we must build our analyses and projections--be they for states, regions, metropolitan areas or even cities. This is particularly true if, at some time in the future, net migration is to be treated explicitly, since migration has no meaning apart from explicit geographic boundaries.

In my opinion, the usefulness of the NPA metropolitan projections is seriously impaired by this procedure. Because area studies are usually carried out for well-delineated administrative units, with boundaries held constant during the study period, the regional analyst--demographer or economist--is constrained to work within fixed boundaries. The effects of the NPA "boundary problem" are reflected in the NPA projections, where, for a number of metropolitan areas, erratic population changes can be observed during the projection period. According to the projection report this can be explained as follows: "when a county is added to a metropolitan area (during the projection period),(8) it usually means a proportionately greater increase in population than employment. This is because much of the work force in the county would have been employed in the metropolitan area prior to annexation."(9)

Other Projection Models

A number of other regional-growth models are in development or operation, but none, to my knowledge, has been used to generate data in such geographic detail as the NPA procedures. With the exception of one, all of these models--whether couched in terms of the interregional input-output framework or in terms of regional accounts--develop population using the employment-to-population ratio or some simple variant thereof. The one exception, worth mentioning here, is a model developed by Stanislaw Czamanski for projecting employment and population in the Baltimore SMSA.(10) Czamanski attempts to incorporate into his model interaction effects between population and employment in a time-lagged model. In this model, the employment of certain basic industries is projected exogenously, the employment of complementary industries is related to the basic industries functionally (lagged relationship), and employment in industries whose main locational factors depend on the existence of the central city is related functionally to population (lagged); total population is a lagged function of employment in all industries combined.

This is not the place to evaluate the various methods for making regional employment and population projections. The great value of the more complex models, it seems to me, has been in elucidating relationships and in demonstrating the critical parameters and possible sources of error in the more aggregative models. In a sense, the complex economic models are to simple models as the component-projection methods for projecting population are to the logistic curve or other simple extrapolative methods. The component model enables demographers to view and control the interaction of births, deaths, and net migration; input-output analyses enable economists to view and constrain flows of goods among sectors of an economy where aggregate models would completely obscure these important relationships.

Whether the more complex methods will become practicable for making extensive regional economic projections is not clear at present. There are a number of conceptual and practical problems that may limit their use. These relate partly to data availability, partly to establishing future demand levels that are inputs to such models, partly to projecting critical coefficients in the models, and partly to constraining the models so that the internal relationships, as they are generated within the model, retain a semblance to the real world. Linking the complex economic models to population growth is a subject that can be considered independently; I feel that this linkage is of sufficient importance in regional studies to warrant considerable study.

Evaluative Studies

It is clear that, to use Mr. Hollman's colorful phrase, the "promised land of econometric models" is not near at hand, at least as far as small-area population projections are concerned. For the demographer in pursuit of a "best" method for making small area projections, there are still no guidelines. A paucity of evaluation studies on various projection methods has not been remedied during the 15 years since Siegel called it to our attention. His own work on evaluating projections for small areas(11) and that of Helen White on evaluating the accuracy of various methods for making state projections(12) still stand alone. Their results were not very comforting, but they show that, on the average, cohort-survival projections are better than other methods. In this regard, Ira Lowry has made a contribution by showing that an objective function linking net migration and employment yield more reasonable projections of net migration than the usual assumptions of constant net migration, used in short-term cohort-survival methods, if the employment projections are "good".

Conclusion

I have discussed Mr. Kupinsky's paper at length because I think that, as Mr. Hollman suggests, the NPA projections represent the beginning of a new period in history of demographic techniques. That demographic projections cannot be made in vacuo is beginning to be appreciated by demographers; the NPA reports represent a first major attempt to blend economic and demographic projections into a coherent and internally consistent whole. I consider their work very important in this respect, although I feel strongly about some of the shortcomings of their projections, particularly the problem of geographic boundaries.

Mr. Hollman's work is also important for it illustrates how the present state of the demographic art can be fruitfully applied at the county level. My hope and perhaps his is that the U.S. Bureau of the Census will soon undertake the task of making county projections using the cohort-survival method, with several assumptions about net migration, similar to the procedures they use in making state projections. If we are to progress in the small-area-projection field in the immediate future, I feel that evaluative studies are a first order of business. Perhaps the Bureau of the Census and the National Planning Association will train their extensive capabilities in this area.

---

(1) Siegel, Jacob, "Forecasting the Population of Small Areas", Land Economics, 29 (1953), 72-88.

(2) Pfouts, Ralph W., "An Empirical Testing of the Economic Base Theory", Journal of the American Institute of Planners, 23 (1957), 64-69.

(3) Blumenfeld, Hans, "The Economic Base of the Metropolis", Journal of the American Institute of Planners, 21 (1955), 114-132.

(4) Thompson, Wilbur R., A Preface to Urban Economics, Wiley, Baltimore (1965), 30.

(5) Gillies, James and Grigsby, William, "Classification Errors in Base Ratio Analysis", Journal of the American Institute of Planners, 22 (1956), 17-23.

(6) Ibid., Isard, Walter, Methods of Regional Analysis, Wiley, New York (1960), 200.

(7) Lowry, Ira S., Migration and Metropolitan Growth: Two Analytical Models, Chandler, San Francisco (1956).

(8) Parenthesis inserted.

(9) National Planning Association, Economic and Demographic Projections for Two Hundred and Twenty-Four Metropolitan Areas, Regional Economic Projection Series, Report No. 67-R-1, National Planning Association, Washington (1967), 39.

(10) Czamanski, Stanislaw, "A Method of Forecasting Metropolitan Growth by Means of Distributed Lag Analysis", Journal of Regional Science, 6 (1965), 35-49.

(11) Siegel, op. cit.

(12) White, Helen R., "Empirical Study of the Accuracy of Selected Methods of Projecting State Population", Journal of the American Statistical Association, 49 (1954) 480-498.

DISCUSSION

Richard Irwin, U.S. Bureau of the Census

My comments will relate chiefly to the paper "Economic Projections for Local Areas" by Graham, Garnick, and Olson. The paper is a worthwhile contribution to a growing literature covering projections of employment and income as related to projections of population. The paper falls into three main parts, the first describing the delineation of 165 economic areas of the United States for statistical purposes, the second describing a model for projecting employment, and the third presenting the results of empirical tests of personal income projections.

The 165 economic areas are combinations of complete counties grouped around an important city, with no attempt to observe State boundaries. It must be admitted that the States are not ideal divisions of the U.S. for economic analysis. For the purpose of projecting employment and income, the economic areas are very suitable since the input of employment and earnings by industry is available annually by county. However, at some point it is usually desirable to consider statistics from other sources, such as the decennial census. Only through special arrangement can data for these areas be developed from the decennial census. Furthermore, estimates and projections by the Bureau of the Census and other agencies often develop figures for States before developing figures for counties. Since the 165 economic areas cannot be grouped into States any comparison between projections for States with projections for the 165 areas would have to wait until county projections had been developed, and reassembled into the 165 economic areas.

It is a central thesis of my remarks that greater integration of demographic and economic projections by various agencies is desirable. Such integration in this case between the Census Bureau and the Office of Business Economics will be difficult or impossible if the economic areas bear no correspondence to States or even Standard Metropolitan Statistical Areas. We at the Bureau are now developing projections of population for all metropolitan areas to 1975 using a fairly sophisticated cohort-component model, projecting gross out-and gross in-migration separately. In doing this, we are projecting the areas according to the 1960 geographic definition. Mr. Kupinsky's paper discusses another set of projections for SMSA's for which the National Planning Association has introduced a flexible definition of metropolitan area boundaries, implicitly assuming that the geographic boundaries of the SMSA's will expand with expanding population. Thus we have three important sets of projections for economic areas smaller than States which disagree fundamentally with each other with respect to the geographic areas used, making comparison difficult or impossible. Is it possible to agree on a set of economic areas which cross State lines only where considerations of economic integration are overriding, as in New York, Philadelphia, Washington, D.C., and a few other places? Then some State data could be used, making special adjustments for these areas, and comparisons with data on widely varying subjects would be facilitated. Another alternative is a built-in two-stage operation, where all economic areas crossing State lines are calculated as a whole, but are automatically computed also as two parts, which are then forced into agreement with the whole. The various parts could then be reassembled into States.

The second major portion of the paper presents a model for projecting employment for the 165 economic areas. Employment is divided into two main categories, basic and residentiary. The former constitutes all of those activities "the products of which either flow in interregional trade or are otherwise determined outside of the region". These include mainly the products of agriculture, mining and manufacturing. Residentiary employment is that directed toward supplying local businesses and households with commodities and services which do not enter interregional trade. In the model basic employment is projected by a shift-share technique, while residentiary employment is developed as a function of total employment in the area.

This approach seems to be reasonable and worthy of analysis. If it is assumed that national employment totals by industry can be projected with reasonable accuracy, then the ability to predict the share which each area will enjoy of the nation's employment in a particular industry will yield an accurate projection of employment in that industry in that particular area. I would make one minor dissenting observation in that the model apparently uses one definition of basic employment for all areas. However, it is clear that industry sectors which are residentiary to one area are clearly basic to another. The examples that spring to mind occur in the field of entertainment, recreation, and education. However, this is not an overriding consideration and the model is well worthy of development and testing.

The third major portion of the paper describes the results of empirical tests of five sets of personal income projections by State. The implication is that these tests bear on the suitability of the employment model previously discussed. The tests are by State and not by economic areas, but this is by no means an insurmountable obstacle. A more important obstacle is that the model deals with employment, while all five sets of projections are of personal income, and only set No. 1 uses employment as an input. For this set the projection of employment by State, by industry, developed from the Harris shift-share model serves as an input in developing income from wages and salaries.

The projections of income from this rather complex model are out-performed by a relatively simple model using ratio techniques. Two questions suggest themselves. 1) Do the authors suggest that the empirical tests do reflect on the employment model previously discussed, and 2) does the relatively poor performance of the most complex shift-share income model as compared with a simple model suggest that shift-share analysis is not promising for employment or income projections?

These are questions of detail. The paper raises by implication broader questions concerning the relationship between projections of employment and population. The authors state that population projections should be made dependent on employment projections, since "the major factor underlying migration is economic opportunity or the lack thereof". Yet they make the point that the several sets of projections they have developed, while varying a great deal in the amount of detail introduced, all rest fundamentally on a technique of extrapolating past trends. The demographic population projections we have developed at the Census Bureau, although very detailed, also extrapolate past trends of migration, fertility and mortality. A strong correlation between employment change and net migration has been established by Lowry and Blanco. In the presence of such correlation, separate projections of employment and population based on the extrapolation of past trends should be highly correlated, assuming the same historical base period is used.

It would seem therefore, that as long as the basic projection technique consists of the extrapolation of past trends, less emphasis should be placed on which projection is developed first, and more on the manner of linking the two projections. Here it is clear that age-sex-specific labor force participation rates are indicated. These participation rates vary widely by age and sex, but the age-sex specific rates show surprisingly little variation regionally. Furthermore, long term trends in these rates can be readily discerned. Therefore, it should be possible to project these rates at least as effectively as migration and employment. These projected rates would then be applied to a demographic projection by age and sex to develop a projected labor force, and provide a satisfactory link with a projection of employment.

In conclusion, I believe it would be worthwhile to consider alternatives to the basic technique of extrapolating past trends. Rapid technological change in the past 30 years has drastically altered man's power to influence his physical environment. Future technological change should even more drastically increase this power. In this event, human attitudes and intentions as to place of residence will become more and more important in determining regional location of population. These attitudes and intentions should be investigated. As a beginning, we should consider jobs and migrants not only as numbers in a table, but also as <u>persons</u> who will behave in a certain way for certain reasons. These persons can be asked what they expect or plan to do in the future, and thus provide an alternative to the extrapolation of past statistical trends in making projections. Evaluation of the answers to such questions will not be easy, and years will be required in evolving a satisfactory use of such data in making projections.

II

ASPECTS OF THE 1966 AND 1971 CENSUS PROGRAMMES IN CANADA

Chairman, W. D. PORTER, Dominion Bureau of Statistics

THE TESTING PROGRAMME FOR THE 1971 CENSUS IN CANADA

I.P. Fellegi and K.J. Krótki
Dominion Bureau of Statistics, Canada

1. INTRODUCTION

Data on Canadian population have four somewhat unique, or at least rare, characteristics. First, they are based on the longest series of modern censuses stretching back just over 300 years.1/ Of uneven territorial coverage, of uneven subject-matter content, without legislative basis until 1841, the censuses were nevertheless being taken with remarkable frequency and make up to date an impressive list of 441 enumerations (7, p. vii). Secondly, the data refer to a population which among all the Western or developed populations had for most, probably all, of the historical period the highest fertility and consequently the highest proportions of children (21, p. vii). Thirdly, Canada conducted in 1956 and in 1966, between the main decennial censuses, national enumerations, even if with a rather limited subject-matter content.2/ Finally, all this wealth of demographic material has been used and analyzed to an extent which until a few years ago could be described only as modest, though the recent quickening of interest in problems of demography and social survey methodology in federal government departments, provincial governments, universities and research institutions, will soon, no doubt, fill this gap.

The testing programme for the next censuses in Canada should be viewed as part of this general development. Many previous censuses had their own test operations, particularly when departures from earlier content or earlier procedures were considered (e.g. 8, p. 17). These tests were held typically some 18 months before the census day. On occasions they could involve numbers as high as 100,000 persons. They were largely in the nature of dress rehearsals after which only minor changes were possible, i.e. they were not investigational. In particular three innovations should be mentioned which required extensive testing: the predistribution of the Agriculture Questionnaire in 1951, the partial completion of which must have lessened the enumerative burden on the enumerators; the introduction of mark sensing in 1951 with consequent machine reading and the postal check on a national scale in 1961 in urban areas which had appropriate mailing service.

In any case, these and other tests made their contribution to the gradual improvement of the census procedures and census results without, however, departing from the basic method of a door-to-door canvass and enumeration through interviews. An exception to this statement is the sampling form in 1961 which was dropped off by enumerators, completed with regard to the sensitive subject of income (it also included fertility and migration) by the respondent and then picked up by the enumerator. One of the most important improvements was due to the establishment of eight permanent Regional Offices after the last war to conduct the monthly Labour Force Survey. At the same time these offices became the focal points of the census-taking activity as well, resulting in an undoubtedly better selection, training and control of the census enumerators. Another major improvement was due to the gradual emergence of a more conscious and better-oriented training programme.

1.1 Motivation of experimentation

The first serious effort to measure errors in a Canadian Census was made in 1956. Much more ambitious programmes were mounted in 1961 and 1966 and more realistic estimates of coverage and content errors were obtained. Projects designed to evaluate the 1966 Census are the subject of the second paper of this session. Hence, they will not be discussed in the present paper. It should be noted, however, that the evaluation programmes of the 1956 and 1961 Censuses had a considerable impact on our thinking (9, 10, 13, 18). We were surprised to find that our censuses, after all the hard work that went into improving them, failed to enumerate 2.5-3.0 per cent of the population and that, in fact, the per cent under-enumeration went as high as 10 per cent for a few critical age-sex groups. We were surprised to find that the response variance of the census statistics on such questions as education, labour force status, industry and occupation was about as high as the sampling variance would have been with a 25 per cent sample (13). In addition, of course, these statistics were also subject to response biases as well. By far the largest portion of this response variance was accounted for by the so-called correlated response variance (16), i.e. roughly speaking the component due to the effect of an enumerator interviewing a substantial number of households (about 150 households were enumerated by one enumerator). In the 1961 Census we have used sampling to a limited extent, but it is substantially true that most of our census statistics had no sampling error but substantial response errors. Any cost-benefit analysis would indicate, that if we were able to reduce the response errors substantially at the price of introducing some controllable amount of sampling error, we would be better off.

We need not emphasize to this audience that the last ten years saw not only the development of some key experiments and important mathematical models leading to a better understanding of the limitations of our censuses, but it was at the same time a period of enormously increasing utilization of census statistics. Further, not only have our users become more numerous — they have become more sophisticated as well. Their needs are: more precise census statistics (even for relatively small areas and/or special sub-populations), measuring more characteristics, available sooner, available in various different forms (in published table form, on tape, on punch cards, in the form of graphs), capable of being followed up by special surveys and linkable with other data. It was considered essential to ensure that the 1971 Census incorporate a number of methodological changes to meet the requirements of census users during 1972-76.

Two of the impulses leading to our programme of experimentation have been mentioned so far: efforts to improve understanding of the census-taking process and the needs of users. A third important impetus should be mentioned: the experiments carried out by the Bureau of the Census. These experiments demonstrated that:

(a) it is feasible to establish an urban address register;

(b) the Post Office can collaborate in materially improving such a register;

(c) such a register can effectively be used for sampling purposes;

(d) if a census form is mailed to addresses on the register, then a large proportion of householders will complete the questionnaires in a machine-readable form, with relatively little follow-up; and

(e) that such a mail census with follow-up holds out important possibilities for reducing the response errors and at the same time making some gains in reducing the coverage errors.

The advantages and disadvantages of self-enumeration have been widely discussed (2, 33). We would like to emphasize only those of the advantages which loomed high in our thinking:

(i) the enumerators' contribution to the response variance should decrease sharply since the role of enumerators is restricted to follow-up;

(ii) the early return of mailed questionnaires to a central office 3/ permits an independent edit in time to trigger off an early follow-up, where necessary; this, we think might reduce both the response variance and the response bias;

(iii) each adult member of the household is able to answer the census questions for himself;

(iv) respondents are able to consult records;

(v) the publicity campaign can be made to "peak" during the mail-back period;4/

(vi) the cases of non-contact might be reduced for people who are difficult to find at home but whose mail will reach them, a consideration particularly important in view of the increasing proportion of women participating in the labour force;

(vii) coverage errors might be reduced, since each household would have several chances of being included in the census: during the preparation of the address register, during its improvements through subsequent (mainly, post office) checks and through the intensive probing during the enumeration process;

(viii) the address register, in machine-readable form, can facilitate geographic tabulations in table or graph form, it can facilitate linkages of census data with data from other sources and it can facilitate the taking of special follow-up surveys.

There was, clearly, a method of census-taking emerging, which held out important promises, at least in urban areas.5/ Throughout the first half of 1966, discussions were held which led to the decision about midway through the year to conduct an experiment using this method.

1.2 Some "boundary conditions"

The problem of developing, testing and implementing a new method of census-taking had some restrictions and conditions attached to it. To borrow a mathematical phrase, these were the "boundary conditions" of our problem.

The first unalterable condition was, naturally enough, that a census will have to be taken on June 1, 1971. Working backwards from this date, it appeared that by the middle of 1969 the method of census-taking, the content of questionnaires and all the important features of the field work will have to be "frozen". At this point a dress rehearsal will be held, but no major changes in procedures. The developmental and testing process will have to be carried out, therefore, during a two-year period.

The second "boundary condition" was the limitation of staff. In mid-1966 when the green light was flashed for testing, there was no organization or personnel available for full-time work on the test programme. The regular staff of the Census Division was working full steam on the processing and publication of the 1966 mid-decade census. A nucleus of full-time staff was borrowed, others had to squeeze in some part-time work. This staff situation in a very real sense determined the pace of our testing programme. The first test could not be scheduled for earlier than the fall of 1967, with a second series of tests in 1968, and finally the dress rehearsal of 1969. This schedule put a very heavy burden on the 1967 test. In fact, the 1968 tests will have to be planned largely, without the benefit of the results of the 1967 test being available.6/

1.3 Plan of the paper

After these introductory remarks the remainder of the subject will be discussed under four headings: an outline of and comments on the specific method of census-taking which was tested in September 1967; an outline of the evaluation programme carried out in conjunction with this test; comments on alternative methods of producing an address register; and a brief discussion of some of our future plans.

2. REASONS FOR AND OUTLINE OF THE LONDON TEST

2.1 Reasons for the specific type of test

The methodological changes which have to be tested before they could be relied upon include:

(a) self-enumeration with or without use of mail;

(b) use of address registers;

(c) automatic geographic coding;

(d) new computer hardware, including especially input devices; and

(e) new computer software.

Long discussions were held, and several position papers were written, covering a very large number of logically possible alternative methods of conducting and analyzing census tests. Without closing our eyes entirely to other methods, alternatives compatible with the main aspects listed above were chosen on grounds of intuition, common sense and experience elsewhere, mainly in the U.S.A. The impact of any major methodological changes on operational procedures and subject-matter content also needed testing. Such impact would be felt on field edits, the flow and handling of questionnaires, production of manuals, training and organization; all creating new problems and requiring new attention. Not all of them could, or indeed should, be tested in one 1967 test.

There was however no difficulty in selecting the obvious corner-stones for the structure of the testing programme. They were four in number:

(i) construction of an address register;

(ii) mailing-out and mailing-back of questionnaires;

(iii) self-enumeration on a questionnaire of sufficient length and complexity; and

(iv) locally organized editing and follow-up procedures to deal with total non-response and partially (and/or inconsistently) completed questionnaires.

Naturally, these corner-stones would provide the opportunity for testing a host of other aspects, some of which have already been indicated briefly in paragraph 1.1 above, but it was understood that they would have to give way, if pressure of work and need for other attention, did not permit going outside these four main purposes.

Clearly, such purposes excluded testing in remote areas with no mail delivery, in areas where the construction of address registers would be prohibitively expensive, and in areas with population centres so small that it would be uneconomic and unrealistic to organize a local office for the centralized edit and follow-up.

The questionnaire eventually used was a householder questionnaire, rather neat and FOSDIC readable. It was of two kinds: short with only basic questions for every household and a long one for one in every four households with two facing pages per individual and well over 70 questions, many with several sub-questions.

2.2 The London Test

The town of London in Ontario was selected out of eight candidates of comparable population size because of several characteristics which made it attractive for testing purposes. Its size (about 200,000 persons), its owner-tenant ratio (close to the national average), its considerable industry and occupation diversity, and its higher-than-average proportion of converted dwellings rendered it an attractive site for the test. Its low ethnic diversity made it unrepresentative, but it was decided to take on this problem at the next round of tests.

In the event the preliminary and impressionistic opinion is that we have chosen "too well". The co-operation of the public and local authorities was of a high order. The publicity given and received was favourable and positive. There was an eagerness on the part of the public to help. Just over 85 per cent of questionnaires were (3) returned by mail, the great majority within a few days. This compares favourably with results obtained in similar tests in the U.S.A. (4, 31). The Telephone Answering Service gave assistance to almost 7 per cent of householders (3) which is higher than expected (6, 17). This again can be taken as an indication of the public eagerness to do well. The selection of an "orderly" city was intended to give the test a better than a fair chance to become an operational success. A very low response rate would not only knock out the cost-benefit basis for a mail questionnaire, but would also probably knock out the whole idea altogether by showing that the Canadian public is not prepared to deal with a mail census.

Within the proportion returned by mail the proportion acceptable without further field work is the next factor determining the economics of the new method.

Within the proportion which fails edit specifications, the proportion of incomplete or inconsistent questionnaires which can be cleared through the telephone and which consequently requires no costly personal visit follow-up is the third important factor determining the economics of the new method. There are proportionately fewer telephones in Canada than in the U.S.A., but according to the telephone companies Canadians speak more and on the average, longer than natives south of the border. It cannot *a priori* be said whether the lower density will be made up by the apparently greater volubility or talkativeness and, therefore, ensure an economic follow-up by telephone.

It is not the purpose of this paper to dwell at any length on the complexity of the operations in the local office. To raise a large labour force at short notice for short periods, to train it in a large number of varied and complex tasks under supervisors who themselves are unfamiliar with the tasks (consider the setting of rates of pay as an example of the variegated problems), to design and work an organization which will process tens of thousands of pieces of paper, each with hundreds of entries, through many different steps, in many different places — these were all no mean tasks to be performed. The satisfaction that these unaccustomed labours can be carried out satisfactorily had to be secured.

It goes without saying that however successful a test may be as an operation and however promising its economics, it must first of all give satisfaction on two points: there must be improved coverage and a higher quality of subject-matter content.

2.3 Important differences between Canada and the U.S.

In a number of ways Canada is similar to the U.S. This is a source of great help to us, since it enables us to learn effectively from the numerous experiments conducted by the Bureau of the Census in the field of census-taking. We can avoid proven pitfalls and follow up the avenues that appear promising in the US experiments. Our colleagues in the Bureau of the Census have given us of their time and experience unstintingly, invited us as observers to their field trials and discussed with us with complete frankness both their successes and their failures. However, there are, in spite of the great similarities, important differences between Canada and the U.S. We shall list a few of the differences which are most significant from the point of view of census methods.

Our census has to be bilingual. Every citizen is entitled to complete his questionnaire in either of the two official languages. This means that if we want to have a mail census, then in certain parts of the country we have to mail out in the same envelope both an English and a French questionnaire.7/ Doubling the amount of paper to be addressed and mailed causes operational problems but, more importantly, it may cause some response problems as well. The long form is long enough as it is and it may well frighten some potential respondents. Two long forms, even if one of them can be thrown away, may be the straw that breaks the camel's back. There may well be some problems in the office as well since questionnaires have to be sorted by language for edit as well as telephone or field follow-up.

A second problem relates to the Canadian winter and to the fact that there is no commercial mailing list. There is no commercial need for it since our Post Office is willing to distribute unaddressed advertising material to householders. We have to prepare therefore our own mailing list. This means that we have to have in the urban areas two distinct field operations: first a listing job,8/ then the mail-out, mail-back census with field office edit and follow-up. One operation cannot smoothly blend into the other since after the listing of addresses and before the mail-out we need several months to key-punch these addresses, prepare the appropriate workloads, implement one or two Post Office checks. For a census date of June 1 the ideal listing time would therefore be some time in February or March. The Canadian winter being what it is, one tries to avoid extensive field work during these months. In fact, the last convenient opportunity is during October of the previous year. This has its disadvantages, however, since our list likely becomes more outdated during those eight months than it would be during three or four months.9/ We have to rely on other sources, primarily the Post Office to update our register, although some of the probing questions on the questionnaires themselves will hopefully improve coverage.

A third important difference relates to scale which this time works primarily to our advantage. The permanent field force of the Bureau of the Census has, as its prime function, the conduct of the Current Population Survey. Similarly, the main task of our permanent field force is to take the Labour Force Survey. These two household surveys have similar objectives and standards. Sampling being the undemocratic discipline that it is, we need about the same sample size as our American colleagues in spite of the fact that the Canadian population is only about one-tenth as large. This means, however, that we have about ten times as high a sampling ratio, ten times as large an experienced field force per capita. This field force can do the bulk of the October listing job, minimizing the problem of hiring and training for this additional task that we have. Equally as important, we may well be able to rely on the permanent field supervisory personnel to run the census field offices.

A fourth important difference relates again to scale, but this time it is disadvantageous. Planning top management and analysis is very little related to the scale of operations. Consequently, the per capita investment on this type of activity has to be much higher in Canada than in the U.S.A. This difficulty, as indicated earlier, is somewhat relieved by the possibility of drawing on the experiences of our professional colleagues elsewhere, particularly south of the border.

3. EVALUATION PROGRAMME OF THE LONDON CENSUS TEST

The evaluation of the London Test will, of course, be carried out on many fronts. Very important aspects of it will be based on judgement and observation: how orderly was the operation, were instructions followed, was it possible to adhere to the timetable of operations, could this timetable be condensed in some fashion, could such an operation be carried out on a national scale, etc. This type of evaluation provides the earliest assessment of the operation. Another key indicator, also available very early, is the response rate on short and long forms as well as the number of telephone and personal visit follow-ups on each.

The more formal evaluation of the test will be carried out under four headings. These are the evaluation of coverage, local office procedures, computer editing and content.

3.1 Coverage

A person or household can get into the count under the London Test procedures through five streams: the field listing, the quality control of field listing operation and some of the relisting triggered by the quality control, the advance Post Office check, the final Post Office check and follow-up which may be triggered off by responses to certain probing questions on mailed questionnaires.10/ One of the objectives of the coverage evaluation programme is to estimate the additions and deletions of basic addresses, subaddresses or households 11/ and persons from each of the five sources. Provision has been made to identify on the relevant records the source of additions and deletions. The total number of additions and deletions by source can also be tabulated by size of basic address, size of household, type of area (e.g., downtown versus suburban), etc.

An important aspect of evaluating the contribution to coverage of some of the operations is afforded by the quality control operations. The original field listing (15) was quality controlled as well as the advance Post Office check. The field listing was quality controlled through the relisting by supervisors of a sample of the more difficult blocks and the comparison of the supervisor's list with the lister's list (12). The advance Post Office check was quality controlled by withholding a sample of addresses originally listed and by checking whether the Post Office made up a "missed address" card for them (1).

The key measures from the coverage point of view will, of course, be the proportion of persons and households finally added by the combined procedures listed above and estimates of proportions missed in spite of these procedures. The total number of households or persons added will be broken down into the number of households in added basic addresses, as well as the number added in partially enumerated basic addresses. We shall also estimate the impact of definitional errors on the count of households. Similarly, the number of all persons in added basic addresses, added households in partially enumerated basic addresses and partially enumerated households will be estimated separately. We shall also attempt to determine some of the characteristics of the added persons.

Parallel to the measures of basic addresses, households and persons added through the five main streams, estimates of basic addresses, households and persons missed in spite of these procedures will be given. Some of these estimates will be based on the Post Office quality control, but the basic tool will be the Post-Enumeration Survey (PES). A brief description of the design of the PES might be in order.12/

The blocks of London were stratified according to the likely deficiency of the address register. The information for stratification was obtained by comparing the address register with other lists and noting the number of addresses on the other lists which were not on our address register. A stratified sample of 20 per cent of the blocks was selected. Within the selected blocks every second basic address was selected to yield a 10 per cent sample of basic addresses.

The first step of the Post-Enumeration Survey (coverage) was a very thorough field relisting of the basic addresses within the selected blocks about two weeks before the census date. This relisting of basic addresses will form the basis of our estimates of the number of basic addresses missed by the census as well as of the number of households and persons missed in completely missed basic addresses.

The next step in the operation was that of the reenumeration. The sample for the reenumeration consisted of the 10 per cent sample of basic addresses originally selected from the census address register plus all basic addresses which, on the basis of the relisting described above, appeared to have been missed by the census. Although from the point of view of sampling efficiency it would clearly have been advantageous, no subsampling was carried out within the basic addresses, since we felt that subaddresses and households are not sufficiently unambiguous to be used as sampling units (an exception to this rule was made in large, regular apartment houses). In fact one of the objectives of the reenumeration is to learn something about the definitional problems. The interviews and questionnaires were highly probing on coverage: both on coverage of persons within the households and on coverage of households at the same basic address. A series of questions were asked probing the de facto population as well.13/

On the content side the only questions asked of everyone in the PES sample were related to age, sex, marital status and relationship to the head of the household. Two housing questions of some coverage importance were also asked. A substantial amount of reconciliation was carried out to clear up discrepancies on both coverage and content between the census and the reenumeration survey. In addition, a few weeks after the reenumeration all persons who were apparently missed by the census were reenumerated once again on the long questionnaire. The objective of this second reenumeration was to find out more about the characteristics of persons who were missed by the census.

3.2 Local office procedures

The following major activities are carried out in the local office or directed from the local office: mail check-in, edit of questionnaires, telephone follow-up, personal follow-up and coding. Ideally we would like to quality control all of these operations, but at any rate at least we would like to evaluate them.

The check-in of the mail returns, since it was obviously an absolutely key operation, was 100 per cent controlled, i.e. all mail returns were checked and all discrepancies were reconciled.

The quality control of the edit operation took the form of acceptance sampling at the work unit level, i.e. all rejected work units were re-edited and the editors concerned were retrained. Only six editors out of 70 were affected by the quality control intervention and it is doubtful whether this operation was worthwhile in terms of its impact on quality. The speed of editing was too great to be caught by control. At best of times, through quality control operations we can control and estimate the average outgoing quality. However, since this operation was carried out in the field office in the heat of operations, the records of the quality control operation do not provide us with refined enough tools to evaluate in detail the edit operation and its impact on the final product. For purposes of evaluation, therefore, a 10 per cent sample of the "short" part of the questionnaires was selected and reproduced (actually this was necessary anyway for the Post-Enumeration Survey as described above). These reproduced questionnaires were edited by the Head Office Staff after the close of the London Office. These specially edited questionnaires were then matched with the "short" pages of the original questionnaires which, by this time, were through the regular processing. A comparison of the original questionnaires with the duplicate copies will enable us to evaluate in respect of the "short" pages the work of the various sections of the local office since the editors, the telephone follow-up enumerators and field follow-up enumerators were all using pencils of different colour and their contributions can be distinguished from each other.

There is, at the time of writing this paper, no plan to evaluate the impact of editors on the long questionnaires and those housing parts of the short questionnaires which will be reproduced for PES purposes, though no doubt some manual study of their interventions will be evolved sooner or later. However, the study of the multi-coloured dots (the "tallies") described in paragraph 3.4, although directed mainly towards assessing partial non-responders, may coincidentally provide for each, or some questions estimates of the proportion of entries made by respondents, by editors, by the two kinds of follow-up and the proportion of entries left blank (the proportion of entries completed by respondents has an important impact on the response variance). It may thus provide estimates for each, or some questions on the questionnaire of the effect of the various operations in reducing the N.A. rate (i.e., the proportion of blanks which should not be blank).

3.3 Computer editing

No firm plans have been made for a formal evaluation of the set of computer programmes used for edit and imputation. We created some possibilities, however, which we intend to follow up and which might help us in assessing these programmes. While the editing programmes are being readied, specifications for the comparison of the questionnaire tapes before and after are being worked out. The intention is to obtain a count of the interventions of the editing programme.

The set of programmes will automatically and without manual intervention create a fully edited file and it will make all the imputations in the case of missing or inconsistent information. A summary will be obtained at the work unit level of all imputations made by the computer. On the basis of these summaries the subject-matter statisticians can decide to pass the work unit or to intervene manually. The programmes will be assessed from several different points of view. The first point is: do they deliver as good a job as clerks? Other important questions are: in what ways should the programmes be altered? How much imputation, at what area level, for what questions should subject-matter professionals let pass without review and at what point should they review the imputations? We think that the process itself of looking for answers to these questions will be very useful on account of its educational effect.

A more formal evaluation of the effect of these programmes on the basic demographic variables of age, sex, marital status and relationship to head of household will take place through a matching of the Post-Enumeration Survey records with the edited census file.

An interesting and very important problem relates to the computer assignment of family codes. A household, which typically will report on one questionnaire, may contain more than one family. The information on the questionnaire does not explicitly reflect all the possible complex family relations, only the relation of each member of the household to one person: the head of the household. Names would provide a useful indication but names are not available in machine-readable form. They will be used by manual coders in coding families 14/ and then checked with the mechanical coding of families based on the few relevant pieces of information available for everyone: age, sex, marital status and relation to head of household.

3.4 Content evaluation

A considerable number of tests of the content of the questionnaire have been considered. They can be conveniently described by being grouped under three headings: those which are not likely to be undertaken (e.g., 5), those which will be undertaken with high priority in order to provide data for the decisions affecting the remaining parts of the testing programme, and those which will be done more thoroughly but which are unlikely to have an early impact on the future testing programme. We shall only outline the projects which will or are likely to be undertaken.

There are several urgent, high priority projects. A quick tally will be made, question by question, of blanks on a small sample of questionnaires (up to five were permitted by editing instructions on non-essential questions), of imputations by manual editors, of completions by telephone follow-up, of completions by personal visit follow-up, of completions by non-response follow-up. The proportions under each question will be combined with observers' impressionistic reports. This will be the immediate basis on which questions will be redrafted, rearranged and relevant parts of the questionnaire design changed. A larger sample of questionnaires will be analyzed more intensively. Summaries from the editing at the microfilming stage may throw some modest light on some content issues. Questionnaires obtained in the monthly Labour Force Survey in London will be matched, item by item, with the questionnaires obtained from the same households in the Census Test (about 400 households are affected of whom 100 were enumerated on a long form in the Census Test). As indicated earlier the coverage PES questionnaire has a few entries of interest from the content point of view. They will be compared with the answers of the self-respondents and analyzed for their content implications. Questions which have shown themselves to be particularly difficult and ambiguous are likely to be tested in the field, with two or three alternative wordings, early in 1968 to give some input for the questionnaire content in the 1968 test(s). All these are modest endeavours but limitations of time and personnel make any wider action inadvisable.

A more ambitious programme is envisaged for the later part of 1968 and for 1969 with long-term implications, but it is unlikely that its results would have much impact on the remaining parts of the testing programme, except possibly on the last dress rehearsal in 1969.

4. EXPERIENCE WITH ADDRESS REGISTERS

As explained earlier the commercial address registers, such as are available in Canada, are not likely to be comparable in comprehensiveness with lists available in the United States. However, investigations into the possibility of existence of city directories and their qualities continue. A commercial offer has been received to build up a list but at a unit price so low that the possibility arises that the firm does not appreciate the high standards expected from such a list.

A preliminary enquiry directed to over 50 urban centres with a 1966 population of almost 7 million (or some 35 per cent of the total population of Canada) revealed that most of them maintain an assessment roll or an electric utility billing list, usually both (28). A fifth of the lists (counting by population size) is in a state not easily transferable into machine-readable form and a quarter is "not available to outside users". Only half of the managers of the lists are "willing to supply an extra copy". It is not necessary to enquire into the extent of overlap between the fifth, the quarter and the half, because it is not unreasonable to expect that should the purpose be explained more fully, most of the lists would become available. It remains to be determined experimentally whether these lists can be used for the purposes of the 1971 censuses either as the core of the address register or only to support and strengthen a register produced otherwise.

Two address registers have been built up in the Bureau mostly from administrative sources (20), and the evidence available with regard to these two address registers is described in the following two paragraphs. Inasmuch as both registers have been constructed in the towns of Ontario and inasmuch as the municipal assessment rolls are important elements in both lists it is not certain how far the experiences are valid for assessment rolls in other provinces with different legislative requirements for these rolls.

While these rolls appear to promise the biggest immediate pay-off, our eyes are not closed to the possibility of utilizing some other sources as well, such as voters' lists, building permits, demolition permits, completion records of the Central Mortgage and Housing Corporation, City directories, postal lists of deliverable addresses (11).

In connection with address registers, reference should be made to their great potential for uses other than as a vehicle to mail out census forms. At least two important potential applications should be mentioned. One relates to the automatic assignment of geographic co-ordinates to the addresses in the register and through the register to the census documents. This capability might add new dimensions to our ability to retrieve census data for user-specified areas. Geocoding is the topic of another paper presented at this session (14). The other important potential application of address registers is during the intercensal period as a sampling frame for current surveys. This latter potential application is, of course, contingent on our ability to keep the register up-to-date at a reasonable cost.

4.1 Kitchener-Waterloo address register

The address register for the two neighbouring towns of Kitchener-Waterloo was built up in 1966 from the 1961 Census lists of households, the current electricity billing lists and the current municipal assessment rolls. The detailed technical operations which led to it and the results have been reported upon in several memoranda (11, 19, 24, 25, 26, 27). It was no big discovery that the out-of-date 1961 Census lists of households contained only 65 per cent of the addresses on the joint list (26).

Of the two other sources the assessment rolls are clearly superior (90 per cent as against 83 per cent) but they still miss proportions too high to leave to the postal check to make up. We intend to investigate why some addresses on the electrical billings and in the census list were missed from the assessment rolls.

The 37,000 addresses in Kitchener-Waterloo were checked by the Post Office letter carriers against the slots which they have on their sorting tables. Although some action or another by the letter carriers was required with regard to 6,000 addresses (e.g., there were apparently 2,000 duplicates), the genuine additions were a mere 2 per cent. However, because a third of about 1.3 per cent of addresses which were withheld from the letter carriers for quality control purposes were not reported as missing during the postal check it can be assumed that another 1 per cent has not been discovered (27).

4.2 London address register

The September 1966 listing of households in London already mentioned earlier provided not only the main source for the distribution of questionnaires in September 1967, but served also for the purposes of the study of address registers. The listing was carried out on principles very close to those of the monthly Labour Force Survey (15, 29).

Whenever possible the listing was to be done from external appearance, and enquiries inside households were not encouraged. The operation was quality-controlled in the hard-to-enumerate central part of the town (12). In such areas the over-all "error" rate was 12 per cent and some observers felt that this high rate could be combated only through enquiries inside the household.15/

In May 1967, the list was postal-checked and 4 per cent of new addresses were gained. Of addresses withheld from the Post Office (a sample of 1,319) almost a quarter was not reported as missing by the Post Office (1). It can, therefore, be assumed that there was another 1 per cent of addresses not found by the Post Office in this advance postal check.

The second postal check, conducted just before the D-day of September 12, 1967, added another 905 addresses, but missed again some addresses judging from the 264 householders who reported not having received mailed questionnaires and who were not on the address register.16/

A direct measurement of the completeness of the list will be attempted in the Post-Enumeration Survey. The address register has also been studied through a comparison with other lists in a manner similar to the study of the Kitchener-Waterloo address register (22, 30).

5. PRESENT PLANS

5.1 Further methodological tests

Address registers being the very cornerstone of any mailing operation, investigations into their reliability, alternative modes of building them up and their costing will continue. In some areas mailing is not feasible. In the absence of other suitable lists and on account of low population density, listing in such areas could be so expensive that only simultaneous enumeration could be considered with respect to short forms. In these areas we may have some experiments to drop off the long forms and ask respondents to mail them back.

In areas where mailing is possible, but a centralized operation to control editing and follow-up is not, a substitute method to carry out the editing and other operations from the back of the enumerator's car would have to be designed and tested. In such areas, the arrangements would have to be made through local talent and it would not be possible to reap the benefit of organization by our regional office personnel. The population concerned might be as much as 35 per cent of the country, if we exclude the 45 per cent in the 17 metropolitan areas and the agricultural or rural areas with, say 20 per cent where the need to take a Census of Agriculture simultaneously creates special conditions and requirements.

There is a host of methodological tests which, though important, are not likely to be undertaken because of shortage of personnel: publicity (the public relations circumstances of a national census cannot be satisfactorily simulated for a local test), questionnaire format and design (linear, columnar, page-per-person, etc.), influence of training (which type is effective with what kind of people?).

5.2 Content of questionnaires and mode of controlling its quality

Alternative wordings for several questions, where there is reasonable hope for some tangible results, are likely to be tried. A few new questions are likely to be tried on an experimental basis. These endeavours will be carried out in the two official languages of the country.

The attempts to discover the influence of editing instructions (both manual and mechanical) have already been described earlier, as well as other investigations leading to the assessment of the meaning of census questions and answers (e.g., 23). Another important question to which no answer will probably be available in time to formulate the 1971 plans is how to strike the balance between the expense of editing and the resulting content of the questionnaires. High-quality editing triggers off follow-up directly and proportionately.

5.3 Census users and their contributions

The main endeavours in the determination of questionnaire content are user-oriented. Same goes for the outputs, but these are too important topics to be treated at the end of an already too long paper. A separate section in the Census Division has been recently established to develop the understanding of users and our understanding of their needs.

FOOTNOTES

1/ The first nominal census of Canada was taken in 1666 and covered the then European population of 3,215 persons enumerated (8, p. 9).

2/ In the middle of the four earlier intercensal periods such censuses were held only in the Prairie Provinces experiencing then high migration.

3/ At least where the functioning of such a central office is a feasible operation. Where due to low density it is not practicable, edit decisions and follow-up decisions have to be taken in less controlled and less independent circumstances.

4/ In areas where a mail census is not possible, questionnaires have to be picked up or actually completed through canvassing by enumerators. In such areas the publicity in nearby large cities is likely to be a source of disquiet to conscientious respondents.

5/ We did not expect the new method to be a cure for all our problems. We were aware that the improvements in the 1960 US census were smaller than hoped for (32). Then there are always unexpected teething troubles, such as overimputations of large numbers (32) on a new mechanical device.

6/ Results of the 1967 test will begin to become available late in 1967, with most of the results becoming available in 1968, too late for the planning of the 1968 tests.

7/ But see footnote 15 for an alternative solution.

8/ Assuming that a field list of households is to be the prime source of the address register. As explained later, this is by no means certain. And, of course, this argument would be quite invalid should some kind of list-drop-off (at least the long questionnaire) -mail-back be adopted.

9/ Whether the winter months while the building activity is at its lowest make really much difference to an address register, remains still to be determined.

10/ Actually in London there was a sixth source: publicity. Some 264 respondents telephoned that they had not received questionnaires and were found not to have been on the list (another 243 were found to be on the list and were given duplicate questionnaires).

11/ In long-standing Canadian census parlance, households are identical with dwellings. It is hoped with the new concepts relating to address registers they will also be identical with subaddresses. A basic address is one street address which may have one or several subaddresses (apt. 1, apt. 2, etc.) within the same basic address.

12/ Before the description of the Post-Enumeration Survey, a qualification should be stated: it is the experience both in the United States and Canada that such evaluation surveys appear to be more successful in uncovering missed subaddresses and even more so basic addresses, but seem to be less successful in finding persons missed in partly enumerated households.

13/ This appears to be a less successful part of the main London questionnaires, at least it seems to be inferior to the back page of the New Haven questionnaire.

14/ The office coding of families (and the manual coding of the more intricate relations to head of household, as well as some other minor manual entries) delay the processing of the "short" part of the questionnaire. In the processing of the 1966 Census, families were created mechanically as they will be in 1971.

15/ No figures are available on the increases in costs due to such procedure. Enquiries inside households, if uniformally imposed, would incidentally provide an opportunity to determine whether households should be sent an English or a French questionnaire.

16/ See footnote 10 for other details.

REFERENCES

(1) Ashenkampf, H.V., "Postal Checks of Addresses (May 1967), London Test". Technical Memorandum (Census Test Series) No. 12. Ottawa: Dominion Bureau of Statistics, Census Division. (In print)

(2) Bogue, D.J., "The Pros and Cons of Self-Enumeration" Demography 2: pp. 600-627, 1965.

(3) Bulletin from the London Census Test, Dominion Bureau of Statistics, 1967 Census Test, September 26, 1967.

(4) Bureau of the Census, U.S., Daily mail receipts. Results Memorandum No. 18, Louisville Special Census, Washington, D.C., September 10, 1964.

(5) Bureau of the Census, Accuracy of Data on Population Characteristics as Measured by Reinterviews. Series ER 60, No. 4, Washington, D.C., 1964.

(6) Bureau of the Census, U.S., Report on Field Operations. Results Memorandum No. 31, Cleveland Special Census, Washington, D.C., May 1966.

(7) Census of Canada, 1931. Volume I: Summary. Ottawa: Dominion Bureau of Statistics, 1936.

(8) Census of Canada, Ninth, 1951. Volume XI: Administrative Report. Ottawa: Dominion Bureau of Statistics, 1955.

(9) Census of Canada, 1956. Volume III, Part II, Administrative Report. Dominion Bureau of Statistics, 1960.

(10) Census of Canada, 1961, Administrative Report of the 1961 Census, Bulletin 7.2-12, Ottawa: Dominion Bureau of Statistics. (In print)

(11) Chesney, C.M., Address Register Project: Visit to Kitchener-Waterloo. Technical Memorandum (General Series) No. 4. Dominion Bureau of Statistics, Census Division, September 1965.

(12) Dunn, M.R., The Quality Control of the London Listing of Addresses. Technical Memorandum (Census Test Series) No. 3. Dominion Bureau of Statistics, Census Division, February 1967.

(13) Fellegi, I.P., " Response variance and its estimation". Journal of the American Statistical Association 59 (308): 1016-1041, December 1964.

(14) Fellegi, I.P. and John Weldon, "Computer methods for geographical coding and retrieval of data in the Dominion Bureau of Statistics, Canada". Paper presented to the session on Aspects of the 1966 and 1971 Census Programmes in Canada at the annual meeting of the American Statistical Association, Social Statistics, to be held in Washington, D.C., on December 27-30, 1967.

(15) Gilfillan, C., 1967 Census London Listing Operation - An Account of Planning and Execution. Technical Memorandum (Census Test Series) No. 8. Dominion Bureau of Statistics, Census Division, June 1967.

(16) Hansen, Morris H., William N. Hurwitz and Max Bershad, "Measurement errors in censuses and surveys" Bulletin de l'Institut International de Statistique (Tokyo) 38 (2): 359-373, 1960.
(17) Kempster, A.J., Report on Telephone Assistance Service, Ottawa Pre-test. Technical Memorandum (Census Test Series) No. 4. Dominion Bureau of Statistics, Census Division, February 1967.
(18) Krôtki, K.J., C. Muirhead and R. Platek, "Evaluation Programme of the 1966 Census in Canada". Paper presented to the session on Aspects of the 1966 and 1971 Census Programmes in Canada, at the annual meeting of the American Statistical Association, Social Statistics Section, to be held in Washington, D.C., on December 27-30, 1967.
(19) Krôtki, Karol J., A match of the Kitchener-Waterloo Address Register against the 1966 Census Visitation Records. Technical Memorandum (Address Register Series) No. 6. Dominion Bureau of Statistics, Census Division. (In print).
(20) Krôtki, Karol J., Address Registers for a Mailing Census in Canada: sources, experiments, investigations. Technical Memorandum (Address Register Series) No. 7. Dominion Bureau of Statistics, Census Division. (In print)
(21) Kuczynski, Robert R., Birth registration and list of statistics in Canada, Washington, D.C.: The Brookings Institution, 1930.
(22) Methods Research Section, A match of the Field-listed London Address Register with 1966 Census VR's. Technical Memorandum (Census Test Series) No. 5. Dominion Bureau of Statistics, Census Division, June 1967.
(23) Mimoto, Hideo, Ottawa Pre-test: Housing and Household Questions. A Study of 479 Respondent-filled Returns. Technical Memorandum (Census Test Series) No. 8. Dominion Bureau of Statistics, Census Division. (In print)
(24) Portigal, A.H., A Description of the Address Register Project. Technical Memorandum (Address Register Series) No. 1. Dominion Bureau of Statistics, Census Division, July 1966.
(25) Sehdev, T.R., An Account of Time and Cost (of the Production of the Kitchener-Waterloo Address Register). Technical Memorandum (Address Register Series) No. 2. Dominion Bureau of Statistics, Census Division, September 1966.
(26) Sehdev, T.R., Evaluation of Sources in the Kitchener-Waterloo List. Technical Memorandum (Address Register Series) No. 3. Dominion Bureau of Statistics, Census Division, November 1966.
(27) Sehdev, T.R., Postal Check of the Kitchener-Waterloo Address Register. Technical Memorandum (Address Register Series) No. 4. Dominion Bureau of Statistics, Census Division, December 1966.
(28) Sehdev, T.R., Preliminary Enquiries into Existing Sources of Addresses. Technical Memorandum (Address Register Series) No. 5. Dominion Bureau of Statistics, Census Division, September 1967.
(29) Sehdev, T.R., The Ottawa Pre-test of the London Listing of Addresses. Technical Memorandum (Census Test Series) No. 2. Dominion Bureau of Statistics, Census Division, December 1966.
(30) Sehdev, T.R., A match of the London Municipal Assessment Roll with the Field-listed Address Register and the 1966 Census Visitation Records. Technical Memorandum (Census Test Series) No. 9. Dominion Bureau of Statistics, Census Division. (In print)
(31) Spear, Jean, Report on the New Haven Special Census Resulting from a Week of Observations (from April 9 till April 15, 1967). New Haven Special Census. Results Memorandum No. 16. 5 October, 1967.
(32) Taeuber, Conrad and Morris H. Hansen, "A Preliminary Evaluation of the 1960 Censuses of Population and Housing", Demography (Chicago) 1:1-14, 1964.
(33) Taeuber, Conrad and Morris H. Hansen, "Self-enumeration as a census method". Demography (Chicago) 3(1): 289-295, 1966.

EVALUATION PROGRAMME OF THE 1966 CENSUS OF CANADA

K.J. Krótki R.C. Muirhead R. Platek
Dominion Bureau of Statistics, Canada

1. INTRODUCTION

1.1 Summary

There are various ways in which methods of measuring errors and evaluating quality of censuses can be arranged. From the point of view of operations and execution, it is convenient to think of four groups:

(i) re-enumeration of a sample;

(ii) matching on a one-to-one basis against some other records;

(iii) comparison with summaries and profiles from other records on a macro basis;

(iv) analytical techniques based on internal consistency and method (iii).

The programme designed for the purposes of evaluating the 1966 Censuses of Canada 1/ has at least one enquiry under each of the four headings. Projects selected for the programme and subject to brief reports in this paper are:

1. Reverse Record Check (RRC); a match of a sample of 1961 Census records, immigration documents, birth certificates and records of persons missed in 1961 with the corresponding 1966 Census records.

2. Match with the Labour Force Survey (LFSM); a one-to-one comparison of the records of the monthly Labour Force Survey (LFS) with 1966 Census records.

3. Agriculture Quality Check (AQC); a sample re-enumeration.

4. Demographic Analysis (DEM); application of the techniques developed and used by Coale, Zelnik, Akers, Bogue and others.2/

5. Check of Lists of Households (VR-AR); match of 1966 Census Visitation Records (VR's) with available or specially constructed Address Registers (AR).

6. Study of Postal Change-of-Address Cards (C-of-A); a study of households reporting to the Post Office a change of their addresses around the time of census-taking.

The purpose of this paper is to summarize the six parts of the evaluation programme and present such results as may be available at the time of drafting this paper.

1.2 Purposes of the evaluation programme

The different parts of the evaluation programme have been constructed with different purposes, sometimes purposely overlapping, and of uneven importance. The purposes of the programme are:

(a) to provide estimates of coverage errors;

(b) to provide estimates of content errors or subject-matter errors;

(c) to locate areas of weaknesses in Canadian census methodology and, possibly, indicate means of strengthening them.

To recapitulate: three purposes have been stated, four groups of methods have been listed and the six studies of the evaluation programme have been very briefly described. This approach must be viewed from the twin and over-all purpose of assessing the reliability of the 1966 Census and providing evidence and ideas for improving, designing and carrying out of the 1971 Census. The programme can be summarized in tabular form using abbreviations suggested earlier:

METHODS OF EVALUATION	PURPOSES OF EVALUATION		
	(a) Coverage	(b) Content	(c) Operations
(i) re-enumeration	AQC	AQC	-
(ii) matching of records	RRC LFSM VR-AR C-of-A	- LFSM - -	- - VR-AR C-of-A
(iii) macro comparisons	DEM	DEM	-
(iv) analytical techniques	DEM	DEM	-

It will be seen from the summary that the purposes of the evaluation programme were oriented towards evaluation of the results rather than concerned with obtaining operational lessons for 1971. When the six studies have been completed, their results in related aspects will be compared and cross-analyzed.

1.3 Kinds of errors measurable in 1966

The 1966 Censuses of Canada were so-called full-count censuses. The usual problem of sample estimation and the estimation of sampling errors did not therefore arise. However, all other errors did arise. They can be shown schematically as follows with the contribution of the six studies indicated with again the same abbreviations:

Biases:						
(a) coverage biases	RRC	LFSM	AQC	DEM	VR-AR	C-of-A
(b) content biases	-	LFSM	AQC	DEM	-	-
Response variances:						
(a) simple	-	LFSM	-	-	-	-
(b) correlated	-	-	-	-	-	-

The blanks in this summary are due to the fact that the content of the 1966 Census of population was very modest (relation to head, age, sex, marital status) and enquiries into response variances on such limited and standard data would be unprofitable. It has been repeatedly shown, by many studies, including our own 1961 evaluation programme (24), that the four characteristics used in the 1966 Census have exceedingly small correlated response variances. With respect to simple response variance some estimates are likely to be provided by the LFSM.

2. REVERSE RECORD CHECK (RRC)

2.1 Objectives of the RRC

The RRC may be defined as a study which attempts to measure the proportion of persons not enumerated in the 1966 Census, by using a sample of persons selected from independent sources. These estimates are to be obtained for urban and rural areas at national and regional levels and by broad age-sex groups at the national level.

2.2 Sample design

Assume there was a complete list of all persons in Canada at June 1, 1966, independent of the 1966 Census. A sample could be selected from this list, the addresses of each selected person established, and by matching this sample of persons with the persons enumerated in the 1966 Census it could be ascertained whether each selected person was or was not enumerated in the 1966 Census. Using this sample an unbiased estimate of the number of persons missed in the 1966 Census could be derived.

Unfortunately no such list is available but it is possible to construct a close approximation to one. Taking the list of persons enumerated in the 1961 Census, plus the list of immigrants who arrived during the June 1961 - May 1966 intercensal period, plus the list of registered births for the same intercensal period, and finally the list of persons missed by the 1961 Census but detected by the 1961 Evaluation Programme (25), one would then approximate the Canadian population as of June 1, 1966.3/ It was from these lists that an independent sample of the Canadian population was constructed. This sample excluded the Yukon and the Northwest Territories.

Specifically the sample for the RRC was selected from the following lists:

LIST	Sample size
1. 1961 population enumerated at their usual residence (persons selected at random in two or three stages from non-self-representing units — NSRU,4/ from self-representing units — SRU,4/ from special areas and from Indian Reserves).	20,804
2. 1961 population enumerated away from their usual residence (but not enumerated at their usual residence; about one third of a sample enumerated away from their usual residence).	1,304
3. Births, 1961-1965 (sample selected systematically from microfilm copy of birth registrations for calendar year 1961 through 1965 5/).	2,632
4. Immigrants, 1961-1965 (sample selected systematically from Forms 1000 completed by each immigrant on arrival in Canada).	1,246
5. All missed persons in the 1961 Census according to the 1961 LFS - Census match.	549
	26,535

2.3 Tracing of addresses procedure

The selected persons in the sample will be referred to as "subjects".

The serial numbers, names and addresses of subjects selected from all the sources were transferred to tape and a print-out prepared by province, and within the provinces of Ontario, Quebec and New Brunswick by a French-English separation of addresses.

The second phase aimed at establishing the current address of each subject. A questionnaire with an explanatory letter and return envelope was sent, between June 1 and June 3, 1966, by registered mail to each subject at the last known address.

The questionnaire was either:

(a) answered as requested;

(b) presumably delivered by the Post Office, but not answered;

(c) returned by the Post Office to the DBS undelivered.

A follow-up letter was mailed to all subjects in category (b).

The results of the follow-up mailing were again classified according to the three categories. Over 26,300 letters were sent out on the first mailing of which some 6,800 were returned by the Post Office as non-deliverable. Of addresses apparently reached by the Post Office, 77 per cent replied. Another 20 per cent replied in response to a reminder, a total of 97 per cent. This is quite an extraordinary response rate for a mail operation. Table 1 summarizes the success of the first two mailings. (For Table 1 see end of this text.)

Letters returned by the Post Office as a result of the two mailings were of the following two kinds:

(i) they had an unambiguous street address, or

(ii) the street address was ambiguous.

In all cases under (i) a letter was sent addressed 'to the householder' at that address asking for some information about the subject in the study. A few more subjects were located through this mailing.

Some 3,600 subjects were turned over to the Bureau's Regional Offices for further tracing. As a result of this tracing, slightly more than 2,900 subjects were located or about 80 per cent. Supplementing mailings to subjects who did not respond and for letters returned by the Post Office, a search of some large administrative files was instituted to obtain the current address.

Table 2 gives the success rates for the three mailings and for the regional and administrative methods of searching. It is important to note that positive replies do not indicate that the subject was enumerated in the census. They merely provide addresses which can be used for searching of 1966 Census documents. (For Table 2 see end of this text.)

It will be seen that of the 26,535 subjects only 738 or less than 3 per cent could not be located. It should be understood, however, that the success rate varied substantially from list to list. The next stage was the location of the 1966 Census records of the (26,535 - 738 =) 25,797 whose addresses were found.

2.4 Searching among census records

The searching of the census records for the 97 per cent of subjects whose addresses had been ascertained was very intensive.

First, the enumeration area (EA) containing the given address was located and searched. If this failed, all surrounding EA's were also searched. If the address was not specific enough (such as a rural route), all EA's which might possibly contain the given address and all surrounding EA's were searched.

The search currently continues and to date (early December 1967) a high percentage of the 25,797 subjects have been identified among the 1966 Census records. It is anticipated that by February 1968 the search will be concluded.

3. MATCH WITH LABOUR FORCE SURVEY (LFSM)

3.1 Design of the project

The project consisted of matching one-half or about 17,500 of the LFS households, with the corresponding 1966 Census households. For coverage of households, it was a one-way match, i.e., for each LFS household, the corresponding census household was searched for and either matched or not. However, within the matched households, persons were matched both ways, i.e., census enumerated members of a household were matched with members of an LFS household and vice versa. The match was undertaken for the four 1966 Census characteristics (relationship to head, age, sex, marital status).6/ Any discrepancies between the number of persons enumerated and between the characteristics reported were reconciled by enumerators in the field.

The attempt to find LFS households among the census households was carried out in stages using such characteristics as the street name and house number; the name of head of household, particularly important in small villages and on rural routes; composition of household, and other similarities in households in the EA. Households which could not be located in the given EA were searched for in adjacent EA's.

To reconcile differences, the enumerator was instructed to read the discrepancy aloud to the respondent and record his answer. The reconciliation form was mailed back to the Regional Office where it was checked and coded.

The data obtained from the match and the reconciliation of discrepancies was weighted to obtain estimates of the number of households and persons by various characteristics for each province. The weighting used was, generally speaking, the one employed in the LFS (9).

3.2 Coverage of households

The purpose was to estimate the coverage of households in the census. The data was tabulated by size of households and such characteristics as farm households, non-farm owned households, non-farm rented households and family households, non-family households. Four such tabulations were produced:

(a) all households in the LFS;

(b) LFS households matched with the census completely;7/

(c) LFS households matched with the census partially;8/

(d) LFS households not found in the census.

Separate tabulations were formed for self-representing units (SRU) 4/ and non-self-representing units (NSRU) 4/ of seven regions and Canada. From these tabulations Table 3 with "blown up" values was formed for each region. (For Table 3 see end of this text.)

The value b is not obtainable from the match as households were matched one way only. It has been estimated from (a + b) - a = b where (a + b) equals all households enumerated in the census and a equals all households enumerated in the census and in the LFS.

The value d can be estimated through the Chandra-Deming formula (5) where

$$d = \frac{b\,c}{a}$$

The applicability of this formula involves assumptions, the strength of which depends, among others, on the quality of the field work and the success of the matching operation. This is not only apparent after a moment of perusal, but has also been shown empirically (17, 18) for unimportance of d; (20, 21) for the importance of d.

After forming the (a, b, c, d) table, as explained above, the following measures can be estimated for households

1. rate of under-enumeration in census $= \frac{c + d}{n}$
2. rate of under-enumeration in LFS $= \frac{b + d}{n}$
3. net rate of under-enumeration in census and LFS $= \frac{b - c}{n}$

3.3 Coverage of persons

To estimate coverage of persons in the census and in the LFS, tabulations similar to those produced for households were produced for persons:

a_1 enumerated for both the census and the LFS in completely matched households;

a_2 enumerated for both the census and the LFS in partially matched households;

b enumerated in the census, but not in the LFS;

c enumerated in the LFS, but not in the census.

From these tabulations Table 4 with "blown up" values was formed for each region. (For Table 4 see end of this text.)

Values a, b and c in Table 4 are obtainable from the tabulations described above. Other values from Table 4 can be obtained as follows:

d can be estimated in the same way as in Table 3;

e can be obtained in one of two ways; let H be the average size of household in corresponding region and let B be the value of b in Table 3, then e = BH; alternatively, e = (a + b + e) — (a + b), where (a + b + e) is a census value, a is obtainable from the household value of a in Table 3, b has just been obtained from the tabulations for Table 4;

f can again be obtained in a manner similar to e;

k can be obtained in one of two ways; let D be the value of d in Table 3, then k = DH; alternatively, apply the following approximation to the Chandra-Deming formula (5):

$$k = \frac{(b + e)(c + f)}{a} - d$$

The writers of this paper prefer the alternative ways of estimating e, f and k because the use of average household size to estimate persons omitted in enumeration must be very deceptive.9/

From Table 4 various measures of under-enumeration can be obtained for persons similar to the three measures arising out of Table 3 for households. In addition, measures of under-enumeration can be obtained separately for persons missed (i) within enumerated households and (ii) in entirely missed households.

To estimate the content error in household or personal characteristics, only households and persons in a of Tables 3 and 4, respectively, can be used. These households and persons depending on their success in reconciliation fell into the various categories of Table 5. (For Table 5 see end of this text.)

From Table 5 the following measures can be obtained (cf., (13, 23)):10/

1. rate of misclassification of characteristic "i" by LFS

   $(b_2 + c_1)/2n$ or b_2/n or c_1/n

2. rate of misclassification of characteristic "i" by census

   $(b_1 + c_2)/2n$ or b_1/n or c_2/n

3. gross misclassification of characteristic "i" by LFS and census

   $(b_1 + b_2 + c_1 + c_2)/2n$

4. net misclassification of characteristic "i" by LFS and census

   $(b_1 + b_2) - (c_1 + c_2)\ /2n$

4. AGRICULTURE QUALITY CHECK (AQC)

4.1 Objectives of the AQC

The objectives of the Agriculture Quality Check were:

(a) national and regional measures of the bias of the Census of Agriculture results for a restricted number of items; these were selected on the basis of their importance and their prevalence in Canadian agriculture;

(b) the study of the characteristics of farm holdings missed by the census;

(c) the study of the magnitude and direction of errors in reporting at the individual farm level.

4.2 Sample design

A full description of the sample design has been provided elsewhere (12). Here a brief summary will suffice. Rural municipalities containing farms according to the 1961 Census of Agriculture were listed in each province within type-of-farming strata. A municipality belonged to a stratum if 70 per cent or more of the commercial farms received 51 per cent or more of the farm income from the given stratum product-type. Within the strata so delineated, a substratification was imposed recognizing the importance of secondary income sources, similarity in the size of the farm enterprises, and geographic contiguity.

Independent selections of area segments were made within each major type-of-farming stratum, the allocation of the sample within a stratum being proportional to the 1961 Census farm count. The sampling ratio was constant for each of the provinces within a region but variable among the four regions recognized for estimation purposes. Table 6 provides a summary of the sample size by region. (For Table 6 see end of this text.)

Two rules were formulated: the open segment rule and the weighted segment rule. Both are described in (12).

4.3 Field and office procedure

The actual field operation began on July 11, 1966 and was completed in all provinces by August 15, 1966. The specially trained AQC enumerator located the boundaries of a segment using a topographical map and the latest available aerial photograph. He identified all agricultural operations carried on in the segment and sketched directly on the photograph all relevant boundaries within the segments.

In order to develop efficient ratio estimates of the net census errors, it was necessary first to compile the census statistics for the sample segments. This was affected by superimposing the segment boundaries onto the appropriate census EA maps. Then an intensive match operation followed to ensure that all related AQC and census schedules were brought together. Where a pronounced difference between the AQC and census questionnaires (according to an arbitrary rule) existed, a reconciliation was conducted by mail with a second mailing for non-response. Any outstanding delinquencies were then followed up by DBS regional personnel.

4.4 Some numerical results

Estimates of the net error in 1966 Census statistics with their sampling errors are shown in Table 7 for a selected group of items in the Maritime Provinces. Data for other provinces will be published as it becomes available. (For Table 7 see end of this text.)

The AQC estimated that there were 25,900 agricultural holdings meeting the census-farm definition compared with the 24,700 holdings actually enumerated in the 1966 Census enumeration of the Maritimes. Therefore, the net error in the census-farm count is estimated to be 1,200 farms, a 4.6 per cent under-enumeration.

The standard deviation of the estimates for eight characteristics are presented in Table 7. Only acreage owned, oats for grain and potatoes have census totals which are not statistically significantly different from the corresponding AQC estimates.

The relatively larger under-enumeration of 7.2 per cent in unimproved land and woodland indicates that there is considerable reporting or content error for this item. The census definitions are vague and require tightening up if more accurate responses are to be obtained by enumerators in future censuses.

It appears that significant under-enumeration arises when the characteristic in question is not readily identifiable, i.e., hayland and cropland. A tendency to omit reporting doubtful cases is evident. A superior enumeration is apparently performed on higher value crops such as potatoes where the acreages are small and less likely to be forgotten by the respondent.

5. DEMOGRAPHIC ANALYSIS

This section of the paper gives a preliminary view of errors in the census by age and sex and describes the difficulties of applying in Canada the (standard by now) techniques of measuring such errors:

(i) the components-of-change method;

(ii) Coale's iterative method (6);

(iii) the survivors-of-births method;

(iv) the mortality-rate method.

The work of investigating the 1956 and 1961 distributions has now been almost completed. The analysis of 1966 is slowed down by a delay in the appearance of single years of age distribution. Many of the results or findings parallel those of other countries, but there are distinct Canadian features to most of them.

The curve of sex ratios, measured here by masculinity ratio by age, which should bend downwards towards the right because of the force of mortality, does so slowly, uncertainly, and in some provinces less than in others; slowly, because of the male immigrants some decades ago who still survive in ages where women would be otherwise predominant; uncertainly, because more recent irregularities could be due to either continuing poor age and sex reporting during census or to continuing importance of immigration and age selective emigration (19). The curve of masculinity ratio in the Prairie Provinces, particularly in the case of Saskatchewan and Alberta, actually rises except for the very oldest ages. These provinces were, of course, in receipt of particularly pronounced in-migration (that is marked relatively to their population size) in the past. One is reluctant to make the usual hypothesis about age and sex selective under- and over-enumeration when various parts of the curves behave unusually.

There seems to have been hardly a lowering in the high censal survival ratio for ages 10-14 and 30-34 reflecting little improvement in the under-enumeration of children and persons aged 20-24 in the previous decennial census. Actually the most severe undercount appears to be among persons aged 18 through 23 (judging by single years of age distribution) (19).

We are spared in Canada the embarrassing (1, 11) balance based on compensating errors of the components-of-change method, because of the great and decisive uncertainty concerning migration. Instead, we have a healthy discrepancy, which can lead to high (15) or low (22) estimates of emigration as well as to interesting discussions in learned publications (3). The method in Canada also suffers from the absence of a correction factor comparable to the USA estimate of under-registration of births.

The reliability of Coale's iterative method is also limited by the lack of such an estimate of birth under-registration. It is further limited in Canada by the fact that Coale's assumptions are particularly strong in Canada. They involve alternatives so different that they qualify the exercise decisively and lead eventually to quite different results. Errors become cumulative as one moves along the age scale.

Until very recently births in Canada were markedly higher than in the USA and as such a more volatile element in estimating the net census undercount than in the USA. We already indicated that immigration and emigration are also proportionately more volatile in Canada. Although this enquiry is not yet completed, it can be reported that there has been no obvious improvement in the apparent net undercount of, say, 2.5 per cent over all ages, which has risen to much higher levels of possibly 15 per cent in some age groups (20-24, male). It is startling to see that these independently arrived estimates are quite close to estimates from a direct re-enumeration exercise, at least in 1961, when such a study was carried out (8, 25).

In view of the difficulties of using the standard techniques with Canadian populations, greater reliance will be laid on hybrid approaches involving both analytical methods and direct measurements.

6. CHECK OF LISTS OF HOUSEHOLDS

6.1 Objectives of the check of lists of households

The lists of households produced by enumerators in the course of their enumeration for the purposes of the 1966 Census in the form of books of Visitation Records (VR's) were checked in three cities of Ontario (Kitchener, Waterloo and London) against several other lists of different kinds available for these three cities for about the same point of time. The other lists were produced for the purposes of testing new methods for the 1971 Census and are described in another paper presented to this session (10).

The purpose of this section of the present paper is to report upon findings of this investigation inasmuch as it casts light on the quality and nature of the 1966 VR's. There is no reason to think that the other lists with which the VR's were compared are in any way inherently superior to the VR's. As is usual with a matching operation there was in each exercise with matching two lists the very large number of households common to both lists and the two other much smaller groups of households contained in one list only.

There was, of course, the fourth category of households missing from both lists, but no estimate is suggested for this fourth category for a number of reasons. There are some slight differences in the timing of the compilation of the lists (26). There are some slight differences in the timing of the postal checks and the burden thrown on them (2, 27). There are doubts about the uniformity of definitions of households used by the various lists. Sometimes all the additions suggested by letter carriers were made, sometimes only after a scrutiny. (Deletions were never made without a scrutiny.)

The three cities from Ontario will be supplemented by further investigations, including field work outside Ontario which will afford an opportunity to enquire into the 1966 Census listings of households in other provinces.

6.2 Some numerical results

The results of the first match show rather high rates of omissions of households in the 1966 Census. The omissions from the address register (AR) based mainly on MAR all apparently lower, but those from a listing conducted for the purposes of a recent text are even higher. The omissions of households from the census, particularly when they involve whole structures, will be investigated in the field in January 1968. Disaggregated, the omissions in the suburbs appear to be lowest in the VR's, highest in the MAR's, while our own field listing falls in between. For the centres of the towns the order appears to be reversed.

The comparison is made difficult by a number of reasons described, but particularly by the advisability of not taking seriously postal advice about deletions as a safety measure against losing "true" addresses. This safety measure is sensible with any one list. It avoids losing a true address at the cost of the unimportant inconvenience of keeping in the list some non-existing addresses. It is, however, disastrous, even if only in an apparent way, for any other list with which the comparison is being made and which shows as misses in the other list the non-existing addresses from the first list.

The purpose of Table 8 is to show the differential impact of postal checks. (For Table 8 see end of this text.) From other experimental work we know that the postal checks are about two-thirds effective (2, 4, 27). In row 2 (iii) of Table 8 one half of one per cent additions were made which suggest three quarters of a per cent for all omissions. From other studies we are conditioned to expect postal corrections of the order of two per cent. Four alternatives arise. The postal check during a census is less effective than during a census test for some inherent reasons not yet understood. Or VR's, have genuinely a better coverage. Or, judging from row 4 of Table 8, the postal suggestions of additions arising during a census test are taken literally and too seriously. That is to say, during a test, the census check which is being applied to postal advice is less effective. Consequently, the weeding out in rows 2 (i), 2 (ii) and 2 (iii) of Table 8 is not available to tests resulting presumably in padding with non-deliverable addresses. Fourthly, there is under-enumeration during a traditional census due to reluctance to accept postal advice. Field work now carried out should throw light on these questions in the near future.

7. STUDY OF POSTAL CHANGE-OF-ADDRESS CARDS

7.1 Summary of the study

When a household moves from one location to another it is expected that a Change-of-Address Card will be completed by some member of that household. The card indicates both the old and new addresses and the expected date of change of residence. When the card is deposited with the Post Office the information on it is used in the transfer of mail from the old to the new address.

The purpose of this study is to find the extent to which the coverage among households that moved during a period, which included the census enumeration date, was different from the rate among all other households in the population and to study whether in future censuses the coverage of these households could be improved.

Change-of-Address Cards covering the Ottawa-Hull metropolitan area were obtained from the Post Office for the period May 15, 1966 to June 15, 1966. These cards were then matched with the census population documents to ascertain whether the relevant households were enumerated in the 1966 Census.

Of the 2,443 cards obtained, 170 were either duplicates or contained addresses in a form which could not be searched for in the census records, for example, business addresses.

7.2 Results of the study

The under-enumeration rate among movers was more than five times greater than for the population as a whole, and although the over-enumeration rate among movers was also greater than for the population it was approximately one-sixth the magnitude of the under-enumeration rate (cf., 14). The net error in counting movers is, therefore, highly biased in the direction of under coverage. (For Table 9 see end of this text.)

For the 363 movers who were not enumerated by the census at either address, a study of the effective date of change of address revealed that 171 of them should have been enumerated at the old address and 192 at the new address. Approximately 59 per cent of the dwellings at the old addresses were enumerated as occupied but not by the mover, while 80 per cent of the dwellings at the new addresses were so enumerated. In the remaining dwellings which were mostly classified as vacant, the chance of the mover being included in the census at either address was, of course, low.

To measure the timing accuracy of the census enumeration of movers a check was carried out to determine how many movers were enumerated but at the wrong address.

A study of the effective date of change of address indicated that eleven movers who moved before the census date were enumerated at their old addresses, although they should have been enumerated at their new address, and 487 movers who moved after the census date were included at their new addresses although they should have been enumerated at their old addresses. (For Table 10 see end of this text.)

As the census enumeration begins on June 1 and continues for approximately three weeks it must be assumed from a reading of Table 10 that persons who move after June 1 are frequently enumerated at the address at which they were residing at the time of interview whether or not this was their usual place of residence at June 1. The conclusion suggests itself that enumerators do not enquire of respondents where they were residing on June 1 but assume that the current address was also the address where they were residing on the census date.

FOOTNOTES

1/ The census date was June 1, 1966 and the censuses of population, housing and agriculture were conducted simultaneously. As explained later the population procedure was limited to a few basic characteristics only.

2/ While the best known work here has been done in the USA, it will be recalled that quite startling results were obtained for India (e.g., 7) and Pakistan (e.g., 16).

3/ Such a list, of course, included emigrants and persons who died during the intercensal period. In subsequent matching "emigration" and "death" were sufficient explanations for non-matches.

4/ SRU's are towns with 1961 population of 15,000 and over and NSRU's are areas other than SRU's.

5/ The documents for the first five months of 1966 were not available. The first five months of 1961 received double coverage.

6/ Minor differences in age groups were ignored. Marital status reconciliation was limited to single — all other.

7/ Completely matched households are defined as households which contain identical persons.

8/ Partially matched households are defined as households which do not contain equal numbers of persons nor are all the persons reported necessarily the same.

9/ On the other hand, a bias may have been introduced due to the correlation between weighting in the LFS and the characteristics estimated.

10/ In the estimation procedure, it was assumed that when reconciliation was not possible, the LFS was correct, i.e., in each case c_2 should be interpreted as $c_2 + c_3$ and b_1 as $b_1 + b_3$. The actual figures involved were quite small.

REFERENCES

(1) Akers, Donald S., "Estimating net census undercount in 1960 using analytical techniques", Paper presented at the annual meeting of the Population Association of America on May 5, 1962. Summary of paper appeared in Population Index (2): 236, 1962.

(2) Ashenkampf, H.V., Postal Checks of Addresses (May 1967), London Test, Technical Memorandum (Census Test Series) No. 12. Ottawa: Dominion Bureau of Statistics, Census Division (in print).

(3) Buckley, Kenneth, "Population and migration", Section A in Historical Statistics of Canada, ed. by M.C. Urquhart and K.A.H. Buckley, Cambridge: at the University Press, Toronto: the Macmillan Company of Canada, Ltd. 1965, pp. 1-13 and Series A 244-253 on p. 22.

(4) Bureau of the Census, U.S., Tests of Use of Post Office Resources to Improve Coverage of Censuses. Working paper No. 19. Washington, D.C., 1965.

(5) Chandra, C., Sekar (now: Chandrasekaran) and W. Edwards Deming, "On a method for estimating birth and death rates and the extent of registration", Journal of the American Statistical Association, 44(245): 101-115, March 1949.

(6) Coale, Ansley J., "The population of the United States in 1950 classified by age, sex and colour - a revision of census figures", Journal of the American Statistical Association, 50 (269): 16-54, March 1955.

(7) Coale, Ansley J. and Edgar M. Hoover, Population Growth and Economics Development in Low-Income Countries, Princeton, New Jersey: Princeton University Press, 1958.

(8) Fellegi, I.P., "Response variance and its estimation", Journal of the American Statistical Association, 59 (308): 1016-1041, December 1964.

(9) Fellegi, I.P., G.B. Gray and R. Platek, "The new design of the Canadian labour force survey", Journal of the American Statistical Association, 62 (318): 421-453, June 1967.

(10) Fellegi, I.P. and K.J. Krótki, "The testing programme for the 1971 Census of Canada", Paper presented to the session on Aspects of the 1966 and 1971 Census Programmes in Canada at the annual meeting of the American Statistical Association, Social Statistics Section, to be held in Washington, D.C., on 27-30, December 1967.

(11) Good, Dorothy, "Preliminary results of 1960 and 1961 Censuses", Population Index 28 (1): 3-12.

(12) Graham, J.E. and R.C. Muirhead, "Evaluating the 1966 Agricultural Census", Canadian Journal of Agricultural Economics, Volume 15, No. 2 (in print).

(13) Hansen, M.H., W.N. Hurwitz, and L. Pritzker, "The estimation and interpretation of gross difference and the simple response variance", Contributions to Statistics. Ed. C.R. Rao. Oxford, N.Y.: Pergamon Press, 1965, p. 111.

(14) Jones, Charles and Walter M. Perkins, Coverage Evaluation for Persons Who Moved Between May 8 and May 25 Within Louisville SMSA. Results Memorandum No. 43, Bureau of the Census, Washington, D.C.

(15) Keyfitz, Nathan, "The growth of Canadian population", The Population Studies (London) 4(1): 47-63, June 1950.

(16) Krótki, Karol J., "Population size, growth and age distribution: fourth release from the 1961 Census of Pakistan", The Pakistan Development Review (Karachi) 3(2): 281-305, Summer 1963.

(17) Krótki, Karol J., "First report on the population growth estimation experiment", International Population Conference, Ottawa 1963, Liège: International Union for the Scientific Study of Population, 1964, pp. 159-174.

(18) Krótki, Karol J. and Ahmed Nazir, "Vital rates in East and West Pakistan: tentative results from the PGE experiment", The Pakistan Development Review (Karachi) 4(4): 734-759, Winter 1964.

(19) Krótki, Karol J. and Evelyn Lapierre, Age and Sex Selective Net Under-enumeration in Four Canadian Censuses (1951, 1956, 1961 and 1966). Paper under preparation, Ottawa: Dominion Bureau of Statistics, December 1967.

(20) Lauriat, Patience, "Field experience in estimating growth", Proceedings of the Social Statistics Section, 1966, Washington, D.C.: American Statistical Association, pp. 250-262.

(21) Lauriat, Patience and Anuri Chintakananda, "Technique to measure population growth: survey of population change in Thailand", World Population Conference Belgrade, Yugoslavia, 1965, Paper B6/ V/E/507.

(22) McDougall, Duncan M., "Immigration into Canada, 1851-1920", The Canadian Journal of Economics and Political Science (Toronto) 27(2): 162-175, 242-243, May 1961.

(23) Pritzker, L. and R.H. Hanson, "Measurement errors in the 1960 Census of Population", Proceedings of the Social Statistics Section, 1962, Washington, D.C.: American Statistical Association, pp. 80-90.

(24) Sampling and Survey Research Staff, Final Report on Projects I and II on the Quality Analysis of the 1961 Census, Ottawa: Dominion Bureau of Statistics, 1963. Processed.

(25) Sampling and Survey Research Staff, Final Report on Project III on the Quality Analysis on the 1961 Census, Ottawa: Dominion Bureau of Statistics, 1963. Processed.

(26) Sehdev, T.R., Evaluation of Sources on the Kitchener-Waterloo List. Technical Memorandum (Address Register Series) No. 3. Dominion Bureau of Statistics, Census Division, November 1966.

(27) Sehdev, T.R., Postal Check of the Kitchener-Waterloo Address Register. Technical Memorandum (Address Register Series) No. 4. Dominion Bureau of Statistics, Census Division, December 1966.

TABLE 1. Success of the First and Second Registered Mailings

	No.	%
Response to first letter	14,473	76.6
No response to first letter	5,021	
Response to second letter	3,696	19.6
No response to second letter	691	3.7
Sub-total of subjects apparently reached by letter	18,860	100.0
Letters returned by Post Office in any of the two mailings	7,467	
Total	26,327	

TABLE 2. Results of Tracing by Method Used, Canada, 1966

	Total addresses mailed or searched for	No. of replies rec'd	Per cent of total sample located
	(1)	(2)	(3)
First mailing	26,327 a/	14,473	54.6
Second mailing	5,021	3,696	13.9
Householder letter	2,714	261	1.0
Regional Office	3,637	2,918	11.0
Administrative files b/	8,796	4,449	16.8
Unable to locate	738		2.7
Total sample	26,535		100.0

a/ Less subjects imputed by the 1961 Census and illegitimate births for which no mailing was attempted.
b/ Subjects were searched for simultaneously in several administrative files. The numbers in column(1) therefore contain some duplicates.

TABLE 3. Coverage of Households
(No. of households)

<table>
<tr><th>LFS
Census</th><th>Enumerated in LFS</th><th>Not enumerated in LFS</th><th>Total</th></tr>
<tr><td>Enumerated in census</td><td>a</td><td>b</td><td>a+b</td></tr>
<tr><td>Not enumerated in census</td><td>c_1 } c</td><td rowspan="2">d</td><td rowspan="2">c+d</td></tr>
<tr><td>Not enumerated in census - Reconciliation not possible</td><td>c_2</td></tr>
<tr><td>Total</td><td>a+c</td><td>b+d</td><td>n=a+b+c+d</td></tr>
</table>

TABLE 4. Coverage of Persons
(No. of persons)

Census \ LFS		Within enumerated households: Enumerated in LFS	Within enumerated households: Not enumerated in LFS	Not enumerated in LFS due to missed hhld.	Total
Within enumerated households	Enumerated in census	a	b	e	a+b+e
	Not enumerated in census	c	d		c+d
Not enumerated in census due to missed hhld.		f		k	f+k
Total		a+c+f	b+d	e+k	n

TABLE 5. Content Error in Household or Personal Characteristics

Census \ LFS		Enumerated with characteristic "i"	Enumerated with characteristic other than "i": LFS is correct	LFS is wrong	Reconciliation not possible	Total
Enumerated with characteristic other than "i"		a	b_1	b_2	b_3	a+b
Enumerated with characteristic other than "i"	Census is correct	c_1		b		
	Census is wrong	c_2	c	d		c+d
	Reconciliation not possible	c_3				
Total		a+c		b+d		n= a+b+c+d

TABLE 6. Population and Sample Sizes Relating to the 1966 AQC

Region	1961 Census-farms	1961 large census-farms a/	AQC sampling rate (%)	Number sample segments	Number 1966 sample farms	Number 1966 sample farms per segment	1966 Census-farms
Maritimes	31,639	333	4.13	218	908	4.2	24,684
P.Q. - Ont. ...	217,110	1,539	0.60	325	1,390	4.3	190,181
Prairies	210,442	1,579	0.60	323	1,427	4.4	194,844
B.C.	19,934	384	0.54	18	88	4.9	19,085
Canada	479,125	3,835	—	884	3,813	—	428,794

a/ Because the listing of large census-farms was given special attention by 1966 enumerators it was assumed they made no contribution for coverage and content errors.

TABLE 7. Comparison of Census and 1966 AQC Estimates

Maritime Provinces

Item		AQC estimate	Census estimate	Net error		Sampling error	
				Amount	% of AQC	Amount ± 1S	% of AQC
Farms	No.	25,900	24,700	1,200	4.6	410	1.6
Total acreage	ac.	4,778,200	4,590,600	187,600	3.9	89,300	1.9
Acreage owned	"	4,522,900	4,295,900	227,000	5.0	115,400	2.6
Cropland	"	1,197,000	1,140,300	56,700	4.7	19,400	1.6
Oats for grain	"	187,200	184,200	3,000	1.6	5,300	2.8
Hayland	"	703,700	650,000	53,700	7.6	16,000	2.3
Potatoes	"	123,900	123,300	600	0.5	2,000	1.6
Unimproved and woodland	"	3,122,300	2,896,300	226,000	7.2	98,200	3.1

TABLE 8. The Impact of the Postal Check on the 1966 Census Lists of Households and Other Lists in Three Towns of Ontario

	Kitchener	Waterloo	London
1. Households in 1966 VR's	26,696	8,689	59,902
2. "Missed" cards made out by letter carriers (i + ii + iii)	358	86	685
(i) Households found to have been included in enumeration	161	27	351
(ii) Non-residential	71	28	206
(iii) Genuine finds ...	126	31	128
3. (iii) as a % of 1	0.5%	0.4%	0.2%
4. 2 as a % of 1	1.3%	1.0%	1.1%
5. Households in AR a/ before postal check ...	31,466	9,522	57,703 b/
6. Postal deletions	3,117	664	603
7. Postal additions	512	228	2,859
8. Households in AR after postal changes (5-6+7)	28,841	9,086	59,959
9. 7 as a % of 8	1.8%	2.5%	4.8%

a/ In Kitchener and Waterloo the AR's were based mainly on MAR's. In London they were field listings by test enumerators.

b/ The large difference between row 5 and row 1 is due to non-mail delivery areas having been excluded from row 5 (some 1,200 addresses).

TABLE 9. Enumeration Status of Persons Who Completed a Change-of-Address Card for the Period May 15, 1966 to June 15, 1966, for the Ottawa-Hull Metropolitan Area by Two-day Intervals

1966 Census status	Total		Date of reported change of address by two-day intervals																	
			May									June								
	No.	%	15,16	17,18	19,20	21,22	23,24	25,26	27,28	29,30	31	1	2,3	4,5	6,7	8,9	10,11	12,13	14,15	No date
All movers	2,273	100.0	164	86	134	71	84	106	223	311	152	525	94	67	109	46	50	16	20	2
Not enumerated at either address	363	16.0	26	17	25	13	13	19	22	26	21	81	23	21	27	9	6	2	2	
Enumerated at one address only	1,618	71.1	122	59	98	54	61	117	169	236	115	385	46	33	57	25	36	12	11	2
Enumerated at both addresses	56	2.5	1	1	3		2	1	3	6	2	10	6	4	8		6		3	
One address not in a form to be searched a/	236	10.4	15	10	8	4	8	7	19	23	14	49	19	9	17	12	12	2	6	

a/ 178 movers were not enumerated at the good address and 58 movers were enumerated at the good address. In the 236 cases one address was not in a form which could be searched.

TABLE 10. Movers Enumerated at Wrong Address. Status of Dwellings at which they should have been Enumerated

Census status		Enumerated at wrong address	Correct dwelling enumerated as		Not determined a/
			Occupied	Vacant	
Incorrectly enumerated at:					
old address	No.	11	6	5	
	%	2.2	1.2	1.0	
new address	No.	487	262	184	41
	%	97.8	52.7	36.9	8.2
Total	No.	498	268	189	41
	%	100.0	53.8	38.0	8.2

a/ Unable to locate apartment or room number or address not found in census.

COMPUTER METHODS FOR GEOGRAPHICAL CODING AND RETRIEVAL OF DATA IN THE DOMINION BUREAU OF STATISTICS, CANADA

I.P. Fellegi and J.I. Weldon
Dominion Bureau of Statistics

A large up-surge in the collection and use of statistics has been experienced in recent years. It is reasonable to expect that the explosion in statistical activities will continue. We shall have to make sure however that future growth will be controlled, well coordinated and that it will be achieved by efficient utilization of the financial and manpower resources.

In view of these considerations and of the recent technological and scientific developments it is important that developmental work should get underway towards the creation of some general tools applicable to several surveys or data files. Such general tools, such automatic survey systems may represent important means to achieve economies to extend our processing and retrieval capabilities, to enable us to deal with massive volumes of data and to build into our data processing systems important elements of standardization. As such, these general survey systems may be the basic technological prerequisities of large-scale national statistical information systems [13].

The present paper will describe briefly the developmental work underway in the Dominion Bureau of Statistics towards the creation of an automatic geographic coding and retrieval system in larger urban areas. Although the system is expected to be of general utility, we shall discuss it in the context of the 1971 Population Census which, we expect, will be the first large-scale application of it.

The major system features which are to be examined in more detail are as follows:

- data retrieval by user specified areas in larger urban municipalities;
- automatic assignment of geographic location identifiers to urban addresses;
- acceptance and recognition of addresses in free format, automatic correction of the spelling and key punching type errors;
- effective data retrieval and tabulation techniques;
- geocoding and geographic retrieval outside of the large cities;
- reliability of data and disclosure.

Conceptually the function of this system is to retrieve and tabulate geographically coded census data for any arbitrarily defined urban area. Geographic coding is achieved by automatic conversion of addresses to unique geographic coordinates. The entire system operation in a nutshell can be described as follows:

- an address conversion file which can convert urban addresses to unique geographic coordinates is to be produced;
- addresses of the enumerated urban households are to be put into a predetermined standard format, verified and corrected;
- the addresses are to be substituted by their respective geographic coordinates (geocoding);
- the geocoded census data is to be stored for future retrieval;
- tabulations by user specified areas are produced on demand, subject to considerations of statistical reliability and confidentiality.

Data Retrieval by User Specified Areas in Larger Urban Municipalities

In the present context a larger urban municipality is tentatively defined as a city or metropolitan area with a population of 50,000 or over. The important consideration is that a municipality must be a certain size, or part of a large urbanized area to be in a position to take advantage of small area information. The user will be able to delineate on a map the area for which he needs statistical tabulations. Such user-specified areas should preferably not cut through block faces and must be sufficiently large to permit the provision of statistical tabulations without violating the principles of confidentiality. Another problem in connection with statistical small area tabulation which will have to be kept in mind relates to sampling and non-sampling errors.

The user specified retrieval areas are conceived as polygons and are described by the coordinate values of the polygon vertices. Data retrieval for the user specified polygon is done by computer. The programme first selects all the block faces (sides of city blocks between neighbouring street intersections) represented by their midpoint coordinates which are within the user specified area, then retrieves and tabulates the census data for the selected block faces. Characteristically the system approximates the arbitrarily specified areas by using block faces as building blocks. The technique enables us to retrieve by streets or street segments as well [1, 3].

Automatic Assignment of Geographic Location Identifiers to Urban Addresses

This operation is commonly referred to as geocoding. The assignment of geographic coordinates to urban addresses enables us to retrieve by the arbitrarily specified areas. The geographic coordinate of an urban address is that of the block face within which the address is located.

Geocoding is performed with the aid of the address conversion file. This file contains street names, address ranges by block faces and the corresponding block face center point coordinates. The geocoding operation is carried out by computer, which tests address ranges block face by block face until it finds the one which encompasses the submitted household address. After having determined the block face identity the corresponding block face centroid coordinates are added to the household address and merged with the census data. It appears that this method may result in a very efficient computer operation since a full tape reel of census data can, we think, be geocoded in about 10 minutes.

The work required for establishing an address conversion file represents a major effort at the present time. It is estimated that the preparation of the conversion file for a city of 1,000,000 people currently would take three clerical man years work. It is very likely, however, that this time will be halved by improved system design, methodology and by gaining on the job experience. The creation of the conversion file requires the selection of an accurate map of the municipality; the updating of it; preparation and key punching of a street index; the digitization of strategic points along all streets representing beginnings, ends, intersections and changes in direction; and the preparation, coding and key punching of address ranges by block face. The input data are edited, verified and processed by computer; block face center points are calculated and the address conversion file is produced. A by-product of the operation is a plotted street map for the municipality. Having produced the address conversion file its periodic updating will require a few days clerical work at a time. The real problem in updating the address conversion file is to obtain street data update information. We hope to get this directly from the respective municipalities, since they might be important beneficiaries of the system. We hope to make the address conversion capability available to interested municipalities to permit them to geocode their locally collected data. The address conversion file can also be used for geocoding any survey data containing addresses. We are currently completing the preparation of an address conversion file for the City of London, Ontario. This work was performed in conjunction with our 1967 Census Test for London, Ontario.

Acceptance and Recognition of Addresses in Free Format, Automatic Correction of the Spelling and Key Punching Type Errors

In free-form address neither the fields of the address components (such as house number, street name, street type, city name, etc.) nor their position sequence have to be specified. The identification of address components is performed by separating the words of the address into numeric and alpha fields and by relating the positions of the words of the address to recognizable key words (e.g. "Street", "Ave.", "Apt.", "County", "Rural Route", etc.). The resulting pattern of numeric fields, alpha fields and key words is unique enough to identify the address components in a large proportion of cases (present limited experience indicates that this proportion is well over 90%). Upon recognition of the address components it is necessary to verify at least the street and city names by comparing them with a file of "correct" names.

The census operation will have to deal with some three million urban addresses. These addresses may be obtained from existing lists in machine readable form or they may be key punched from field listings prepared by the Bureau, or both. These addresses will contain errors and they may be in different formats as well. The rewriting of these addresses on coding sheets in fixed format would require hundreds of clerks for many months. This operation, besides being error-prone, is impractical because of manpower, space, equipment and other limitations. To overcome these difficulties, we have developed a computer operation to accept and to recognize addresses in free-form. It is estimated that there are some three million addresses in Canada, which may require 100 magnetic tape reels for recording. It is reasonable to expect that 10% of the addresses will contain spelling or key punching type errors, especially if the addresses are produced without key punch verification. These errors would amount to some 300,000 address rejects requiring further manual intervention in the form of correction, key punching and reintroduction.

This major clerical operation might be substantially reduced by developing a computer programme for automatic error correction of the key punching and spelling type errors. A good proportion of these types of errors is due to a few different, missing or surplus characters. The error correction logic of the system is based on checking street or city names of similar (but not necessarily identical) lengths and on finding the name which produces the smallest number of discrepancies. The maximum allowable number of discrepancies is some variable function of the name length. The recognition and error correction of three million addresses would probably require about three to five days of continuous processing on a large scale computer. Judging from a performance of a similar system, it is to be expected that 1 1/2 or 2 1/2% of the addresses would still be rejected, amounting to 45 to 75,000 rejects. These address rejects would then have to be processed manually. We have an operational computer program now to decode addresses in free-form with certain restrictions. The complete system described above should be operational by summer of 1968. The concepts described in the following sections are still in the planning stage at the present time.

Effective Data Retrieval and Tabulation Techniques

The proposed census file would contain enumeration data with urban addresses organized in block face sequence. We expect to produce from this file the traditional tabulations by census tracts and enumeration areas, as well as tabulations by any combination of characteristics by ad hoc user specified areas. The difficulty in providing tabulations by user specified areas is, of course, that the requirements cannot be known in advance, yet the Statistical Bureau has to satisfy these demands without much delay and at a reasonable cost. These restrictions will quite possibly necessitate that census data be organized in random access storage. We also hope to be able to satisfy at least the simpler types of special tabulations by using a generalized, efficient retrieval and tabulation program.

Random Access Storage of the Census Data

The random access file organization appears to hold out several promises for storing and retrieving census data on users' requests which we intend to carefully investigate. If we shall use randomly accessible storage devices we may be able to compress the data to the extent where little storage will be wasted; and we may be able to increase the efficiency of retrieval to the extent that only data required for retrieval would be accessed. This type of file might consist of two modules, which are the data file and the index file. The organization of the records in the data file would be by cities or metropolitan areas and by block faces within them. A record in the census file today typically contains all the characteristics relating to one person. The records of the proposed random access file would be organized by characteristics in a string form, each string containing one of the characteristics for all the enumerated persons. This means that if in a metropolitan area there are one million persons enumerated and there are, say, 50 characteristics reported per person, the proposed file will consist of 50 strings, each one of them one million characters, digits or bits long depending on the data content.

The other file module mentioned was the index file. The index file might be organized in a hierarachical fashion in list mode. The first level of this hierarchy might contain the list of province names and address pointers which are directing to the list of city names within the respective provinces. The second level of the hierarchy might contain the list of city names by provinces and address pointers which are directing to the list of block faces within the respective cities. The third level of the hierarchy contains the list of block faces for each of the cities in the form of block face centroid coordinates and address pointers which are directing to the first sequential appearances of block faces in the various census data characteristic strings.

Retrieval by arbitrary areas can be achieved by listing the coordinate points of the retrieval polygon vertices, accessing the block face centroid list for the requested municipality by descending through the hierarchy of the index file, determining the block face centroids which are contained within the arbitrarily specified retrieval polygon, retrieving the desired characteristic data string portions for the selected block face groups from the census data file, and performing the requested tabulation. The entire operation is an integrated computer process.

Generalized Retrieval Programme

An important aspect in providing fast turn around time at a low cost to users is the availability of a generalized retrieval programme. Input to such a retrieval programme requires the designation of the province, municipality, the listing of the desired characteristics and retrieval conditions for tabulation or cross-tabulation, and the coordinate points of the vertices for the requested retrieval polygon. The significance of such a generalized programme would be that at least simpler types of special tabulations could be specified through the use of the programme without extensive training in programming. The data file organization by characteristic strings and the index file organization in hierarchical structure would greatly facilitate the utilization of such a generalized programme. The most significant advantage of such high level retrieval languages is that they permit the description of the retrieval and tabulation requests in some restricted English language form, which then can be used as an input to a computer programme to perform the designated operations.

The system must also be designed to facilitate an inverse retrieval function. This refers to the type of request which seeks the delineation of an area (or areas) which satisfy some stated conditions. After having determined the desired area its boundary could be mapped by means of computer graphics.

The Problem of Geocoding and Geographic Retrieval Outside the Larger Cities

Automatic geocoding assigns to postal addresses, with a minimum of manual intervention, the location-specific coordinates of the center point of the block face in which the address is located. In this fashion the traditional coding is carried out in that the address is identified as belonging to a particular pre-designated standard area, the block face in the present case. Automatic geocoding differs, however, from the traditional geographic coding in three important ways. First, it carries the coding to much smaller areas (block faces) than would be conceivable using manual methods, hence it provides very small building blocks for future aggregations. Second, it provides a reasonably error-free general tool that can be applied to data files, whatever their origin, as long as the data fields are identified by postal addresses*. Third, the codes which identify the "building blocks" of the coding system directly identify their location as well, hence future aggregations of contiguous "building blocks" into larger areas are greatly facilitated, no matter how the larger areas are specified.

The system, as outlined above, is designed to handle addresses in the larger cities or metropolitan areas. In smaller urban areas it would not be economically feasible. In rural areas it is not even conceptually feasible since rural addresses often are not specific enough to determine their location (e.g. the address John Smith, 29 Bank Street, Ottawa is location specific even without the name of the occupant; John Smith, R.R. 2, Cornwall, Ontario is not location specific). It is unlikely, therefore, that we shall have in these areas an automatic and general geocoding system capable of coding addresses to sufficiently small areas**. Present plans for the 1971 Census in these areas [6] indicate that enumerators will be canvassing the addresses and will, at the same time, code them in the traditional fashion to Enumeration Areas (similar to the Enumeration Districts in the U.S.). In a separate manual operation the coordinates of the center points of these Enumeration Areas will be determined. Each census record in an Enumeration Area (about 150 households) will carry the coordinates of this center point. The number of affected Enumeration Areas is expected to be less than 20,000. Of the three benefits of automatic geocoding two will be lost using this rather primitive method: it will not carry the coding process to building blocks smaller than the traditional ones and it will not provide a general tool applicable to data files other than the census. It will, however, retain the third important advantage, namely it will facilitate the aggregation of contiguous Enumeration Areas into larger areas, no matter how the larger areas are specified provided they do not cut across Enumeration Areas.

The Problems Relating to Reliability of Data and Disclosure

References have been made to many remaining unsolved problems in system design and programming, as well as to others on which developmental work is well under way towards a satisfactory solution. In conclusion two different and very important problems should be at least briefly mentioned.

* Such a general tool has an important unifying influence in that it facilitates the production of comparable and compatible geographic tabulations from different data files.

** A solution might be found for the problem if an appropriate ZIP-code type system was adopted by the Canadian Post Office.

The first problem relates to the reliability of data. It is well known to this audience that census data, whether they are based on a full count or on sampling, are subject to potentially large measurement errors and in the latter case also to sampling errors. It is important to understand that although in the larger urban areas we will code to the block face level this will not be a level at which data can be made available (except possibly some very simple counts). The purpose of coding to the block face level is to achieve a degree of flexibility in aggregating to larger areas which was not open to us before. Just the same, the temptation will be substantial to ask for data for areas smaller then, say, census tracts. Also undoubtedly there will be a greater number of people using census data than before. This would make it very important to be able to associate with census tabulations, particularly with those referring to smaller areas, a measure of the reliability of the data, including the contributions of measurement errors as well as sampling errors where applicable. This will pose several serious problems. It is well known that the measurement of such errors is notoriously difficult [2, 4-12, 14-16]. It would be even more difficult to measure them early enough to be available when the main census publications are prepared. And even if they were available for some of the counts in time, there would be some difficulty in imputing them for the remaining counts (since direct calculation for all items would be inconceivable) and presenting them in a readable and useful form. A series of special problems arise related to sample estimation but these will not be discussed here.

The second problem area relates to disclosure of individual information. The flexibility of retrieval of information for areas which are not the standard, pre-coded ones greatly increases the danger of the so-called residual disclosure. The monitoring of all the previously produced small area tabulations appear to be the best solution to protect against residual disclosure, at this time. Whether or not this monitoring function can be performed automatically by computer, we are going to ensure the confidentiality of census data.

REFERENCES

[1] Calkins, H. W., Research report No. 2, Operations Manual for street address conversion system.

[2] Cochran, W. G., Sampling Techniques, Second Edition. New York: John Wiley and Sons, 1963. Pp. 141-2.

[3] Dial, R. B., Research report No. 1, street address conversion system.

[4] Fellegi, I. P., "An analysis of response variance", Bulletin of the International Statistical Institute, 34th Session, 40 (1963), 758-9.

[5] Fellegi, I. P., "Response and its estimation", Journal of the American Statistical Association, 59 (1964), 1016-1041.

[6] Fellegi, I. P., and Krotki, K. J., "The Testing Programme for the 1971 Census in Canada", Proceedings of the Social Statistics Section, American Statistical Association (1967).

[7] Hansen, M. H., Hurwitz, W. N., and Bershad, M. A., "Measurement errors in censuses and surveys", Bulletin of the International Statistical Institute, 32nd Session 38 (1959), 359-74.

[8] Hansen, M. H., Hurwitz, W. N., and Pritzker, L., "The estimation and interpretation of gross differences and the simple response variance", Unpublished report, Bureau of the Census, U.S.A. (1963),(honouring Professor P. C. Mahalanobis).

[9] Hanson, R. H., and Marks, E. S., "Influence of the interviewer on the accuracy of survey results", Journal of the American Statistical Association, 53 (1958), 635-55.

[10] Kish, L., and Lansing, J. B., "Response errors in estimating the value of homes", Journal of the American Statistical Association, 49 (1954), 520-38.

[11] Kish, L., "Studies of interviewer variance for attitudinal variables", Journal of the American Statistical Association, 57 (1962), 92-115.

[12] Mahalanobis, P. C., "Recent experiments in statistical sampling in the Indian Statistical Institute", Journal of the Royal Statistical Society, 109 (1946), 325-70.

[13] Nordbotten, S., "Automatic files in statistical systems".

[14] Pritzker, L., and Hanson, R. H., "Measurement errors in the 1960 Census of Population", Proceedings of the Social Statistics Section, American Statistical Association (1962), 80-90.

[15] Sukhatme, P. V., Sampling Theory of Surveys with Applications, Ames, Iowa: Iowa State University Press and New Delhi, India: Indian Society of Agricultural Statistics, 1954. Pp. 445-6.

[16] Sukhatme, P. V., and Seth, G. R., "Nonsampling errors in surveys", Journal of the Indian Society of Agricultural Statistics, 4 (1952), 5-41.

Discussion

Conrad Taeuber, Bureau of the Census

It is well-established at Suitland, near Washington, and at Tunney's Pasture, near Ottawa, that the Census Bureau and the Dominion Bureau of Statistics share experiences, fears, successes, frustrations, and the very considerable amount of critical evaluation of the work which is carried on at both locations. Our Canadian colleagues tend to see in our decennial census a pretest for theirs. We took a census in April 1960; they took theirs in June 1961. On the other hand, recognizing the lead time that is a necessary part of all such activities, we find that the census of 1966 provides some useful lessons for our census of 1970. Actually, the sharing is a continuous one and it is a 2-way street which is heavily traveled. Their observers at our pretests have given us some very incisive and useful observations and criticisms. When we saw their questionnaires for the London Pretest, we gave our printer a hard time, telling him that if the Canadians could print a relatively uncluttered questionnaire, we could do so also. I mention this only to emphasize that the sharing of experience goes far beyond an occasional scanning of the formal papers or an occasional conference. From the standpoint of the Bureau of the Census, these papers are welcome, for they give us in systematic fashion an account of work in progress and some leads for further joint exploration.

In view of our close association, I was startled to learn from the paper by Fellegi and Krotki that the Canadian data are based on the "longest series of modern censuses stretching back just over 300 years." I hasten to set the record straight. Even though we don't want to claim the "longest continuous census"--(our critics tell us it is much too long), we like to point out that ours is the longest series of periodic nationwide censuses. In addition, some 38 censuses of individual colonies had been taken prior to the first national census in 1790, chiefly at the instigation of the British Board of Trade. The first of these was taken in Virginia in 1635.

From the papers presented here this morning, it is easy to see that we share many common approaches to our problems. We are agreed on the need of pretests, on the need for evaluation of the census as it is taken, on the need for quality control in all phases of the work, and on methods of achieving such control. Our experiences, too, have been quite similar. Thus, in the case of the re-enumeration in connection with the Census of Agriculture, as reported by Krotki, Muirhead and Platek, the total underenumeration of farms seems relatively large, but the undercount of the important commercial production is far less. In other words, much of the underenumeration occurs at the margins where the determination that a unit belongs in the universe becomes difficult to make. The units which are clearly to be included in the universe have a much lower likelihood of being missed. We are agreed also on the need to investigate coverage and content errors and on the relative importance of sampling and response errors and of methods for dealing with the reduction of response variance.

That same paper offers an interesting demonstration of why experts are often charged with making any apparently simple problem complex. Matching two sets of records seems like a simple operation, especially when both records were secured by the same organization within a short time span. But this simplicity is apparent only to the uninitiated. Matching two records to determine whether they relate to the same individual turns out to be a very difficult operation, requiring careful specification of when a match has actually occurred, what differences in spelling or characteristics to accept as not violating the match, what degree of field reconciliation is feasible and what to do with the apparent failures to match. The whole field of research in census methods offers another case in point, for no one with experience in the field would agree with the comment of an uncritical observer who said with some surprise, "I would have thought that census-taking is least in need of research." It seemed to him that nothing could be simpler than counting such discrete units as people. At least there can be little question whether a unit once located belongs in the national inventory, even though it may be debatable that his attachment to a given locality is such that he should be enumerated there.

Although the conditions under which a census count is made differ somewhat in the two countries, there is a remarkable similarity in the degree of the undercount, and in its incidence by age and sex.

Response rates to the mail questionnaire in London at 85 percent were only slightly higher than the experience of the Census Bureau in its pretests. The experience in London was like that in the States, that the great majority of the returns were in the office within a few days of the census date.

On one point our experience would provide a clear answer to a question posed in the paper by Fellegi and Krotki. They appear to question whether the computer delivers as good a job of editing as clerks. That there is a tendency at the outset to ask the computer to do much more editing than was expected of editing clerks is a common experience. That some clerks will challenge the reasoning behind an instruction, whereas the computer is not likely to do so is also common experience. But granting this, it is clear that the computer can edit more consistently and take into account a more complex set of factors than is the case with clerks. It can also be depended upon to reject for manual intervention more consistently than was the case with clerks.

The paper by Fellegi and Weldon outlines an ambitious program of geo-coding which we will be watching with a great deal of interest. They are motivated by the same conditions which lead the Census Bureau to make efforts along this line; i.e., the need for quick and inexpensive tabulation for areas required to meet special needs. The users of census data are no longer content with having to take standard small area units as building blocks and laboriously allocating segments and adding the

on persons who submitted change of address cards to the post office. We are also in essential agreement in regard to the significance of response variance in relation to sampling variance as applied to the social and economic characteristics which have been mentioned. It is not clear from the scheme presented that adequate provision has been made for special tabulations involving a number of characteristics; for example, the number of preschool age children living with both parents in families with an income below a specified level. It is likely that the provision for rapid retrieval of single characteristics will not prove adequate to the needs of users of census data.

These papers are a welcome addition to the exchange of experience and the joint study of census problems which has been going on for many years, and, hopefully, will continue for many more.

bits and pieces to arrive at the distributions and totals which are needed. The use of computers has considerably enhanced the ability to meet these needs and here is one more instance of both agencies working along closely similar lines in order to provide the same basic services.

Our experience has been quite similar to that of our Canadian colleagues in respect to the proportion of free form addresses which can be coded by the computers, in the ability to find persons whose names come from lists of a previous census, birth records or immigration records, in the reliability of the post office check, and in some of the results of a followup

DISCUSSION

J. KANTNER UNIVERSITY OF WESTERN ONTARIO

The three papers presented at this session suggest a sequential order of business. First there is the evaluation of the 1966 Census; second a discussion of plans for the coming decennial census of 1971; and finally an account of the work being done to increase the utility of that census through the introduction of geocoding. I shall take the papers in this order. In addition, since the title of the session "Aspects of the 1966 and 1967 Census Programmes in Canada" tempts a discussant to reflect on those aspects not mentioned by the authors, I do so briefly despite the acknowledged unfairness of such behavior.

Krotki, Muirhead and Platek "Evaluation Programme of the 1966 Census of Canada".

This paper discusses a series of projects designed to evaluate the results of the 1966 Census. These include:

(1) Reverse Record Check -- an impressive effort about which I would raise only two minor questions:

a) What is the logic and purpose of the double coverage of persons under six months of age in 1961? The RRC cannot evaluate coverage of this age group in 1966 and this may as well be accepted as a deficiency of the method. Nothing is gained by the double coverage technique.

b) The case for the application of the Chandra-Deming correction for events missed both in the Census and in other records is not convincing since the independence of the two approaches -- in terms of the types of errors to which each is prone -- has not been demonstrated.

(2) Agricultural Quality Check -- two points which the authors do not discuss are of some interest:

a) the direction of the 'error' is the same for all items compared. Thus the errors would not appear to be due to ignorance about agricultural holdings, as is suggested, but rather, to omissions that can be filled in by additional probing.

b) the net error for number of farms is larger relative to its estimated standard error than is the relative error for total acreage. This suggests something about the location of the error.

(3) Demographic checks -- the authors candidly acknowledge their difficulties in applying standard methodology. It would be ungenerous to add to their woes. There are, however, one or two minor points:

a) the continued high survival ratios for persons 10-14 and 30-34 are noted (p.15) as evidence that there has been little improvement in the underenumeration of children and persons 20-24. Apart from the fact that survival ratios are somewhat dull tools for the analysis of underenumeration, the exposition at this point is confusing perhaps because of the overly condensed references to dates and intervals.

b) in view of all the methodological uncertainty to which the authors confess and the fact that the matching tests have not yet been completed, how do the authors conclude, as on page 15, that "it can be reported that there has been no obvious improvement in the apparent net undercount of, say 2.5 per cent over all ages ..."? Such an important conclusion deserves fuller documentation.

(4) Household check -- theoretically such a check is useful but in practice leads to some ambiguity of interpretation since a good Address Register shows up deficiencies in the Census and vice versa. For the three cities in which the household check was performed the number of households on the A.Rs exceeded the number on the V.Rs by more than 1 percent. The authors seem inclined to dismiss the results for two of these cities but would, I expect, take the view that the A.R. is generally more complete than the V.R. and yet the postal check added to the A.R. by 2 per cent but only by ½ percent to the V.R. It is hard to reconcile this with the idea that the A.R. is the more inclusive

list. Since the A.R. is made up from several sources and involves different conceptual approaches to the same "place" e.g. as an addressable structure and as an assessable property, might there not be an inflationary tendency in the A.R. of serious enough magnitude to compromise this particular validation technique. The very substantial deletions from the A.R. made on the basis of the postal check raises a similar question and also introduces an element of judgement that may be difficult to assess.

(5) Postal-Change-of-Address-Cards -- this postal check, instituted in Ottawa-Hull, is designed to measure underenumeration among persons who changed their residence around Census time. The test of underenumeration consisted of determining whether moves were enumerated either at their old or new address. That is to say, the matching of records involved searching Census records for the addresses in question and then ascertaining whether the family enumerated at that address matched the one which according to the change-of-address-card should have been there. At least this appears to have been the procedure from the description given. A further test taking account of the moving date was used to gauge the extent to which those who were enumerated were assigned the appropriate address as of the critical date of the Census i.e. June 1. From the first test it was learned that "the under-enumeration rate among movers was more than five times greater than for the population as a whole..." This is an important finding but possibly fallacious since in many cases indicated change-of-address may not represent a bona fide residence but merely a place, perhaps the home of a friend or relative, where mail is picked up until the relocation is completed. Unless the entire universe of census documents is searched, a formidable undertaking, the extent of underenumeration of movers could easily be overstated. Perhaps a greater danger is that I have misunderstood the procedure from the abbreviated description given in the paper.

Fellegi and Krotki "The Testing Programme for the 1971 Census in Canada".

The testing programme for 1971 thus far has been confined largely to the Test Census conducted in London in the fall of 1967. Although to some observors the most interesting aspects of this test are the content innovations, it is clear from the remarks of DBS officials as well as from the design of the testing program, that Bureau interest centers in the feasibility of self-enumeration and mail-out-mail-back procedures. The debate over the advantages and disadvantages still rages (this may not be too strong a term) in the United States where the general educational level is significantly higher than it is in Canada. Thus, the DBS concern on this score is fully justified although London, a relatively sophisticated commercial city with a large university and closer to the urban-industrial heartland of the United States than most states, can hardly be regarded as the place for an acid test. London is "average" or "typical" with respect to a number of Census variables - and it is for this reason that it was selected - but it is probably well above average relative with respect to many of the considerations that make for successful self-enumeration. This is not to criticize DBS for its choice of London as a test site but merely to say something about the interpretation of results. If the London test "fails", which appears quite unlikely, the outlook for self-enumeration in 1971 would be bleak. If it "succeeds", well ......

The authors in stating the a priori case for self enumeration, overstate it to some extent by claiming advantages which logically could be had also in connection with conventional enumeration. They site the advantages of (1) early mailed returns (2) reduction of coverage errors through the preparation of address registers (3) questionnaire probes to improve coverage (4) concentrated publicity (5) geographic coding and (6) respondent specificity. The last mentioned which refers to the greater liklihood under self enumeration of getting information from the best qualified respondent in the household is perhaps an advantage that can be granted although there is nothing in the test program as here described that directly tests this point. As for the other points, none seems overwhelmingly wedded to self enumeration. The mails are available to interviewers as well as to other citizens so that field editing of mailed-in daily batches would seem to be a possibility; address registers are better than poor listings but nothing prevents using A.R.s in conjunction with regular enumeration; interviewers can probe as well, possibly, as printed instructions; publicity campaigns may be easier to bring to a pitch on a given census-day but to conclude that this condition is the sine qua non of an effective publicity campaign perhaps gives too little credit to an industry that sells soap and automobiles all year round. The link between self enumeration and geocoding is a derivative of the earlier statements about the advantage of Address Registers. As an argument for self-enumeration it is, by itself, a non sequitur.

The test provisions themselves deal with a limited number of questions. the editing test, the Post Enumeration Survey and the computer programming are confined to the items on the short form questionnaire. Presumably several interesting content innovations, most of which are found on the long form questionnaire, are to be examined in later Census Tests.

The London Test questionnaire carried new items designed to give the Census greater penetration in the areas of education, fertility, language, foreign background, mobility and household structure. There does not seem to be any ready way to validate these new items, at least in the London test. The Post Enumeration Survey would have been the logical vehicle but, as already noted, the P.E.S. was limited to the small list of items collected on a 100 percent basis. It is possible to evaluate the relative troublesomeness of the new items by tallies of omissions, imputations, required follow-up, etc. but no tests of validity are presented. One hopes that the fate of these items will not be left solely to an impressionistic evaluation.

Fellegi and Weldon "Computer Methods for Geographical Coding and Retrieval of Data in the Dominion Bureau of Statistics"

This paper describes an exciting new development for the custom assembly of data by prescribed areal units. If successful in operation, one of the major sources of discontent between producers and consumers of census data will be eliminated. The chief virtues of computerized geocoding to which the paper gives attention are the ability to specify suitable areal units and the efficient access to data stored according to block face coordinates.

The application as well as the problems of this new technique are only dawning. It can be expected that subsequent application will go beyond problems of areal assembly to applications that would treat location as an individual, household or family attribute in deriving journey-to-work configuration, pattern of intercensal residential mobility and so on. If spatial information other than residence and place of work, say the schools attended by children or the place the wife purchased her last basket of groceries were added, we would be well along toward a spatial representation of urban metabolism as it applies to individuals and households. As described in the present paper, the technique appears to be much more supple in arranging space than in filling the resulting parallelograms with information of more than one dimension. One can be confident that the competent ingenuity that has brought the technique this far will soon take it beyond its present limitations.

Returning to what might be regarded as the implicit theme of this session - the capacity of the Census to meet those data requirements for which it is the appropriate vehicle - it is difficult not to remark on the lack of attention on an occasion such as this to the role of the Census in a system of statistical indicators. The Census, as it is now, is the outcome of a large number of games being played by governmental, professional, academic, commercial and industrial gamesmen each with his own requirements for data. There is a growing insistence that this is not good enough in the face of the complex business of diagnosing and prescribing for the ills of a modern society. The development of a comprehensive and meaningful system of social indicators is one of the major tasks before our statistical agencies over the next decade. Significant improvement in the battery of economic indicators is to be hoped for also and indeed the momentum in this field of measurement makes progress highly likely. The succession of social crises that continually assault our Panglossian perception of the world can be expected to spur the development of social indicators in much the same way as the problems of depression and wartime mobilization forced us to a more sophisticated monitoring of the economy.

One should not expect Census personnel to do this job single handedly. This is a task requiring effort and broad commitment from the academic, professional and governmental communities. It can be anticipated however, that just as in the development of economic series, our Census colleagues will contribute in a major way. Let us hope that they will find time soon to give this "aspect" the attention it deserves.

III

METHODOLOGY OF SAMPLE SURVEYS IN SOCIAL STATISTICS

Chairman, JOSEPH STEINBERG, Social Security Administration

THE UNRELATED QUESTION RANDOMIZED RESPONSE MODEL[1/]

D. G. Horvitz, B. V. Shah and Walt R. Simmons
Research Triangle Institute and National Center for Health Statistics

I. Introduction

Warner[2/] has developed a randomized response technique which allows the respondents in personal interview surveys to provide information on sensitive or highly personal questions and yet retain their privacy. The technique requires the respondent to select one of two related questions using a random device. The respondent answers only "yes" or "no" to the chosen question without revealing which question was actually selected. The responses to either question divide the population into the same two mutually exclusive and complementary classes. The proportion of "yes" answers in the sample and the known chance of selecting either question is sufficient to provide an unbiased estimate of the proportion of the population in each of the mutually exclusive classes, provided the respondents answer truthfully.

A variation of the Warner technique, suggested by Walt R. Simmons and designed to increase further the cooperation of the respondents and the veracity of their responses, is reported in this paper. It requires the respondents to randomly select one of two unrelated questions, so that the mutually exclusive and complementary properties of the Warner technique no longer apply. Two samples are required with a different set of selection probabilities for the two questions for each sample. The method for estimating the parameters and variances for this alternative randomized response model are developed also for two independent trials per respondent. Results from two empirical studies concerned with estimating the proportion of illegitimate births from household interviews are reported.

II. The Warner Randomized Response Model

The purpose of the Warner Model is to provide a method for estimating the proportion of persons with a sensitive attribute, say A , without requiring the individual respondent to report his classification (whether it be A or $\bar{A}$) to the interviewer. The respondent is provided with a random device to choose one of two statements of the form:

1. "You have the attribute A "
2. "You do not have the attribute A "

Without revealing to the interviewer which statement has been chosen, the respondent then answers "yes" or "no" according to the statement he has selected and whether he does or does not have the attribute A .

Let

Π = true proportion with attribute A

p = probability that the first statement is selected (the second statement is selected with probability 1-p)

x_i = 1 if the i-th respondent says "yes" to the selected statement

x_i = 0, otherwise

n = sample size

Then, with a single sample and a single trial with respondents who always tell the truth,

$$\Pr(x_i=1) = \Pi p + (1-\Pi)(1-p)$$

$$\Pr(x_i=0) = (1-\Pi)p + (1-p)\Pi$$

It follows that the maximum likelihood estimate of Π is

$$\hat{\Pi} = \frac{p-1}{2p-1} + \frac{n_1}{n(2p-1)}, \quad p \neq \tfrac{1}{2}$$

where $n_1 = \sum_{i=1}^{n} x_i$. This is an unbiased estimate, if all respondents answer truthfully, with variance given by

$$\mathrm{Var}(\hat{\Pi}) = \frac{\Pi(1-\Pi)}{n} + \frac{p(1-p)}{n(2p-1)^2}$$

III. The Simmons Unrelated Question Randomized Response Model[3/]

The Warner technique is designed to elicit truthful answers to questions many respondents would refuse to answer at all, if asked directly. Walt R. Simmons has suggested that the confidence of the respondents might be further increased and hence the likelihood of truthful answers, if two unrelated questions (or statements) are used, one pertaining to the sensitive attribute, say A , and the other to a non-sensitive characteristic, say B .

1/ Research supported by the National Center for Health Statistics, Contract No. PH 86-65-68.

2/ S. L. Warner, "Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias," Journal of the American Statistical Association, 60, (1965), pp. 53-69.

3/ The Simmons single trial unrelated question model has been discussed in some detail by Abdel-Latif A. Abul-Ela, "Randomized Response Models for Sample Surveys on Human Populations," unpublished Ph.D. thesis, University of North Carolina, Chapel Hill, 1966.

It is noted that both of the Warner statements divide the population into the same two mutually exclusive and complementary classes. The unrelated question model uses two statements of the form:

1. "You have the attribute A "
2. "You have the attribute B "

so that some respondents might belong to both groups and some might not belong to either group.

Two independent samples are required with this model. Let

Π_1 = true proportion with sensitive attribute A

Π_2 = true proportion with non-sensitive attribute B (not related to A)

p_1 = probability that the first statement is selected by each respondent in the first sample (the second statement is selected with probability $1-p_1$ by respondents in this sample)

p_2 = probability that the first statement is selected by each respondent in the second sample, $p_2 \neq p_1$

x_i = 1 if the i-th respondent in the first sample says "yes" to the selected statement

= 0, otherwise

y_i = 1 if the i-th respondent in the second sample says "yes" to the selected statement

= 0, otherwise

n_1 = size of the first sample

n_2 = size of the second sample

With a single trial per respondent,

$$Pr(x_i=1) = p_1\Pi_1 + (1-p_1)\Pi_2$$

$$Pr(y_i=1) = p_2\Pi_1 + (1-p_2)\Pi_2 ,$$

provided all respondents answer truthfully. If $n_{11} = \sum_{i=1}^{n_1} x_i$ and $n_{12} = \sum_{i=1}^{n_2} y_i$ are the respective number of "yes" answers in the two samples, then unbiased estimates of Π_1 and Π_2 may be obtained by solving the pair of equations (i.e. by equating observed proportions of "yes" answers to expected proportions)

$$\frac{n_{11}}{n_1} = p_1\Pi_1 + (1-p_1)\Pi_2$$

$$\frac{n_{12}}{n_2} = p_2\Pi_1 + (1-p_2)\Pi_2 .$$

The estimates are:

$$\hat{\Pi}_1 = \frac{1}{p_1-p_2} \{(1-p_2)(n_{11}/n_1)-(1-p_1)(n_{12}/n_2)\}$$

$$\hat{\Pi}_2 = \frac{1}{p_2-p_1} \{p_2(n_{11}/n_1) - p_1(n_{12}/n_2)\}$$

It is noted that if Π_2 is known, then a single sample is sufficient to estimate Π_1 . The estimator is

$$\hat{\Pi}_1 = \frac{(n_{11}/n_1) - (1-p_1)\Pi_2}{p_1}$$

IV. Test of the Simmons Unrelated Question Model (Single Trial Per Respondent)

The Simmons Model was tested in late October 1965 by personal interviews in a total sample of 148 households in which it was known that a birth had occurred during August and September 1965. The sample was selected from birth certificates upon which the marital status of the mother on the date of the birth was recorded. Twenty-eight (28) or 18.9 percent of the 148 mothers were not married. The respondents were asked to select a card from a shuffled deck of 50 cards and to answer "yes" or "no" to the truth of the statement printed on the card. The two statements used in the deck were:

1. "There was a baby born in this household after January 1, 1965, to an unmarried woman who was living here." (Attribute A)
2. "I was born in North Carolina." (Attribute B)

The results obtained in this test of the technique and model are shown in Table 1. The estimated proportion of all households with a birth to an unmarried woman, that is $\hat{\Pi}_1$, is in reasonable agreement with the true proportion. The results are even closer when computed separately for white and non-white households.

It is reasonable to ask whether results as good or better could have been obtained by direct questioning of the respondents. Although this has not been tested, the completeness with which births known to have occurred out of wedlock are reported in household interviews has been found to be somewhat less than for births classified as legitimate on the birth certificate.[4] The latter results indicate that the legitimacy status of births is sufficiently sensitive to warrant use of a technique which respects privacy of the respondent.

Table 1

Parameters and Estimates in Test of Simmons Model (Single Trial per Respondent)

Item	All Households	White Households	Nonwhite Households
p_1	.7	.7	.7
p_2	.3	.3	.3
n_1	63	40	23
n_2	85	64	21
n_{11}	24	10	14
n_{12}	49	31	18
$\hat{\Pi}_1$	.235	.074	.423
Π_1	.189	.077	.454
$\hat{\Pi}_2$	.722	.660	1.043

V. Extension of Simmons Model to Two Trials per Respondent

A simple extension of the Simmons unrelated question model requires each respondent to make two independent selections of the two questions (or statements) using the randomizing device. Let n_{11}, n_{10}, n_{01}, n_{00} be the numbers of individuals answering (Yes, Yes), (Yes, No), (No, Yes) and (No, No) respectively in the first sample and m_{11}, m_{10}, m_{01} and m_{00} be the corresponding numbers for the second sample. The sample sizes are n_1 and n_2. As before the observed proportions of "yes" answers for each sample are equated to the expected values of these proportions and the unknown parameters Π_1 and Π_2 are estimated by solving the resulting pair of equations,

$$\gamma_1 = \frac{2n_{11} + n_{10} + n_{01}}{2n_1} = p_1\Pi_1 + (1-p_1)\Pi_2$$

$$\gamma_2 = \frac{2m_{11} + m_{10} + m_{01}}{2n_2} = p_2\Pi_1 + (1-p_2)\Pi_2$$

4/ Horvitz, D. G. "Problems in Designing Interview Surveys to Measure Population Growth." Proceedings, Social Statistics Section, American Statistical Association, (1966), 245-249.

Thus,

$$\hat{\Pi}_1 = \frac{1}{p_1-p_2}\{(1-p_2)\gamma_1 - (1-p_1)\gamma_2\}$$

$$\hat{\Pi}_2 = \frac{1}{p_2-p_1}\{p_2\gamma_1 - p_1\gamma_2\}$$

These are unbiased estimates if all respondents answer truthfully.

VI. Test of the Simmons Model With Two Trials Per Respondent

Two different randomizing devices were used with stratified cluster samples of North Carolina households in the summer of 1967 to test the two trial model. The first device was the same as used previously, a deck of 50 cards. The second was a sealed plastic box, designed by B. G. Greenberg, containing red and blue beads.[5] When the box is shaken by the respondent, a bead appears in a small window and the respondent answers "yes" or "no" to the statement corresponding to the color of the bead. Both statements are attached to the box. The two statements used with both randomizing devices were:

1. "In the past 12 months there was a baby born in this household to an unmarried woman who was living here at the time."
2. "I was born in North Carolina."

The interview procedure for the deck of 50 cards randomizing device is given in Appendix A. The parameters, distributions of "yes" and "no" answers, and estimates for the respective randomizing devices are given in Tables 2 and 3. In contrast to the first field trial, the estimated Π_1's for both devices bear little relationship to the values expected for this parameter.

VII. Some Speculations

Under the basic unrelated question model, the major field trials for North Carolina yielded as the estimate of the proportion of households with an illegitimate birth, the value $\hat{\Pi}_1 = 0.14$. Vital records for the State indicate a figure $\Pi_1 = 0.008$. Thus the estimate is almost 20 times the criterion value. Why did the basic model fail so badly? This question has even more force when it is recalled that the same model was highly successful in the first (pilot) test, and further

5/ The Department of Biostatistics, University of North Carolina School of Public Health, supplied the plastic boxes and beads for this trial.

Table 2

Parameters and Estimates in Test of Unrelated Question Model with Two Trials per Respondent
Randomizing Device: Deck of 50 Cards

Item	All Households	White Households	Nonwhite Households	All Households Reporting Births
p_1	.7	.7	.7	.7
p_2	.3	.3	.3	.3
n_1	1450	1227	223	88
n_2	1638	1340	298	103
n_{11}	166	137	29	17
n_{10}	323	271	52	16
n_{01}	298	253	45	18
n_{00}	663	566	97	37
m_{11}	636	512	124	47
m_{10}	345	291	54	25
m_{01}	276	215	61	13
m_{00}	381	322	59	18
γ_1	.328	.325	.348	.379
γ_2	.578	.571	.609	.641
$\hat{\Pi}_1$	.142	.141	.151	.183
Expected Π_1	.008	.002	.034	.042*
$\hat{\Pi}_2$	.765	.755	.805	.837

*Based on data from matched birth certificates.

Table 3

Parameters and Estimates in Test of Unrelated Question Model with Two Trials per Respondent
Randomizing Device: Plastic Box of 50 Beads

Item	All Households	White Households	Nonwhite Households	All Households Reporting Births
p_1	.7	.7	.7	.7
p_2	.3	.3	.3	.3
n_1	437	320	117	29
n_2	442	375	67	28
n_{11}	53	37	16	4
n_{10}	76	55	21	6
n_{01}	83	61	22	2
n_{00}	225	167	58	17
m_{11}	166	141	25	6
m_{10}	80	67	13	8
m_{01}	53	48	5	3
m_{00}	143	119	24	11
γ_1	.303	.297	.321	.276
γ_2	.526	.529	.507	.411
$\hat{\Pi}_1$	.136	.122	.180	.174
Expected Π_1	.008	.002	.034	.080*
$\hat{\Pi}_2$	.693	.704	.648	.512

*Based on data from matched birth certificates.

that the direction of the failure might seem on first inspection to imply that far too many--rather than possibly too few--persons replied "yes" to the question of whether there had been an illegitimate birth in their household.

Let it be noted immediately that we do not claim to know the answer to this question. But study of the model and the experiment suggest a number of possibilities, and several significant relevant facts. A review of some of these matters makes it clear that the range and variety of possible hazards is great. It also suggests several types of modification in the basic model or its application to protect against selected hazards. First, one must consider the possibility of a boner in the coding or processing of data. That source of failure is consistent--as are several other hypotheses--with the observation that the scale of error is so very large, and is fairly constant over nearly all of 8 different subsamples, and six different domains of study. The error in result is persistently constant. One can never be absolutely certain that all boners have been removed from data reduction, but diligent audit has failed to uncover residual flaws of that type.

One very important class of causes of model failure is that characterized by the realized p-values being different from the intended probabilities. That is, the actual proportion of respondents being called upon to answer Statement 1 may be different from the a priori probability that Statement 1 will be drawn. Within this class are found several quite separate kinds of possible circumstances. (We'll speak in terms of the decks of cards, although similar remarks might be made for the plastic box randomizing device.) It could be that the deck as used contained, say, 60 percent of the cards with Statement 1 rather than the intended 70 percent--and this could have occurred either because the deck was initially made up incorrectly, or because it changed constitution during the trials, having been dropped or otherwise acquiring an imperfectly operating state or condition. If, for example, Statement 1 had been answered 61% of the time rather than the expected 70% in the first sample (green deck), or 32% rather than 30% in the second sample (yellow deck), the basic model would yield almost perfect results. (The sampling standard error for p is approximately 0.012.)

The deck may have had correct proportions, and the drawing have been truly random, but the effective p-values still incorrect because of respondent reaction. One hypothesis might be that some proportion of respondents was sufficiently confused by the game--or not in sympathy with it--that they made up their mind that they would answer "yes" in any event, and so in fact, among those who drew Statement 1, the effective proportion answering Statement 1 was less than the proportion drawing the statement, the remainder answering no particular question, but just saying "yes." A number of plausible variations of this hypothesis are easily formed.

Another class of possible causes of failure might be termed response errors. This class, too, is broad and heterogeneous. One member of the class is misreading of Statement 1 by the respondent. The statement is

> "In the past 12 months there was a baby born in this household to an unmarried woman who was living here at the time."

It's not hard to imagine that in the mind of some readers the unmarried gets translated married, leading to a corresponding gross overstatement of "yes." It is possible, too, that the manner of the interviewer in presenting or clarifying the game to the respondent, or especially to those respondents who had concern or difficulty in replying, contained a bias that encouraged a "yes" answer. Another type of response error is conscious non-truthful reply by respondents. If the end result had been too few "yeses," this would have been a prime contender for the villain's role. But with the opposite factual result, deliberate untruthfulness appears less likely.

The problem was more difficult in the Statewide study than in the pilot study, because in the latter the true Π_1 was about 0.2, while in the former the true Π_1 was less than one percent--nearly zero. Contrastingly, and for a good reason, Π_2 was very much larger, and thus any operating deviation from expectation, even though slight, had a most substantial relative impact on estimating the near-zero Π_1. This is simply a case of the well-known fact that it is difficult to estimate near-zero parameters by sampling procedures. (We admit with embarrassment that we mistakenly thought in planning the study that the true Π_1 was of the order of 0.08 rather than 0.008. This "boner" arose because we were thinking of the proportion of births which were illegitimate rather than the proportion of households with an illegitimate birth in one year.)

When the basic model was formulated, an unrelated question was chosen for which it was possible to secure a criterion measure from an outside source. The reason for the choice was that use of the outside criterion would permit a modification in the model which might make it more efficient, or alternatively provide a means for validating results. At this point the outside source is mentioned only because we wish to note that some of the alternate models considered are quite sensitive to any error in the outside source and accordingly, it is necessary that the external data be true to within close tolerances if they are to be used.

We turn now to the observation that there are many possible alternate models which are more or less closely allied to the basic 2-deck 2-trial model discussed in Section V above. Some of them are consistent with one or the other of the speculations just offered. Indeed there are a number of these alternative models which are plausible. One example is presented in Section VIII below. Unfortunately we have as yet been unable to explore quantitatively other models. But in Section IX we call attention to several promising routes which illustrate types of models that deserve investigation.

VIII. Alternate Model I

This model assumes that the realized values of p_1 and p_2 are not as expected but are modified by a factor λ . This could occur for any number of reasons; e.g. a certain proportion of the respondents who actually have attribute A and select the statement referring to attribute A respond instead to their status concerning attribute B. The equations become

$$\gamma_1 = \lambda p_1 \Pi_1 + (1-\lambda p_1)\Pi_2$$

$$\gamma_2 = \lambda p_2 \Pi_1 + (1-\lambda p_2)\Pi_2 \quad ,$$

and can be rearranged to show that

$$\lambda(\Pi_1 - \Pi_2) = \frac{\gamma_1 - \Pi_2}{p_1} = \frac{\gamma_2 - \Pi_2}{p_2} \quad ,$$

yielding a solution for Π_1 in terms of λ and Π_2 :

$$\Pi_1 = \Pi_2 - \frac{\Pi_2 - \gamma_1}{\lambda p_1} \quad ,$$

or

$$\Pi_1 = \Pi_2 - \frac{\Pi_2 - \gamma_2}{\lambda p_2} \quad .$$

The estimator for Π_2 is the same as before, namely:

$$\hat{\Pi}_2 = \frac{p_2\gamma_1 - p_1\gamma_2}{p_2 - p_1} \quad .$$

It is assumed next that $\Pi_1 = 0$ in households not reporting any births. When Π_1 does equal zero, and data for these households are used, an estimate for λ is obtained:

$$\hat{\lambda} = \frac{\hat{\Pi}_2 - \gamma_1}{\hat{\Pi}_2 p_1} \quad ,$$

or

$$\hat{\lambda} = \frac{\hat{\Pi}_2 - \gamma_2}{\hat{\Pi}_2 p_2} \quad .$$

Finally, using these values, Π_1 is estimated with

$$\hat{\Pi}_1 = \hat{\Pi}_2 - \frac{\hat{\Pi}_2 - \gamma_1}{\hat{\lambda} p_1} \quad .$$

Results from this model are reported in Tables 4 and 5. The estimates of Π_1 in these tables are in fairly close agreement with the values expected. The estimates of the adjustment factor λ are all in the neighborhood of .82 for the various household groups. Negative estimates occurred in several instances for which the expected Π_1 is close to zero.

Since this model sets $\Pi_1 = 0$ for households not reporting births, and these are the vast majority of the households, it can be argued that the estimated Π_1 for <u>all</u> households will be forced to be close to zero. Despite this, the model behaved rather well yielding sensible estimates of Π_1 for white versus nonwhite households and for households reporting births.

IX. Outline of Other Alternate Models

We note here in outline only several types of alternative models which deserve further investigation. They may suggest still additional ways of utilizing the kind of information obtained in the North Carolina experiment.

<u>Alternate II</u>. Same as Alternate I except that adjustment to the p values is made in an additive rather than multiplicative form.

<u>Alternate III</u>. Utilizing the data in the yes-yes, yes-no, no-yes, and no-no cells, it is possible to set up equations with λ , Π_1 and Π_2 as the unknowns, and to solve simultaneously for all three parameters, using <u>all</u> the data (rather than just the no-birth households for estimating λ).

<u>Alternate IV</u>. Assuming the realized values of p_1 and p_2 are unknown and again using the data in the yes-yes, yes-no, no-yes and no-no cells, it is possible to solve for the four parameters p_1 , p_2 , Π_1 and Π_2 using an iterative procedure.

<u>Alternate V</u>. An estimate of Π_2 (proportion of respondents born in North Carolina) can be secured from an external source. Using that value, and the domain of households with no

Table 4

Parameters and Estimates Obtained with Alternate Model I

Item	All Households	White Households	Nonwhite Households	All Households Reporting Births
	Randomizing Device: Deck of 50 Cards			
$\hat{\lambda}$	.817	.813	.840	.816
$\hat{\Pi}_1$	.002	-.0003	.027	.037
Expected Π_1	.008	.002	.034	.042*
	Randomizing Device: Plastic Box of 50 Beads			
$\hat{\lambda}$	.811	.825	.745	.811
$\hat{\Pi}_1$	.006	-.0007	.020	.096
Expected Π_1	.008	.002	.034	.080*

* Based on data from matched birth certificates.

Table 5

Parameters and Estimates Obtained with Alternate Model I

Combined Data*

Item	All Households Reporting Births	All White Households Reporting Births	All Nonwhite Households Reporting Births
$\hat{\lambda}$	.816	.815	.829
$\hat{\Pi}_1$	.041	-.010	.123
Expected Π_1	.051**	.019**	.140**
$\hat{\Pi}_2$	.770	.711	.990
n_1	117	80	37
n_2	131	100	31

* Data for both randomizing devices were pooled for these estimates.

** Based on data from matched birth certificates.

births, derived values of realized p's can be obtained and then a solution for Π_1 for each of the two samples, the latter being averaged to obtain a final estimate for Π_1 .

Alternate VI. Using a known value of Π_2 from the external source, it is possible to solve the simultaneous system for unknown values of p_1 and p_2 in terms of Π_1 , and then a resulting quadratic equation for Π_1 .

In closing we should like to emphasize our belief that the randomized response method includes a very large family of techniques, which are just beginning to be recognized and explored.

APPENDIX A: INTERVIEW PROCEDURE FOR TWO TRIALS PER RESPONDENT

G. Randomized Card Question

(READ THE FOLLOWING STATEMENT TO THE RESPONDENT, BUT DO NOT HAND CARDS TO RESPONDENT UNTIL YOU HAVE FINISHED)

"Here are some cards, each of which has a statement on it. There are just two different statements. Each statement can be answered simply as 'true' or 'false.' You are to answer only one of the statements by picking a card from the deck at random. We are using this procedure so that I will *not* know which of the two statements you happen to select. Since both of the statements would be true for some people and false for others, I will not be able to tell from your answer which one *you* have chosen.

The cards with the first statement are marked with a circle and the cards with the second statement are marked with a square. When I hand you the cards you may look through them and read the statements if you wish. Then, shuffle the cards thoroughly and, *without looking*, select one of them, being careful not to show it to me. Simply answer 'true' or 'false' to the statement on the card you happen to select. Replace the card in the deck, shuffle the cards, and hand them back to me. (HAND CARDS TO RESPONDENT) (RECORD RESPONSE TO FIRST TRIAL)

Now that you fully understand how the game works, let's play it just one more time. Forget about the question you have just answered. Please shuffle the cards another time and select one of them, without looking or showing the card to me. Again, simply answer 'true' or 'false' to the statement you select, shuffle the cards and hand them back to me." (HAND CARDS BACK TO RESPONDENT. RECORD ANSWER TO SECOND TRIAL.)

G1. First Reply: True ☐ False ☐ Refused ☐

G2. Second Reply: True ☐ False ☐ Refused ☐

G3. Color of random card set used: Green ☐ Yellow ☐

G4. Respondent's name: ______________________ Time interview ended: ____________

ON THE CHARACTER AND INFLUENCE OF NONRESPONSE IN THE CURRENT POPULATION SURVEY

Susan Palmer, The Bureau of the Census

Introduction

An often frustrating component of all sample surveys which rely on public cooperation is the element of nonresponse. 1/ Failure to obtain observations from all members of the sample can lead to appreciable biases in the estimates and faulty inferences about the population. The character and influence of the nonresponse component of sample surveys are as varied as the surveys themselves. Their nature often depends on the subject of inquiry, methods of data collection, interviewer ability and motivation, public opinion, allotted time for survey completion, response rates, and even on the changing seasons. Survey methodologists have adopted a variety of methods to compensate for nonresponses. Some of these methods are rather sophisticated, such as setting up response models which take into account the probability that a household will be contacted and interviewed, or subsampling the nonresponse cases and carrying out intensive follow-up on the subsample to collect data for the nonresponse group. Other methods are more straightforward. The adjustment procedure used most often is an inflation of the interview universe, based on the assumption that the unknown, or nonresponse, elements are like the known, or response, elements.

Whether nonresponse presents a serious limitation on the total estimate is principally related to the nonresponse group and to the accuracy desired in the survey results. For surveys with moderate response rates where only crude measures are desired, the potential biases of nonresponse may be of little importance if what is known about the nonresponse elements indicates that they do not differ markedly from the response elements. Where a high degree of precision is desired, however, even a relatively low nonresponse rate can seriously affect the survey results.

The present paper concerns research conducted in connection with the Current Population Survey, a monthly household survey which is the source of the official estimates of employment and unemployment. As part of the effort to improve the accuracy of the CPS estimates, the Census Bureau has been actively concerned with the nonresponse problem - its size, its effect on the statistics, and how to improve the adjustment procedures for nonresponse. Realizing that serious biases can arise from differences between interviewed and nonresponse households, the Census Bureau has always placed great emphasis on keeping the CPS nonresponse rate at a minimum consistent with budget and time considerations (interviewing must be completed in one week). As a result, the nonresponse rate in the Current Population Survey averages around 4½ percent, and ranges from a low of 3 to 3½ percent in certain spring and fall months to a high of about 6 percent in some summer months--one of the lowest nonresponse rates among surveys where public cooperation is elicited.

As part of a broader study begun in 1963, observations on nonresponse households were obtained by carrying out an intensive field follow-up of such households during the three weeks following the CPS survey week in September 1965. Unfortunately, interviews were obtained for only about one-half of the survey households in the scope of this study. Because such a large proportion of these households remained inaccessible, the data presented here are themselves biased estimates for the nonresponse group. However, the results are consistent with what one would logically expect of CPS nonresponse, and probably understate the true differences between the nonresponse and response households. That is, the "hard core" nonresponse elements probably differ even more markedly, and in the same direction, than those for which observations were obtained.

The Character of Nonresponse Households

The information gathered on the CPS nonresponse households provides some insight into the character of these households and the persons within. Obviously, the probability of finding a qualified respondent at home at any time is related to the size of the household. It is not surprising that more than 1 in 4 of the nonresponse units are single-person households, as compared with a total household figure of only 1 in 7. The average number of persons 14 years and over in the response household in CPS is 2.3, while the nonresponse units average only about 2 such persons (See Table 1). The difference between the two groups is even more striking for total household size--3.4 persons per household for the interviewed units versus 2.6 persons for the nonresponse groups. We are not directly concerned in labor force statistics with children under 14 years of age. However, data on family characteristics, health, etc., from supplemental CPS questions do take into account the younger age groups.

1/ We are not concerned in this paper with the problems of incomplete information (e.g., where a sample unit is interviewed but some of the desired information is omitted, either inadvertently or for other reasons) but with only the problems of nonresponse caused by the refusal of the sample unit to cooperate or by its unavailability to the survey interviewer.

Nonresponse persons show a slightly higher median age than the interviewed groups (43 years as compared to about 40 years), caused by a conspicuous shortage of persons in the younger age categories. Almost 17 percent of the interviewed persons in CPS are between the ages of 14 and 19 years while only 12 percent of the nonresponse persons are of these ages (see Table 2).

Household sizes and population age distributions are relatively static, that is, households will exhibit the same general size and age distributions whether they are response or nonresponse in any month. The important point is that certain households, because of their composition, are more likely to be nonresponse than others.

With these demographic differences it is reassuring to find the labor force participation rate for the response and nonresponse groups about the same, and along with it the overall employment and unemployment rates (see Table 3). Large differences exist, however, in two labor force categories; persons working part time and those with a job, but not at work. About 17 percent of the interviewed labor force are part-time workers, whereas the comparable rate for nonresponse persons is only 12 percent. But even more striking is that while 4 percent of the interviewed labor force participants are temporarily away from their work for the survey period, about 10 percent of the nonresponse persons fall into this category. It is unlikely that these labor force differences can be attributed to the demographic composition of the nonresponse households. By comparing the labor force status of identical persons in response and nonresponse months we discovered that a definite change in the working routine had frequently taken place. 2/ Over half of the persons "with a job, not at work" during the nonresponse month had been working full time during a previous response month. Clearly, a large proportion of the nonresponse universe had shifted from "working" into the rather miscellaneous category composed of vacationers, leave-takers, and the ill.

A similar change is associated with nonresponse persons outside of the labor force. While 22 percent of the non-labor force interviewed persons are attending school during survey week, only 16 percent of the nonresponse persons are thus engaged at the time of the survey. Partially explained by the shortage of school-age children in the nonresponse group, the difference is also attributable to an influx of school-attending persons into another heterogenous category, "other." The "other" category appears to function in much the same way for nonresponse persons outside the labor force as the "with a job, not at work" category does for the nonresponse labor force participants. This non-labor force classification includes students temporarily away from class, housewives whose week's activities were not as they normally are, etc. To illustrate, consider the categorical changes in both labor force and non-labor force status of a family during a vacation month. We conclude, therefore, that certain selective shifts in status are associated with CPS households being nonresponse. They are not like the response households, nor are they like themselves, that is, they do not have the same survey characteristics that they had in the months in which they were interviewed. (This finding has a significant effect on proposed nonresponse adjustment procedures to be discussed later).

2/ The Current Population Survey is a rotation panel operation. Households are in sample for four consecutive months, drop out for the next eight months, and then return for four more months. 75 percent of the households in sample in any one month were in sample the preceding month.

THE UNAVAILABLES. The CPS nonresponse elements are of two basic types--those households which were not interviewed because they were unavailable during the survey time period, and those which, although contacted, would not cooperate with the survey. Mainly for the purpose of interviewing control the unavailable nonresponses are customarily divided into three groups on the basis of their reason for nonresponse. These are the "no one home," "temporarily absent," and "other unavailable" households. Since collectively these households normally account for about two-thirds of the total CPS nonresponse component, their individual natures, to a large extent, are responsible for the differences previously mentioned in the known and unknown universes.

The No One Home Households: This group represents those households which cannot be found at home by the interviewer after repeated calls at varying times. Probably related to their unavailability, the no one home persons 14 years and over have the highest labor force participation rate of any nonresponse group--about 68 percent of this population. 36 percent of the no one home households are single person, and the average number of persons 14 and over in these units is only 1.8.

The no one home rate of nonresponse fluctuates between 1 and 1½ percent of the total CPS households throughout the year.

The Temporarily Absent Households: The households of this type, unlike the no-one-homes, have no probability of being interviewed. The interviewer knows that they will be unavailable for the entire survey period. The temporarily absent households contain only 1.7 persons 14 years and over. In contrast to the no-one-homes, the persons in these households have an uncommonly low labor force participation rate (only 44 percent), and an unusually high percentage of them are retired (reflected in the non-labor force category of "other"). Most important, however, is the large proportion of persons "with a job, not at work"--about 28 percent of all temporarily absent persons in the labor force. These households appear to create the large difference between response and nonresponse with respect to this category. Their influence on the total component is greatest during the summer months when they account for almost ½ of all nonresponse cases.

The Other Unavailable Households: Representing only ½ of 1 percent of the total CPS sample, the other unavailable households have a negligible effect on the character of the nonresponse component. The most heterogeneous of the nonresponse types, this group consists of households which normally would have been interviewed except for extenuating circumstances (impassable roads, a death in the family, etc.). The only noteworthy attribute of its diverse members is their abnormally high "with a job, not at work" rate. Thus they contribute, however slightly, to the character of the total nonresponse universe.

THE UNCOOPERATIVES. Information on the characteristics of uncooperative nonresponse (about 1½ percent of the total CPS sample) was not collected in the manner used for the "unavailables." No attempt was made through follow-up personal interview to persuade the refusing persons to cooperate. Instead it seemed reasonable to assume no inherent labor force change associated with their refusing to be interviewed, but to regard their uncooperativeness as an individual factor--a reflection of the respondent's attitudes and personal feelings, not of his labor force status during the month in question. Therefore, demographic and labor force information for this subgroup of the population was obtained from months in which refusal households did cooperate and used as an indication of refusal household character. 3/

3/ Not reflected in the refusal data are those households which were never interviewed during their 8 months in sample. This necessary omission of information represents only about one-tenth of one percent of the total CPS sample.

The Refusal Households: We have yet to uncover any definitely correlated nonresponse characteristics for the refusal group. Whereas an obvious difference in household size was noted for the unavailable nonresponse types, in particular the no-one-homes and temporarily absents, the refusal household size of 2.3 persons 14 years and over is the same as the interview statistic. Also, the labor force status of the uncooperative persons in months when they were interviewed closely resembles the status of persons who had never refused during their eight months in sample.

The level of refusal nonresponse shows very little seasonal variation. Its effect on the total nonresponse component in CPS is static, and because of its close resemblance to the interviewed universe, actually serves to temper the effect of the unavailables and bring the total character of nonresponse closer to that of the interviewed.

The Influence of Nonresponse on Labor Force Statistics

When the survey results are in, and estimates are being made for the universe, the nonresponse cases must be taken into account either implicitly or explicitly. If the nonresponses are apparently disregarded by inflating the interview data to the sample frame, they have, in effect, been treated as though they were a representative sample of the universe. Thus, even the lack of an explicit method of dealing with nonresponse can be considered an adjustment procedure.

The nonresponse adjustment technique used in the Current Population Survey is a differential weighting system applied to the interview data by color and residence within each of about 70 groups of sample units on the assumption that the average population size and the labor force status of the interview and nonresponse households are identical within these groups. 4/ Although this assumption does not introduce appreciable biases in the major labor force categories such as employed and unemployed, the substitution of actual nonresponse characteristics for those "manufactured" to represent the nonresponse group brings to light troublesome differences in some of the minor categories (see Table 4). For example the difference between the published and the "true" proportion of persons "with a job,

4/ For a detailed explanation of the CPS nonresponse adjustment procedure, see Part VII of Technical Paper No. 7--"A Report on Methodology", Bureau of the Census, Dept. of Commerce, 1963.

not at work" is two-tenths of one percent, or about 5 percent of the published estimate. As the approximate standard error on this rate for the interview cases is about one-tenth of one percent, the imputation of interview characteristics for nonresponse households has substantially affected the reliability of the "with a job, not at work" category.

Nonresponse biases of about the same level as the standard error are noted for part time workers, school attenders, and "other--not in the labor force." Clearly, the unusually low level of nonresponse in CPS does not preclude the introduction of errors by the assumption of response and nonresponse comparability. One way of viewing this effect is to note that if the nonresponse bias could be eliminated, the sample size could be cut in half without increasing the total mean square error for these items (although obviously a decision concerning sample size in CPS would not be based only on consideration of the specific items that are seriously affected by nonresponse).

Several different imputation techniques for nonresponse were explored to determine whether any others would decrease this nonresponse bias in the sample estimates. First, we considered the substitution of labor force data from a previous interview month for current nonresponse households. The assumption here is that a household tends to exhibit similar labor force characteristics over a series of months independent of its interview status. As discussed previously, this assumption for nonresponse households is in error, and labor force estimates derived in this manner do not seem to represent any improvement over those derived by the present adjustment method.

Secondly, it seemed intuitively reasonable to expect the nonresponse units in CPS to resemble more closely the "hard to enumerate" households than the interview universe as a whole, the notion being that the substitution of a portion of the response data for the nonresponse might be more accurate than the present method of inflation. However, when comparisons were made between the labor force status of the households not contacted until the second or third personal visit and those never enumerated, the important differences in the nonresponse-related categories of "with a job, not at work," "school," etc., still existed.

Different statistics are likely to be affected in vastly different ways by nonresponse. As seen in the results of the CPS research, the fact that the estimates of primary importance are not seriously biased did not assure that all labor force statistics would not be affected. Nonresponse households do not constitute a single, homogeneous group. It is therefore difficult to anticipate the problems that they will cause in a particular survey. We do not presume that the findings presented here on the phenomenon of nonresponse apply to the general nonresponse universe, but are offered in an effort to provide some insight into the relationship between the known and unknown in sample surveys.

Table 1: Size of Household for the Total Population of the United States and for Nonresponse Households in the Current Population Survey, September 1965.

Size of household	Total households a/	Total nonresponse households b/	Unavailable nonresponse c/			Uncooperative nonresponse d/
			No one home	Temporarily absent	Other unavailable	Refusals
All households	100.0	100.0	100.0	100.0	100.0	100.0
1 person	13.1	28.5	36.8	34.1	31.2	16.4
2 persons	27.8	35.0	33.3	42.2	31.2	32.4
3 persons	18.9	13.8	10.2	11.6	10.1	19.6
4 persons	17.6	10.9	8.4	7.0	16.5	14.2
5 persons	11.5	5.3	4.2	2.7	4.6	8.3
6 persons or more	11.1	6.5	7.0	2.3	6.4	9.2
Average number of persons per household	3.4	2.6	2.3	2.1	2.5	3.0
Average number of persons 14 years of age and older per household	2.3 e/	2.0	1.8	1.7	2.0	2.3

a/ Source: Current Population Reports, Series P-20, No. 106, Table 3; March 1960.

b/ Estimates obtained by weighting the unavailable and the uncooperative households to the nonresponse sample level for September 1965.

c/ Data obtained from CPS records of households which were unavailable during the month of September 1965 but had been interviewed during a previous month. Unweighted sample totals for No one home - 285; Temporarily absent - 258; Other unavailable - 109.

d/ Data obtained from CPS records of households in sample as of December 1965 which, during their 8 months in sample, refused to cooperate in at least one month, and were interviewed in at least one month. Unweighted sample total - 720 households.

e/ Estimate the same in both March 1960 and September 1965 for total CPS households.

Table 2: Age Distribution for Persons 14 Years and Older in Response and Nonresponse Households in the Current Population Survey, September 1965.

Age	Total households a/	Total nonresponse households b/	Unavailable nonresponse c/			Uncooperative nonresponse d/
			No one home	Temporarily absent	Other unavailable	Refusals
Persons 14 years of age and older	100.0	100.0	100.0	100.0	100.0	100.0
14-19 years	16.7	12.4	13.9	6.7	12.1	14.8
20-24 years	8.8	9.9	15.1	8.2	7.5	8.0
25-29 years	8.0	8.3	12.5	6.9	6.5	6.8
30-34 years	7.9	6.5	7.6	4.8	9.5	6.0
35-39 years	9.1	7.0	7.0	5.6	3.5	8.6
40-44 years	9.5	9.0	8.3	5.9	10.6	10.7
45-49 years	8.4	8.8	9.1	5.0	11.1	10.2
50-54 years	7.7	8.8	7.4	7.8	11.1	9.8
55-59 years	6.7	8.6	7.8	10.2	8.5	8.3
60-64 years	5.4	7.0	5.0	10.6	5.5	6.8
65-69 years	4.8	5.7	2.8	10.2	8.0	4.6
70+ years	8.1	7.8	3.6	18.0	6.0	5.4
Median age	39.7	43.2	35.6	54.4	45.1	42.7

a/ Estimates include only those persons in households which were interviewed in September 1965. The CPS adjustment for nonresponse households is not reflected in the estimates.

b/ Estimates obtained by weighting the unavailable and the uncooperative households to the nonresponse sample level for September 1965.

c/ Data obtained from CPS records of households which were unavailable during the month of September 1965 but had been interviewed during a previous month. Unweighted sample totals for No one home - 503; Temporarily absent - 461; Other unavailable - 199.

d/ Data obtained from CPS records of households in sample as of December 1965 which, during their 8 months in sample, refused to cooperate in at least one month, and were interviewed in at least one month. Unweighted sample total - 1,658 persons.

Table 3: Labor Force Characteristics for Persons 14 Years of Age and Older in Response and Nonresponse Households in the Current Population Survey, September 1965

Labor Force Status	Response persons a/	Total non-response persons b/	Unavailable nonresponse c/			Uncooperative nonresponse d/
			No one home	Tempo-rarily absent	Other unavail-able	Refusals
Persons 14 years of age and older	100.0	100.0	100.0	100.0	100.0	100.0
In labor force	56.3	56.6	67.7	44.0	51.7	57.4
Not in labor force	43.7	43.4	32.3	56.0	48.3	42.6
<u>In labor force</u>	100.0	100.0	100.0	100.0	100.0	100.0
Employed	96.0	96.7	96.2	94.7	98.6	97.3
Working 35+ hours	75.0	74.2	78.8	63.2	68.9	76.7
Working 1-34 hours	17.0	12.1	12.3	3.5	14.9	14.9
With a job, not at work	3.9	10.3	5.2	28.1	14.9	5.7
Unemployed	4.0	3.3	3.8	5.3	1.4	2.7
<u>Not in labor force</u>	100.0	100.0	100.0	100.0	100.0	100.0
Keeping house	59.6	57.8	54.5	62.8	55.1	56.9
School	22.4	16.1	16.8	5.5	24.6	20.5
Other	15.2	22.4	22.8	29.7	15.9	19.3
Unable to work	2.9	3.6	5.9	2.1	4.3	3.3

a/ Estimates include only those persons in households which were interviewed in September 1965. The CPS adjustment for nonresponse households is not reflected in the estimates.

b/ Estimates obtained by weighting the unavailable and the uncooperative households to the nonresponse sample level for September 1965.

c/ Estimates are based on persons in households which were unavailable to be interviewed during survey week in September 1965 but were interviewed during a three week follow-up period (approximately 50 percent of the unavailable nonresponse households were interviewed during the follow-up). Unweighted sample totals for No one home - 313; Temporarily absent - 259; Other unavailable - 143.

d/ Data obtained for households which refused to cooperate at least once during the months of June through November 1964 and were interviewed at least once either before or after the refusal month. Unweighted sample total - 5,967 persons.

Table 4: Labor Force Characteristics for Persons 14 Years of Age and Older Using the Present CPS Nonresponse Adjustment Method, and as Estimated Substituting Actual Data for the Nonresponse Households in the Current Population Survey, September 1965.

Labor Force Status	Estimated using present nonresponse adjustment method a/ (1)	Substituting actual data for nonresponse households b/ (2)	Absolute difference between Columns (1) and (2) (3)	Approximate standard error on level of estimated percent in Column (1) (4)
Persons 14 years of age and older	100.00	100.00		
In labor force	56.24	56.27	0.03	
Not in labor force	43.76	43.73	0.03	
<u>In labor force</u>	100.00	100.00		
Employed	95.97	96.00	0.03	0.1
Working 35+ hours	75.04	75.00	0.04	0.2
Working 1-34 hours	17.02	16.88	0.14	0.2
With a job, not at work	3.92	4.12	0.20	0.1
Unemployed	4.03	4.00	0.03	0.1
<u>Not in labor force</u>	100.00	100.00		
Keeping house	59.57	59.50	0.07	0.2
School	22.38	22.18	0.20	0.2
Other	15.15	15.40	0.25	0.2
Unable to work	2.89	2.92	0.03	0.1

a/ The present nonresponse adjustment method is based on the assumption that nonresponse elements within a specific region-color-residence category are like the response elements in that category, and the response elements are weighted accordingly. For a detailed explanation see Part VII of <u>Technical Paper No. 7</u>, "The Current Population Survey--A Report on Methodology," Bureau of the Census, U.S. Dept. of Commerce, 1963.

b/ Data were collected for nonresponse households in a special study in September 1965. These data were used as an estimate for the nonresponse households and combined with data from response households to obtain this estimate.

IV

STATISTICAL AREAS

Chairman, WALTER F. RYAN, U. S. Office of Statistical Standards

STATE ECONOMIC AREAS -- A REVIEW AFTER 17 YEARS

Calvin L. Beale, U.S. Department of Agriculture

The delineation of the system of State economic areas developed out of the need to devise tabulation areas larger than counties but smaller than States for the migration data of the 1950 Census. Multi-county areas for migration data had been prepared for the 1940 Census by the then prominent geographer and demographer, O.E. Baker. However, because of the intervention of World War II, the 5-year interval migration data of the 1940 Census were neither published nor tabulated, although they were card punched.

In 1949, Donald Bogue of the Scripps Foundation headed a successful effort to salvage these data before the cards were destroyed. But, in connection with the 1950 Census it was decided not to use Baker's areas. Baker -- who had recently died -- probably had known as much as anyone about the regions of the United States. However, there was little information available about the basis of delineation of his areas. Furthermore, since 1940, the Standard Metropolitan Area system had been adopted and generalized to county lines outside of New England. These circumstances led to a decision to re-delineate the nonmetropolitan part of the Nation, using more systematic procedures than those under which Baker's areas were produced, and to use such areas with the SMA's as migration measurement units. Bogue got the assignment.

In the early stages of this work, it became apparent that the Agriculture Division of the Bureau of the Census desired a system of multi-county units for cross tabulations of the 1950 Agriculture Census. The Department of Agriculture had a set of type-of-farming areas available, but they were far too variable in size and consistency from State to State to be suitable. It was demonstrated that a general purpose set of economic units would be superior to the type-of-farming areas for agricultural tabulations even though not delineated solely on the basis of agricultural considerations. Thus, the system of State economic areas was prepared for use in the Censuses of Population, Housing, and Agriculture, with some subdivisions of areas for agricultural purposes being made that were not to be recognized in other tabulations. Program requirements made it necessary that standards of minimum population size and number of farms be observed wherever feasible to permit reliable tabulation of sample statistics, but that the total number of areas be kept within bounds to limit costs of presentation. In general, we endeavored to include at least 100,000 persons in each population-housing area, (although some exceptions were made) and 10,000 farms in each agricultural area (again with some exceptions in practice).

The nonmetropolitan economic areas were designed to be areas of relative homogeneity of resources, economic activity, geographic factors, population and social characteristics, rather than integrated, nodal areas. They were essentially production areas rather than marketing, commuting, or service areas. This is a major feature distinguishing them from the functional economic areas now being developed. The areas were based on social as well as economic criteria. Measures such as race, fertility, infant mortality, cultural history, and housing conditions were used in their delineation, but the intellectual climate of the time did not permit them to be called socio-economic areas.

The areas were announced in 1950 after a fairly extensive review procedure which included consultants in every State. A bulletin was published by the Census Bureau showing the boundaries and giving a number of statistics for the areas. Narrative descriptions of the character and identity of each area together with names for them, were not published until 1961 when the book, Economic Areas of the United States, by Bogue and myself appeared. The areas were not controlled by the Budget Bureau or officially designated by that agency as the Standard Metropolitan Areas were.

From the 1950 Censuses, reports were issued by economic areas (or subregional combinations) showing migration, rural housing, and agricultural data that were not also available for counties. The agricultural census use was repeated in 1954. In the 1960 Census, much more extensive migration materials on the SEA basis appeared and the rural housing report was repeated and improved. The reports of this Census also include a valuable compilation of base data by SEA's, but these figures are summations of county-level statistics and carry no new detail for SEA's. Under a project headed by Donald Bogue, special cross-tabulations of population data from the 1960 Census have been made by SEA's and are to be published by him in 1968. In the Agriculture Census of 1959, a subregional level of detail was substituted for the full set of State areas, and the data were not published but simply distributed to some interested users. Thus far, this is also the procedure used in the 1964 Census of Agriculture.

Use of the concept and of data based on it.--Considerable use has been made (1) of the State economic area census data, (2) of the SEA's as a means of grouping data for analysis, and (3) of SEA's as sampling universes for surveys. The Department of Agriculture has probably been the major single adopter of the system. USDA has computed major statistical series by SEA's on population migration, replacement measures for persons of working age, and farm level-of-living indexes, in addition to numerous studies in the field of agricultural economics. No systematic effort has been made to identify or count all of the uses made of the SEA system. I have encountered many uses of the concept in such a chance and unexpected way that I believe most of the use outside of the Department of Agriculture has not come to my attention and is not susceptible to complete compilation.

When I am asked whether I think that the system and data based on it have been satisfactory for most uses or adequately used in regard to the Census resources put into them, I find it difficult to give a categorical reply. There is an obvious need for a set of areas between the State and county level. Yet an almost infinite number of best areas could be devised for the infinite number of special uses and definitions desired. Area delineation is basically a "to each his own" proposition. Whether or not the SEA system has been used sufficiently is probably answerable only in terms of whether any other system would have been used to a greater extent; and in the absence of another existing system the question remains.

The discussants on the session at which this paper is being given illustrate two levels of criticism that have been directed to the SEA system. Otis Dudley Duncan has expressed the view that the "...agricultural tail was allowed to wag the Census dog in the organization of the SEA system". 1/ The authors of the system certainly tried not to do this, even though one of the purposes for which the areas were intended was agricultural analysis. Considerable weight was given to the major economic activity of the county groupings irrespective of whether or not this activity was farming. But Duncan's opinion is a matter of judgment and clearly an analyst of nonagricultural data might decide not to use the system if he felt it to be overly agricultural.

A second form of criticism is the flat assertion, made several times in recent years by Karl Fox, that the areas have "...been almost totally useless to social scientists". 2/ This is simply not correct, and would seem to reflect either a lack of acquaintance with the relevant literature or an implication that those who have employed the SEA's are not really social scientists. I am not sure which is the greater sin. But the assertion has been made sufficiently often in connection with the campaign for functional economic areas that some comment on the extent and variety of use of the SEA system is necessary.

One of the early uses of the SEA's was for a study of juvenile delinquency and dependency in Iowa. The author of the study (a professor then as now in the department of which Dr. Fox has long been head) commented "We have found (the economic areas methodology) much better than any other method known to us". 3/

In the late 1950's, Allan Beegle and Leo Schnore concluded in a "Memorandum on State Economic Areas" that "In general, State Economic Areas appear to serve the main purpose for which they were designed...." 4/ In 1958, I compiled a list of more than 50 demographic studies alone in which the system of State economic areas had been used. Of particular interest in judging the usefulness of the system to demographers was the fact that every study on the list had required either collection of original data or summation of county data by the authors. None were based on analysis of previously published SEA data, for practically no such data came out of the 1950 Population Census.

In 1964, I compiled an illustrative list of about 80 items that had appeared since 1958, largely demographic and economic (but again with no pretension of completeness). Shryock at this time concluded that "...various exhibits seem to me to demonstrate fairly extensive use of SEA's, in view of the fact that this Bureau has not really published a great many kinds of data for them". 5/

For more current usage -- and in the absence of comprehensive lists -- I should like simply to mention three recent publications. Each has come across my desk quite by chance within the 6 weeks prior to writing this paper, and each represents a different type of use of the SEA system. The first is an article by David Heer of Harvard University on Negro-White Marriage in the United States. 6/ In the study, Heer grouped data on Negro-white intermarriage by State economic areas in California, and then analyzed actual and statistically expected intermarriage percentages for each area. This is an example of the use of the system to aggregate data for meaningful areas where county frequencies are too small for analysis.

A second example is An Exploratory Analysis of the Roles and Role Conflicts of Vocational Teachers in Oklahoma, by Solomon Sutker and associates. 7/ In this work, the State economic areas were used for stratification purposes in the selection of a sample of high schools for study.

A third instance is the recent issuance by North Dakota State University, of a series of circulars on Crop Costs and Returns. 8/ A separate circular has been issued for each economic area in the State. The circulars are essentially work sheets on which the individual farmer can compare his production inputs and labor-management returns, crop by crop, with the usual costs and returns that research has revealed currently pertain for his economic area.

My own view is that the use and usefulness of the State economic areas has been neither exceptionally good nor poor. Personally, I like the system best in the South and Border regions, where a combination of small counties, and sharp changes in physical geography and cultural zones permit delineation of areas with a higher and clearer degree of interareal variation and internal meaning than is possible in some other parts of the nation. Because they consist of county building blocks, the utility of the SEA's is somewhat limited in some of the Western States where population is low and where individual counties are typically too large in area to be relatively homogeneous. But this problem occurs with any county-unit system.

I firmly believe that the areas would have received greater use if the Census Bureau had consented to name them. Although metropolitan areas have always been named, the published Census data for SEA's have been burdened with the mask of numerical anonymity. What does it mean to anyone but the constant user of the system to speak of Nebraska Area 1, West Virginia Area 4, or Texas Area 15? One can much more readily place and visualize data for the Nebraska Sand Hills, the West Virginia Southern Coal Fields, or the Lower Rio Grande Valley. Bogue and I, together with Shryock and Brunsman of the Census Bureau, made a strong plea in 1959 that names (as well as numbers) be employed in the 1960 Census publications, but without effect.

Another distinction between the status accorded the State economic areas and the standard metropolitan areas, is the fact that the metro areas are labeled as "standard" and are controlled and announced by the Bureau of the Budget. Such a designation both engenders and effectively forces greater use of a system. But Budget Bureau designation is anything but an unmixed blessing for an areal system in which statisticians or other researchers have an interest. It subjects a system to constant lobbying and political pressure over boundaries, wherever economic rewards are at stake -- if we can judge from the experience of the metropolitan areas. And the Budget Bureau is not always able to resist pressure for changes that constitute a violation of proclaimed standards, if the requests have sufficient political clout behind them.

The future.--What would we do differently if the areas were being delineated for the first time today, rather than in 1949? Obviously one great difference in the state of the arts is the availability of the computer. It would be practical to consider additional quantifiable variables and to engage systematically in formal tests of adherence to homogeneity or goodness of fit criteria (although such tests are in my opinion, often of more theoretical than real importance).

But perhaps a more crucial question is whether so-called uniform or homogeneous areas would be used again. The 1960 Census made it possible for the first time to seriously consider delineation of job commuting areas. The temptation would have been great, had such data been available in 1949, to forget horizontal similarity and go for vertical integration. Also, it must be recognized that with the decline in primary-industry employment since 1949, the great improvement in highways, and the corresponding extension of commuting zones, the relative logic of integrated commuting areas and interest in them is greater today than earlier. Although the principal proponents of integrated or so-called "functional" economic areas seem to feel it necessary in the advocacy of that system to disparage the State economic area system, I feel no reciprocal antipathy toward the functional approach. I do think there are surprisingly grave defects in the first version of a national functional area delineation submitted to the Budget Bureau earlier this year (1967). But the notion is a logical extension of the metropolitan area system and properly claims a place in the spectrum of area systems.

Must one system preclude the use of the other? I say no. They are not based on the same premises; they are not necessarily in conflict. The need to examine both the horizontal and vertical features of the spatial arrangement and organization of our society needs no defense. Each has its superior uses. This conclusion may come as bad news to the Bureau of the Census, of course, which is faced with the chore of identifying and providing at least minimal tabulations for as many systems of classification as it adopts.

The statistician's stake in the continuation of the SEA system in the population census is not tremendous, except in the case of migration data. Indeed, one of the criticisms made by Duncan in the 1950's was the fact that there had not been programs of cross-tabulation at the SEA level of data that were not otherwise available at the county level. But I think it essential that the 1965-70 migration tabulations be made by the SEA classification, both because of the need for time series comparison with the 1955-60 material and because I consider the SEA's to be superior to functional economic areas as units of migration analysis.

Functional areas tend to merge and obscure the common outmigration pattern of most hinterlands with the inmigration of the central and suburban counties, without there being any functional relationship between the two. For example, the Northern Blackland of Texas has three metropolitan areas within it -- Dallas, Waco, and Austin, of which Dallas in particular is a city of heavy inmigration. The nonmetropolitan part of the Blackland (Texas Area 8) is an area of net outmigration, but despite the proximity and dominance of the three cities mentioned, only a fourth of the gross outmovement from Area 8 went to those metropolitan areas from 1955 to 1960. On a functional economic area grouping, this coexistent condition of inmigration and outmigration would be masked by the inclusion of the hinterland counties with the functional centers. We would know less rather than more about migration patterns. But the State economic area system segregates the migration patterns more meaningfully.

A similar instance is Memphis. Shelby County is an area of net inmigration. The functionally related counties around it are all areas of net outmigration. But do the people from the surrounding counties seek Memphis as the primary destination? No, less than 15 percent of migration from the nonmetropolitan State economic areas contiguous to Memphis went to Shelby County from 1955 to 1960. From such considerations, as well as from the fact that we have not had consecutively comparable migration figures from any two of the last three censuses, it is my strong conviction that the basic unit of tabulation for streams of migration in 1970 should continue to be the State economic area. I am in favor, however, of identifying the 1965 county of origin on the basic tape so that special tabulations of data by combinations of counties other than SEA's would be feasible. This was not done in 1960.

In sum, the State economic area system has its advocates and satisfied users and it has its detractors and nonusers. It probably has not had as extensive a place in the statistical program of its sponsor, the Bureau of the Census, as was generally envisioned at the time of delineation. The general use of data based on the areas would be facilitated if the Bureau of the Census would employ names for them. Principal adoption of the system has come in demographic and agricultural research, although by no means to the exclusion of use in other fields. In recent years, interest in nodal areas -- especially those reflecting work commuting patterns -- has risen more rapidly than interest in areas of comparative homogeneity. These differently premised area systems serve different uses and properly should not be viewed as duplicative or competitive. The conclusion is offered that so far as data of the Bureau of the Census are concerned, the State economic areas are especially useful for tabulation of streams of migration. This quality is enhanced by the opportunity through continued use of SEA's in 1970 to compare migration streams with those of the recent past.

REFERENCES

(1) Duncan, O.D. (Undated). Comment appended to "Memorandum on State Economic Areas" by Allan Beegle and Leo F. Schnore.

(2) Fox, K.A. and Kumar, T.K. (1966). "Delineating Functional Economic Areas", Research and Education for Regional and Area Development. Iowa State University Center for Agricultural and Economic Development, p. 15.

(3) Lunden, W.A. (19[illegible]2). Letter to Howard C. Brunsman.

(4) Beegle, A. and Schnore, L.F. (Undated). "Memorandum on State Economic Areas" prepared for Bureau of the Census.

(5) Shryock, H.S. (1964). Memorandum to Conrad Taeuber.

(6) Heer, D.M. (1966). "Negro-White Marriage in the United States", Journal of Marriage and the Family, pp. 262-273.

(7) Sutker, S., Egermeier, J.C., and Twyman, J.P. (1967). An Exploratory Analysis of the Roles and Role Conflicts of Vocational Teachers in Oklahoma, The Research Foundation, Oklahoma State University.

(8) Rice, B.B. and Paul, R.R. (1967). Crop Costs and Returns, Circular FM North Dakota State University.

GENERALIZATION OF THE METROPOLITAN AREA CONCEPT

Brian J.L. Berry, University of Chicago

This paper presents a brief overview of a study undertaken at the Center for Urban Studies of the University of Chicago for the Committee on Areas for Social and Economic Statistics of the Social Science Research Council, under a contract between the Council and the Bureau of the Census, U. S. Department of Commerce. The study's purposes were:

-- "(to) conduct an examination of existing principles of area classification for Standard Metropolitan Statistical Areas, and an examination of alternative criteria, such as the concept of Functional Economic Areas, in order to formulate new principles of area classification."

-- "(to) examine the effect of applying alternative criteria of integration of central cities and their outlying areas in the delineation of Standard Metropolitan Statistical Areas and their relationship to other classification systems."

-- "(to) classify the entire United States into a hierarchy of urban, metropolitan and consolidated areas using criteria of size and of the linkages between places of work, places of residence, and places of shopping."

In the course of the investigation, five background papers, four large maps, a set of tables classifying the counties of the U. S. into functional economic areas, and a final report were issued. At the time this paper was being written, a few copies of the papers and reports were still obtainable from the Bureau of the Census. Copies of the maps are now being circulated for inspection by members of the audience. A final monograph will be published by the Bureau of the Census early in 1968. The tentative title is <u>Metropolitan Area Definition: A Re-evaluation of Concept and Statistical Practice</u>. Paralleling the contents of the monograph, in this paper I will first review briefly the history of statistical definition of metropolitan areas in the United States, the major types of criticisms of the criteria used to define Standard Metropolitan Statistical Areas in 1960, principal results of a study of the small-area journey-to-work data collected in the twenty-five per cent sample of 1960, and implications of these results for definition of statistical areas in the future.

History of Metropolitan Definition

Clearly, during the twentieth century both the scale and the pattern of the nation's urban growth have been transformed continuously and with increasing rapidity. These changes in the scale and pattern of American life were first recognized by the Bureau of the Census in 1910, when it introduced <u>Metropolitan Districts</u> to its system of area classification. This marked the first use by the Bureau of the Census of a unit other than the corporate boundaries of a city for reporting data on urban population. The Metropolitan District of 1910, defined for every city of over 200,000 inhabitants and reapplied with little alteration by the Bureau of the Census in 1920, 1930 and 1940, served basically to distinguish urban population, whether located within the central city or adjacent to it, from surrounding rural population. The idea behind the definition was in essence that stated in 1932:

> ". . . the population of the corporate city frequently gives a very inadequate idea of the population massed in and around that city, constituting the greater city, . . . and (the boundaries of) large cities in few cases . . . limit the urban population which that city represents or of which it is the center . . . If we are to have a correct picture of the massing or concentration of population in extensive urban areas . . . it is necessary to establish metropolitan districts which will show the magnitude of each of the principal population centers."

Almost as soon as the metropolitan concept was introduced to statistical practice, in the attempt to capture "the greater city," several factors led to dissatisfaction with the criteria and operational definitions used, or the results of their application, however. It is inevitable that any set of statistical areas transcending conventional legal jurisdictions will become the subject of local protest and political pressure. Almost any set of statistics will attract a coterie of users, too, and many of these users find weaknesses in the system for their particular purposes. Criteria used to operationalize something as fundamental and important as the metropolitan concept become the objects of academic evaluation and critique. And society itself continues to change, so even if criteria and areas may have been valid representations of conditions at a given period of time, they just as surely cease to be so in the course of time.

The resulting response has been one of successive modification of the definitional criteria. Metropolitan Districts were defined in 1940 for each incorporated city having 50,000 or more inhabitants, and included adjacent and contiguous minor civil divisions or incorporated places having a population density of 150 persons per square mile or more. In 1940, however, relatively few data were tabulated by minor civil divisions. At the same time, the various government agencies had no set of standardized regions for which they reported statistics. For example, <u>Industrial Areas</u> defined by the Census of Manufacturing, and <u>Labor Market Areas</u> used by the Bureau of Employment Security both differed from the Metropolitan Districts by which the Bureau of the Census reported data.

As a consequence, a further consideration introduced in developing the Standard Metropolitan Areas of 1950 was "so that a wide variety of statistical data might be presented on a uniform basis." The S.M.A. consisted of one or more contiguous counties containing at least one city of 50,000 inhabitants. Additional counties had to meet certain criteria of metropolitan character and social and economic integration with the central city in order to be classified within an S.M.A. Various governmental agencies cooperated to collect and report data by this statistical unit. The S.M.A. was by its very nature a compromise, designed to facilitate uniform reporting of data. It differed from the old Metropolitan District in that it was not defined primarily upon density criteria. The introduction of the Urbanized Area in 1950 provided a unit that fit more closely to the idea of the Metropolitan District.

The Standard Metropolitan Statistical Area of 1960 represents a slight revision of the S.M.A. concept, the word "statistical" being added so that the character of the area being defined might be better understood.

The primary objective of the S.M.S.A. has been stated to be to facilitate the utilization by all Federal statistical agencies of a uniform area for which to publish statistical data useful in analyzing metropolitan problems. The usefulness of the data has been related most especially to the fact that S.M.S.As. take into account places of industrial concentration (labor demand) and of population concentration (labor supply).

Two main claims have been advanced for the S.M.S.A. First, it provides a 'standard' area composed of a large city and its closely integrated surrounding area which can be used by the Bureau of the Census and other government agencies for purposes of data gathering, analysis and presentation. Secondly, the classification provides a distinction between metropolitan and non-metropolitan areas by type of residence, replacing the older rural-urban, farm-non-farm distinctions.

The S.M.S.A. has been used extensively as a reporting unit by many government agencies for publication of statistics. Statistical users outside the federal establishment have included local planning agencies, sales and advertising concerns, while much non-statistical use has been made of the classification by local boosters and political organizations in individual communities. Many of the non-Federal users of the S.M.S.A. data assume that the areas defined as metropolitan represent, in some measure, trading areas for the metropolis. Thus, use of S.M.S.A. data to establish quantitative indices of potential sales market areas, to set comparative guidelines for contrasting markets and market penetration, and to allocate man-power for sales and promotion efforts is common. Local and regional planners find S.M.S.A. data useful especially because of the quantity of information provided that would be unavailable to them otherwise, and because the areas are ready-made planning regions within which they can study broad trends of change relating to mobility, social and economic patterns of the population, and land use consumption. Recently, as an outcome of the Demonstration Cities and Metropolitan Development Act of 1966, many kinds of requests for federal public works monies must first be submitted to regional metropolitan planning agencies designated by the federal government. The Department of Housing and Urban Development has been given the responsibility for selecting the appropriate planning agencies, covering the relevant S.M.S.A., so the statistical units now have an increasing institutional superstructure.

Criticisms of the 1960 S.M.S.A.

At the very time that new legal status has been given to the current set of statistical areas, a wide and substantial volume of criticism is evident, however. The criteria used in 1960 sought to operationalize, in the words of the Bureau of the Budget, a "general concept of a metropolitan area . . . one of an integrated economic and social unit with a recognized large population nucleus." Population criteria (a central city of at least 50,000 people, or qualifying "twin" cities) were used to identify a set of nuclei, for each of which an S.M.S.A. would be created. Criteria of integration then revealed the outlying counties that had qualifying levels of integration with the nucleus (15 per cent of the outlying county's workers are employed in the central county, or 25 per cent of the workers in the outlying county live in the central county). Finally, criteria of metropolitan character eliminated those otherwise integrated counties that did not have 75 per cent of their labor force engaged in non-agricultural activities and failed to meet at least one of three additional considerations:

a) "It must have 50% or more of its population living in contiguous minor civil divisions with a density of at least 150 persons per square mile, in an unbroken chain of minor civil divisions with such a density radiating from a central city in the area."
b) "The number of nonagricultural workers employed in the county must equal at least 10% of the number of non-agricultural workers employed in the county containing the largest city in the area, or be the place of employment of 10,000 non-agricultural workers."
c) "The non-agricultural labor force living in the county must equal at least 10% of the number of the non-agricultural labor force living in the county containing the largest city in the area, or be the place of residence of a non-agricultural labor force of 10,000."

Each of the criteria used to define the S.M.S.A. has been subject to criticism from many points of view, viz:

Population Criteria

Questions have been raised concerning the basis on which the population criteria should be defined, concerning the necessity of a minimum and/or a maximum limit to population, and regarding the county and distance measures established in Criterion 2 for combining adjacent counties, each containing central cities, into a single S.M.S.A. On a more basic level, there is disagreement concerning the relation of population thresholds and economic organization.

Some authors have argued that the urbanized area should be used as the population base instead of the central city. The number 50,000 itself has been challenged on several scores. To some, that number seems too arbitrary and too large since a great many of the smaller centers of local activity in rural areas will be missed, thus over-emphasizing the importance of size in economic organization of space. Others feel that a city of 50,000 is really too small to constitute a metropolitan center, and that larger areas exceeding 250,000 people are most meaningful in an economic context today.

Criteria of Metropolitan Character

The criteria of metropolitan character have been subjected to heavy criticism and question. The criticisms arise, for the most part, from the vague and uncertain understanding of the meaning of this concept. No full or adequate apologia has been enunciated, and the social and economic connotations of the criteria have been subject to much debate. The evident compromise nature of the present definition has contributed considerably to the confusion.

At the most explicit level, questions about the selection of all particular thresholds have been raised. How does one justify a requirement that 50% of the population live in contiguous minor civil divisions with a certain minimum density? Further, how does one define the non-agricultural labor force? Where do part-time farmers fit in? Specific objections have been raised to the unique definition of the New England S.M.S.A.

In reviewing the comments addressed to it, the Bureau of the Budget has found numerous inconsistencies of application and a bewildering variety of choices made possible because of non-conformance to a few criteria by many counties. For example, the Bureau of the Budget found 38 areas in which counties otherwise qualifying as metropolitan have been excluded because of low total population, low total labor force, or insufficiently high population density.

General uncertainties of meaning are accompanied first, by specific questions about the apparent conflicts arising from defining metropolitan character in both economic and social terms. Second, issues of the urban-rural distinction, a distinction long indistinct, still appear to be built into the metropolitan character criteria in the language of density and size introduced by Wirth. Third, the definition ignores, except in the crudest sense, the question of the necessity for some landscape criteria by which to enunciate metropolitan character. The literature on metropolitan areas reveals a basic cleavage between scholars relying on some landscape element to form part of their definition and another group who find it unnecessary to include any specific reference to particular landscape features when discussing the concept.

Definition of the S.M.S.A. with reference both to social and economic criteria has created differing interpretations. It has been implied by some that the county was both a place of work and a home for concentrations of non-agricultural workers while, at the same time, functioning as the primary trading area for the metropolis. Are either or both of these conditions necessary for a county to be metropolitan in an economic sense? Some evidence suggests that wholesale trading territories for large metropolitan areas are coterminous with farm to city migration areas, suggesting a correspondence of boundaries of several indicators of metropolitan economic influence, and that retail trade areas are coincident with commuting areas for smaller places. In agricultural areas and around smaller S.M.S.A. central cities, these findings notwithstanding, others have argued that the general trade area of the central city covers a more extensive terrain than does any kind of extended migration or commutation zone. Further knowledge about commuting patterns will elucidate the unknowns here. It is likely, however, that the patterns will vary for metropolitan areas of different sizes and in different parts of the country. If one refers to a "metropolitan economy," then it is clear that the larger S.M.S.A.'s are underbounded. If one refers to activity patterns of individuals and groups living within metropolitan areas, then it is clear from research that there is little difference between groups included within metropolitan areas and some of those which are excluded. The differences appear to be more distinct between workers engaged in urban pursuits and those engaged in rural agricultural pursuits. If by metropolitan character of an area we mean the use of that land by various groups, then it is clear that the sphere of influence of metropolitan dwellers extends far beyond the counties currently classified as metropolitan. At this point the discussion reverts to the problem of interpreting what is meant by "metropolitan".

Criteria of Integration

The main thrust of criticism of the criteria of integration is to demand that a more precise and detailed statement about economic and social integration within the metropolitan area be made.

The percentage figures established by the Bureau of the Budget have been questioned. The necessity for direct contact with the central county has been questioned by pointing to the lack of unified labor markets within large metropolitan areas. The achievement of maximum accessibility throughout the metropolitan area with reductions in the cost and time required for travel has led to the suggestion that a commuting radius be established on the basis of time taken to reach the central county or its central area.

The whole question of integration without what is commonly thought to be metropolitan character is implicit in several of the classification schemes. The classifications suggested by both Friedmann and Miller, and by Fox revolve around a notion of integration without the accompanying population density criterion now closely associated with metropolitan character. These schemes propose a radical alternative to our present definition of the metropolitan concept. Friedmann and Miller see a changing scale in urban life accompanying technological and economic developments. Such an idea rejects as no longer useful the classification distinguishing metropolitan from non-metropolitan, and it suggests that a new and broad urban realm is significant. The argument rests largely on the claim that the area in which a metropolitan population lives and conducts its social activities now encompasses a broad zone around metropolitan centers. This zone, or realm, extends, perhaps, to about 100 miles from the central city, and is defined as the limits for regular week-end or seasonal use. Within this area, the imprint of the urban dweller is of paramount significance. This realm is largely coincident with areas of general economic health as well, they maintain.

Fox is concerned with small, functionally specialized regions which he considers to be the major facts of economic importance in the regionalization of most of the country. Integration here is often without metropolitan character since many of the smaller centers are too tiny to be classed as metropolitan under present schemes or because population densities may be low. Nevertheless, Fox posits such a system of functional economic areas as the economic building blocks for a regionalization of the United States.

The Journey-to-Work Evidence of 1960

Analysis of the small area commuting data collected as part of the 1960 census shows that a set of urban realms in fact constitutes the nation's functional economic areas. This finding leads to proposals for a revised area classification that lends itself to a range of practical applications within the framework of emerging national urban policy.

In 1960, you will recall that for the first time, the Census of Population and Housing included a question to determine the commuting behavior of the population of the United States. Item P 28 of the Household Questionnaire read:

P28. What city and county did he work in last week?

Individual and household data were assembled into totals for each of the country's 43,000 Standard Location Areas. For each S.L.A. it was then decided what initially appeared to be the thirteen most important workplace locations for residents from among a set of 4,300 possible workplaces. Theoretically, then, the 1960 journey-to-work data were assembled into a 43,000 by 4,300 matrix of from-to journey-to-work information; however, the SLAs were in fact grouped into some 4,300 sub-matrices, each of which had only 13 columns and two balance categories.

Commuting Fields and Labor Markets

For any workplace, a reporting booklet could be prepared listing all SLAs sending commuters to it, and for each of the SLAs showing how many and what proportion of the resident workers travelled to each of its thirteen workplace alternatives or fell in one or the other of the balance categories. With such information in hand it was possible to plot a map for each workplace and surrounding territory, showing the percentage of the workers resident in each S.L.A. commuting to the workplace. Because of the regular decline of the commuting rate with distance, it was also possible to contour the percentages to depict the commuting field of that workplace. The outer limit of this field is described by a zero contour beyond which there is no reported inward commuting; this is the area within which jobs and homes are brought into balance--the area which serves as a bounded "container" for the journey-to-work.

Commuting fields were mapped for every S.M.S.A. central city, for most urban centers in the 25,000-50,000 population range, and many small places. The complete set of commuting fields is depicted on the map Commuting Fields of Central Cities, which has been circulated to the audience.

An immediate contrast may be drawn between the map of the country's Standard Metropolitan Statistical Areas, as defined by the Bureau of the Budget, and the map showing areas within the commuting fields of cities in 1960. Whereas two-thirds of the nation's population resided in the 1960 S.M.S.As., in fact 87 per cent lived within the commuting area of one of the 1960 S.M.S.A's. central cities (many within more than one such area). Another 9 per cent lived in the commuting fields of somewhat smaller centers that filled the populated gaps between metropolitan labor markets.

In fact, then, in 1960 the populated parts of the nation were completely metropolitanized -- covered by a network of urban fields. They were also patterned socially and economically by them. Each of the commuting fields shows a very fundamental property of the country's residential areas: degree of participation in metropolitan labor markets. As degree of metropolitan labor market participation declines with increasing distance from the city, population densities and proportion of the population classified as urban decline, together with average value of farmland and buildings, median family income, median school years completed, rate of population increase (which becomes negative in the peripheries) and per cent gain in the population through migration (which also becomes negative). On the other hand, percentage of the population classified as rural nonfarm rises and then falls, and both the percentage of families with incomes less than $3,000.00 and the unemployment rate increase.

The changes are like a musical score; they are rhythmic, rising and falling in concert. Population response is revealed by relative rates of change at the center and decline at the periphery.

Only urban centers with populations exceeding 25,000 appear to have much peaking effect on the gradients. For larger sizes of central city the "peaks" rise with rank in the urban hierarchy, up to the level of the nation's largest metropoli. But in the latter there is an involution, with gradients dropping in the inner-city ghettos. Also, where labor markets overlap, and substantial cross-commuting results from alternative employment opportunities, regional welfare levels are maintained at high levels at the outer edges of the commuting fields. The journey-to-work data thus indicate patterns of labor market participation in metropolitan commuting fields that are very profound indexes of socio-economic rhythms present in the characteristics of the population of the United States.

Several other things are apparent in the maps of commuting fields and the related socio-economic gradients:

(1) Clearly, the areas socially and economically integrated with given central cities are far more extensive than the 1960 S.M.S.A.s. This should be no surprise, given use of the 15 per cent commuting criterion plus the criteria of metropolitan character to reduce and constrain their boundaries. However, if the commuting patterns and resulting variations in degree of metropolitan labor market participation and related socio-economic gradients are to be considered seriously, these constraints cut accross continous, correlated patterns rather than seeking out real limits such as discontinuities or major transitional zones. In fact, the only such limit evident in the data is where one commuting field leaves off and the socio-economic characteristics begin to respond to the pulls of another central city.

(2) Similarly, in the least densely-populated parts of the nation's settled area, commuting fields focus on urban centers of less than 50,000 population, although sizes must in general exceed 25,000 to have any effects on the socio-economic gradients flowing outwards from larger places.

(3) At the other extreme, particularly in the manufacturing belt, labor markets overlap in elaborate ways. The urban regions of "megalopolis" are highly complex, multi-centered entities.

At least three questions of definitional practice are raised in view of these considerations:

(1) If the intent is to define "economically and socially integrated units with a recognized large population nucleus," are the limiting constraints of the criteria of metropolitan character and the 15 per cent commuting criterion desirable and reasonable? In light of the evidence, we think not.

(2) What is an appropriate size limit for the central city of the statistical area, and indeed, is the size of the central city a valid population criterion? Although one may want to start with the 50,000 size for historical reasons, the total population of the entire region is probably more interesting.

(3) How is the complexity of the most densely-populated parts of the country to be handled? Clearly, no units focusing on single centers will be able to embrace the interdependent labor markets. Is a multi-centered urban region an appropriate substitute? We think so, and suggest a comprehensive set of Consolidated Urban Regions.

The steps taken in 1960 clearly overcame these issues by beginning with a prior definition of a set of centers, for each of which a metropolitan area was to be built of county building-blocks. The commuting criterion then pointed out counties potentially eligible for membership in each of the S.M.S.A., and the criterion of metropolitan character led to elimination of some of the potential candidates. Lip-service was given to interdependencies by creation of the New York and Chicago Consolidated Regions. The whole procedure was simple, straightforward, and easy to apply.

Functional Economic Areas and Consolidated Urban Regions

Are there equally simple alternates that start with the same population criterion, rely on county building-blocks (one of the features of the journey-to-work small area data is that county units may be retained without undue loss of detail) in the same way, but come closer to real areas of daily journey-to-work interdependence?

Proposed Definitions

Considerable experimentation with the journey-to-work data led to the following set of definitions, which goes a long way to providing a viable series of alternates:

1. COMMUTING FIELD

   An area encompassing all standard location areas sending commuters to a designated workplace area. The field varies in intensity according to the proportion of resident employees in each SLA commuting to the workplace, and may be depicted cartographically by contours that enclose all areas exceeding a state degree of commuting.
   Note: There will be as many commuting fields as there are designated workplace areas.

2. LABOR MARKET

   All counties sending commuters to a given central county.

   2a. CENTRAL COUNTY

   The designated workplace area for definition of a labor market.

   2b. CENTRAL CITY

   The principal city located in a central county.
   Note: S.M.S.A. criteria 1 and 2 might be carried through to further specify 2a and 2b.

3. FUNCTIONAL ECONOMIC AREA (F.E.A)

   All those counties within a labor market for which the proportion of resident workers commuting to a given central county exceeds the proportion commuting to alternative central counties.
   Note: There will be as many F.E.A.'s

as there are central counties.

4. CONSOLIDATED URBAN REGION (C.U.R.)

Two or more F.E.A.'s for which at least five per cent of the resident workers of the central county of one commute to the central county of another.
Note: No prior determination of the number of C.U.R.'s is possible, but application of the criterion to the 1960 data produced 31.

Results of applying these criteria are shown in the maps Functional Economic Areas of the United States and Consolidated Urban Regions.

The regionalization used to create the 1960 S.M.S.A.'s and the functional regionalization evidenced by commuting behavior are significantly different. A major choice must be made by the U. S. Bureaus of the Budget and Census, for the 1960 classification does not produce fully-integrated areas with a large population nucleus even though this was the underlying concept. Is the intention to classify areas on the basis of how they look? In this case, continuation of present practice will suffice, and attention should be focused on the criteria of metropolitan character (although continuation of the practice of defining urbanized areas may be a more appropriate substitute). Alternatively, should the areas embrace people with common patterns of behavior? Then, commuting data dealing with daily behavior and the links between place of residence and place of work are relevant.

Comparability is not the issue if county building-blocks are used. Besides, there has been little consistency in definitional practice since inception of attempts to define metropolitan areas. Nor should consistency be expected in a dynamic socio-economy in which patterns of organization and behavior are subject to continuing change.

There is a hard problem of choice, since there is general agreement that some form of area classification will be required for publication of summary statistics for some time to come.

We recommend the following:

1. County building-blocks or equivalent units be retained as the basis of any area classification, in all parts of the country.

2. County-to-county commuting data be the basis of the classification of counties into functional economic areas.

3. Functional Economic Areas be delineated around all central counties satisfying the existing S.M.S.A. criteria 1 and 2, and in addition be created for smaller regional centers in the less densely-populated parts of the country.

4. Where significant cross-commuting takes place, functional economic areas be merged by the creation of a consistent set of Consolidated Urban Regions.

5. Consideration be given, for neatness of social accounting, to allocating all unallocated counties to one of the F.E.A.'s or C.U.R.'s on the basis of additional criteria of regional interdependence.

STATUS OF "ZIP" AS A TOOL OF MARKETING RESEARCH

Dik Twedt, Oscar Mayer & Co.

About 30 months ago, a combined task group of the American Marketing Association and the United States of America Standards Institute (formerly the American Standards Association) began work on a project to evaluate and recommend a single system of standard geographic units --- to be used for purposes of both communication and analysis. One of the first decisions of this task group, on which I was asked to serve as chairman, was that there would be "no further needless proliferation of systems, and that first consideration would be given to any existing system that met the requirements of widespread usage and acceptance, together with provision for maintenance and dissemination of the system."

After careful consideration of the major systems in use, including those of the Bell Telephone System, the transportation industry, Federal agencies, and others --- the task group agreed that the Post Office Department's system of approximately 580 ZIP Sectional Centers probably had the greatest investment (by both Government and business) and the widest public acceptance.

A recent estimate by Bernard Fixler, speaking at the National Postal Forum here in September, was that mailers have already spent more than $200 million to ZIP Code their lists, with sizeable additional expenditures committed to maintenance. As for public acceptance, a study completed last October by Roper Research Associates revealed a 98 per cent awareness level, and 78 per cent actually knew the ZIP Code for their area. A check by the Post Office Department of more than one million pieces of mail in 225 post offices showed that 74 per cent of envelopes actually contained ZIP Codes.

As most of you already know, beginning with the Census of 1970, there will be a built in capability to report Census data by Zip because about 55 of the 70 million addresses to be covered will receive questionnaires by mail, and the ZIP Code will already be incorporated in the source document --- the questionnaire --- as a necessary part of the address. The remaining 15 million addresses in non-metropolitan areas will be canvassed by personal interview and therefore ZIP Code data will not be as easily available.

ZIP Codes are not now a part of the machine readable data record used for making tabulations of Census results. However, Census could make ZIP part of the data record fairly easily for about 45 million of the 55 million mailing addresses in metropolitan areas, because they will be on magnetic tape to be purchase commercially. Census has estimated that incorporation of Zip data into the 10 million metro list would cost an additional $300,000 --- and to incorporate the remaining 15 million non-metro addresses would cost substantially more.

These capital costs, plus processing costs of rearranging 70 million records to permit ZIP tabulations at the detail record level, raises legitimate questions about the extent to which such data are likely to be used. The growing interest in marketing analysis by some form of ZIP classification suggests that it is already making substantial contributions toward the advancement of scientific method in marketing. In a mail survey conducted by the Federal Statistics Users' Conference last July, one in four of the 64 respondents reported present use of some form of ZIP for marketing and statistical analysis. And of those not now using ZIP for analytic purposes, more than half reported interest in the collection and tabulation of Census data by some form of ZIP classification.

For those who want to get started before 1970 data are available on this basis, a commercial organization, the Yuan Liang Marketing Service of Chicago (ZIP Code 60611), has developed "Zip-O-Data" --- a service which provides estimates of income, education and per cent single family dwelling units for each of approximately 40,000 five-digit ZIP Codes, and for Sectional Centers as well.

A proposal by the Marketing Department of Northern Illinois University for research in ZIP Sectional Center coding as a sales research tool resulted in the Goodman Grant for 1967 to the university by the Sales Promotion Executives Association, International. A ZIP Conference was held at Northern Illinois University last October, and the proceedings of this Conference will soon be available to those who are interested.

At this year's annual meeting of the Direct Mail Advertising Association, an example was given of how direct mail of the future may be segmented and personalized by computer. A subscription letter to college students, for example, can include the college attended, the recipient's major, his year of graduation, and even the approximate distance from his school to his home --- and that last part comes right out of the memory bank which contains a table of distances between ZIP Sectional Centers.

Martin Baier's fine article in the HARVARD BUSINESS REVIEW for January-February, 1967 describes how ZIP is being used by a large insurance company for marketing analysis.

For print advertising media, particularly those with large propostions of their audience delivered by mail, ZIP provides an obvious method of segmentation by geographic markets. LOOK magazine, which was one of the first national magazines to embrace the geographic editions concept that comes rather logically out of ZIP, has recently added a new dimension, the "Geodemic Edition," which begins with the February 20, 1968 issue. The Geodemic Edition will contain all regular advertising and editorial material, plus appropriate regional advertising and editorial pages, plus "Top/Spot" advertising. Out of the 40,000 ZIP Codes, LOOK has selected the 1600 with the highest median incomes based on Census tract data --- so that it is now theoretically possible to reach the one million richest readers of LOOK for about $10,000 a page.

Now what kinds of data will actually be available from the next Census? The present plans of the Bureau of the Census for 1970 are best describes by the following quotation from "Plans for the 1970 Census of Population and Housing" --- an article by Census' David Kaplan, in the November, 1967 issue of the Bureau of the Budget's Statistical Reporter (p. 76):

> "The Bureau proposes to provide for the capability to produce statistics for zip codes areas. This is likely to be on the basis of the 3-digit sectional centers generally, and perhaps the full 5-digit areas in the larger cities. To accomplish this purpose, each enumeration district (or block in some cases) will be assigned its appropriate zip code. Where a zip code boundary cuts through an enumeration district, determination will be made as to which single zip code to give it. The result will, therefore, be adequate for the purposes of the many prospective users who have requested that this capability be created. The data produced by this approach would not be published, but would be made available on a reimbursable basis."

In the short time we have available today, I have obviously covered only a few of the current applications of ZIP to marketing analysis and experimentation. Many other large companies and trade associations are also working in the same area. For example, a Cleveland company --- Spade Drills Incorporated --- has found a need for a listing of ZIP Codes by state, then county, with all three-digit ZIP Codes for each county. Copies of this directory are available from Spade Drills Incorporated for $20 (ZIP Code 44105).

I would like now to discuss briefly a model with which I am quite familiar, since it is one on which our own company has been working for the past several months. Perhaps you will find it useful in your own efforts.

In an attempt to provide a relatively stable geographic model for sales analysis, records were set up so that sales data could be collected by county and consolidated into 200-plus television marketing areas decribed by the American Research Bureau (based on a mail ballot of TV viewing). These areas have the advantage that each surrounds a major city, and each is covered by major media out of a central source. Since county lines are respected, the usualy demographic data are readily available.

This concept was helpful in setting sales goals, measuring relative volume, allocating promotional charges to sales units, and evaluating the effectiveness of alternative programs --- but there were four problems associated with this approach:

1. ARB markets change annually, and are completely resurveyed every five years.

2. Though broadcast cast media may reach all parts of each market, markets are not always tied together as well by highways --- a primary means of distribution.

3. Sales measurement by county calls for a special state and county code in all pertinent records. Look-up maintenance of this code became a very considerable chore.

4. There are more areas than were needed for sales analysis at a national level.

DEVELOPMENT OF ZIP MARKETING AREAS

Following careful consideration of existing geographic models, a decision was made that ZIP offered the most usable base for a system requiring minimum look-up and maintenance. Since ZIP Sectional Centers do not follow county lines, it was necessary to "force a fit." ZIP Sectional Center areas that split counties were extended to include all of each county in which Sectional Centers included the majority of the population. Next, ZIP Sectional Center areas were grouped to match TV market (as defined by ARB) as closely as possible. Finally the resulting areas were grouped into 136 new "ZIP" Marketing Areas (see map) by overlaying known distribution patterns of major food chains.

As with most working models, the results had both advantages and disadvantages. Briefly, the advantages were these:

1. The ZIP number is widely understood and accepted by all segments of business and the general public.

2. At the ZIP Sectional Center level, the system is relatively stable.

3. Since ZIP is increasingly considered to be an essential part of a deliverable mail address, the user's look-up problem is minimized.

4. Because the ZIP Code is already in many addresses, additional place codes are unnecessary, which makes for more efficient data processing.

5. Maintenance (including dissemination of changes) is greatly simplified through availability of up-to-date Post Office Department directories.

6. ZIP is geared to the problem of delivering mail, and there is a fairly close parallel between highway networks and mail distribution, and the normal channels of distribution used by marketing organizations.

7. By aggregating counties into broad groups, it is possible to make historical comparisons (although a fair amount of effort is required to do so).

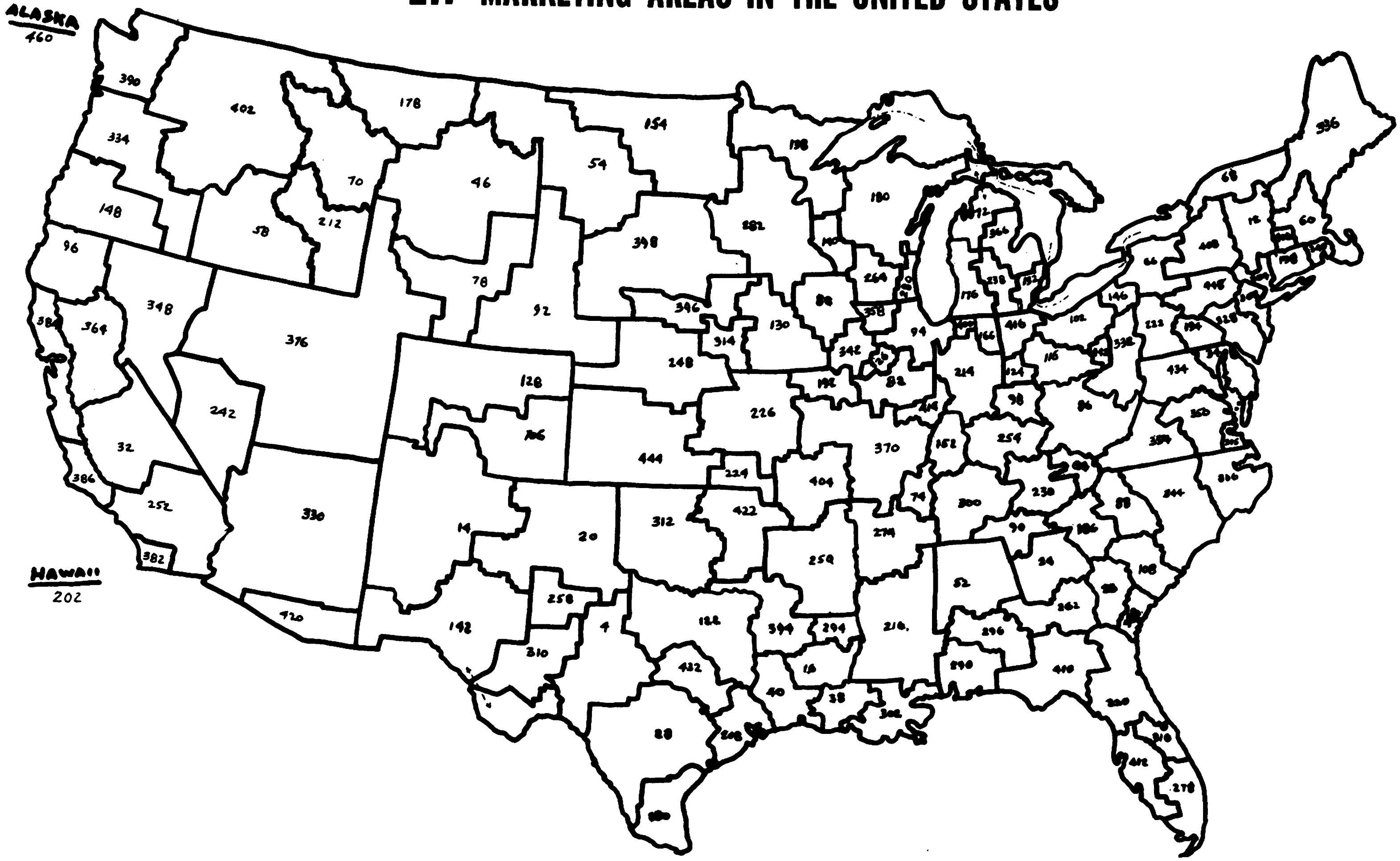
But there were also some disadvantages:

1. In the arbitrary allocation of "split" counties to a given ZIP Marketing Area, the possibility obviously exists that a part of the split county (even though a minority) actually belongs in a neighboring marketing area. Although this does happen, fortunately it tends to be a rather minor problem. In practice, the question arose in about 10 percent of the counties, and most of the time the "split" was so disproportional that that allocation was obvious. In every case where the problem existed, it was in fringe population areas (rural and small town).

2. Some counties fell in one ZIP Marketing Area, but in another area of dominant television influence. Again, this situation occurred only in fringe population areas. In the actually application of this system, it was found that less than four per cent of the U.S. population lived in counties which ARB classified in one area, but our system placed in a different ZIP Marketing Area. Two out of three of these "switched counties" actually seemed to fit better (in terms of our existing distribution patterns) in the ZIP Marketing Areas, than they did in the ARB area of dominant television influence. This is understandable of course, in view of the fact that ZIP was planned in accordance with actual distribution traffic patterns, whereas ARB data are subject to normal sampling error.

In summary, we believe that ZIP provides a widely accepted, well understood basis for a relatively stable model of geographic analysis of marketing activity. Installation and maintenance cost of such a system is minimized (since ZIP is already being used for a quite different reason --- efficient mail delivery). Although the system of ZIP Marketing Areas described in this paper falls short of complete rigor, it has served as a working model of considerable empirical value.

ZIP MARKETING AREAS IN THE UNITED STATES



Ref. No.	Area & Counties	Number of Households 1967-01-01 (000)
004	**Abilene, Texas** Texas: Brown, Callahan, Coke, Coleman, Concho, Crockett, Fisher, Haskell, Irion, Jones, Kent, Kimble, Knox, McCulloch, Mason, Menard, Mills, Mitchell, Nolan, Reagan, Runnels, San Saba, Scheicher, Scurry, Sterling, Stonewall, Sutton, Taylor, Tom Green, Val Verde	127
012	Albany, New York New York: Albany, Columbia, Fulton, Greene, Hamilton, Montgomery, Rensselaer, Saratoga, Schenectady, Schoharie, Ulster, Warren, Washington Vermont: Addison, Bennington, Rutland	492
014	Alberquerque, New Mexico Colorado: Alamosa, Archuleta, Conejos, Costilla, Dolores, La Plata, Mineral, Montezuma, Rio Grande, Saguache New Mexico: Bernalillo, Catron, Colfax, Harding, Los Alamos, McKinley, Mora, Rio Arriba, Andoval, San Juan, San Miguel, Santa Fe, Sierra, Socorro, Taos, Torrance, Union, Valencia	187
016	Alexandria, Louisiana Louisiana: Avoyelles, Grant, LaSalle, Natchitoches, Rapides, Sabine, Tensas, Vernon	89
020	Amarillo, Texas New Mexico: Curry, De Baca, Guadalupe, Quay, Roosevelt Oklahoma: Beaver, Cimarron, Texas, Texas: Armstrong, Briscoe, Carson, Castro, Childress, Collingsworth, Cottle, Dallam, Deaf Smith, Donley, Floyd, Foard, Gray, Hale, Hall, Hansford, Hardeman, Hartley, Hemphill, Hutchinson, King, Lipscomb, Moore, Motley, Chiltree, Oldham, Parmer, Potter, Randall, Roberts, Sherman, Swisher, Wheeler	172
024	Atlanta, Georgia Georgia: Banks, Barrow, Bartow, Butts, Carroll, Cherokee, Clarke, Clayton, Cobb, Coweta, Dawson, DeKalb, Douglas, Elbert, Fannin, Fayette, Floyd, Forsyth, Franklin, Fulton, Gilmer, Greene, Gwinnet, Habersham, Hall, Haralson, Hart, Heard, Henry, Jackson, LaMar, Lumpkin, Madison, Meriwether, Morgan, Newton, Oconee, Oglethorpe, Paulding, Pickens, Pike, Polk, Rabun, Rockdale, Spalding, Stephens, Taliaferro, Towns, Troup, Union, Upson, Walton, White, Wilkes	611
026	Augusta, Georgia Georgia: Bulloch, Burke, Candler, Columbia, Emanuel, Evans, Glascock, Jefferson, Jenkins, Lincoln, McDuffie, Montgomery, Richmond, Screven, Tattnall, Toombs, Treutlen, Warren South Carolina: Aiken, Allendale, Barnwell, Edgefield, McCormick	134
028	Austin, San Antonio, Texas Texas: Atascosa, Bandera, Bastrop, Bee, Bexar, Blanco, Burnet, Caldwell, Calhoun, Colorado, Comal, DeWitt, Dimmit, Edwards, Fayette, Frio, Gillespie, Goliad, Gonzales, Guadalupe, Hays, Jackson, Karnes, Kendall, Kinney, Kerr, LaSalle, Lavaca, Lee, Live Oak, Llano, McCullen, Maverick, Medina, Real, Travis, Uvalde, Victoria, Webb, Williamson, Wilson. Zapata, Zavala	389
032	Bakersfield, California California: Fresno, Inyo, Kern, Kings, Madera, Mono, Tulare	316
034	Baltimore, Maryland Delaware: Kent, Sussex Maryland: Anne Arundel, Baltimore, Baltimore City, Caroline, Carroll, Cecil, Dorchester, Hartford, Howard, Kent, Queen Annes, Somerset Talbot, Wicomico, Worcester	698
038	Baton Rouge, Louisiana Louisiana: Acadia, East Baton Rouge, East Feliciana, Evangeline, Iberia, Iberville, Jefferson Davis, Lafayette, Livingston, Pointe Coupee, St. Landry, St. Martin, St. Mary, Vermilion, West Baton Rouge, West Feliciana	219
040	Beaumont, Texas Louisiana: Allen, Beauregard, Calcasieu, Cameron Texas: Angelina, Hardin, Jasper, Kefferson, Nacogdoches Newton, Orange, Polk, Sabine, San Augustine, Shelby, Tyler	220
046	Billings, Montana Montana: Big Horn, Carbon, Carter, Custer, Fallon, Garfield, Golden Valley, Musselshell, Park, Petroleum, Powder River, Prairie, Rosebud, Still Water, Sweet Grass, Treasure, Wheatland, Wibaux, Yellowstone, Wyoming: Big Horn, Hot Springs, Johnson, Park, Sheridan, Washakie	71
052	Birmingham, Alabama Alabama: Bibb, Blount, Calhoun, Cherokee, Chilton, Clay, Cleburne, Colbert, Coosa, Cullman, DeKalb, Etowah, Fayette, Franklin, Greene, Jackson, Jefferson, Lamar, Lauderdale, Lawrence, Limestone, Madison, Marion, Marshall, Morgan, Pickens, Randolph, St. Clair, Shelby, Sumter, Talladega, Tuscaloosa, Walker, Winston	576
054	Bismarck, North Dakota Montana: Daniels, Dawson, McCone, Richland, Roosevelt, Sheridan, Valley North Dakota: Adams, Billings, Bowman, Burleigh, Dunn, Emmons, Golden Valley, Grant, Hettinger, Logan, McIntosh, McKenzie, McClean, Mercer, Morton, Oliver, Sioux, Slope, Stark, Williams South Dakota: Campbell, Corson, Dewey, Perkins, Walworth, Ziebach	78
058	Boise, Idaho Idaho: Ada, Adams, Blaine, Camas, Canyon, Cassia, Elmore, Gem, Gooding, Jerome, Lincoln, Minidoka, Owyhee, Payette, Twin Falls, Valley, Washington Oregon: Malheur	109
060	Boston, Massachusetts Connecticut: Windham Massachusetts: Barnstable, Essex, Middlesex, Norfolk, Plymouth, Suffolk, Worcester New Hampshire: Belkamp, Cheshire, Hillsborough, Merrimack, Rockingham, Strafford Vermont: Windham	1436
064	Bristol, Tennessee Kentucky: Knott, Leslie, Letcher, Perry, Tennessee: Carter, Johnson, Sullivan, Unico, Washington Virginia: Dickenson, Lee, Russell, Scott, Washington, Wise	144
066	Buffalo, New York New York: Allegany, Cattaraugus, Chautauqua, Erie, Genesee, Livingston, Monroe, Niagara, Ontario, Orleans, Wayne, Wyoming, Yates Pennsylvania: McKean, Potter, Tioga	836
068	Burlington, Vermont - Watertown, New York New York: Clinton, Essex, Franklin, Jefferson, St. Lawrence, Vermont: Caledonia, Chittenden, Essex, Franklin, Grand Isle, Lamoille, Orleans, Washington	168

Ref. No.	Area & Counties	Number of Households 1967-01-01 (000)
070	Butte, Montana Montana: Beaverhead, Broadwater, Deer Lodge, Gallatin, Granite, Jefferson, Lake, Lewis & Clark, Madison, Meagher, Mineral, Missoula, Powell, Ravalli, Sanders, Silver Bow	74
072	Cadillac - Traverse City, Michigan Michigan: Alpena, Antrim, Benzie, Charlevoix, Cheboygan, Chippewa, Crawford, Emmet, Grand Traverse, Kalkaska, Leelanau, Mackinac, Manistee, Missaukee, Osceola, Otsego, Presque Isle, Wexford	78
074	Paducah, Kentucky Kentucky: Ballard, Calloway, Crittenden, Fulton, Graves, Hickman, Livingston, Lyon, McCracken, Marshall Tennessee: Carroll, Henry, Obion, Weakley	88
078	Casper, Wyoming Wyoming: Campbell, Carbon, Converse, Crook, Fremont, Natrona, Weston	37
080	Cedar Rapids, Iowa Iowa: Allamakee, Benton, Blackhawk, Bremer, Buchanan, Butler, Cedar, Chickasaw, Clayton, Delaware, Dubuque, Fayette, Floyd, Grundy, Howard, Iowa, Jackson, Johnson, Jones, Linn, Tama, Washington, Winneshiek Wisconsin: Crawford, Grant	253
082	Champaign, Springfield, Illinois Illinois: Cass, Champaign, Christian, Coles, DeWitt, Douglas, Edgar, Livingston, Logan, McLean, Macon, Macoupin, Mason, Menard, Morgan, Moultrie, Piatt, Sangamon, Schuyler, Scott, Shelby, Vermilion	315
086	Charleston, West Virginia Kentucky: Boyd, Breathitt, Carter, Elliott, Floyd, Greenup, Johnson, Lawrence, Lee, Lewis, Magoffin, Martin, Morgan, Owsley, Pike, Wolfe Ohio: Adams, Athens, Gallia, Jackson, Lawrence, Meigs, Pike, Ross, Scioto, Vinton, Washington Virginia: Buchanan, Tazewell, West Virginia: Boone, Braxton, Cabell, Calhoun, Clay, Fayette, Greenbrier, Jackson, Kanawha, Lincoln, Logan, McDowell, Mason, Mercer, Mingo, Monroe, Nicholas, Pocahontas, Putnam, Raleigh, Roane, Summers, Wayne, Wyoming	596
088	Charlotte, North Carolina North Carolina: Alexander, Alleghany, Anson, Ashe, Avery, Burke, Cabarrus, Caldwell, Catawba, Cleveland, Gaston,Iredell, Lincoln, Mecklenburg, Rowan, Rutherford, Stanly, Union, Watauga, Wilkes South Carolina: Chester, Chesterfield, Lancaster, York	388
090	Chattanooga, Tennessee Georgia: Catoosa, Chattooga, Dade, Gordon, Murray, Walker, Whitfield North Carolina: Cherokee, Clay Tennessee: Bledsoe, Bradley, Coffee, Franklin, Grundy, Hamilton, Lincoln, McMinn, Marion, Meigs, Monroe, Morre, Polk, Rhea, Sequatchie	212
092	Cheyenne, Wyoming Nebraska: Arthur, Banner, Box Butte, Cheyenne, Dawes, Deuel, Garden, Grant, Hooker, Kimball, Lincoln, Logan, McPherson, Morrill, Scotts Bluff, Sheridan, Thomas South Dakota: Butte, Custer, Fall River, Harding, Lawrence, Meade, Pennington, Shannon Wyoming: Albany, Goshen, Laramie, Niobrara, Platte	119
094	Chicago, Illinois Illinois: Bureau, Cook, DeKalb, DuPage, Ford, Grundy, Iroquois, Kane, Kankakee, Kendall, Lake, LaSalle, McHenry, Putnam, Will Indiana: Jasper, Lake, LaPorte, Porter, Starke	2421
096	Chico-Redding, California California: Del Norte, Humboldt, Shasta, Siskiyou, Tehama, Trinity	89
098	Cincinnati, Ohio Indiana: Dearborn, Franklin, Ohio, Ripley, Switzerland Kentucky: Boone, Bracken, Campbell, Carroll, Fleming, Gallatin, Grant, Harrison, Kenton Mason, Pendleton, Robertson Ohio: Brown, Butler, Clermont, Hamilton, Highland, Warren	545
102	Cleveland, Ohio Ohio: Ashland, Ashtabula, Carroll, Columbiana, Crawford, Cuyahoga, Erie, Geauga, Holmes, Huron, Lake, Lorain, Mahoning, Medina, Portage, Richland, Seneca, Stark, Summit, Strumbull,Tuscarawas, Wayne Pennsylvania: Lawrence, Mercer	1452
106	Colorado Springs, Colorado Colorado: Baca, Bent, Chaffee, Cheyenne, Crowley, Custer, El Paso, Fremont, Gunnison, Hinsdale, Huerfano, Kiowa, Kit Carson, Las Animas, Lincoln, Otero, Prowers, Pueblo, Teller	134
108	Columbia, South Carolina South Carolina: Bamberg, Berkeley, Calhoun, Charleston, Clarendon, Colleton, Dorchester, Fairfield, Kershaw, Lee, Lexington, Newberry, Orangeburg, Richland, Saluda, Sumter	248
116	Columbus, Ohio Ohio: Champaign, Coshocton, Delaware, Fairfield, Fayette, Franklin, Guernsey, Hardin, Hocking, Knox, Licking, Logan, Madison, Marion, Monroe, Morgan, Morrow, Muskingum, Noble, Perry, Pickaway, Union, Wyandot	468
120	Corpus Christi, Texas Texas: Aransas, Brooks, Cameron, Duval, Hidalgo, Jim Hogg, Jim Wells, Kenedy, Kleberg, Nueces Refugio, San Patricio, Starr, Willacy	205
122	Dallas, Texas & Fort Worth, Texas Texas: Anderson, Archer, Baylor, Cherokee, Clay, Collin, Comanche, Cooke, Dallas, Delta, Denton, Eastland, Ellis, Erath, Fannin, Franklin, Freestone, Grayson, Henderson, Hood, Hopkins, Houston, Hunt, Jack, Kaufman, LaMar,Leon, Montague, Navarro, Palo Pinto, Parker, Rains, Red River, Rockwall, Shackelford, Smith, Somervell, Stephens, Tarrant, Throckmorton, Titus, Trinity, Van Zandt, Wichita, Wilbarger, Wise, Wood, Young	923
124	Dayton, Ohio Ohio: Clark, Darke, Greene, Miami, Montgomery, Preble, Shelby	303
128	Denver, Colorado Colorado: Adams, Arapahoe, Boulder, Clear Creek, Delta, Denver, Douglas, Eagle, Elbert, Garfield, Gilpin, Grand, Jackson, Jefferson, Lake, Larimer, Logan, Mesa, Moffat, Montrose, Morgan, Ouray, Park, Phillips, Pitkin, Rio Blanco, Routt, San Juan, San Miguel, Sedgwick, Summit, Washington, Weld, Yuma	464

Ref. No.	Area & Counties	Number of Households 1967-01-01
150	Des Moines, Iowa Iowa: Adair, Adams,Appanoose, Audubon, Boone Buena Vista, Calhoun,Cass, Clarke, Dallas, Davis, Decatur, Greene, Guthrie, Hamilton, Hardin, Humboldt, Jasper, Jefferson, Keokuk, Kossuth, Lucas, Madison, Mahaska, Marion, Marshall, Monroe, Palo Alto, Pocahontas, Polk, Poweshiek, Ringgold, Isac, Story,Taylor, Union, Van Buren, Wapello, Warren, Wayne, Webster, Wright	327
132	Detroit, Michigan Michigan: Macomb, Monroe, Oakland, St. Clair, Washtenaw, Wayne	1247
138	Duluth, Minnesota - Superior, Wisconsin Minnesota: Beltrami, Carlton, Clearwater,Cook, Itasca, Koochiching, Lake, Lake of the Woods, St. Louis, Wisconsin: Ashland, Barron, Bayfield, Burnett, Douglas, Rusk, Sawyer, Washburn	159
140	Eau Claire - La Crosse, Wisconsin Wisconsin: Buffalo, Chippewa, Dunn, Eau Claire, Jackson, La Crosse, Monroe, Pepin, Trempealeau, Vernon	91
142	El Paso, Texas - Roswell, New Mexico New Mexico: Chaves, Dona Ana, Eddy, Grant, Hidalgo, Lea, Lincoln, Luna, Otero, Texas: Brewster, Culberson, El Paso, Hudspeth, Presidio, Terrell	190
146	Erie, Pennsylvania Pennsylvania: Crawford, Erie, Forest, Venango, Warren,	132
148	Eugene - Roseburg - Klamath Falls, Oregon Oregon: Coos, Curry, Douglas, Jackson, Josephine, Klamath, Lake, Lane	161
152	Evansville, Indiana Indiana: Daviess, Dubois, Gibson, Knox, Martin, Perry, Pike, Posey, Spencer, Vanderburgh, Warrick Kentucky: Caldwell, Daviess, Hancock, Henderson, Hopkins, McLean, Mulenberg, Ohio, Union, Webster	196
154	Fargo, North Dakota Minnesota: Becker, Clay, Grant, Kittson, Mahnomen, Marshall, Norman,Otter Tail, Pennington, Polk, Red Lake, Roseau, Wilkin North Dakota: Barnes, Benson, Bottineau, Burke, Cass, Cavalier, Dickey, Divide, Eddy, Foster, Grand Forks, Griggs, Kidder, La Moure, Mc Henry, Mountrail, Nelson, Pembina, Pierce, Ramsey, Ransom, Renville, Richland, Olette, Sargent, Sheridan, Steele, Stutsman, Towner, Traill, Walsh, Ward, Wells	193
166	Fort Wayne, Indiana Indiana: Adams, Allen, De Kalb, Huntington, Lagrange, Noble, Steuben, Wabash, Wells, Whitley	136
176	Grand Rapids, Michigan Michigan: Allegan, Barry, Berrien, Branch, Calhoun, Cass,Kalamazoo, Kent, Lake, Mason, Mecosta, Muskegon, Newaygo, Oceana, Ottawa, St. Joseph, Van Buren	476
178	Great Falls, Montana Montana: Blaine, Cascade, Chouteau, Fergus, Glacier,Hill, Judith Basin, Liberty, Phillips, Pondera, Teton, Toole	54

Ref. No.	Area & Counties	Number of Households 1967-01-01
180	Green Bay, Wisconsin - Marquette, Michigan Michigan: Alger, Baraga, Delta, Dickinson, Gogebic, Houghton, Iron, Keweenaw, Luce, Marquette, Menominee, Ontonagon, Schoolcraft Wisconsin: Brown, Clark, Door, Florence, Fond du Lac, Forest, Green Lake, Iron, Kewaunee, Langlade, Lincoln, Manitowoc, Marathon, Marinette, Oconto, Oneida, Outagamie, Portage, Price, Shawano, Taylor, Vilas, Waupaca, Waushara, Winnebago, Wood	387
186	Greenville, South Carolina - Asheville, North Carolina North Carolina: Buncombe, Graham, Haywood, Henderson, Jackson, McDowell, Macon, Madison, Mitchell, Polk, Swain, Transylvania, Yancey South Carolina: Abbeville, Anderson, Cherokee, Greenville, Greenwood, Laurens, Oconee, Pickens, Spartanburg, Union	322
192	Quincy, Illinois Illinois: Adams, Brown, Hancock, Pike Missouri: Adair, Clark, Knox, Lewis,Macon, Marion, Putnam, Ralls, Schuyler, Scotland, Shelbv, Sullivan	82
194	Harrisburg, Pennsylvania Pennsylvania: Adams, Cumberland, Dauphin, Juniata, Lancaster, Lebanon, Mifflin, Perry, York	348
198	Hartford, Connecticut Connecticut: Fairfield, Hartford, Litchfield Middlesex, New Haven, New London, Tolland	864
202	Honolulu, Hawaii Hawaii: Kauai, Maui	177
208	Houston, Texas Texas: Austin, Azoria, Chambers, Fort Bend, Galveston, Harris, Liberty, Matagorda, Montgomery, San Jacinto, Walker, Waller, Wharton	604
212	Idaho Falls, Idaho Idaho: Bannock, Bear Lake, Bingham, Bonneville, Butte, Caribou, Clark, Custer, Franklin, Fremont, Jefferson, Lemhi, Madison, Oneida, Teton	59
214	Indianapolis, Indiana Indiana: Bartholomew, Benton, Blackford, Boone, Brown, Carroll, Cass, Clinton, Decatur, Delaware, Fayette, Fountain, Fulton, Grant, Greene, Hamilton, Hancock, Hendricks, Henry, Howard, Jackson, Jay, Jefferson, Jennings, Johnson, Kosciusko, Lawrence, Madison, Marion, Miami, Monroe, Montgomery, Morgan, Newton, Orance, Owen, Pulaski, Putnam, Randolph, Rush, Shelby, Tippecanoe, Tipton, Union, Warren, Wayne, White	757
216	Jackson - Meridian, Mississippi Alabama: Choctaw Mississippi: Adams, Alcorn, Amite, Attala, Bolivar, Calhoun, Carroll,Chickasaw, Choctaw, Claiborne, Clarke, Clay, Copiah, Covington, Forrest,Franklin, George, Greene, Grenada, Hinds, Holmes, Humphreys, Issaquena, Itawamba, Jasper, Jefferson, Jefferson Davis, Jones, Kemper, Lamar, Lauderdale, Lawrence, Leake, Lee, Leflore, Lincoln, Lowndes, Madison, Marion, Monroe, Montgomery, Neshoba,Newton, Noxubee, Oktibbeha, Pearl River, Perry, Pike, Pontotoc, Prentiss, Rankin, Scott, Sharkey, Simpson, Smith, Sunflower, Tallahatchie, Tishomingo, Walthall, Warren, Washington, Wayne, Webster, Wilkinson, Winston, Yalobusha, Yazoo	492

Ref. No.	Area & Counties	Number of Households 1967-01-01 (000)
220	Jacksonville, Florida Florida: Alachua, Baker, Bradford, Citrus, Clay, Columbia, Dixie, Duval, Flagler, Gilchrist, Hamilton, Lafayette, Levy, Marion, Nassau, Putnam, St. Johns, Suwanee, Union, Volusia Georgia: Appling, Bacon, Brantley, Camden, Charlton, Glynn, Pierce, Ware, Wayne	357
222	Johnstown, Pennsylvania Pennsylvania: Bedford, Blair, Cambria, Cameron, Centre, Clearfield, Elk, Franklin, Fulton, Huntingdon, Indiana, Jefferson, Somerset	264
224	Joplin, Missouri Kansas: Allen, Bourbon, Chautaucua, Cherokee, Crawford, Elk, Labette, Montgomery, Neosho, Wilson, Woodson Missouri: Jasper, McDonald, Newton	111
226	Kansas City, Missouri Kansas: Anderson, Atchison, Chase, Coffey, Doniphan, Douglas, Franklin, Geary, Jackson, Jefferson, Johnson, Leavenworth, Linn, Lyon, Marion, Marshall, Miami, Morris, Nemaha, Osage, Pottawatomie, Riley, Shawnee, Wabaunsee, Wyandotte Missouri: Andrew, Atchison, Barton, Bates, Benton, Buchanan, Caldwell, Carroll, Cass, Clay, Clinton, Daviess, DeKalb, Gentry, Grundy, Harrison, Henry, Holt, Jackson, Johnson, Lafayette, Linn, Livingston, Mercer, Nodaway, Pettis, Platte, Ray, St. Clair, Saline, Vernon, Worth	725
230	Knoxville, Tennessee Kentucky: Bell, Casey, Clay, Clinton, Harlan, Knox, Laurel, McCreary, Pulaski, Russell, Wayne, Whitley Tennessee: Anderson, Blount, Campbell, Clairborne, Cocke, Cumberland, Grainger, Greene, Hamblen, Hancock, Hawkins, Jefferson, Knox, Loudon, Morgan, Roane, Scott, Sevier, Union	285
238	Lansing - Jackson, Michigan Michigan: Clinton, Eaton, Gratiot, Hillsdale, Ingham, Ionia, Isabella, Jackson, Lenawee, Livingston, Montcalm, Shiawassee	234
242	Las Vegas, Nevada Nevada: Clark, Lincoln, Nye	83
248	Lincoln - Hastings, Nebraska Kansas: Cloud, Jewel, Republic, Smith, Washington Nebraska: Adams, Blaine, Boone, Buffalo, Butler, Cass, Chase, Clay, Colfax, Custer, Polk, Dawson, Dundy, Fillmore, Franklin, Frontier, Furnas, Gage, Garfield, Gosper, Greeley, Hall, Hamilton, Harlan, Hayes, Hitchcock, Howard, Jefferson, Johnson, Kearney, Keith, Lancaster, Loup, Merrick, Nance, Nemeha, Nuckolls, Otoe, Pawnee, Perkins, Phelps, Platte, Polk, Red Willow, Richardson, Saline, Seward, Sherman, Thayer, Valley, Webster, Sheeler, York	222
250	Little Rock, Arkansas Arkansas: Arkansas, Ashley, Bradley, Calhoun, Chicot, Clark, Cleburne, Cleveland, Columbia, Conway, Crawford, Dallas, Desha, Drew, Faulkner, Franklin, Fulton, Garland, Grant, Hot Spring, Independence, Izard, Jackson, Jefferson, Johnson, Lincoln, Logan, Lonoke, Monroe, Montgomery, Ouachita, Perry, Pike, Polk, Pope, Prairie, Pulaski, Saline, Scott, Sebastian, Sharp, Stone, Union, Van Buren, White, Woodruff, Yell,	379
252	Los Angeles, California California: Imperial, Los Angeles, Orange, Riverside, San Bernardino	3083
254	Louisville-Lexington, Kentucky Indiana: Clark, Crawford, Floyd, Harrison, Scott, Washington, Kentucky: Adair, Anderson, Bath, Bourbon, Boyle, Breckinridge, Bullitt, Clark, Cumberland, Estill, Fayette, Franklin, Garraro, Grayson, Green, Hardin, Hart, Henry, Jackson, Jefferson, Jessamine, Larue, Lincoln, Madison, Marion, Meade, Menifee, Mercer, Montgomery, Nelson, Nicholas, Oldham, Owen, Powell, Rockcastle, Rowan, Scott, Shelby, Spencer, Taylor, Trimble, Washington, Woodford	479
259	Lubbock, Texas Texas: Bailey, Cpchran, Crosby, Dawson, Dickens, Gaines, Garza, Hockley, Lamb, Lubbock, Lynn, Terry, Yoakum	100
262	Macon - Columbus, Georgia Alabama: Chambers, Lee, Russell, Tallapoosa, Georgia: Baldwin, Bibb, Bleckley, Chattahoochee, Crawford, Crisp, Dodge, Dooly, Hancock, Harris, Houston, Jasper, Johnson, Jones, Laurens, Macon, Marion, Monroe, Muscogee, Peach, Pulaski, Putnam, Schley, Stewart, Talbot, Taylor, Twiggs, Washington, Webster, Wheeler, Wilcox, Wilkinson	194
264	Madison, Wisconsin Wisconsin: Adams, Columbia, Dane, Dodge, Green, Iowa, Jefferson, Juneau, Lafayette, Marquette, Richland, Rock, Sauk	198
274	Memphis, Tennessee Arkansas: Clay, Craighead, Crittenden, Cross, Greene, Lawrence, Lee, Mississippi, Phillips, Poinsett, Randolph, St. Francis, Mississippi: Benton, Coahoma, DeSoto,LaFayette, Marshall, Panola, Quitman, Tate, Tippah, Tunica, Union Tennessee: Benton, Chester, Crockett, Decatur, Dyer, Fayette, Gibson, Hardeman, Hardin, Haywood, Henderson, Lake, Lauderdale, McNairy, Madison, Shelby, Tipton	480
278	Miami, Florida Florida: Broward, Dade, Glades, Hendry, Martin, Monroe, Okeechobee, Palm Beach, St. Lucie	671
280	Milwaukee, Wisconsin Wisconsin: Calumet, Kenosha, Milwaukee, Ozaukee, Racine, Sheboygan, Walworth, Washington, Waukesha	557
282	Minneapolis, Minnesota Iowa: Cerro Gordo, Franklin, Hancock, Mitchell, Winnebago, Worth, Minnesota: Aitkin, Anoka, Benton, Big Stone, Blue Earth, Brown, Carver, Cass, Chippewa, Chisago, Dakota, Dodge, Douglas, Faribault, Fillmore, Freeborn, Goodhue, Hennepin, Houston, Hubbard, Isanti, Kanabec, Kandiyohi, Lac Qui Parle, LeSueur, Lyon, McLeod, Martin, Meeker, Mille Lacs, Morrison, Mower, Nicollet, Olmsted, Pine, Ramsey, Redwood, Renville, Rice, Scott, Sherburne, Sibley, Stearns, Steele, Stevens, Swift, Todd, Traverse, Wabasha, Wadena, Waseca, Washington, Watonwan, Winona, Wright, Yellow Medicine, Wisconsin: Pierce, Polk, St. Croix	906
290	Mobile, Alabama Alabama: Baldwin, Clarke, Conecuh, Covington, Escambia, Mobile, Monroe, Washington, Florida: Escambia, Okaloosa, Santa Rosa	238
294	Monroe-Alexandria, Louisiana Louisiana: Caldwell, Catahoula, Concordia, East Carrol, Franklin, Jackson, Lincoln, Madison, Morehouse, Ouachita, Quachita, Richland, Union, West Carroll, Winn	82

Ref. No.	Area & Counties	Number of Households 1967-01-01
296	Montgomery, Alabama Alabama: Autauga, Barbour, Bullock, Butler, Coffee, Crenshaw, Dale, Dallas, Elmore, Geneva, Hale, Henry, Houston, Lowndes, Macon, Marengo, Montgomery, Perry, Pike, Wilcox	133
300	Nashville, Tennessee Kentucky: Allen, Barren, Butler, Christian, Edmonson, Logan, Metcalfe, Monroe, Simpson, Todd, Trigg, Warren Tennessee: Bedford, Cannon, Cheatham, Clay, Davidson, De Kalb, Dickson, Fentress, Giles, Hickman, Houston, Humphreys, Jackson, Lawrence, Lewis, Macon, Marshall, Maury, Montgomery, Overton, Perry, Pickett, Putnam, Robertson, Rutherford, Smith, Stewart, Sumner, Trousdale, Van Buren, Warren, Wayne, White, Williamson, Wilson	386
302	New Orleans, Louisiana Louisiana: Ascension, Assumption, Jefferson, Lafourche, Orleans, Plaquemines, St. Bernard, St. Charles, St. Helena, St. James, St. John the Baptist, St. Tammany, Tangipahoa, Terrebonne, Washington Mississippi: Hancock, Harrison, Jackson, Stone	483
303	New York (New York portion) New York: Bronx, Dutchess, Kings, Nassau, New York, Orange, Putnam, Queens, Richmond, Rockland, Suffolk, Sullivan, Westchester	3899
304	New York (New Jersey portion) New Jersey: Bergen, Essex, Hudson, Hunterdon, Middlesex, Monmouth, Morris, Ocean, Passaic, Somerset, Sussex, Union, Warren	1604
305	Norfolk, Virginia Virginia: Accomack, Hampton City, Gloucester, Isle of Wight, James City, Nansemond, Norfolk, Northampton, Princess Anne, Southampton, York	295
306	Wilmington, North Carolina North Carolina: Beaufort, Bertie, Camden, Carteret, Chowan, Craven, Currituck, Dare, Edgecombe, Gates, Greene, Halifax, Hertford, Hyde, Jones, Lenoir, Martin, Nash, Northampton, Onslow, Pamlico, Pasquotank, Perquimans, Pitt, Tyrrell, Washington, Wilson	212
310	Midland - Big Spring, Texas Texas: Andrews, Borden, Crane, Ector, Glasscock, Howard, Jeff Davis, Loving, Martin, Midland, Pecos, Reeves, Upton, Ward, Winkler	97
312	Oklahoma City, Oklahoma Oklahoma: Alfalfa, Beckham, Blaine, Caddo, Canadian, Carter, Cleveland, Comanche, Cotton, Custer, Dewey, Ellis, Garfield, Garvin, Grady, Greer, Harmon, Harper, Hughes, Jackson, Jefferson, Johnston, Kay, Kingfisher, Kiowa, Lincoln, Logan, Love, Mc Clain, Major, Marshall, Murray, Noble, Okfuskee, Oklahoma, Pontotoc, Pottawatomie, Roger Mills, Seminole, Stephens, Tillman, Washita Woods, Woodward	462
314	Omaha, Nebraska Iowa: Carroll, Crawford, Fremont, Harrison, Mills, Montgomery, Page, Pottawattamie, Shelby Nebraska: Burt, Dodge, Douglas, Sarpy, Saunders, Thurston, Washington	219
316	Orlando, Florida Florida: Brevard, Indian River, Lake, Orange, Oscheola, Seminole	215
326	Peoria, Illinois Illinois: Fulton, Marshall, Peoria, Tazewell, Woodford	121
328	Philadelphia, Pennsylvania Delaware: New Castle New Jersey: Atlantic, Burlington, Camden, Cape May, Cumberland, Gloucester, Mercer, Salem Pennsylvania: Berks, Bucks, Chester, Delaware, Lehigh, Montgomery, Northampton, Philadelphia, Schuylkill	1999
330	Phoenix, Arizona Arizona: Apache, Coconino, Gila, Graham, Greenlee, Maricopa, Mohave, Navajo, Pinal, Yavapai, Yuma	340
332	Pittsburgh, Pennsylvania Pennsylvania: Allegheny, Armstrong, Beaver, Butler, Clarion, Fayette, Greene, Washington, Westmoreland West Virginia: Barbour, Doddridge, Gilmer, Harrison, Lewis, Marion, Monongalia, Pleasants, Preston, Randolph, Ritchie, Taylor, Tucker, Tyler,Upshur, Webster, Wetzel, Wirt, Wood	874
334	Portland, Oregon Oregon: Benton, Clackamas, Clatsop, Columbia, Crook, Deschutes, Harney, Hood River, Jefferson, Lincoln, Linn, Marion, Multnomah, Polk, Sherman, Tillamook, Wasco, Washington, Yamhill Washington: Clark, Cowlitz, Klickitat,Skamania, Wahkiakum	489
336	Portland, Maine Maine: Androscoggin, Aroostook, Cumberland, Hancock, Kennebec, Knox, Lincoln, Oxford, Penobscot, Piscataquis, Sagadahoc, Somerset, Waldo, Washington, York New Hampshire: Coos, Grafton, Sullivan Vermont: Orange, Windsor	348
340	Providence, Rhode Island Massachusetts: Bristol Rhode Island: Bristol, Kent, Newport, Providence, Washington	399
342	Quad Cities, Iowa Illinois: Henderson, Henry, Knox, Mc Donough, Mercer, Rock Island, Stark, Warren Iowa: Clinton, Des Moines, Henry, Louisa, Muscatine, Lee, Scott	214
344	Raleigh - Durham, North Carolina North Carolina: Alamance, Bladen, Brunswick, Caswell,Chatham, Columbus, Cumberland, Davidson, Davie, Duplin, Durham, Forsyth, Franklin, Granville, Guilford, Harnett, Hoke, Johnston, Lee, Montgomery, Moore, New Hanover, Orange, Pender, Person, Randolph, Richmond, Robeson, Rockingham, Sampson, Scotland, Stokes, Surry, Vance, Wake, Warren, Wayne, Yadkin South Carolina: Darlington, Dillon, Florence, Georgetown, Horry, Marion, Marlboro, Williamsburg	741
348	Reno, Nevada California: Alpine, Lassen, Modoc Nevada: Churchill, Douglas, Esmeralda,Humboldt, Lyon, Mineral, Ormsby, Pershing, Storey, Washoe	64

Ref. No.	Area & Counties	Number of Households 1967-01-01 (000)
350	Richmond, Virginia Virginia: Albemarle, Amelia, Brunswick, Buckingham, Caroline, Charles City, Chesterfield, Cumberland, Dinwiddie, Essex, Fluvanna, Goochland, Greene, Greensville, Hanover, Henrico, King and Queen, King George, King William, Lancaster, Louisa, Lunenberg, Madison, Middlesex, Nelson, New Kent, Northumberland, Nottoway, Orange, Powhatan, Prince Edward, Prince George, Richmond, Spotsylvania, Stafford, Surry, Sussex, Westmorland	296
354	Roanoke, Virginia Virginia: Alleghany, Amherst, Appomattox, Augusta, Bath, Bedford, Bland, Botetourt, Campbell, Carroll, Charlotte, Craig, Floyd, Franklin, Giles, Grayson, Halifax, Henry, Highland, Mathews, Mecklenburg, Montgomery, Radford City, Patrick, Pittsylvania, Pulaski, Roanoke, Rockbridge, Smyth, Wythe	299
358	Rockford, Illinois Illinois: Boone, Carroll, Jo Daviess, Lee, Ogle, Stephenson, Whiteside, Winnebago	150
364	Sacramento, California California: Amador, Butte, Calaveras, Colusa, Dorado, Glenn, Mariposa, Merced, Nevada, Placer, Plumas, Sacramento, San Joaquin, Sierra, Stanislaus, Sutter, Tuolumne, Yolo, Yuba	524
366	Saginaw-Flint, Michigan Michigan: Alcona, Arenac, Bay, Clare, Genesee Gladwin, Huron, Iosco, Lapeer, Midland, Ogemaw, Oscoda, Roscommon, Saginaw, Sanilac, Tuscola	289
370	St. Louis, Missouri Illinois: Alexander, Bond, Calhoun, Clay, Clinton, Edwards, Franklin, Gallatin, Greene, Hamilton, Hardin, Jackson, Jefferson, Jersey, Johnson, Madison, Marion, Massac, Monroe, Montgomery, Perry, Pope, Pulaski, Randolph, St. Clair, Saline, Union, Wabash, Washington, Wayne, White, Williamson Missouri: Audrain, Bollinger, Boone, Butler, Callaway, Camden, Cape Girardeau, Carter, Chariton, Cole, Cooper, Dunklin, Franklin, Gasconade, Howard, Iron, Jefferson, Lincoln, Madison, Miller, Mississippi, Moniteau, Monroe, Montgomery, Morgan, New Madrid, Osage, Pemiscot, Perry, Pike, Randolph, Reynolds, Ripley, St. Charles, St. Francois, St. Louis, St. Louis City, Ste. Genevieve, Scott, Stoddard, Warren, Washington, Wayne	1099
376	Salt Lake City, Utah Nevada: Elko, Eureka, Lander, White Pine Utah: Beaver, Box Elder, Cache, Carbon, Daggett Davis, Duchesne, Emery, Garfield, Grand, Iron, Juab, Kane, Millard, Morgan, Piute, Rich, Salt Lake, San Juan, Sanpete, Sevier, Summit, Tooele, Uintah, Utah, Wasatch, Washington, Wayne, Weber Wyoming: Lincoln, Sublette, Sweetwater, Teton, Uinta	299
382	San Diego, California California: San Diego	380
384	San Francisco, California California: Alameda, Contra Costa, Lake, Marin, Mendocino, Monterey, Napa, San Benito, San Francisco, San Mateo, Santa Clara, Santa Cruz, Solano, Sonoma	1559
386	Santa Barbara, California California: San Luis Obispo, Santa Barbara, Ventura	201

Ref. No.	Area & Counties	Number of Households 1967-01-01 (000)
388	Savannah, Georgia Georgia: Bryan, Chatham, Effingham, Liberty, Long, McIntosh South Carolina: Beaufort, Hampton, Jasper	93
390	Seattle, Washington Washington: Clallam, Grays Harbor, Island, Jefferson, King, Kitsap, Lewis, Mason, Pacific, Pierce, San Juan, Skagit, Snohomish, Thurston, Whatcom	652
394	Shreveport, Louisiana Arkansas: Hempstead, Howard, Lafayette, Little River, Miller, Nevada, Sevier Louisiana: Bienville, Bossier, Caddo, Claiborne, DeSoto, Red River, Webster Oklahoma: Bryan, Choctaw, McCurtain Texas: Bowie, Camp, Cass, Gregg, Harrison, Marion, Morris, Panola, Rusk, Upshur,	269
396	Sioux City, Iowa Iowa: Cherokee, Clay, Dickinson, Emmet, Ida, Lyon, Monona, O'Brien, Osceola, Plymouth, Sioux, Woodbury Nebraska: Antelope, Boyd, Cedar, Cuming, Dakota, Dixon, Holt, Keya Paha, Knox, Madison, Pierce, Rock, Stanton, Wayne	128
398	Sioux Falls, South Dakota Minnesota: Cottonwood, Jackson, Lincoln, Murray, Nobles, Pipestone, Rock Nebraska: Brown, Cherry, South Dakota: Aurora, Beadle, Bennett, Bon Homme, Brookings, Brown, Brule, Buffalo, Charles Mix, Clark, Clay, Codington, Davison, Day, Deuel, Douglas, Edmunds, Faulk, Grant, Gregory, Haakon, Hamlin, Hand, Hanson, Hughes, Hutchinson, Hyde, Jackson, Jerauld, Jones, Kingsbury, Lake, Lincoln, Lyman, McCook, McPherson, Marshall, Mellette, Miner, Minnehaha, Moody, Potter, Roberts, Sanborn, Spink, Stanley, Sully, Todd, Tripp, Turner, Union, Washabaugh, Yankton	185
400	South Bend, Indiana Indiana: Elkhart, Marshall, St. Joseph	121
402	Spokane, Washington Idaho: Benewah, Bonner, Boundary, Clearwater, Idaho, Kootenai, Latah, Lewis, Nez Percé, Shoshone Montana: Flathead, Lincoln Oregon: Baker, Gilliam, Grant, Morrow, Umatilla, Union, Wallowa, Wheeler Washington: Adams, Asotin, Benton, Chelan, Columbia, Douglas, Ferry, Franklin, Garfield, Grant, Kittitas, Lincoln, Okanogan, Pend Oreille, Spokane, Stevens, Walla Walla, Whitman, Yakima	354
404	Springfield, Missouri Arkansas: Baxter, Boone, Carroll, Marion, Newton, Searcy Missouri: Barry, Cedar, Christian, Crawford, Dade, Dallas, Dent, Douglas, Greene, Hickory, Howell, LaClede, Lawrence, Maries, Oregon, Ozark, Phelps, Polk, Pulaski, Shannon, Stone, Taney, Texas, Webster, Wright	165
406	Springfield, Massachusetts Massachusetts: Franklin, Hampden, Hampshire	186
408	Syracuse, New York New York: Broome, Cayuga, Chemung, Chenango, Cortland, Delaware, Herkimer, Lewis, Madison, Oneida, Onondaga, Oswego, Otsego, Schuyler, Seneca, Steuben, Tioga, Tompkins	567
410	Tallahassee, Florida - Albany, Georgia Florida: Bay, Calhoun, Franklin, Gadsden, Gulf, Holmes, Jackson, Jefferson, Leon, Liberty, Madison, Washington, Taylor, Wakulla, Walton, Washington Georgia: Atkinson, Baker, Ben Hill, Berrien, Brooks, Calhoun, Clay, Coffee, Clinch, Colquitt, Cook, Decatur, Dougherty, Early, Echols, Grady, Irwin, Jeff Davis, Lanier, Lee, Lowndes, Miller, Mitchell, Quitman, Randolph, Seminole, Sumter, Telfair, Terrell, Thomas, Tift, Turner, Worth	271

DISCUSSION

Karl A. Fox, Iowa State University

As chairman of the Social Science Research Council's Committee on Areas for Social and Economic Statistics, I have been in close touch with the study which underlies Brian Berry's paper. The SSRC Committee was appointed in November 1964 and terminated in September 1967 upon completion and review of Berry's study. The Committee's final report was summarized in the December 1967 issue of Social Science Research Council ITEMS. 1/

As I agree with Berry's recommendations, which were also adopted by the SSRC Committee, I will simply add some comments of my own on the background and potential uses of the functional economic area concept.

To the best of my knowledge I originated the functional economic area concept in a March 1961 paper entitled "The Concept of Community Development." I developed the concept further in a series of invited papers during 1962-64. Their titles suggest some of the practical concerns to which I was addressing myself: "The Study of Interactions Between Agriculture and the Non-Farm Economy: Local, Regional and National" (February 1962); "Delineating the Area" (January 1962); "The Major Problem of Rural Society" (December 1962); "On the Current Lack of Policy Orientation in Regional Accounting" (December 1962); "Economic Models for Area Development Research" (May 1963); and "Integrating National and Regional Models for Economic Stabilization and Growth" (March 1964).

These early papers stated a number of propositions about functional economic areas:

1. An FEA is a home-to-work commuting field;
2. An FEA is a relatively self-contained labor market in a short-run, economic stabilization context;
3. In FEA's with populations of less than half a million or thereabouts, the "regional" shopping area tends to coincide with the commuting field.
4. Because of its relative closure (self-containment) with respect to the home-to-work commuting trip and to most resident-oriented retail and service activities, the FEA should be an ideal unit for regional social accounts;
5. An FEA can be approximated reasonably well by a cluster of contiguous whole counties;
6. In contrast with most individual counties, an FEA lends itself to area development planning by virtue of its larger population base, wider array of leadership and professional talent, and relative closure as a commuting and retail trade area (so that most of the benefits from programs initiated in the FEA will accrue to its own present and future residents and taxpayers);
7. FEA's approximated by clusters of contiguous whole counties could evidently be delineated for all but the most sparsely populated regions in the United States; hence,
8. FEA delineations could be used to disaggregate national economic and social statistical magnitudes into a set of areas each of which would have desirable properties with respect to social accounting, employment stabilization policies, economic development planning, and the provision of a wide array of public services.

In October 1964, I formulated the functional economic area concept in a more rigorous way. Many policy-oriented people had found the concept appealing on a pragmatic basis. My October 1964 formulation made it more appealing to location theorists, regional scientists and quantitative geographers. At first glance, some action-oriented people are startled by the compass orientation of the squares in Figures 1 and 2, but their confidence returns when they recognize (1) that the areas can still be represented rather well by clusters of whole counties and (2) that the orientation of the squares reflects an "ideal" rectangular road grid with properties which are approximated only crudely by any real road system.

The significance of the rotated squares in Figures 1 and 2 is as follows:

1. Iowa, like some other Midwestern states, has an almost complete grid of "section roads" one mile apart and oriented either east-west or north-south. Each mile of road forms one side of two adjacent square miles of land; a complete grid would contain two miles of road for every square mile of area. Iowa approaches perfection in this respect; the area of the state is about 56,000 square miles and the Iowa Highway Commission presides over approximately 112,000 miles of open-country roads! There are relatively few diagonal roads in Iowa.
2. If we start at the center of a rotated square, we can reach any of its four corners by traveling 50 miles over an actual road. If we want to reach some point on (say) the northwest side of a square we must also travel 50 miles--for example, 40 miles west and 10 miles north or 25 miles west and 25 miles north. Hence, given the rectangular road grid, each square is the locus of points 50 highway miles from the center of the square.
3. If we can travel at an average speed of 50 miles an hour over every segment of the road grid, the boundary of a square is also the locus of all points from which the commuting time to the center of the square is 60 minutes.
4. Few people are willing to commute more than 60 minutes each way for long periods.

Thus, a 60-minute perimeter may serve as a reasonable first approximation to a commuting field. The area of each square in Figures 1 and 2 is 5,000 square miles, equivalent to 8 or 9 counties of the sizes usually found in the Midwest.

In today's session, we are mainly concerned with functional economic areas as alternatives to SMSA's. Figure 1 indicates that an FEA system can readily incorporate the existing system of SMSA's. The shaded areas in Figure 1 are the seven 1960 SMSA's which were wholly or partly in Iowa. In each case, a 50-mile commuting radius around the central city of the SMSA completely encloses the SMSA itself. As both SMSA's and FEA's are made up of whole counties, continuity with 1960 and earlier data could be achieved for either type of area.

Seven of the squares in Figure 2 are identical with those in Figure 1. In addition, Figure 2 shows 50-mile squares around several smaller cities in Iowa (and around some cities in adjoining states). The smaller FEA central cities in Iowa include Fort Dodge, Mason City, Ottumwa, Burlington and Spencer. Collectively, the 50-mile squares in Figure 2 include about 80 percent of the area and 90 percent of the population of Iowa. We should remember also that these squares are only first approximations to the actual commuting fields.

To achieve completeness in a national system of social and economic data, counties in the gaps between commuting fields could be allocated to one FEA or another on the basis of retail trade patterns or highway travel times. Alternatively, the policy implications of the larger gaps could be highlighted by showing them as separate interstitial areas. Figure 2 shows that there is room for an FEA centering on Taylor County in southwest Iowa. However, Taylor County is nearly 100 highway miles from either Des Moines or Omaha-Council Bluffs. This is too far for long-term commuting. The obvious alternative is out-migration, and indeed, between 1950 and 1960, the number of young men in Taylor County aged 25 to 34 decreased by 41 percent! Presumably, most of them moved into the commuting fields of FEA central cities in Iowa and elsewhere.

Some potential uses of functional economic area delineations. - The relative closure of FEA's as commuting fields makes them ideal units for a national system of regional social accounts. A national system of FEA's could also be used to appraise the regional impacts of Federal economic policies and programs. For example, short-run employment targets could be established for each FEA and backed up by facilities in each FEA for vocational education, training and retraining. A national employment service regionalized on an FEA basis could also promote mobility to actual jobs (rather than merely prospective ones) in other FEA's with tighter labor markets and stronger growth trends.

Consideration might reasonably be given also to stimulating the growth of some FEA's from their current populations of 100,000 or 150,000 to figures of 250,000 or more. The contiguously built-up area of the central city might include only 100,000 or so people, but the whole multicounty area might be designed and zoned as an urban-regional entity with a central city, satellite towns, and open-country components.

Some functions often carried on at present by individual counties or towns might be better handled on a multicounty FEA basis. The following functions might be considered in this light:

School districts including public junior colleges and four-year colleges; centers for vocational education, training and retraining; university extension and adult education programs; police and fire protection; public health services; social welfare services; the maintenance and construction of local or "secondary" streets and roads, as distinct from those connecting major population centers and maintained by state highway commissions; regional zoning; and public library services.

Functional economic areas have a strong affinity with Economic Development Districts, which are clusters of counties centered on a city of 20,000 population or more considered to be an actual or potential growth center. FEA's also have considerable affinity with ZIP code or mail delivery service areas.

In Iowa, the FEA concept has been applied by the Iowa State University Extension Service to extricate itself from the traditional pattern of 99 separate county administrative units. Its field operations have, since 1966, been reorganized on the basis of 12 multicounty areas, following FEA outlines so far as possible. (A state agency must, of course, serve residents of the gaps between commuting fields; also, when a central city lies on a state boundary, the state administrative area can include only one-half of the commuting field.)

Also in Iowa, 16 area vocational-technical school districts have been organized, following FEA outlines with minor modifications. The governor of Iowa has recommended that the field operations of state agencies be organized on the basis of 16 multicounty areas, again following FEA outlines quite closely apart from problems posed by state boundaries and gaps between commuting fields.

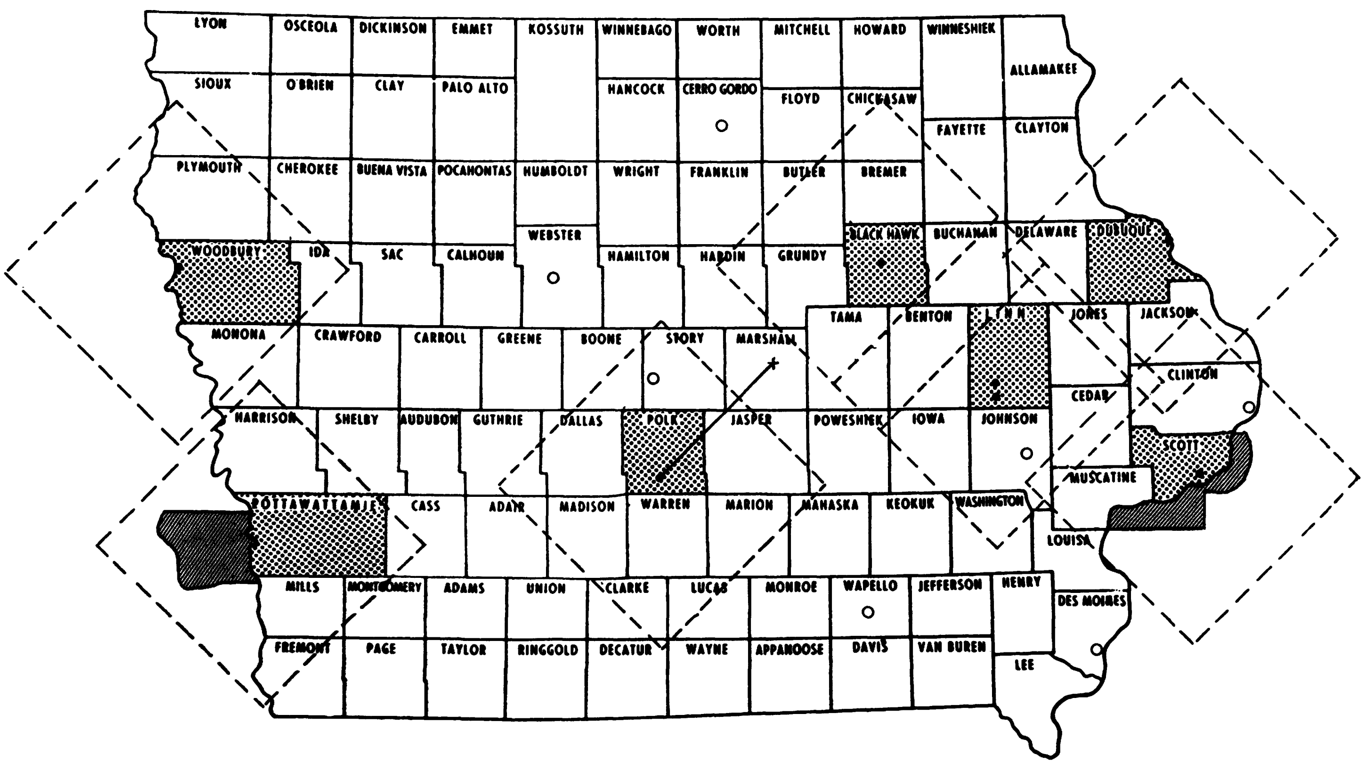
The spatial organization of the United States economy as indicated in Brian Berry's studies and my own leaves no room for the traditional rural-urban dichotomy. What we see today is a new synthesis of rural and urban society, predominantly urban in tone. As an economic and cultural entity, the city has surrounded the country and rural poverty is largely concentrated in the interstices between urban commuting fields.

A national system of social and economic data based upon functional economic areas can greatly clarify public perception of the structure of our society and facilitate the solution of many problems which are erroneously dichotomized into urban and rural segments and agency jurisdictions. A major impediment to sound economic and social policy is our institutionalized belief that a rural society exists, and can be manipulated successfully, apart from the society as a whole.

1/ Karl A. Fox, "Functional Economic Areas and Consolidated Urban Regions of the United States," Social Science Research Council ITEMS, Vol. 21, No. 4, December 1967, pp. 45-49.

10-3-64
K. Fox

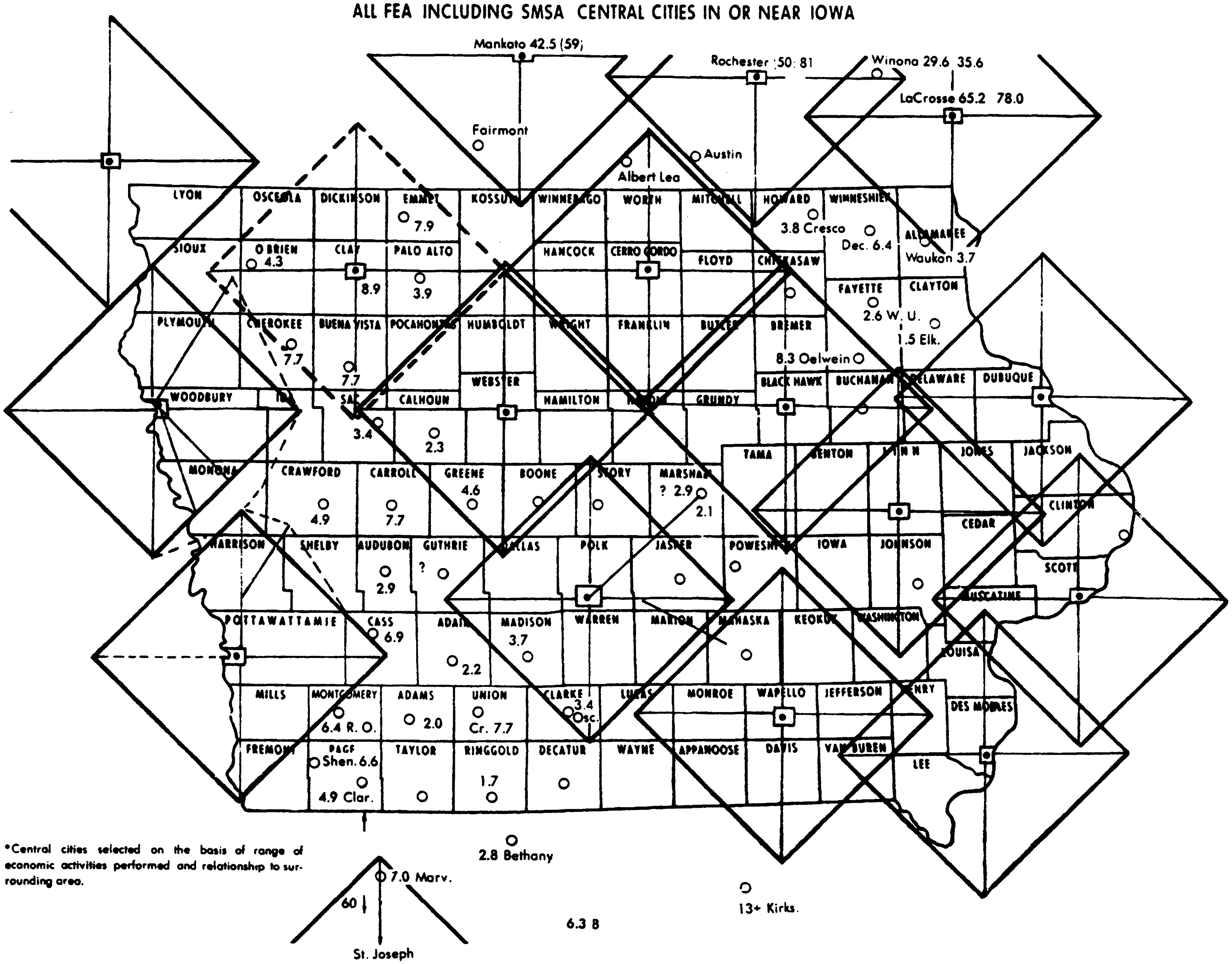
FIGURE 1. 50-MILE COMMUTING DISTANCES FROM THE CENTRAL BUSINESS DISTRICTS OF IOWA SMSA CENTRAL CITIES*



*Central cities of 50,000 people or more in 1960. Each shaded county or pair of shaded contiguous counties are SMSA's.

10-2-64
K Fox

FIGURE 2. 50-MILE COMMUTING DISTANCES FROM THE CENTRAL BUSINESS DISTRICTS OF ALL FEA INCLUDING SMSA CENTRAL CITIES IN OR NEAR IOWA



*Central cities selected on the basis of range of economic activities performed and relationship to surrounding area.

V

FAMILY STATISTICS

Chairman, HENRY S. SHRYOCK, U. S. Bureau of the Census

STATISTICS NEEDED TO MEASURE AND PROJECT CHANGES IN MARRIAGE AND FAMILY LIFE AND THEIR DEMOGRAPHIC CONSEQUENCES

Frederick F. Stephan, Princeton University and the Population Council

Introduction

During the next ten or twenty years, we may expect to see a great development of social statistics, particularly the measurement and statistical analysis of family formation and family life. There are several reasons why such a development may be anticipated:

1. We have cultivated this field very inadequately in the past and hence much needs to be done just to catch up with the progress of statistical work in other fields.

2. There is increasing urgency in the need to cope with certain problems, both in the United States and abroad, which involve conditions of, and changes in, aspects of human life that can be better understood with the aid of social statistics. Among them are population growth, poverty, unemployment, education, race relations, and housing.

3. In response to these problems, national governments are devoting greater resources to the collection of data and support of research in this field, though their progress is slow and uneven and much of the originating and experimenting is still being done with support from private foundations and donors.

4. While these problems typically center in individuals or in large groups of people, they have very important connections with family life and hence, whether the question is one of causes, effects, or cures, they call for a better understanding of what is happening to people in families and what people are deciding and doing in their family life.

This fourth reason for anticipating a great expansion of work in social statistics is exhibited in many particular interactions of family life and the larger web of life in the entire society. The marriage rate is influenced by the condition of the economy and in turn the resultant formation of new households affects the demand for housing, house furnishings, and various goods and services. The birth rate is also responsive to economic and social influences and in turn affects the demand for various goods and services, as well as the economics and politics of urban and suburban development. Migration involves family life and relationships in many complex ways. It is clear that, in a certain sense, poverty is transmitted through families from generation to generation almost as if it were an hereditary disease. So, too, in the case of race prejudice, juvenile delinquency, adult crime, and other problems some of the processes that generate, aggravate, and maintain these problems are constituent parts of family life.

These relations between family life and the whole society have been recognized for many decades but now they are brought into prominence by the crucial problems of the day. The controversial Moynahan report is but one evidence of concern about inadequacies and malfunctioning of the family institution which in many ways is the foundation of the whole society. Indeed, an understanding of the constructive functioning of the family is more than a background for analyses of its involvement in social problems; it is an essential part of these analyses and hence the statistical contributions to such research will extend broadly into all the major phases of family life.

As the development of social statistics proceeds, much of the new information will be obtained first of all in small-scale special studies, often based on existing records and on interviews conducted in conjunction with other activities. Later there will be further exploration of similar questions in local surveys. After these exploratory studies some types of information will be obtained by well-established national surveys and finally a few statistical series will be established and maintained to measure trends and variation in rates. We can find examples of this evolutionary pattern in the now familiar statistical series on births, deaths, employment, unemployment, education and income. Many analogous developments can be observed in the field of health statistics.

In the case of statistics of marriage and the family, as in other branches of social statistics, demand for new and more detailed information will surely continue to exceed what can be supplied even by a considerable addition to currently available data. It is not too early to start a sustained discussion and review of the more attractive lines of development, the major obstacles and difficulties each would encounter, and the likely benefits each might yield for better understanding and better control of the urgent problems of our time.

Present Sources of Statistics of Marriage and the Family

In 1959, Jacobson compiled a very useful compendium of marriage and divorce statistics for the United States. In his introductory comments he stated, "With respect to marriage and divorce, the information available for our country is woefully inadequate". A similar statement could be made for the greater part of the world's population outside Europe.

Substantial progress has been made in family statistics, especially with respect to the simpler facts of family composition, the characteristics of the heads of households, income, education, and several other major demographic variables. Thirty years ago almost none of this information was available. The advent of sampling for cross-tabulations as well as for data acquisition and the evolution of computer technology have contributed greatly to the progress since then and should continue to facilitate much more.

There are still weaknesses and gaps in the knowledge of family structure as well as a dearth of data on family functioning. We should not expect that all aspects of family life can or should be treated statistically. Indeed, there is some portion of merit to the otherwise extreme claims of those who attack statistics, such as the claim made for the author of a recent book-club selection "With gusto and bite he snatches the subject of marriage away from the adjustment engineers, the sex technicians, the whole army of today's statistical desplendorizers". Yet when all necessary concessions have been made for the limitations of statistical inquiry, there remains a great deal of knowledge and probably quite a little wisdom to be gained from further progress in statistical studies of family life. We find no reason to be complacent about the current output of family statistics.

After a long struggle the United States has attained a high statistical standard in birth and death registration, aided greatly by the establishment of Social Security and other programs that strengthen incentives. Many nations preceded us in this attainment and many are still far from the goal. We have had three successive national sample surveys centered on family size and contraceptive practices. Other countries have conducted notable studies of population growth and related family variables. Yet we know all too little about the processes that generate decisions and behavior relative to marriage and childbearing when adequate means of contraception are readily available. We were puzzled by the Baby Boom after World War II and we lack a firm base from which to project the fertility of the cohorts of young people now approaching parenthood - the Baby Boom infants now becoming adults. A large and increasing number of studies contribute to our understanding of household finances, housing problems, educational trends, social stratification and other features of our society that are closely linked with family life. More are needed. Ultimately they must be fitted together in a coherent analysis revealing the complex system of interacting influences that underlies observable social change. Looking ahead, we should prepare a schedule of further exploration, well-designed experiment, and constructive development necessary to overcome the inadequacies of current sources of statistics and meet as well as we can future needs. Several observations may be offered now on what remains to be done.

Marriage

Marriage rates have been fairly stable except for the influence of wars and periodic changes in economic conditions. Nonetheless, they remain vulnerable to new influences and major cultural changes. A bride's age at marriage has in the past been closely related to her subsequent fertility. With ready access to contraception this may be less important and new statistical relations of the two variables may emerge. Changes in the participation of women in the labor force and entry into occupations previously filled by men may well be associated with changes not only in childbearing but in other aspects of family life. Changes have been observed in the financial contributions made to young married couples by their parents and many arrangements are being made that facilitate higher education for married couples.

Attitudes toward marriage and divorce are changing and will continue to change in the future. The direction of change and the consequences are difficult to foresee. Hence measurements of trends and variation are needed to provide up to date information. Discovery and analysis of the factors that determine marriage rates and changes in marital status will be an increasingly important task for social statisticians, working with other social scientists.

Households

Households are important as residential groups and economic units apart from their family aspects. Much of the statistical data on families is limited to the household group. Members of the family who reside elsewhere are simply ignored. This reflects a traditional concept, expressed, for example, on Civil Service Form 61 as follows: "A family is defined by the Attorney General as persons who live under the same roof with the head of the family and form part of his fireside. When they branch out and become heads of new establishments, they cease to be part of the old family". In the past thirty years the Census Bureau has distinguished groups of persons living together and related by blood, marriage, or adoption as families, noting that more than one family may reside in the same household. The relationship of each member of the family to the head is reported but no further analysis is made of the network of relationships. Some development would seem desirable to extend our knowledge of family composition, possibly from the Current Population sample. The effect of births and children on household arrangements is also an important subject for future analytic studies using new statistical data.

Kinship

The neglect of kinship relations and omission of family members not residing with the "head" of a family group should be corrected. Several recent studies point to the importance of studying families as they are defined in terms of kinship relations, as well as their more extended kinship networks. One such is Kinship and Casework by Hope J. Leichter and William E. Mitchell (Russell Sage Foundation, 1967). Marvin B. Sussman and others have written on the neglect of kinship in family studies (Kinship and Family Organization, edited by Bernard Farber, Wiley, 1966). Kinship relations form a central part of Elizabeth Bott's study of urban family roles, norms and external relationships in Family and Social Network (Tavistock, 1967). Anthropologists have wrestled with the problems of analyzing data on kinship; social statisticians will find fascinating as well as frustrating problems in extending traditional concepts and analyses into this domain.

Siblings

Very little has been done to provide data on the similarities and differences of siblings, including data on their life careers after they "leave home". Much more attention has been given to peer group interaction without distinguishing peers who are relatives from peers who are not. This reflects, no doubt, the effect of school organization where siblings usually are separated by the class structure of the school or attend different schools. The interplay of school influences and family is coming to the forefront in the discussion of educational problems of the underprivileged. It may be of comparable importance for other children. Statistical analyses are difficult but may contribute much to the larger research activity on this aspect of family life.

Future Development of Statistical Research

These examples may be supplemented by others such as statistical analysis of the roles of husband and wife, father and mother, in the direction of the family, the care of children and management of the home, of travel and migration, and of participation in community affairs. Clearly there are limits to what can and will be done. The choice of what will be attempted will be determined by the strength of competing interests and other factors. Differing emphases may be given to child development, health, education, fertility and similar major programmatic interests. Some support will be given to more basic research problems expecting ultimately to gain greater understanding and control of problems by digging deeper into their roots. We may expect more emphasis to the examination of families as dynamic systems, responding to a complex and changing environment, maintaining some internal conditions in spite of interference from without, and going through a family cycle or metamorphosis in somewhat predictable ways.

Conceivably the direction of social change will be such that families will be regulated more from within than from without, paralleling the expansion of individual opportunity and freedom in many societies. Such a change makes more important the study of attitudes and motivations at the same time that it makes even average behavior more difficult to predict.

Problems

The greatest problem, it seems, is to determine what new variables will be valuable in measuring and explaining the most salient aspects of family life and how complex patterns and syndromes can be managed effectively in statistical analyses. The latter are presented descriptively in case studies but more than intuition is needed to establish the deeper systematic basis for what is observed. Some systematic relations will only be revealed by statistical analyses and others only confirmed by statistical tests. Relatively larger samples will be required to detect changes in rates than were required to estimate the level attained by a rate.

Respondent cooperation will be difficult to obtain on some questions and for especially lengthy interviews. Many inquiries will be misunderstood and attacked as invasions of privacy or contrary to the public interest. We will find that for many variables even the best respondents are unable to provide the information that is needed. We can expect that many exciting clues and attractive hypotheses will prove useless in the pursuit of better explanations of what we observe. In spite of these and other difficulties, however, we can be confident that the further development of social statistics centering in family life will be well worth the sustained, intensive, exploratory effort that now seems to be greatly needed.

THE ROLE OF THE FEDERAL GOVERNMENT IN THE COLLECTION OF STATISTICS ON THE FAMILY

Arthur A. Campbell, National Center for Health Statistics

The relevance of social statistics to the formulation of national policy is receiving wide recognition. This is particularly true of statistics on marital status, household composition, and fertility. The Department of Labor's report on the Negro Family, which was released in 1965, brought many of these data to widespread public attention. The hearings of the Gruening Committee have made policy-makers more aware of the importance of family size to national welfare. The development of "social indicators" in the Department of Health, Education, and Welfare is another auspicious sign. Recent research on the characteristics of persons living in poverty has also emphasized the importance of facts about the family. In many parts of the federal government, we find an increased awareness of such topics as illegitimacy, average number of children per family, the incidence of divorce, and median age at marriage. Administrators are not only becoming more conscious of these variables, but are eager to use them to develop and guide programs and evaluate their results.

In brief, the current trend is toward an increased recognition of the policy and program relevance of statistics on the family. Therefore, the plea for the development and improvement of statistics on the family comes not only from the traditional sources - the professional statisticians, sociologists, and demographers - but also from administrators and policy-makers.

The responsibility for responding to this challenge rests primarily with two federal agencies: the Census Bureau and the National Center for Health Statistics. Speaking very broadly, the Census Bureau is concerned mainly with statistics on the prevalence of various states, while NCHS is concerned mainly with statistics on incidence. For example, the Census Bureau estimates the distribution of the population by marital status at single points in time, while NCHS estimates the number of people marrying, divorcing, or dying during specific intervals of time. However, this distinction between the functions of the two agencies is by no means rigid. It is determined mainly by the data-collection systems each agency uses. Actually, there is some overlap in the statistics provided by the two agencies, and a great deal of overlap in the interests and objectives of their staffs.

Inasmuch as I am associated with NCHS, I shall deal largely with the present programs and future plans of that agency. In following this course, I do not want to suggest that the Census Bureau is doing less than the Center, but only that I know less about it.

In considering how to improve statistics on the family and fertility, we may think of two directions in which we might proceed.

The first is toward the more intensive development of family-related statistics themselves. This would include, for example, the development of improved measures of fertility, the construction of useful analytical devices, such as cohort tables for first marriage rates, and the intensification of efforts to improve the quality of all our statistics, especially those on marriage and divorce.

The second possible line of development is extensive, rather than intensive, and would involve new research programs to relate statistics on the family to a wide range of social and economic variables. The goal of this line of development is to identify more precisely the links between family-related variables and other aspects of the social and economic system.

If we wanted to interpret our responsibility narrowly, we could confine attention largely to the first line of development: the refinement and improvement of our basic statistical measurements. Certainly, we would all agree that this is highly important, particularly in areas where the data are seriously deficient.

But I believe that we must also follow the second, more extensive, line of development. We are led to this conclusion not only by personal inclination, but also by the needs of the administrators and policy-makers. Obviously, they are not using these measures in a vacuum, they are relating them to changes in other aspects of the social and economic system. So, I think that we have a broad responsibility to study the socioeconomic determinants and consequences of changes in family size and composition.

These, then, are the general principles that are being used to guide the development of statistics on the family in the National Center for Health Statistics. Now I shall describe the specific ways in which we are trying to apply these principles in the three subject-matter areas for which we have major responsibility: marriage, divorce, and fertility.

Although we publish total numbers of marriages for counties, States, and the entire United States, tabulations by age, race, and other characteristics of the bride and groom are shown only for those States included in the Marriage Registration Area. This comprises the States and independent registration areas that report marriages with a sufficient degree of completeness and in sufficient detail to warrant their inclusion in the national reporting system. At the present time it includes 39 States and the District of Columbia. This gives us coverage of about 78 percent of the total number of marriages occurring in the United States. Our most urgent task in the immediate future is to bring more States into the Registration Area.

Recent efforts to improve data on marriages include the development of a new standard certificate, which is being recommended for use by the States. In addition to items that appeared on the former version of this document, the revised version asks for the educational attainment of the bride and the groom, so that we will be able to develop new information on the socioeconomic correlates of marriage, using educational attainment as an indicator of socioeconomic status. The new version also asks for the date the previous marriage ended for persons who had been married before. This will enable us to develop data on the timing of remarriage.

The staff of the Marriage and Divorce Statistics Branch is now working on a method for estimating the total number of marriages in the United States classified by age, race, and order of marriage. This will make it possible to produce such long-needed statistics as age-specific first marriage rates. This project is still at an early stage of development, so we cannot be sure when it will be ready.

A still longer-range goal is the development of a set of tables showing central and cumulative first-marriage rates for birth cohorts over a relatively long time-span. Such a set of tables covering the period since 1917 was developed seven years ago by the Scripps Foundation for Research in Population Problems, but was not kept up to date. We hope to be able to improve on the original version of this table and to provide for regular updating. In preparing estimates of this kind, it is our intent to provide a model of what has happened over a long period of time and to make this model as consistent as possible with the few facts that we have to guide us.

Finally, in order to supplement the few statistics we are able to obtain from marriage records themselves, we are pretesting a mail survey designed to obtain information from recently married couples. Three versions of this survey are being tested: one to obtain data on the health of the couple, another to obtain information on migration patterns following marriage, and another to ask about numbers of children desired and expected and the methods of family planning being used.

The reporting of divorces in this country is far less satisfactory than that of marriages. At the present time, the Divorce Registration Area includes only 22 States, which account for 36 percent of all the divorces occurring in the country. Even within the Divorce Registration Area the reporting of certain characteristics is deficient. For example, the ages of husband and wife at the time of decree is reported for only 54 percent of the divorce records that we receive. In effect, this means that we can tabulate divorces by age of husband or wife for only 19 percent of all divorces occurring in the United States.

So, we are concentrating our major efforts on the expansion of the Divorce Registration Area and the improvement of reporting within the Area itself.

The new version of the U.S. Standard Certificate of Divorce, like the marriage certificate, requests information on the educational attainment of the husband and the wife. This should eventually provide some very useful information on the socioeconomic factors associated with divorce, in view of Census Bureau data showing that divorce is more common among couples with less education.

In spite of the fact that the registration of divorces is far from satisfactory, we still do research in a relatively wide framework with the few data that we have. For example, Dr. Plateris of the Marriage and Divorce Branch is preparing a monograph on the number of children affected by the divorce of parents. He finds that the average number of children per divorced couples has risen markedly over the years - a social fact that has wide ramifications.

The coverage of birth reporting has included the entire country since 1933, and the completeness of reporting has risen to a high level. At the present time, we believe that reporting is nearly 99 percent complete.

Our most recent technical contribution in the field of fertility statistics has been the updating of the cohort fertility tables, which were originally developed by the Scripps Foundation. At the present time, we are working on a revision of the cohort tables that will enable us to publish rates separately for the white and nonwhite populations from 1917 forward.

However, the bulk of our effort in the Natality Statistics Branch has been to relate statistics on births to a wider social and economic context. This has led to the publication of reports on recent trends in fertility, which contain discussions of the factors associated with the recent decline in fertility. Another report describes the relationship between fertility and educational attainment in Puerto Rico, and a report that will soon be issued descrites recent trends and differentials in illegitimacy.

In order to supplement the information obtained from birth records, the Natality Survey was established in 1963. This is a mail survey of 1 in 1,000 women giving birth to legitimate children. So far, only a methodological report has been published, but we will begin to issue reports of substantive findings soon.

At the present time, we are particularly interested in developing ways of measuring the incidence of unwanted childbearing in the United States. In spite of the fact that there is a great deal of concern about the birth of unwanted children, we do not have a very good idea of how severe the problem is, and we do not yet have a data-collection mechanism that will permit us to see whether it is becoming more or less severe. Recently we did a small pretest of a mail survey in a large city to see whether mothers who had recently had a baby would be willing to answer a question on whether or not they had wanted another child when they became pregnant. We were surprised to find that in this small and unrepresentative sample, approximately 20 percent said they had not wanted another child and an additional 30 percent said that they had wanted another child, but not right away. We feel that these results are sufficiently encouraging to warrant further methodological investigation, which we hope to begin soon.

The Center's major effort to strengthen its research capabilities in statistics relating to the family has been its attempt to establish the National Survey of Family Growth. This would be a program for the scientific investigation of demographic, sociological, and health-related aspects of fertility. Data would be collected at regular intervals through surveys of women in the reproductive years of life. The survey would cover such topics as past and expected childbearing, physiological limitations on fertility, the use of various methods of contraception, and the effectiveness of efforts to space children and limit family size.

Previous surveys of family growth, conducted by universities and private foundations, have demonstrated the feasibility and value of such research, and the proposed survey program would be built on the foundations they have provided. The latest project of this kind is the National Fertility Survey of 1965, conducted by Princeton University and supported by funds from the U.S. Public Health Service. The establishment of a National Survey of Family Growth would assure the continuation of the valuable series of data that the earlier studies began.

In many ways, the interests of the research-oriented demographer and the program-oriented administrator coincide in the proposed family growth survey. Without such a survey, it would be impossible to measure the overall success of the government's efforts to help couples control their fertility. At the same time, it will provide a great deal of information needed for the scientific investigation of the relationship between fertility and the socioeconomic system.

In summary, I believe that we have a unique opportunity at the present time to meet the government's immediate needs for better and more meaningful statistics on the family, and at the same time to provide data that will carry us further toward our long-range scientific goal of better understanding the relationship between the family and the social order.

VI

MEASURES OF MANPOWER

Chairman, DONALD CULLEN, Cornell University

ON ASPECTS OF UNDERUTILIZATION OF HUMAN RESOURCES

Harold Goldstein, U.S. Department of Labor

We worry about underutilization of human resources for different reasons at different times. During a war, when the military build-up takes large numbers of young workers out of the civilian labor force, we beat the bushes for additional workers and try to get the most out of the available human resources. In more peaceful circumstances, when unemployment is low and we are concerned about economic growth without inflation, we again want to get the fullest utilization of manpower. I suspect that it is precisely at such periods--when demand for labor is pressing against limited resources--that labor is most fully utilized, while good utilization is neglected when labor is plentiful. Thus we have the paradox that those of us who make speeches are more concerned about underutilization of human resources when it is at a minimum than when it is really rampant. However, both in good times and in bad, we are also concerned about underutilization in relation to poverty. Finally, it is viewed not only as an economic phenomenon but also in social and psychological terms--the denial of opportunity to serve at one's maximum capacity is seen as contributing to anomie, to the alienation from and anger at society that breed riots. A phenomenon that has all these effects is worth serious study.

One can identify several major forms of underutilization of human resources:

1. Unemployment - which we usually define as a situation in which people look for work but can't find it.

2. Involuntary part-time employment - a situation in which a worker wants full-time work, but can only get part-time work.

3. Involuntary nonparticipation in the labor force - a failure to seek work by people who really want it but don't look for it because their health or personal circumstances don't permit it, or because they think the search is hopeless.

4. Underemployment - situations in which a worker is employed below his highest skill, capacity, or potential.

In an attempt to identify and measure the total impact of these four aspects of underutilization of human resources, the Department of Labor, in reporting on a November 1966 survey of the employment problems in the 10 seriously disadvantaged poverty areas of 8 large cities, coined the term "sub-employment" to describe the total number of persons falling in any of these groups. Low income was used as a surrogate for underemployment. Also, an attempt was made to estimate how many underutilized persons were missed in the population count. The estimated number of persons "sub-employed" by these standards equalled from one-fourth to nearly one-half of the civilian labor force in these poverty areas 8/. *

While the aggregate impact is formidable in poverty areas, far more human resources are under utilized outside such areas, simply because the population outside is so much larger. In order to get better insights into both the geographical locations of people hit by employment problems and into the nature and causes of their problem, as a basis for developing programs to deal with them, the Department of Labor is undertaking a number of steps:

1. The Bureau of the Census will soon be making separate tabulations for the Bureau of Labor Statistics each month of the data collected in the Current Population Survey for the aggregate of the poverty neighborhoods of the large Standard Metropolitan Statistical Areas and for the balance of these areas. The poverty neighborhoods are defined as contiguous groups of census tracts in which the population meets certain criteria associated with poverty. A total of some 4,660 census tracts are included, located in 100 SMSA's. These data were first compiled for March 1966 6/. Analysis of all the data regularly collected in the Current Population Survey, including its supplements, by poverty and nonpoverty areas in large cities will show the total effect of changing economic conditions and government programs on the status of the residents of both types of areas.

2. The Department of Labor is now designing a program of in-depth studies of employment problems in the poverty areas of each of ten large cities to probe more deeply into the reasons for the existence of these problems, the attitudes associated with them, etc. In preparation for these surveys, an experimental survey program was conducted in the spring of 1967 to study the problems of communication with slum residents in surveys, ways of getting insight into the characteristics of the large number of persons not enumerated in household censuses and surveys, and to develop and test questionnaires on reasons for nonparticipation in the labor force, on how people look for work, and on attitudes.

*References are listed at the end of the paper.

Unemployment

Of the four aspects of underutilization listed above, two are relatively well documented in the statistics and in the literature--unemployment and part-time employment. On unemployment and the factors associated with it, we have, in addition to many studies going back more than a century, the wealth of insights to be derived from the monthly Current Population Survey (now to be supplemented by data on the poverty areas as noted above), and information on the insured unemployed from the unemployment insurance records. These data usually deal with a snapshot of all individuals unemployed at one time and their characteristics. However, the work and unemployment experience of individuals over time, which gives better measures of the burden of unemployment, can be seen in the annual work experience survey made as a supplement to the Current Population Survey. The most recent in this series that has been published 2/ shows, for example, that although the average number of persons unemployed at any one time in 1965 was 3 million, there were 11.4 million persons who were unemployed at some time during the year. Similarly, although the number unemployed for 15 weeks or more as of any one time was only 500,000, the number who accumulated 15 or more weeks of unemployment during the year was 2.7 million. Thus the burden of unemployment falls on many more people in the course of a year than the monthly data will show.

Involuntary part-time employment

Involuntary part-time employment is also fairly well measured and documented. Defining part-time work as less than 35 hours a week, we find that there were 2 million workers involuntarily on part-time work for economic, as distinct from personal, reasons, in the average week of 1966. Two groups of workers are affected--those who regularly work full-time but are working part-time temporarily, and those who regularly work part-time.

The former, averaging one million in 1966 at any one time, present a transitory problem. They suffer temporary short work weeks as a result of bad weather, slack work, etc. A certain amount of this is inevitable; some workers are protected against the worst impacts of such temporary part-time work by measures such as union contract rules calling for a minimum payment if the worker is called in to work and by partial unemployment coverage under unemployment insurance systems. The impact on any individual is relatively small, since this group regularly works full-time--but the loss to the economy of one million workers being on short time in any one week is substantial.

The more serious problem is that of the group that usually work part-time although they want full-time work--a group that is seriously and persistently underutilized. In 1966 this group averaged about one million, with average hours of 17.7 per week or a bit more than two days' work. Some of them were engaged in activities that by their nature afford only part-time work, such as domestic service. Others were in depressed industries, such as coal mining.

The combined effect of total unemployment and involuntary part-time employment traceable to economic reasons is measured by a figure regularly reported by the Bureau of Labor Statistics--"Labor Force Time Lost." This is a measure of the total work time offered to the economy by workers but unutilized. In making the estimate it is assumed that the unemployed and the part-time workers who want full-time work would be prepared to work a standard work week 1/. This figure exceeds the total unemployment rate somewhat more when the latter is high than when it is low--that is, the extent of involuntary part-time employment rises as the general demand for labor drops. For example, in 1966 when the overall unemployment rate averaged 3.8 percent, the estimate of labor force time lost amounted to 4.2 percent. On the other hand, in 1958 when the total unemployment rate was 6.8 percent, the labor force time lost amounted to 8.0 percent. Thus the combined measure runs from 10 to 20 percent above the unemployment rate alone.

Involuntary nonparticipation in the labor force

About 4 out of 10 people of working age are neither working nor looking for work at any one time. Among them are housewives, students, retired persons, and those unable to work for reasons of health. Most of these people are outside the labor force because they don't want to work at present, although they form a large potential labor reserve.

There are among them, however, some whose nonparticipation is involuntary--who want to work but are not looking for work because they believe the search would be useless or because some factor outside their control prevents them from working. In many cases the barrier to their working is one that society could eliminate if it set about doing so, such as lack of skill, discrimination in employment, a remediable physical handicap, or a family responsibility that could be dealt with by such facilities as day care centers for children.

Of these involuntary nonparticipants, the men of working age awaken the most concern. In a society in which the accepted role of men as workers is so unambiguous, the man who has no job and isn't busy looking for one becomes the object of indignation or compassion, depending on one's social philosophy. There has, in fact, been a clearly rising trend in labor force nonparticipation among men between 55 and 64 years of age, since 1961, and for nonwhite men since the early 1950's. The increase occurs earlier and has been more marked for nonwhite than for white men--for the former the rate has gone from about 13 percent in 1953 to 20 percent in the mid-sixties; for the latter, from about 12 percent in 1961 to 15 percent 4/.

In part the decline in labor force participation among men over 55 reflects earlier retirement, particularly in the more recent years when disability retirement (1957), and early retirement on reduced benefits (1961) under the Social Security Act became possible. At about the same time private pension plans were widely adopted. But it is hard to distinguish between voluntary retirement and the involuntary retirement of men who lost their jobs and had trouble getting new ones because of discrimination in employment, or low skill or educational level. Similarly, when ill health or a partial physical disability is present, in addition to difficulty in finding a job, some men may give up more easily and seek the "out" afforded them by retirement from the labor market.

For men below age 55 an increasing rate of nonparticipation in the labor force is more difficult to explain. The rates for white men have not changed significantly over the post-war period, but those for nonwhite men aged 45-54 rose from about 5 percent in the late forties to about 8-1/2 percent in the mid-sixties. The proportion at age 35-44 doubled, from nearly 3 to nearly 6 percent. These trends took place in periods of improving employment opportunity as well as during economic downturns. One factor in the long-term trend may have been the shift of many Negroes from farm (where underemployment is endemic but unemployment less common) to city, where their skills were not adaptable to the needs of industry.

For women, the factors are equally complex. The labor force participation rates for women in every age have been rising fairly steadily for many years, and for most of the women who are not in the labor force the choice of not working is voluntary. For those who want or need work but do not look for it, the reasons may be a mixture of inability to find work, poor health, and home responsibilities.

Some students have argued that persons who drop out of the labor force because they cannot find jobs should be considered unemployed even though they do not meet the usual test of unemployment in that they are not currently looking for work. In its report in 1962 the President's Committee to Appraise Employment and Unemployment Statistics, after considering the question carefully, reaffirmed the principal of labor force classification on the basis of current activity--i.e., working or currently looking for work--but recommended that those counted as not in the labor force by this procedure should be the subject of continued study and that statistics about their numbers and characteristics should be published regularly.

To make possible a regular count of some of the major groups of nonparticipants, a few questions were added to the monthly interview in the Current Population Survey, to be asked of all persons who were not working or looking for work. These were, in addition to the traditional sorting question on major activity in the survey week, questions on present desire for work, reasons for not looking for work, and intention to look for work in the next 12 months. The first analysis of results of these questions, asked in September 1966, has been published including some data for later months in 1966 and early 1967 5/. The data will be published regularly beginning in the near future.

It seems clear, however, that, useful as it will be to get a regular series to measure the major dimensions of nonparticipation as they change under different economic circumstances, the subject requires more subtle and complex analysis. The problems involved in studying involuntary nonparticipation were well summarized by Robert L. Stein:

> "The challenge to the researcher in this area is to develop objective methods for measuring what are mainly subjective phenomenon. While most of our labor force concepts are based on objective, overt actions (e.g., working, having a job, seeking work in a specific way and within a specified time span), the data on reasons for nonparticipation are subjective based on desire for work, attitudes, perceptions, and opinions. These more elusive data require careful probing and cross-checking, to explore the depth of a reported attitude or the reality of a reported reason " 5/.

In recognition of these complexities, a number of special studies are being made to explore the problems of these workers more intensively. A special intensive survey of nonparticipation among adult men was made in February 1967 as a supplement to the CPS. Data from the National Health Interview Survey have been analyzed for the relation of health to nonparticipation 7/. And, as noted above, the experimental survey program on urban poverty area employment problems in the spring of 1967 included a pilot questionnaire on nonparticipation.

Some of the magnitudes of involuntary nonparticipation are beginning to emerge from analysis of these data. For example, in the survey made in September 1966, 5.3 million persons who were not currently in the labor force said they wanted a job now. Of these 1.1 million were not looking for work because of ill health or disability; 1.2 million because they were in school; 1.5 million (women) because of family responsibilities or inability to arrange for child care; .4 million for personal reasons--such as a death in the family; and 750,000 because they believed it was impossible to find work 5/. (When more than one reason was given for not looking for work, as happened in one case out of four, people who reported they had ill health or disabilities, school or home responsibilities were classified in those groups, rather than in the group who believed no work was available, since the former reason was so overwhelming.)

Each of these groups presents its own problems that call for various remedial programs. Of particular interest is the 750,000 who don't look for work because they believe they cannot find it--the group sometimes referred to as "discouraged workers." One-third of them were men, and over half of these (140,000) were under 65 years of age.

These data suggest the types and approximate sizes of programs needed to help these people find employment. The greatest potential payoff is in programs to remedy health problems or rehabilitate the handicapped, and programs to provide day care centers for children--but before launching or expanding such programs we would need additional information on the skill resources of, and employment opportunities for, those who say they are kept from working by health or family problems. They may need training or special placement services, for example. On the other hand, smaller numbers, and especially few men, are nonparticipants because they think they couldn't find a job, but, having no health or family problems, these are most readily employable with the help of training and placement programs, and an investment in helping them might have the most immediate results.

Nonparticipation rates are significantly higher for both Negroes and whites in the poverty neighborhoods of large cities than in other areas. In the middle-age group, 55-64, white men in poverty areas have far higher rates than those in other areas, while for Negroes the differences between poverty area residents and others is less 6/.

One of the significant characteristics of nonparticipants is that a great many of them have a real attachment to the labor force. Some of them participate in the labor force in peak seasons (e.g., students in the summer, housewives in the Christmas shopping season) or enter intermittently. During the course of a single year some have a more permanent change of status--e.g., a student who finishes school and goes to work, a housewife whose children grow up, enabling her to take a job, an older person retiring. Thus there is considerable movement between nonparticipant and participant status.

This is true not only of students and housewives, but also of adult men. While we find significant numbers of adult men not in the labor force currently, many had worked in the recent past or intend to look in the near future.

For example, of 4-1/2 million men who were not in the labor force in the first 6 months of 1966, one-third had been employed in the preceding 6 months and nearly half had been employed in the preceding 18 months 3/. One million of the 4-1/2 million were reported as being unable to work at the time of the survey; of the 3.4 million who were able to work, 57 percent said they intended to seek employment within the next 12 months.

Even more striking in its indication that nonparticipation is a transitory phenomenon for many adult men is the fact that, in 1964, of the 5.2 million men 25 to 64 years of age who were out of the labor force some time during the year, only 350,000, or one-sixth of the total were out of the labor force the whole year 3/.

In summary, the increasing amount of research on involuntary nonparticipation in the labor force has given us some broad dimensions of the problem--showing that it is not as vast as had been imagined, though still of substantial size--but has left many questions unanswered. The simple questions to which the monthly CPS questionnaire is perforce limited do not give enough insight into the complex of past experience, present needs, and motivations that determine job-seeking behavior, nor do they provide enough information to shape the variety of remedial programs needed.

Underemployment

It is safe to say that of all the aspects of employment problems the phenomenon of underemployment is one that, despite widespread and continuing concern, not only in the United States but in other countries, has yielded least to analysis and measurement. The problem arises in many different contexts. In developing countries, of which India is a notable example, there has been continued concern about the existence of a large group of university graduates for whom no really professional work can be found and who are forced to work in clerical capacities far below their highest putative skills. In the United States, particularly during the great depression of the 1930's there were many reports of doctors driving taxis and musicians selling dry goods. Even in relatively prosperous times, members of groups subject to discrimination are found to be employed in jobs well below the level for which they are qualified.

We have some insights into the extent of this phenomenon in the depression of the 1930's through data collected in the population census of 1940, in which people were asked not only for the occupation in which they were currently employed but also for their "usual occupation." In the aggregate about 950,000 men and 150,000 women (or 2.8 percent and 1.4 percent, respectively of all those employed at the time) were employed in occupations at a lower level of skill than their usual occupation 9/.

There is considerable evidence in more recent statistics for the underutilization of highly educated people. For example, the following tabulation shows the number of people in the labor force in 1960 who were 25 years of age or older and had 4 or more years of college education, and the number of these employed in occupations that do not require a college education: 10/

training, on wage rates, or on some other value system? Or is it a matter of the values of the worker himself? We don't worry about a dentist who drops his profession and goes into the real estate business, if we know it was a voluntary choice. Are we then reduced to saying that a worker is underemployed if he considers himself underemployed?

	Total	Employed in occupations not requiring college education	
		Number	Percent
Total	6,077,588	802,203	13.2
White males	4,219,540	510,988	12.1
White females	1,571,645	237,183	15.1
Nonwhite males	145,944	33,239	22.8
Nonwhite females	140,426	20,793	14.8

The higher underutilization rate for Negro men represents, in part, discrimination in employment. Negro women, with professional employment opportunities in the schools, apparently do better than Negro men.

These data, applying as they do only to persons with college education, probably understate the total incidence of underutilization, which is likely to be prevalent among workers with less obvious skill credentials.

In trying to develop clear-cut concepts of underemployment we may distinguish several aspects:

1. Workers employed in occupations below their highest skill.
2. Workers functioning below their capacity, either because they work in inefficient enterprises or because they are prevented, sometimes for reasons of personal inadequacy, from functioning at their highest capacity.
3. Workers employed at their highest present skill but below their potential if they had more education or training.

Each of the kinds of underemployment listed above presents problems of concept that involve questions of public policy and value systems. They also present serious problems of measurement.

We generally think we know what we mean by a worker being employed below his highest skill. But is the skill hierarchy implicit in this judgment one based on length of education or

Public policy in this area is undergoing evolutionary change. The principle that the worker has some right to maintain his skill status was recognized more than 3 decades ago when the unemployment insurance system was set up. In the administration of this system, a qualified worker gets compensation if no suitable work can be found for him; suitable work, differently interpreted in actual practice in different States, generally means work in the person's own occupation rather than one of lower skill level or earnings.

Giving the worker unemployment compensation is one thing; creating a job for him is quite another. Even this we do when there is a clear social benefit. For example, tariff protection is invoked to save the jobs of skilled watch factory workers who can also make fuses for artillery shells. We don't--yet--create or protect job opportunities when all that is at stake is the frustration of the individual who has, at great pains, acquired a skill that he cannot market. As long as he has a job--any job--we consider him taken care of.

Our concepts of what constitutes the social good may and do change, however; for example, the creation of the National Foundation on the Arts and Humanities in 1965 shows a willingness to subsidize musicians and artists for the public good. And it is not inconceivable that the frustrations of underemployment--expressed through riots and civil commotion--may create a willingness to deal with this problem--not for the sake of the individual, but for social peace!

In fact, we already have adopted some general measures that deal with major causes of the underemployment resulting from employment below a worker's highest skill level. Such employment can occur when more workers acquire a particular high-level skill than the economy demands, when jobs are reduced by technological change, or when workers who have a skill are kept out of jobs by discrimination on the basis of race, sex, religion or age. One safeguard against excessive numbers of workers entering an occupation is a good program of vocational guidance in the schools which provides young people with information on prospective employment opportunities, such as that contained in the Occupational Outlook Handbook and other publications of the Department of Labor. Vocational guidance would be more effective if it were accompanied by planning of vocational school programs in line with manpower needs. The incidence of underemployment resulting from technological change can be reduced by retraining workers whose skills do not match the requirements of the economy, as is done under the Manpower Development and Training Act. This helps the workers who chose an occupation unwisely as well as the victims of structural change. Finally, discrimination in employment is being approached through fair employment practices and "equal employment opportunities" legislation. While we cannot be complacent about the results so far achieved, we do know a few things that can be done about this form of underemployment, and a beginning has been made.

In addition to the conceptual and public policy problems involved in underemployment of skilled workers there is the problem of measurement. If we use household surveys to get information about the characteristics of individuals, we would have to rely on the person's own statement as to the highest level of skill in which he is competent. Even academic credentials don't prove a level of skill, as this audience knows; how much less can we rely on the individual's own appraisal? A more objective test would be previous work experience in the occupation in question; but even this is limited proof: the person may have been incompetent in the judgment of objective and unprejudiced supervisors, or the standards of work in the occupation may have risen and he may not have kept up. Thus, while in an individual case it would be possible to evaluate a person's highest skill on careful investigation, the methods available for general statistics would not necessarily give us clear results.

The second aspect of underemployment is the situation in which the worker is operating below his capacity, either because he works in an inefficient enterprise or marginal farm, or for personal reasons. Within any one industry there may be a wide range in efficiency among enterprises. In a sense, all workers in an industry employed in plants other than the technologically most advanced, and whose output is lower because of this, may be considered underemployed in terms of known and available technology. By a less rigid standard, plants so inefficient that they cannot pay minimum or prevailing wages and remain in business may be said to be underemploying their workers.

This is not entirely a theoretical concept. The relationship of low wages to plant efficiency and the need to do something about it has been recognized. For example, the International Ladies' Garment Workers' Union has for many years provided an engineering service to companies with which it has collective bargaining agreements, to increase their productivity and thus make possible higher wages for the workers and greater economic viability for the companies who deal with the union. The technical help in farming provided by county agents is another attack on this aspect of underemployment.

Working below one's capacity is not only a function of inefficient enterprises; there are also inefficient workers, who are prevented from working up to capacity by poor organization of their own work habits or by neurotic blocks or other psychological problems. Who among us is working at full capacity? The prevalence of this problem is recognized by industrial psychologists, but only the most severe cases receive any treatment or attention.

The third type of underemployment--workers employed at their highest present skill but below their potential if they had education or training--is also a real problem. Efforts to deal with it have included improvement in general education and the extension of vocational education, the development of training programs both on the job and in institutions, scholarships, fellowships and other aids to education, and programs to motivate workers and youth to get additional education and training or to "stay in school."

The measurement problems in connection with this kind of underemployment are extremely difficult. It is not easy to identify the potential of an uneducated or an under-educated person. Standard intelligence tests offer some clue but they have been rightly criticized because they are not free of the influence of language and cultural factors.

In summary, the various forms of underemployment, while not easy to conceptualize or measure, affect a great many people, and have been the source of a great deal of concern and the focus of specific measures and legislation to deal with them.

It can be said in general of the four aspects of underutilization of human resources that they appear to affect a substantial proportion of the labor potential of the United States; that we now measure with reasonable accuracy only two of them--unemployment and part-time employment--and will soon have more insight into nonparticipation, but that we are a long way from understand ing or measuring underemployment; that even without such knowledge we have adopted measures to deal with it; but that more sharply focused programs would require better information than we now have.

1/ Gertrude Bancroft, "Some Alternative Indexes of Employment and Unemployment," Monthly Labor Review, February 1962, page 167.

2/ U.S. Department of Labor, Bureau of Labor Statistics, "Work Experience of the Population in 1966," Advance Summary, October 1967.

3/ Susan S. Holland, "Adult Men Not in the Labor Force," Monthly Labor Review, March 1967, Volume 90, No. 3, pages 5-15.

4/ "Manpower Report of the President," April 1967, Appendix Table A-4.

5/ Robert L. Stein, "Reasons for Nonparticipation in the Labor Force," Monthly Labor Review, July 1967, pages 22-27.

6/ James R. Wetzel and Susan S. Holland, "Poverty Areas of Our Major Cities," Monthly Labor Review, October 1966, (Reprinted as Special Labor Force Report No. 75).

7/ Carl Rosenfeld and Elizabeth Waldman, "Work Limitations and Chronic Health Problems," Monthly Labor Review, January 1967, page 41.

8/ U.S. Department of Labor, "A Sharper Look at Unemployment in U.S. Cities and Slums," 1967.

9/ U.S. Department of Commerce, Bureau of the Census, "16th Census of the United States: 1940. Population. The Labor Force (Sample Statistics): Usual Occupation," U.S. Government Printing Office, 1943, Table 9. In making this computation, occupation groups were assumed to conform to the following rough hierarchy of skill which was determined on the basis of average earnings of year-round, full-time workers in 1939, plus some consideration of educational levels required: (1) Professional and semiprofessional, proprietors, managers and officials, farmers and farm managers; (2) clerical and sales, craftsmen and foremen; (3) Operatives; (4) Service Workers (except protective and domestic); (5) Laborers (farm and nonfarm), and domestic service workers.

10/ U.S. Department of Commerce, Bureau of the Census, Census of Population, 1960. Educational Attainment (Final Report, PC(2)-5B). Table 8 is the source of the data used. Occupation of unemployed persons was the last occupation in which they were employed. Occupations not requiring college education were taken to be the following: Clerical, craftsmen and foremen, operatives, service workers, and laborers (farm and nonfarm). Note that the occupation groups in this list include a few individual occupations that require college training, such as FBI agents, in the service occupations; on the other hand, all sales occupations and farm owners and managers are excluded even though many of them do not require college training.

MANPOWER IN CONSTRUCTION: NEW METHODS AND MEASURES[1]

Daniel Quinn Mills, Harvard University

Adjustments to Labor Shortages

Problems of labor stringency in construction are the source of continual public comment. President Johnson has referred to the paradox of simultaneous unemployment and labor shortages in construction.[2] In 1966, for example, there were, on annual average, 297,000 construction unemployed, or 7.1 percent of the wage and salary work force in the industry. Construction trade journals continually refer to serious shortages of building tradesmen.[3] The adequacy of trained manpower has become an issue of importance in the prospects for a significant campaign to rebuild the cities. The level of public interest in the manpower situation in construction requires that some analysis of manpower patterns be attempted.

The labor shortages which occur in construction are generally confined to certain crafts and to specific areas. The simultaneous existence of areas of labor shortage and of significant unemployment results primarily from less than perfect geographic mobility of craftsmen, and from the differing occupational requirements of jobs. The construction industry is characterized by remarkably swift variations in the geographic locus and composition of building activity. Thus, in construction a high degree of worker mobility and labor force elasticity to variations in demand, relative to other industries and occupations, might be inadequate to prevent transitional shortages and surpluses.

In the short run, the construction labor force appears considerably more flexible than is commonly supposed. Indeed, flexibility of the labor force as compared with other industries is an outstanding characteristic of construction.[4] A considerable degree of flexibility is due to the seasonal expansion of employment which occurs in construction. For example, for construction craftsmen other than carpenters, expansion in the construction labor force has averaged some 76.5 percent of the net increase in employment in spring months, 1961-1966, (reduction in the numbers of the construction unemployed has accounted for the remaining 23.5 percent of the increased number of the employed). For carpenters the statistic is identical (76.5); for laborers it is 161.9 percent.[5] However, the large non-seasonal variations in construction demand which occur in local areas and rapid shifts in the composition of output also contribute to variations in the industry's labor force. Thus the ratio of persons employed at some time during the year to annual full-time jobs is quite high for construction. In 1963, some 5.4 million workers were employed in contract construction to fill 3.0 million full-time jobs -- a ratio of 1.8 workers to jobs. In manufacturing the ratio was about 1.30 for that year.[6] Certain building trades skills seem widely distributed throughout the economy,[7] and the construction industry seems able to increase its work force in brief periods at a fairly rapid rate and to a very large degree.

The number of man-hours available to construction firms may be increased within a brief period either by expanding the number of persons available for employment or by more intensive utilization of the existent labor force. This paper is concerned with increasing the supply of labor. However, the number of man-hours obtained by better utilization of the work force (for example, continuing operations further into the winter than is customary)[8] might be large. In fact, it appears that one might distinguish two construction labor forces.[9] The one consists of journeymen fairly regularly employed by a contractor or homebuilder -- his key men. The contractor is often quite reluctant to allow these employees to go elsewhere, and may even be willing to undertake jobs simply to keep them on his staff during slack in the building market.[10] Alternatively, there are craftsmen who follow the traditional pattern of the journeyman, following the work and going from employer to employer (and, perhaps, industry to industry). It is possible, if a man is a desirable employee and fortunate, to accumulate considerable working time in this manner -- it is also possible to be out of work a good part of the year. Department of Labor surveys indicate, for example, that a construction worker (defined by industry of longest job during the year) who has experienced any unemployment is very likely to have had two or more spells of unemployment during the year. (Only in agriculture is the likelihood of repeated spells of unemployment sometimes higher than in construction.)[11] Consequently, work scheduling, with a view toward reduced seasonality of employment, for example, could significantly increase the degree of utilization of the existent labor force either in an area or nationally.

Expanding the Supply of Labor

The number of employees of a particular craft working on a particular type of construction in a given area might be expanded by attracting workers from other types of construction, from non-construction industries, from other occupations, or other areas.[12] For example, the work force of carpenters on residential jobs might be increased by drawing carpenters from employment on commercial construction in the same locality. Or carpenters might be drawn from the maintenance crews of local manufacturers. Again, carpenters from residential construction in other geographic areas might arrive for employment; or persons employed in other occupations might be induced to take carpenters' jobs. An occupational change might involve mobility from other construction crafts, or from non-construction occupations. Workers might have had prior experience with some or all of the skills required of a residential construction carpenter -- or they may have had none. Thus, the mobility of workers may include any or all of the four basic dimensions listed above. In the usual case, the shift of workers into construction jobs probably includes movement along more than one of these dimensions. In analyzing the mobility of construction workers, the most interesting cases involve the interactions among geographic, occupational, and industrial mobility.

Currently we possess only the most rudimentary information concerning the size and character of manpower flows to and from construction. At best, we are able in some cases to measure flows along a single dimension. A very simple calculation indicates that there are some 15 combinations possible of the four basic types of flows mentioned.

The importance of certain types of manpower flows to construction is apparent even from the primitive sources of information now available. For example, the inter-industry mobility of construction workers is relatively high. During 1962, men who were employed in contract construction at some time during the year averaged employment in 1.204 industry divisions.[13] Those employed in manufacturing averaged employment during the year in only 1.090 divisions; those in mining, 1.008 divisions; those in wholesale and retail trade (second to contract construction in this measure), 1.114 divisions.[14]

The industrial distribution of earnings among contract construction workers also indicates considerable mobility. In 1957, only 72.3 percent of the approximately five million male wage and salary workers employed in contract construction earned most of that year's income from contract construction employment. With the exception of the service industries, this was the lowest reported percentage.[15] In 1963, for persons whose major source of earnings was general building construction, mean annual earnings from all employment were 13 percent higher than mean annual earnings in general building; for heavy construction the figure was 14 percent; for highway and street construction, 13 percent. For workers in blast furnaces and steelworks, the statistic described above was 2 percent; in motor vehicles and equipment, 2 percent. The extent to which construction workers report earnings outside a specific construction industry is more apparent when all employees who worked in the industry (not simply those the majority of whose earnings are from the specified industry) are considered. In 1963, for all who worked for general building contractors, median earnings in general building were exceeded by all reported earnings by 62 percent; in highway and street construction by 67 percent; in heavy construction by 72 percent; in blast furnaces,/steelworks, etc., by 4 percent; in motor vehicles and parts, by 8 percent.[16]

Apparently there is a great deal of occupational mobility for construction craftsmen. A survey of job shifts by men in 1961[17] indicated that approximately one-third of job shifts by persons initially employed as carpenters were to non-construction occupation groups (e.g., to occupations other than those of construction craftsmen or laborers). Similarly, of 562,000 shifts from construction craft occupations other than carpenters, 25 percent were into non-construction occupations. Conversely, over one-quarter of shifts into carpentry were from non-construction occupations, and one-quarter of shifts into other craft occupations in construction were from non-construction occupations. On the other hand, only 1.5 percent of shifts from carpenters' jobs were to other construction occupations; and 1.2 percent of shifts from non-carpenter construction trades were into carpentry.[16] This single survey, virtually all the available information on the occupational mobility of construction craftsmen, suggests that occupational mobility within the building trades is significantly less important than inter-industry movement.

Considerable research into manpower flows in construction is needed. Investigators might begin by specifying the factors which determine the adequacy of the labor force nationally or in a particular area to an expected level of construction activity. Among the more important are the composition of construction demand, the level of employment in non-construction industries, the relative wage between construction and other industries (both among and within occupations), traditional patterns of labor mobility, and the size of the labor pool possessing construction skills.[19] Pressures on the supply of manpower in construction undoubtedly affect the channels and determinants of manpower flows. It is likely, for example, that the tightening of construction labor markets in recent years is directly due to the improved unemployment situation nationally. The 3.2 percent general unemployment rate for males 16 years and older in 1966 undoubtedly reduced the supply of manpower to construction. Presumably, tightened labor markets have been partially responsible for recent rapid wage increases negotiated in the industry. However, models specifying the character of the wage-unemployment connection in terms of craft and locality cannot now be constructed.

The Limitations of the Data

Virtually any hypothesis offered concerning manpower patterns in construction industries may find support somewhere in the practices of the industry. Consequently, data on manpower become essential to a critical analysis of the industry. Only with reliable data can the relative importance of different labor market patterns and their determinants be assessed. Unfortunately, the existent data are inadequate for many important analytic purposes.

Most importantly, data on construction employment and unemployment by detailed craft are remarkably sparse. In the absence of these data it is virtually impossible to have a quantitative impression of the state and direction of the construction labor market. The only important current sources are the Current Population Survey and the decennial censuses.[20] The Current Population Survey is too limited in sample size to bear extensive disaggregation by craft, locality, demographic features, etc. The Census, of course, occurs only infrequently.

Employment data by detailed craft for persons employed in contract construction are generated by the records of trust funds established under collective bargaining agreements in construction. These data include counts of the number of hours worked by craftsmen covered by the funds, hours worked by individual employees, and demographic information concerning these craftsmen. At present, neither the government nor any other institution samples these data on a continuing basis. Such sampling could usefully supplement existing employment data on construction by providing: (1) employment data by detailed occupation; (2) information on the degree of utilization of the employed (in terms of hours worked); and (3) estimates of employment disaggregated by geographic locality, craft, and, in some cases, type of contractor.

Sampling of the records of private health, welfare and pension funds for manpower information would require extensive work in sample design. Most importantly, efforts are necessary to extend estimates based on fund data to the non-union sector of construction. Only about half of the employees in construction are currently covered by the funds.[21] However, coverage in major cities is virtually complete, and the number of funds is increasing. Second, the records of the fund would require adjustment to a wage period basis. Currently employers are often in arrears to funds, and when payment of contributions is made, the contribution hours are included in the current month's total. The simple stratagem of requiring contractors, if errant, to specify the earnings period of contributions could correct the misallocation of hours worked by time periods.

In addition, current employment data do not allow adequate analysis of manpower flows in construction and between construction and other industries. The Social Security Administration's Continuous Work History Sample is the most promising source of inter-industry and geographic mobility data.[22] Unfortunately, the Social Security Administration collects no occupational information. However, extensive data on industrial and geographic mobility by craft for construction workers might be obtained by utilizing the records of private trust funds and the Social Security Administration in concert. Private fund records normally contain the Social Security number of persons reported to the funds. With these numbers, the work experience, in terms of industry of employment, number of employers, earnings, etc., of craftsmen could be traced in the records of the Administration. A final possibility is to tap the records of the Internal Revenue Service (IRS) for occupational information. The IRS collects on an annual basis the occupation of taxpayers as well as considerable information as to their sources of income. Currently, however, IRS does not publish data on an occupational basis.

Conclusion

Future research into manpower in construction must undertake to describe and analyze the flow of craftsmen among industries, occupations, and areas. Models relating the flow of construction manpower to the all-industry rate of unemployment, relative wages among crafts and industries, traditional patterns of mobility, training and education efforts, and the level and composition of construction demand are urgently needed. In order to facilitate this research, measures of the flow of manpower must be developed. Data currently available are inadequate to isolate the direction, magnitude, or determinants of these flows. Both the federal government and the industry itself are collecting, in one form or another, a large body of information relating to manpower in construction. Considerable effort is now required to make these data accessible to scholars in a form usable for economic analysis.

Footnotes

[1]I should like to thank John T. Dunlop, Donald E. Cullen, and Joe Russell for their comments on earlier drafts of this paper.

[2]See the President's remarks at the appointment of Stanley Ruttenberg to be Assistant Secretary of Labor, June 17, 1966.

[3]See, for example, "Manpower Crisis Ahead?" American Builder, 99, 1 (January, 1966), 70-71; "The Building Trades Shortage," Practical Builder, 31, 4 (April, 1966), 65-67; and "AGC Rises to Labor Challenge," Engineering News-Record (October 5, 1967), 17-18.

[4]There are, of course, limitations to entry to building trades employment in certain areas and among specific crafts. The recent controversy over the proportion of Negroes in building trades employment has been concentrated on the specialty trades, especially the plumbers, sheet metal workers, electricians and ironworkers. See Ray Marshall and Vernon M. Briggs, Jr., The Negro and Apprenticeship (Baltimore: John Hopkins University Press, 1967).

[5]These figures are from the Monthly Survey of the Labor Force and refer to the all-construction definition rather than to contract construction. Wage and salary workers in construction include those employed both in contract construction and government construction agencies. The labor force is defined, of course, as the sum of the employed and the unemployed attributed to the industry.

[6]These figures are extensions of a series first presented by David Farber and Elsa Loewenstein in Annual Paid Man-Hours of Employment and Annual Wages, 1946-54 (Washington, D.C.: U.S. Department of Health, Education and Welfare, Social Security Administration, 1962),

40-41; and updated by Farber, "Apprenticeship in the United States: Labor Market Forces and Social Policy," Journal of Human Resources, 2, 1 (Winter, 1967), 88. The full-time job measure is actually average annual employment, which Farber argues effectively to be a job -- not persons -- concept. See Farber and Loewenstein, op. cit., 29ff. (The employment figure is derived from the records of the Social Security Administration.)

[7]At any given moment, a considerable number of persons trained in construction occupations are employed in non-construction industries. For example, among carpenters, on annual average for 1966, some 30 percent were employed in non-construction sectors of the economy. And, in 1966, on annual average, over 30 percent of all construction craftsmen other than carpenters were employed in non-construction industries. These data are estimates of the number of persons employed in non-construction industries doing work of comparable skill and job content to the work of a building tradesman in construction. They do not include, of course, those persons with building trades skills employed in other occupations.

These data are from unpublished estimates of occupation by industry made from the Monthly Survey of the Labor Force by the Bureau of Labor Statistics. I am most grateful for their being made available to me for this paper.

See also Allan F. Salt, "Estimated Need for Skilled Workers, 1965-75," Monthly Labor Review (April, 1966), 365-71. Mr. Salt reports 1965 annual average employment in construction and non-construction industries for eleven building trades (368).

[8]The degree of employment seasonality in construction nationally has remained virtually unchanged in the post-World War II period. See Robert J. Myers and Sol Swerdloff, "Seasonality and Construction," Monthly Labor Review, 90, 9 (September, 1967), 1-8.

[9]This dichotomy is not original. See A. H. Belitsky, "Hiring Problems in the Building Trades," unpublished Ph.D. dissertation, Harvard University, 1960; also Maurice Parodi, "Wage Drift and Wage Bargaining: A Case Study of the Building Industry in Marseilles," British Journal of Industrial Relations, 1, 2 (June, 1963), 213-227.

[10]See Belitsky, op. cit., for a discussion of the efforts of contractors to keep their key people.

[11]See the annual Work Experience Surveys of the Bureau of Labor Statistics, "Extent of Unemployment by Industry: Wage and Salary Workers, by Longest Job in 19__." These data are published annually in the Monthly Labor Review or are available from the Bureau.

[12]The process of attracting and holding workmen in construction jobs may be accompanied by wage increases offered by contractors. Non-wage incentives are also offered. The nature and relative importance of special incentives in attracting labor are not described in the text. This paper attempts to describe the flows into the industry, the current pattern of incentives taken as given.

[13]This figure is obtained by dividing the number of men for whom earnings were reported in contract construction into the total number of industry divisions (two-digit SIC classification) in which they had earnings reported. The use of industry divisions as the unit of measurement here suppresses the considerable rate of job transfer which occurs between branches of contract construction itself.

[14]These data are from tabulations made from the Social Security Administration's One Percent Sample. They are published in Sebastia Svolos, "Measures of Labor Mobility and OASDHI Data," Social Security Bulletin (April, 1966), 42.

[15]See the Handbook of Old-Age, Survivors, Disability Insurance Statistics: Employment, Earnings and Insurance Status of Workers in Covered Employment, 1957 (Baltimore, Md.: Social Security Administration, 1965), 34 and 42.

[16]These estimates were made for the Bureau of Labor Statistics from the Social Security Administration's One Percent Sample.

[17]Gertrude Bancroft and Stuart Garfinkle, "Job Mobility in 1961," Monthly Labor Review (August, 1963). Note: one person could have had several job shifts, as the terms are used in this study.

[18]These statistics were tabulated by the author from Bancroft and Garfinkle, op. cit., Table A-9.

[19]See, e.g., Donald E. Cullen, "Labor Market Aspects of the St. Lawrence Seaway Project," Journal of Political Economy, 68, 3 (June, 1960), 232-251.

[20]In 1966, the Bureau of Employment Security discontinued collection of monthly employment figures by detailed occupation from the records of the Unemployment Insurance System. Currently, therefore, the only central source of unemployment estimates by craft are those of the Current Population Survey.

[21]This is the author's estimate, based in part on Daniel M. Holland, Private Pension Funds: Projected Growth (New York: Columbia University Press [NBER Occasional Paper 97], 1966), 27.

[22]Social Security data are seriously deficient, however, with respect to analysis of both industrial and geographic mobility. The classification of firms by type of contractor (following the Standard Industrial Classification) restricts analysis by type of construction done, such as residential construction, demolition, commercial construction, etc. With respect to geographic mobility, our sources of employment data are biased in an unknown manner by the character and location of the reporting establishments. Employment is attributed to the place of the reporting unit, which may or may not be near the construction site.

THE MILITARY AS A TRAINER: A STUDY OF PROBLEMS IN MEASURING CROSSOVER

Paul A. Weinstein, University of Maryland
Eugene L. Jurkowitz, Standard Oil of New Jersey

Introduction

Problems of government employment continue to be of relevance for study by those interested in industrial relations and human resources. Especially vital is a study of the most pervasive as well as the largest manpower activity, the military.

This report is concerned with the impact of military manpower training on the individual. The economic benefits of military service might take either of two forms; general capital advantage, which would increase the marginal productivity of the veteran vis-a-vis the non-veteran, regardless of civilian occupation, and specific capital advantage, which is related to the set of occupations defined as having this capital as a component.

This research is a part of a larger study which treats the military as a variable set and examines the impact of it on a set of dependent factors.[1] Whereas the purpose of this part of the study is to examine the partial effect of the military on the economy through the direct training system, the method employed may be useful in evaluating alternative training institutions.

I. The Problem

A design was chosen to test the factors accounting for the utilization or non-utilization of skills acquired in the military. It was assumed that age at entry, education prior to service, region, and interest in the military for training purposes are significant in the utilization of particular occupational training. The age at entry was interpreted as an index of occupational attachment that varied inversely with transfer. Thus, draftees, who have traditionally been older than enlistees, were thought to make less use of their training. It was assumed that component of entry, independent of age, similarly varied inversely with transfer; that is, volunteers leaving the service, who had a wider selection of military occupations, would be more likely to apply their chosen experience. The veteran's interest in the military for training purposes as well as the relation of the military occupation to pre-military experience were both assumed to affect transfer positively, although not necessarily additively. Aptitude, education, and region are variables with a more complex impact on crossover. The set of variables used in the study and the hypotheses with respect to each appear in our forthcoming study.

Both prior research and casual observation suggested that the findings would indicate non-transfer of human capital and the reasons for non-transfer should be probed. Specifically, do the reasons suggest policy parameter adjustments on either supply or demand factors and, in turn, are the variables perceptual or existential?

II. Study Design

Basic to the overall study design is the idea of convergence of occupational structures. Setting aside activities or occupations which are purely military, another set of activities exists in the military which is obviously transferable to the civilian sector. As a consequence of technological developments in both the military and civilian employment sectors, the occupational structures have converged, with less emphasis on "people-killing" and more on technical, administrative, and service activities. The hypothesis, therefore, was the more alike the occupations, the greater the probability that skill would be transferred. This assumed that convergence is real and that capital holders would pursue the use of the capital and employers would accept and ascribe economic value to a veteran's experience.

The design then became fixed upon a categorization of military skills on the basis of apparent convergence (See Table I). The ten groups reflect significant areas of the military training effort, yet a variety of occupations, employers, and prospects in addition to varying degrees of openness and trade influence.

We assumed, following custom, that the military may have a value (positive) abstracted from any occupational benefit that may accrue through the training.[2] For this reason, infantry was chosen as a benchmark and a proxy for the military.

While it would be valuable to generalize results, we are unable to do this because of the self-selective nature of the military. The differences between the services are significant for such manpower variables as procurement and training.

We opted to limit the analysis to the Army and the Navy, which substantially cover all reasonable occupational categories, and to consider only enlisted male non-careerists.

Time out of service was maximized, and it was assumed that utilization as a function of time would vary by occupational category, the constraint on time being the increased likelihood of response loss, hence potential bias in the sample.

As a consequence, the standby reserves were chosen for the defined population. This minimized the task of sample selection. In procedures to be described elsewhere, we sampled, located, and interviewed the population. A telephone survey was chosen because of population characteristics, the nature of the questions asked, and the higher cost of data acquisition by other means. The response is shown in Table I.

The data for the study were derived from three sources, two of which are relevant to this report. The Army record was used to supply data on pre-employment education and work experience, aptitude and work experience in the military, and basic demographic data. In the survey, relations between pre-service and service experience were ascribed in dichotomized form, as were the relations between pre-service and post-service occupation, both primary and moonlighting. These relationships were set on a three-point scale. The post-service occupations were not necessarily complete, but covered a maximum of three positions, including the first, current, and prior to current position. The occupations were coded according to the Dictionary of Occupational Titles (1966).

The specific problem discussed below, is part of the second stage of the analysis for the Army group. It is intended to illustrate the technique used as well as to illuminate the problem.

III. The Procedure for Analysis

While the impact of military training might take many economic forms, we concentrated on the differential effects of such training on adjusted annual income. The hypothesis was that investment in human capital in military vocational fields increased post-military individual earnings. The ten occupational groupings referred to above were used to test this hypothesis. Each of ten groups (except infantrymen) represented potential streams of lifetime earnings in excess of the increment yielded by military life *per se*.

This hypothesis was first tested by analysis of covariance. In that stage of analysis, we compared the income of each occupational group with that of the infantrymen, cleansed of six independent variables: Educational level, number of dependents at time of exiting service, age, ability index, number of post-service jobs held, and relation of post-service occupation to military vocational training. Two conclusions emerged: (a) The difference between the occupational groups and the base group was nil, and (b) within an occupational group, income was positively related to occupational crossover. These conclusions are inconsistent if the military training caused the post-service earnings to expand. Alternatively, the response may be attributed to other variables which were correlated with (b).

The inconsistency was examined with additional explanatory variables. One set of variables involved pre-military occupational experience and raised the question of effects due to draftee versus volunteer mix. Draftees, as opposed to enlistees, are older and more likely to have had pre-service job experience at the time of induction. It was also observed that the infantry group contained the highest percentage of draftees. This data led to a test of the relation between pre-service occupational experience and post-military economic activity.

This hypothesis was examined for overall patterns by pooling all occupational groups. Then the data was disaggregated for the ten occupational groups to clarify relations which may emerge from the aggregate.

Four variables directly or indirectly related to pre-service occupational experience used in the test are:

- Z_1 = A dummy variable representing the existence or absence of any pre-service job. Coded 1 for affirmative and 0 for negative.
- Z_2 = A dummy variable indicating the relation of the pre-service job to military activity. Coded 1 if related and 0 if not.
- Z_3 = A continuous variable ranging directly from 0 to 3 according to the percentage of post-service time in occupations held before the service.
- Z_4 = A continuous variable ranging from 0 to 6. Summing the post-service experience weighted by time and the degree of ascribed relation to the military vocational training (0=no relationship; 1=somewhat related; 2=highly related).[3]

The test was performed with least squares regression lines on the ten pooled groups and then separately for the various occupational groupings. Each of these four variables was inserted separately into four different regressions, in which the effects of the five independent variables had been nullified using the linear effect for the variable. The dependent variable, average annual post-service earned income, was adjusted to comparable 52-week periods. These four variables were inserted into separate regressions because of their relatively high correlation to each other and the consequent statistical problems associated with multi-colinearity and interaction.

IV. All Group Results

The corresponding means, partial regression coefficients, and t-values for each of the four variables is shown in Table II. We note the following:

1. Pre-service occupation was highly significant (t=3.5), positively

related to income (b=460), and held by 88 per cent ($\overline{Z}_1$=.88).
2. For 26 per cent ($\overline{Z}_2$=.26) military vocational training was related to a pre-service job, highly significant (t=3.1), and was positively related to income (b=287).
3. One-third of the post-service time ($\overline{Z}_3$=1.0)[4] was spent in occupations held prior to military service. This variable is significant (t=3.6) and positively related to income. The b-value indicated that income would increase $402 (3 x $134) if all post-service employment were related to pre-service jobs.
4. This variable was significant only at the 80 per cent level and was positively related to income (b=29). The b-value of 29 indicated that if all post-service time were in a job highly related to military vocational training, earnings would be increased by $174 more per year (6 x $29).

That individuals are benefitting in post-military economic life from their pre-service occupation was indicated by the positive relationship of Z_1 to income. This suggestion was reinforced when we observed that Z_1 was not correlated with the other variables (age, ability index, etc.) used in the analysis.

Two explanations for the high correlation shown in Z_2, between pre-service occupation and military vocational assignment are:

1. Individual preference for a related occupation.
2. Military preference for pre-service experience.

The positive relation of the variables to income is reasonable. The relation of military to pre-service employment, Z_1, points toward pre-military service as an explanation for the income gain. This is reinforced by Z_3, the distribution of post-service time in occupations related to pre-military activity. This is specially true for infantrymen where Z_1 was only significant at the 85 per cent level, but the partial regression (b=628) was larger than the pooled groups. This may explain the lack of an income difference between the infantry and other groups.

The partial regression coefficients for Z_1 and Z_3 were approximately equal (b=400). The income effect appeared invariant to whether or not individuals pursue their pre-service activity; but veterans with related military vocational training. This confirms our conclusion (a) that military vocational training was without economic benefit.

Despite this, a positive income effect was evidenced for individuals with post-service time in jobs highly related to their military vocational training (Z_4). This positive effect led to conclusion (b), although the magnitude of the Z_4 effect was relatively low compared with Z_1 and Z_3 (174 versus 400).

V. Disaggregation Results

Disaggregation at the level of the ten occupational groups yielded the results shown in Table III.

1. Pre-service job experience ($\overline{Z}_1$) varied little among the occupational groups. Z_1 was significant and positively related to income as observed in the aggregate.
2. The relation between military service and pre-military occupation showed considerable variety among the groups. The statistic, as in the total group, was positively related to income except in II and X and was significant.
3. Post-service time spent in pre-service occupations ($\overline{Z}_3$) also demonstrated little variation among the groups, and was both significant and positively related to income.
4. The proportion of individuals who accepted a post-service job related to their military vocational training ($\overline{Z}_4$) was variant among the groups. Z_4 showed a positive relation to income, and was significant.

The problem of the relation of pre-military and post-military experience was broached by dichotomizing the total sample by Z_2 and Z_4 to observe the pre-service experience of these two groups. For $\overline{Z}_4$, a continuous variable, arbitrarily, respondants were assigned to the non-transfer group if their value was 0 to 2.

The comparison of regressions from the two Z partitions revealed that those with a pre-service job related to military vocational training earned significantly higher incomes. However, they were also a little older and had fewer jobs. A higher percentage of veterans with pre-service related jobs were draftees; but draftee status was significant only for those with a pre-service related occupation. The relation of Z_2 to income for the group was large and positive -- more than $400 per year.

The regression comparisons highlighted that a larger percentage of individuals with pre-service jobs related to their military vocation had a post-service job related to their military vocation (Z_4). Moreover, being in a related post-service job was significant and positively related to income for those with pre-service related jobs and the reverse for those without a pre-service related job. Furthermore, individuals without a pre-service related job demonstrated a significant and positive income effect from obtaining a post-service job related to prior service. Thus, these individuals (similar to those who had a pre-service related job) are not benefitting but are using their pre-service occupational experience.

We dichotomized Z_4 at 2 -- the group that was 2 or greater was called the high-transfer group ($\bar{Z}_4$=3.8), and below 2, the low-transfer group ($\bar{Z}_4$=.18). Average income was greater for the former group -- more draftees than volunteers, with a high percentage of pre-service occupational experience. The percentage of high-transfer group with a pre-service job related to their military vocation was three times as great as the low-transfer group. Relatively more of the high-transfer group had vocational preferences and were trained and served in these preferences. Finally, a larger percentage of the high-transfer group returned after the service to pre-service jobs (Z_3).

These observations strongly reinforce the argument of the previous section concerning the post-service impact of pre-service occupational experience.

Summary and Conclusions

It is unquestionable that little use is being made of service-related experience, but the cause is moot. While for some the mere association of an occupation with the military is noxious, this does not appear to be significant. Similarly, the argument that the experience is valueless would appear wrong. *Ex post* the experience is not deleterious nor is the training poor, unplanned, or ineffective. For methodological reasons, the explanation should be integral to economic analysis for relevance and to enrich the theory. This is possible in the case at hand.

In other phases of the study, we observed that a primary reason for non-transfer in all occupational groups was low pay, while the vast majority of those who had jobs related to their service indicated they had not received any economic advantage, either through pay or title, from their service experience. The explanation may lie in the hiring practices of firms. Recent studies indicate that employers do not hire semi-finished workers. Whether it is the cost of determining how to use these individuals or the cost of further training them to the level of an internally produced worker, there are costs which, at the margin, including the calculation of uncertainty, diminish the rate of return for a semi-employed veteran compared with the internally produced factor. At the same time, the seniority system encourages employees to continue with their previously set occupations, and discourage outsiders. Further, the training for this group may be below a minimum threshold level.

A major obstacle in this study is the tenuous design by which tasks are aggregated into occupations. Despite efforts to identify occupations of respondents, the use of this data is of little significance. The difficulties lie in the categories *per se*. The degree of crossover and the direct monetary returns are continuous rather than dichotomous variables. Both at a moment of time and through time, transfer and benefits can vary in intensity. Thus, a descriptive variable to cover a period of time is difficult to produce and interpret.

This illuminates a general problem between data gathering and processing on one hand and analytical manipulations and analysis on the other. Whereas the technique employed in this study allowed one to design the data, it also created problems. The intensity of detail was bought at the price of significance. Our ability to collect and maneuver data is below our capacity to digest the surfeit of material.

Informational Footnotes

[1] This paper was supported by grants from the Ford Foundation and the U.S. Office of Education. It is preliminary to a larger study covering a wide array of topics.
The authors wish to thank the above organizations and the Computer Science Center at the University of Maryland for their support. We also wish to express appreciation to William B. Clatanoff, Jr. for his aid and comments.

[2] We have no evidence on the sign or magnitude of the general capital effect as it relates to occupational or other groupings. It is unlikely to be invariant in the groupings chosen. In this paper benefits when conjectured could be negative or positive.

[3] $Z_4 = 3\left[\Sigma \frac{R_i M_i}{M_t}\right]$

i=1, 2, or 3
R_i=0, 1, or 2; Relation of the i^{th} job to the service
M_i=Months in the i^{th} job
M_t=Months covered

[4] When referred to below b should be interpreted as (bZ_i).

Table I. OCCUPATIONAL CLASSIFICATION, SAMPLE SIZE [a]

AND RESPONSE Rate, 1966

	Career Group[b]	Number in Sample	Percent Interviewed
I.	Combat-Infantry	712	28%
II.	General Military - Duty Soldier	173	40%
III.	Police	478	45%
IV.	Electronic Data Equipment Skills[c]	209	55%
V.	Esoteric Skills[d]	179	55%
VI.	Radio, Radar, TV, and Auto Repair	504	46%
VII.	Trades Related to Telephone	160	49%
VIII.	Operatives-Construction and Repair	560	43%
IX.	Teamster and General Warehousing	218	38%
X.	Business and Service Activities	1,228	46%

a - Universe drawn from Standby Reserves, with indicated PMOS 2-5 years from active duty. (Non-career)
b- These are composed of series of occupations.
c - Operation and repair
d - High formal training on sophisticated military equipment

TABLE II

The Magnitude and Significance of Four Variables Related to Pre-Service Occupation for Total Sample

5847 = Average Income

Variable	Mean	(b) Regression Coefficient	$t_{.95}=1.96$ t-value
Z_1	.88	460	3.5
Z_2	.26	287	3.1
Z_3	1.0	134	3.6
Z_4	1.2	29	1.3

TABLE III

Mean, Partial Regression, and t-value
for Four Variables by Occupational Groups

Occupational Groups[a]	Z_1			Z_2			Z_3			Z_4		
	$\bar{Z}$	b	t	$\bar{Z}$	b	t	$\bar{Z}$	b	t	$\bar{Z}$	b	t
I	0.9	628	1.4	---	---	---	1.1	28	.2	.5	25	.8
II	0.9	566	.9	0.1	-1257	1.4	.9	30	.7	.7	-55	.5
III	0.9	541	1.6	0.1	368	0.9	1.1	69.7	.7	.7	153	✓2.3
IV	0.9	562.7	1.1	0.6	550	✓2.1	.8	-261	.7	2.3	-31	.5
V	0.8	-314	.5	0.1	749	1.5	1.2	19	.1	.7	-13	.1
VI	0.8	481	1.3	0.3	169	.6	.9	173	1.5	1.3	38	.5
VII	0.9	1407	✓2.2	0.4	347	.6	1.3	406	✓2.1	2.1	79	.9
VIII	0.9	-740	1.5	0.4	503	✓2.1	1.2	8.9	.1	1.9	131	✓2.[illegible]
IX	0.8	121	.2	0.2	152	.4	1.1	90	.7	.9	-89	.9
X	0.8	622	*2.8	0.2	-104	.5	.9	193	*2.8	1.1	-91	✓2.[illegible]
Total	0.8	460	*3.5	0.2	287	*3.1	1.0	134	*3.6	1.2	28.5	1.[illegible]

[a]See Table I for definition of groups.
*significant to .99 level
✓significant to .95 level

DISCUSSION

T. Aldrich Finegan, Vanderbilt University

Since my own current research efforts are largely concerned with underutilization of manpower, I should, perhaps, have confined my remarks to Harold Goldstein's stimulating paper. However, I cannot resist the temptation to comment briefly on the other two papers as well.

The main thrust of Mr. Mills' paper is that before much headway can be made in analyzing the sources and flows of manpower in the construction industry, new data -- some of which are already being collected -- must be made available to interested researchers. As an omniverous consumer of manpower statistics in related fields, I can only applaud his efforts to increase our knowledge about this important industry.

At the same time, it must be added that economists have always wanted more and better data, and that the prospects of getting them have often turned on how much the missing information might contribute to policy issues of great moment. Mills mentions one such issue -- the extent to which the supply of manpower to construction would support a large scale program to rebuild our cities. Let me mention another one. In comparing the economic effects of monetary and fiscal policies to combat inflation, the point is generally made that reliance on tight money imposes a disproportionate burden on the construction industry. Yet we know very little about the extent to which the various groups of workers in this industry are able to find temporary jobs elsewhere when a decline in building construction occurs. The data Mills seeks ought to yield some tentative answers to these and other important questions.

I would also urge investigators in this field not to neglect the theoretical framework for their research. Constructing models with testable implications should help us to select the key empirical relations to be explored and should also help to clarify what particular kinds of data are needed most. Mills' already persuasive case for more data would have been even more persuasive had he developed more fully some of the specific propositions he wishes to test.

I turn now to the challenging paper by Weinstein and Jurkowitz. As the authors suggest, the problem here is not lack of data but what to make of the intriguing results that we observe. This paper strikes me as a valiant effort to estimate the impact of military vocational training on the subsequent civilian earnings of veterans -- an inquiry of major importance; but I am not entirely convinced by the authors' findings that this impact is nil.

While I do not fully understand all of the intricacies of the statistical methods employed, certain aspects of the model are unsettling. One is the technique of interviewing by telephone. Another is the extremely low response rate (28 percent) for former infantrymen, who serve as the "control group" in the interoccupational analysis, and the bias that this low rate of response may have generated. (In this connection, I wonder whether a sample of men with no military service at all would not have constituted a better control group -- to guard against the possibility that even infantry training conveys some lasting economic benefits. But this alternative approach may not have been feasible.)

I am also concerned about the authors' choice of control variables in the multi-group analysis (p. 7). Why, for example, was it necessary to control for the relation between the veteran's post-service earnings and (1) the number of dependents he had at the time he left the service, or (2) the number of post-service jobs he had held, or (3) the relation of his post-service occupation to his military vocational training? The flow of causation between family size and subsequent earnings is, at best, obscure; and the other two variables are likely to be by-products, at least in part, of whatever occupational skills the individual may have acquired during military service. At the same time, the apparent failure to include color as a control variable is also puzzling.[1] For all of these reasons, I am reluctant to accept at face value the absence of any significant difference between the adjusted civilian earnings of veterans in the infantry group and the earnings of veterans with some in-service vocational training.

---

[1]After the session, the authors indicated that color had been explicitly taken into account in other phases of their research, but that this characteristic had proved to be statistically nonsignificant. An evaluation of this surprising result must be deferred until the more detailed findings of this study have been published.

Later on in the paper Weinstein and Jurkowitz observe that those veterans who had a pre-service job and related military vocational training tended to have lower earnings (on balance) than their counterparts with a pre-service job but no related military vocational training; and the authors interpret this result as "confirming" the conclusion "that military vocational training was without economic benefit" (p. 12). I am not so sure. The second of these two groups contains those veterans who received military vocational training that was unrelated to their pre-service work experience, but it scarcely follows that the latter set of skills would be less valuable in the long run than those skills that were related to previous civilian jobs. Indeed, just the opposite might be true.

Finally, the authors indicate some lack of confidence in the occupational categories and measures of crossover employed -- a lack of confidence I share. Specifically, variable Z_4 may underestimate the extent to which veterans with low scores on this characteristic were actually using skills acquired during their period of military service. At any rate, a more precise measure of the amount of occupational training received by servicemen while on active duty is needed, in my view, before the impact of this training on subsequent earnings can be confidently appraised.

Mr. Goldstein's paper offers a comprehensive and perceptive discussion of the major forms of underutilization of human resources in our economy, as well as some recent BLS data on the quantitative dimensions of the problem. In the balance of my allotted time, I wish to comment on only one kind of underutilization -- specifically, what Goldstein refers to as "involuntary nonparticipation in the labor force."

Goldstein mentions a number of programs that the government might adopt to reduce the amount of this kind of economic loss, and they have much to commend them. Let me simply mention another policy whose benefits in this regard should not be overlooked, and that is the macro-policy of maintaining a high rate of growth of aggregate demand. A substantial body of recent research by Glen Cain, Alfred Tella, Jacob Mincer, and William Bowen and myself (among others) leaves little doubt that the labor force participation rate of most major demographic groups is inversely related to the overall rate of unemployment in the economy. In short, a policy that succeeds in keeping labor markets relatively tight will not only reduce the level of reported unemployment (especially among disadvantaged groups) but will also raise the overall labor force participation rate.

In this connection, Goldstein cites the results of a survey made in September 1966 indicating that 750,000 persons were not in the labor force that month because they believed it was impossible to find work. The figure is instructive, if only because it shows that the number of "discouraged workers" does not fall to zero once the unemployment rate reaches 4 percent. But I am confident that a similar survey conducted three or four years earlier would have revealed a much larger number of discouraged workers.

Research on the reasons that persons give for nonparticipation holds considerable promise, and current efforts of the Bureau of Labor Statistics to increase our knowledge in this area are certainly to be commended. But this line of inquiry also has its limitations, in that the reasons people offer for their current labor force status may be related in rather complex ways to the labor market conditions prevailing at the time. One might very well find, for example, that housewives who report they are not seeking market work for personal (non-economic) reasons during a recession year would gladly accept such work if wages were higher or jobs of the desired kind were more plentiful. Similarly, the same chronic health condition which is cited as the reason for not seeking work in a recession may prove to be less of a handicap when job prospects are brighter.

Thus, I hope that major efforts will be made to relate the findings of these household interviews with the results obtained from statistical analyses of ex-post relationships between participation rates and objective measures of personal characteristics and labor market conditions.

DISCUSSION

Felician F. Foltman, Cornell University

These three papers represent an important contribution to the burgeoning manpower field, but are too dissimilar to be compared except in a very general fashion. In both comparative and absolute sense I find I am in substantial agreement with the authors of these papers. Having duly performed this rite of expiation I am now free to enter my reservations, observations and caveats in the time honored manner of discussants. Since I am not a statistician either by training or inclination, my comments will necessarily represent a "consumer's" viewpoint.

Weinstein and Jurkowitz have made a substantial contribution to a facet of manpower development that to date has not received broad study; the presumed convergence of military and civilian occupational structures. This convergence, has resulted from important changes that are occurring in both the military and civilian sectors. Some have observed that the military has undergone a "civilianization" and that, by the same token, there is a militarization of society.

They make a number of assumptions which need to be more explicit in order to determine the validity of their conclusions. Firstly, they assume that income gain in post-military experience is derived from pre-service occupational experiences. This suggests persons are paid on the basis of skill or merit alone without reference to labor market factors. In this same connection their aggregated career groups tend to dampen out the often important labor market factors for particular occupations at particular moments. Thus, in Category III, Policy, presumably one would find the person whose military training prepared him to be primarily a "traffic cop" as well as other more skilled crime investigators. I would not consider such occupations homogeneous; neither have they received the same amount and type of military training. I do not think these two illustrative occupations fare equally well in post-service labor markets.

It may well be that the scope, amount and type of military training is more important than pre-service training. One could have wished that they had controlled for military training.

The relatively short time span after military service is another facet of the study that might lead to invalid conclusions. Is post military income derived in the first two or three years significantly different from that obtained after 10 years? Is employer liability and responsibility to ex-servicemen a unique factor in these early post-service years?

Furthermore, I think it particularly important to carry this type of study forward over a longer time span to test the assumption that the infantry is a proper benchmark against which to measure other military occupations. It may well be that the infantry is composed of untrained and uneducated or those who don't qualify for more esoteric skills. It may also be true that the infantry contains many generalists who were assigned to that branch because of the service demands at the time. In many cases it is simply a matter of when you arrive at the distribution point rather than what training you have had. This may simply reflect my myopia acquired in World War II.

Other methodological considerations bear mention. Does the fact that the universe is drawn from the standby reserve affect either the type of service experience or the characteristics of people? What were the questions asked of respondents during the telephone interview? Since the adequacy of the sample is crucial in this type of study one would have preferred more information. As a matter of preference I would have elected to explore more of the qualitative aspects rather than confine myself exclusively to this more limited design. The authors have produced a rigorous methodology but in the process have overlooked important nuances and overtones.

Finally, the cost benefit technique as applied to this and other studies should be mentioned. Such studies are currently in vogue. They have forced a quantification of many phenomenae which have often been treated on a general or impressionistic basis. But simply to total up annual income is not cost-benefit analysis nor should one infer that all this is human capital and as such related to human investment theory. Too many technical and other considerations are simply not treated.

Goldstein's paper reviews the conceptual, theoretical and practical problems associated with "under-utilized" human resources. I was impressed reviewing this paper, that we will continue to have considerable difficulty measuring the extent and quality of under-utilization. Although there has been some experimentation we have not yet developed even partial measures of underemployment - something the "Gordon Committee" recommended in 1962.

Involuntary part time employment and unemployment are fairly well measured. In 1968, Goldstein notes there were about one million persons who averaged a little over two days work per week although they wanted full time work. We have less of a measurement problem here and more of a public policy problem.

I would underscore his concern with involuntary non-participants in the labor force. He emphasizes repeatedly that men of working age without jobs face particularly difficult problems in our society In a job oriented economy where role, status, rewards, and even emotional and mental health depend on one's job it is disheartening among both white and nonwhite men under 55.

Of all the categories of "sub-employed" or underutilized that of "underemployment represents the most challenging phenomenon to both data producers data consumers. To begin with, there is the tricky problem of whether to measure underemployment against actual or potential skill level. for example, there are people in the labor force who have four or more years of college education, but are employed in occupations that do not require a college education. Very appropriately, this is a matter of concern especially when we suspect the reason for such under use of college training may be discrimination in employment. On the other hand should this be a matter of concern if there is no discrimination? Goldstein doesn't quite take a stand here, but seems to imply that this is under utilization and a matter for concern. I am not that sure.

His discussion of underemployment highlights another important policy problem, that is, should our active manpower policy strive to create just jobs for our citizens or should we embark on the more complicated task of providing jobs which will challenge individuals to realize their every potentiality? I would agree with what I consider the implied thrust of his analysis - we must provide jobs that will use all human capacities.

Mill's paper is less a report of new methods and measures for construction industry than a careful delineation of the deficiences of data currently available. He has performed a very useful service by reminding scholar and practicioner that interpretations of the construction industry work force are made on the basis of very skimpy data. It is remarkable how much we do not know about construction industry manpower.

Questions concerning the adequacy of our skilled manpower are rather continuously surfaced. Although it is true the construction labor force is quite flexible, there are many employers who cannot obtain qualified craftsmen when they need them in some occupational categories. What we don't know, is the magnitude of the shortage that may be related to qualitative factors. It is within the realm of possibility that the paradox of simultaneous unemployment and labor shortages may be at least partially explained by the fact that many of the construction unemployed do not measure up to the minimum standards employers specify.

Reviewing expansion and contraction in the construction industry Mills points out that many persons in the industry do not spend full time there, and we do not yet use this force efficiently throughout a calendar year. He speculates there may be two construction labor forces; one fairly regularly employed as a cadre and another following the work wherever it is. Following this reasoning I think there may be a third force consisting of those partially trained persons who enter the industry by obtaining a work permit in very tight labor market situations. It would be interesting to know to what extent expansion in the construction industry may be due to changing standards and the use of temporary work permits.

Limitations of existing data sources are adequately portrayed but the reader who is looking for new methods will be disappointed. One new source which has great promise he feels, is collectively bargained trust funds. Although only about half the employers in construction are covered by such funds he suggests that by coupling fund data to social security data much could be learned about construction work force mobility. In the event that trustees of such funds could be persuaded to make records available for analysis this might prove to be a fruitful source.

I cannot argue with Mill's conclusion that we need to tap new sources of data at the same time that we extract even more from data already available. Data now inaccessible to scholars should be made available. No one can argue against sin. It would be useful, however, to specify more precisely why we need these data; are we interested in changing public policy; are we concerned with efficient allocation of manpower within our labor markets; or are we interested in evolving new manpower theory. Sharpening the focus would have added considerably to the value of this excellent paper.

VII

MORE LIGHT ON THE DEFINITION OF POVERTY

Chairman, MARGARET E. MARTIN, U. S. Office of Statistical Standards

Page

THE MEASUREMENT OF POVERTY -- AN EXPLORATORY EXERCISE

Harold W. Watts, University of Wisconsin

Introduction

Our basic notions about the extent, distribution and urgency of the problems of poverty, despite many fundamental disagreements about what constitutes the appropriate concept of poverty, are dominated by the measurements produced by a specific, and by no means self-evident procedure. And these basic notions about where poverty is located, what kinds of persons are afflicted, etc., have a profound effect both on the choice of policies to combat poverty and on the allocation of resources among programs serving different parts of the poverty population.

It is, then, of substantial importance to find out how sensitive our practical measures are to variations in the specification of the measurement function. If the relatively minor changes introduced below seem to yield a substantially altered picture of the poverty problem then, at the very least, we must consider carefully the merits of alternative specification. If, as intended, the modifications serve to bring the measure closer to an ideal construct, then the implication of replacing the current measures with revised ones should be assessed.

The rule of measurement currently used whenever the basic data permit determines poverty status as a function of: annual money income of the family unit, the number of persons in the family unit, and farm residence status. This is a function which takes on two values--poor and non-poor--recognizing no further gradations within each category. The current practice of relying on annual money income as the indicator of economic status is maintained in what follows—as, indeed, is the implicit choice of economic status as the essential element of poverty. There are of course persuasive arguments for use of a more comprehensive measure of economic resources of families.[1] There is also a wide range of radically different concepts of poverty that rely on essentially non-economic criteria.[2] But in this paper, attention will be focused on less drastic departures from current practice. The measures used below are of substantial interest in their own right and have, moreover, the inestimable advantage of being applicable to available data.

It is useful to decompose the current poverty function into two sub-functions. The first defines a poverty threshold income or "poverty-line" as a function of family size and farm residence. This threshold value, together with actual income of a family, enters the second sub-function to determine the poverty measure. To be more specific, the function:

$$(1)\quad \hat{y}_1(t) = [\$1000 + \$500n(t)] \cdot (.75)^{f(t)}$$

where $\hat{y}_1(t)$ = current poverty threshold for the t^{th} family

$n(t)$ = number of persons in the t^{th} family

$f(t)$ = 1 if the t^{th} family lives on a farm
= 0 otherwise

provides a very close approximation to the "poverty-lines" developed by Mollie Orshansky--which have been adopted by the Office of Economic Opportunity as the official standard. The second sub-function which completes the current measure of poverty is simply:

$$(2)\quad P_{11}(t) = \begin{cases} n(t) & \text{if } y(t) \leq \hat{y}_1(t) \\ 0 & \text{if } y(t) > \hat{y}_1(t) \end{cases}$$

where $P_{11}(t)$ = current measure of poverty, i.e., number of persons below poverty

$y(t)$ = annual money income of the t^{th} family

In terms of this decomposition, one may consider changes in the threshold function, $\hat{y}$ (n, f, ...), and changes in the poverty indicator, $P(\hat{y}, y, n)$. In the case of the former, changes will take the form of additional variables, i.e., differences in need beyond those related to family size and farm status. Changes in $P(\hat{y}, y, n)$ will be reflected by altered functional form.

The consequences of altering the rules of measurement are observed by applying each variant to the 1/1000 sample of census returns from the 1960 population census. Each of 15 different poverty measures defined below (including P_{11} above) are evaluated in terms of the total amount and severity of poverty (poverty per capita) found in separate demographic groups, occupational groups, and in distinct geographic categories. Since there is no precise basis for standardizing the aggregate measures of poverty, the comparisons will be in terms of the distribution of poverty among different groups--e.g., does one measure find a larger fraction of poverty in central cities or among female-headed families than another?

Variation in the Threshold Function

Two different formulations of the threshold function are proposed for comparison with $\hat{y}_1$, defined in (1) above. The first elaborates the function by taking account of the age structure of the household in addition to the number of persons in it. To be specific, the function proposed is:

$$(3)\quad \hat{y}_2(t) = \left\{ \begin{aligned} \$550 &+ \$750\, n_1(t) \\ &+ \$600\, n_2(t) \\ &+ \$350\, n_3(t) \end{aligned} \right\} \cdot (.75)^{f(t)}$$

where $\hat{y}_2(t)$ = Age-structure poverty threshold for the t^{th} family

$n_1(t)$ = number of persons age 17 or older in the t^{th} family

$n_2(t)$ = number of persons age 6-16 years in the t^{th} family

$n_3(t)$ = number of persons under age 6 in the t^{th} family

and $f(t)$ = farm dummy (defined above).

The choice of allowances for persons of different age was guided by the recently revised equivalence scales estimated by the Bureau of Labor statistics.[3] Their estimates are based on data from the 1960-61 Survey of Consumer Expenditures and use the basic notion that families spending equal fractions of their income on food enjoy equivalent levels of well-being.[4] The above approximation to their equivalence scales was adjusted to place the threshold for a non-farm family of four composed of two adults, one school-age child and one pre-school child at \$3000--the same as for a four-person family in $\hat{y}_1(t)$.

The second modification allows for differences in need and/or cost according to region and size of place of residence of the family. It can be written:

$$(4) \quad \hat{y}_3(t) = (\$1000 + \$500\, n(t)) \cdot I\,(\text{Region } (t), \text{ Place } (t))$$

where $\hat{y}_3(t)$ = geographical poverty threshold for the t^{th} household

I (Region (t), Place (t)) = a geographical equivalence index (tabulated in Table I)

and $n(t)$ = family size (defined above).

The values for the Geographical Index are shown in Table 1. They have been estimated in much the same way and from the same data as the B.L.S. family equivalence scales discussed above. In this case, however, the estimation is based on individual household records for households within a band around the official poverty lines. The methodology is based on fitting constant elasticity curves[5] to expenditures on food and a more inclusive category of necessities. The estimates provided by the regression have been rounded and simplified to produce the index shown in Table 1. The values are normalized, largely by conjecture, to equal 1.000 for the location of an "average" poor person, i.e., to produce roughly the same number of total persons below the threshold $\hat{y}_3$, as below $\hat{y}_1$ and $\hat{y}_2$. In comparison to a flat \$3000 in the case of $\hat{y}_1$ (for non-farm families of four persons), $\hat{y}_2$ ranges from \$4500 in large Northeastern cities to \$2475 for families of four in hamlets and rural areas of the South.

Variations in the Form of the Poverty Function

The function P_{11} specified in (2) above is a very simple kind of poverty measure. It could be termed a "head-count" measure since it simply counts the number of persons in families below the poverty line. This same form of function can be used with both of the revised threshold functions defined above to produce P_{12} and P_{13}. In more general terms, we may write this first poverty function as:

$$(5) \quad P_{1j}(t) = \begin{cases} n(t) & \text{if } y(t) \le \hat{y}_j(t) \\ 0 & \text{otherwise} \end{cases}$$

A closely related function is introduced to investigate the consequences of raising the poverty thresholds by 50%:

$$(6) \quad P_{2j}(t) = \begin{cases} n(t) & \text{if } y(t) \le 1.5\hat{y}_j(t) \\ 0 & \text{otherwise} \end{cases}$$

Another sort of measure of poverty has been used in the literature--namely, the "poverty gap." This measure accords a greater weight to a family's poverty if it is far below the threshold than if it is close to it. In fact it measures the poverty of a family in terms of the dollar distance of family income below the poverty threshold.

Stated precisely, let:

$$(7) \quad P_{3j}(t) = \begin{cases} \hat{y}(t) - y(t) & \text{if } y(t) \le \hat{y}_j(t) \\ 0 & \text{otherwise} \end{cases}$$

A fourth poverty function is derived from $P_{3j}(t)$ by increasing the threshold by 50%:

$$(8) \quad P_{4j}(t) = \begin{cases} 1.5\hat{y}_j(t) - y(t) & \text{if } y(t) \le 1.5\hat{y}_j(t) \\ 0 & \text{otherwise} \end{cases}$$

Finally, a non-linear function of the poverty gap is proposed. It also gives greater weight to poorer families, but at an increasing rate the poorer they get--in contrast to the constant rate implicit in P_{3j} and P_{4j}. It is based on the plausible notion that the personal and social cost or pain increases, not only in proportion to the deficiency of income below some standard, but more than proportionately. Such an assumption is implicit in the argument that it is more important to add \$500 to the income of someone \$2000 below the poverty threshold than it is to add \$500 to the income of someone who is only \$500 short of that same threshold. P_{3j} and P_{4j} imply that it is equally important to add to anyone's income as long as they are below poverty. P_{1j} and P_{2j} give no credit at all for increases in income except when an increase pushes total income over the poverty threshold. The explicit form for this non-linear measure is:

$$(9) \quad P_{5j} = n(t)\, \frac{\log\,(1.5\hat{y}_j(t) - \log\,(y(t) + \$100)}{\log\,(1.5\hat{y}_j(t) - \log\,(\hat{y}_j(t) + \$100)}$$

if that expression is positive

= 0 otherwise

This function, ignoring the \$100 added to income to prevent zero incomes from producing infinite values, equals zero for incomes at or above $1.5\hat{y}_j(t)$. It equals $n(t)$ when income equals $\hat{y}_j(t)$; $2n(t)$ when income is 2/3 of $\hat{y}_j(t)$; $(k + 1)n(t)$ when income is $(2/3)^k\hat{y}_j(t)$. Another way of describing this measure is to say that it specifies equivalence between the poverty of 1000 persons at $\hat{y}_j(t)$ and 500 persons at 2/3 of $\hat{y}_j(t)$ or 333 persons at 4/9 of the poverty line, etc. Economists will recognize the origin of this function in notions of diminishing marginal utility of income. It will, in what follows, be termed the "disutility function."

The Application of the Alternative Measures to 1960 Census Data

Combining the three threshold functions with the five variants of the poverty function produces 15 different combinations--or 15 different poverty measures. These will be denoted P_{ij}(i = 1,2,3,4,5; and j = 1,2,3). Each measure can be evaluated for any family (or individual) for which we have data on annual money income ($y(t)$), family size and age composition ($n_1(t)$, $n_2(t)$, $n_3(t)$, $n(t)$), farm status ($f(t)$), and location by region and place size. The $P_{ij}(t)$ can be summed over all t in a national sample of families and individuals to produce estimates of the total amount of poverty as variously measured. They can also be summed over sub-groups to produce corresponding sub-totals.

The 1/1000 sample of the 1960 census provides the necessary information for evaluating the P_{ij} along with a large and representative sample from which to generalize the results. Its primary disadvantage is that its data are now seven years old. While this may not seriously impair the value of the study for comparison among the P_{ij}, it does reduce the interest one might have in what the various measures indicate about poverty as we are faced with it today. It is hoped that any improvements in our measures of poverty resulting from the analysis here can be applied to more timely data in the near future.

Recognizing that each of the 15 measures of poverty will come up with a different total amount of poverty--some indeed are measured in different units--one needs a basis for comparison among them. By finding the sub-totals of P_{ij} for families and individuals classified by one or two characteristics, and dividing these by the grand total of P_{ij}, the percentage distribution among components of the population is produced. These distributions can be compared and are an important basis for evaluating the several alternatives. Denote by $p_{ij}(k)$ the percentage of total P_{ij} found in class k:

$$p_{ij}(k) = \sum_{t\epsilon k} P_{ij}(t) \Big/ \sum_{\text{all } t} P_{ij}(t)$$

(The notation $\sum_{t\epsilon k}$ denotes summation over all t belonging to the $k^{\underline{th}}$ group).

Another sort of measure often used to assess the severity of poverty is "incidence,"* a per capita measure, and "relative incidence," the level of per capita poverty in some sub-group relative to its level for the general population. Denote by $\rho(k)$ the fraction of the population falling in the $k^{\underline{th}}$ group:

$$\rho(k) = \sum_{t\epsilon k} n(t) \Big/ \sum_{\text{all } t} n(t) \quad .$$

Now an indication of relative incidence can be obtained for poverty measured by P_{ij} by taking the ratio:

$$r_{ij}(k) = p_{ij}(k) \big/ \rho(k) = \frac{\sum_{t\epsilon k} P_{ij}(t)}{\sum_{t\epsilon k} n(t)} \div \frac{\sum_{\text{all } t} P_{ij}(t)}{\sum_{\text{all } t} n(t)}$$

If $r_{ij}(k)$ is greater than one, then incidence of poverty measured by P_{ij} is greater in the kth group than in the general population, and conversely if $r_{ij}(k)$ is smaller than one. In the tables which follow r_{ij}'s will not be calculated, but the ρ's will be provided for each classification so that the r_{ij} can be calculated by the reader.

Table 2 displays the distribution of poverty between the white and non-white parts of the population according to each of the 15 different measures of poverty. Also shown are the several grand totals for the P_{ij}. It will be noted that all measures indicate incidence rates for non-whites more than twice those for the general population. Looking more closely, one finds that the non-white share and relative incidence falls sharply when the threshold is increased by 50%. Compare P_{1j} with P_{2j} and P_{3j} with P_{4j}. This is simply explained by the fact that unlimited increases in the the threshold would eventually include all the population, and the shares would necessarily approach the P's. Aside from that variation, the measures are very similar in their distribution by race. The geographical thresholds yield a somewhat smaller non-white share, especially for the "gap" type measure, but no drastic change is induced.

Table 3 displays the distribution of poverty by family type for a selected group of four measures, along with the basic population distribution. P_{2j} and P_{4j} (j = 1,2,3) were eliminated since they generally showed regression from P_{1j} and P_{2j} respectively toward ρ as was noted in Table 2. Among the "head-count" measures P_{1j} (j = 1,2,3), there was very little variation by threshold function. Consequently only the distribution for P_{13} is shown. In the case of the "gap" measures, P_{32} (shown) showed more of the poverty among husband-wife families and less among individuals than did P_{31} or P_{33} (shown). Again, the P_{5j} were very similar and only P_{53} is shown. As compared to the head count and disutility measures, the gap measures show less poverty in husband-wife families and more among individuals. Of total P_{33}, 22.3% is found among primary individuals comprising 4.6% of the population, as compared with 9.5% and 11% for P_{13} and P_{53} respectively. Both the gap and disutility measures show almost 20% of the poverty among female-headed households, compared to 16.6% for the head-count measure P_{13}. Finally, it is interesting to note that the incidence of poverty for young husband-wife families is above the population incidence as measured by P_{13}, but below it for the others shown.

Table 4 shows the distribution of poverty by gross occupational categories and work experience of the head. Distributions are shown for P_{13},

*Technically, prevalence is the appropriate term, but "incidence" has been given a currency that will be respected here.

P_{33}, and P_{53}. Since very little change was produced by variations in the threshold formula, tables are produced here only for the geographical thresholds. But given the threshold function, there are striking differences in the allocation of poverty between the Head-count (P_{13}), Gap (P_{33}), and Disutility (P_{53}) measures. Sixty-three percent of P_{13} poverty is found in households with a head possessing a definite non-farm occupation, in comparison with 58% for P_{53}, and only 54% for P_{33}. Only 15% of the gap-type poverty is found among farm occupations, compared to 17% for the P_{13}, and 18.5% for P_{53}. With regard to employment in 1959 (the year to which income data pertains), only 43% of all poor persons are in households headed by a person with less than a half-year of work, but 59% of the gap and 50% of the disutility is found in such households.

Table 5 shows the distribution of poverty by census region and by urbanization, with separate urbanization distributions for the North-East and South. The poverty measures P_{i1}(i = 1,3,5), which use the Orshansky thresholds, and P_{i3}(i = 1,3,5), which use the geographically revised thresholds, are shown. The expected reduction in the South's share of total poverty is shown for all three variations of the poverty function, with the sharpest reduction (from 49.5% to 35.9%) occurring in the gap measurement. The North-East and, to a much smaller extent, the West received the balancing increases in shares of poverty. It is, however, the South which remains the location of a disproportionate number of the poor even after a fairly radical adjustment of the threshold levels.

Turning to urbanization, it can be seen that the geographical thresholds serve to reduce the poverty share in rural and non-SMSA urban areas, with the share in central cities of SMSA's receiving the offsetting increase. The relative incidence in the rural areas is clearly the highest for the P_{i1} measures, but this picture is altered when geographical thresholds are used in P_{i3}. The center cities have the highest incidence for head count and gap measures and are not far behind the rural areas for the disutility measure. Note also that the urban parts of SMSA's outside the center city contain 30% of the population but only 9% of the poverty--a very low incidence rate.

Within the North-East, which received a substantial increase in its regional share by introduction of geographical thresholds, the shift in distribution by urbanization is toward the center cities of SMSA's. These cities now appear to have more than half of the poverty in the North East region--particularly if the income gap measure is used. Within the South, on the other hand, relatively little change in the distribution by urbanization is induced by the geographical thresholds. The rural areas remain the high-incidence areas and the location of more than half of the South's poverty--and 25% of the nation's, as measured by P_{53}.

The substantial shift of poverty out of all areas of the south induced by the geographical thresholds, coupled with the earlier finding of little change in the share for non-whites, suggests that the shift is largely explained by finding more poor Negroes in cities outside the South (especially the North East) and fewer in the South.

Each of the measures of poverty produces a grand total of poverty as shown in the last column of Table 2. These figures, together with some additional totals calculated in the process of adjusting the level of the geographical thresholds, enable one to calculate the elasticity of the totals with respect to changes in the threshold. The head-count measures each have an elasticity of around 1.4--e.g., a one-percent increase in all thresholds will increase the number of poor persons by 1.4%. The gap measures, in contrast, increase by 2.1 or 2.2% with a 1% increase in the threshold. The disutility measure has an elasticity of 1.2 and is thus the least sensitive of the three to variations in the level of the threshold.

Summary

Of the two basic changes in the poverty threshold function, only the geographical variant showed much consequence in terms of the distributions of poverty considered here. Indeed only the quite obvious and expected change in geographical distribution was noted for it. No doubt the age-structure variant would have produced an equally obvious shift in the age distribution of the poor (away from children), but it did not affect distributions examined here.

While neither of these changes in the threshold produced remarkable or surprising effects--perhaps because they didn't--it is important to consider carefully the implications of the effects they do have. Certainly the urgency and magnitude of the poverty problem in our large cities has impressed itself on everyone, including policy makers, perhaps beyond its importance as measured by the currently used measures. The geographical revision provides some support for our extra-statistical senses about the importance of urban, non-Southern, poverty. Although the consequences of the age-structure thresholds for the age distribution of poverty were not computed, it can be expected that the disproportionate share of poverty suffered by the young--according to current thresholds--would be somewhat reduced. And with that reduction, some of the concern about the next generation would be reduced (though certainly not eliminated).

The variations in the shape of the functions measuring poverty for a given threshold appeared to be of some importance for all of the distributions studied here. In view of these differences, and with an inclination in favor of some degree of convex (to the origin) nonlinearity, it would appear that further analysis--both theoretical and empirical--would be useful. The particular nonlinear function used here is only one of the possibilities, and not one that has been chosen for its demonstrable superiority over others with roughly similar shapes. But, having shown that such changes can make an appreciable difference, it becomes doubly important to investigate the alternatives more fully.

As mentioned in the introduction, the measures applied here are all limited to current money income as the indicator of economic status. More comprehensive measures of the level of command over goods and services are desperately needed, and may be available in the near future. Most of the basic ideas introduced above would be fully applicable to a more adequate measure of economic level. Thresholds could be defined in terms of such a measure, and the rest of the analysis could be carried out without change.

Table 1: Geographical Equivalence Index, I(Region, Place)

	No. East	No. Central	South	West
Over 1 million	1.500	1.275	1.050	1.200
.25 to 1.0 million	1.500	.975	.900	1.200
.05 to .25 million	1.125	.975	.900	1.050
2,500-50 thousand	1.125	.975	.900	1.050
Under 2500 & Rural Non-farm	1.125	.975	.825	1.050
Rural Farm	1.012	.878	.742	.945

Table 2: Distribution of Poverty by Race

	Type of Measure	Percentage Distribution by: White	Non-White	Total Absolute Measure
Total population	ρ	89.1	10.9	172.2 million persons
Number below threshold:	P_{11}*	72.3	27.7	37.2 million persons
	P_{12}	72.2	27.8	38.9 million persons
	P_{13}	73.1	26.9	39.1 million persons
Number below 1.5 times threshold:	P_{21}	78.9	21.1	65.3 million persons
	P_{22}	79.2	20.8	68.1 million persons
	P_{23}	79.7	20.3	69.4 million persons
Income gap below threshold:	P_{31}	72.2	27.8	$13.75 billion
	P_{32}	71.5	28.5	13.89 billion
	P_{33}	74.1	25.9	14.55 billion
Income gap below 1.5 times threshold:	P_{41}	75.4	24.6	$33.79 billion
	P_{42}	75.2	24.8	34.81 billion
	P_{43}	77.1	22.9	36.66 billion
Disutility function**	P_{51}	71.8	28.2	133.8 million disutility units
	P_{52}	71.8	28.2	137.4 million disutility units
	P_{53}	72.8	27.2	137.0 million disutility units

*The second subscript denotes the threshold function as follows: If 1, the Orshansky approximation (1) above; if 2, the age-structure threshold (3) above; if 3, the geographical threshold (4) above.

**Equals zero above 1.5 times threshold.

Table 3: Distribution of Poverty by Family Type

Type of Measure	P_{13}	P_{32}	P_{33}	P_{53}	ρ
Type of Family:					
Husband-wife					
Head under 25 years	4.3	3.1	3.3	3.7	3.9
Head 25 to 64 years	56.9	46.9	42.3	54.3	74.7
Head 65 or over	10.2	11.4	10.2	9.0	7.0
Male head without spouse	2.5	2.6	2.5	2.4	2.1
Female head without spouse	16.6	19.9	19.4	19.6	7.7
Primary Individual					
Under 65	3.9	7.0	9.6	4.9	2.8
65 or over	5.6	9.1	12.7	6.1	1.8
TOTAL	100.0	100.0	100.0	100.0	100.0

Table 4: Distribution of Poverty by Work Experience in 1959 and Class of Worker

	Weeks Worked:	0	1-26	27-47	48-52	TOTAL
	Class of Worker					
ρ	Farm	.44	.67	1.00	5.15	7.26
	White-collar	1.30	1.24	2.70	26.34	31.58
	Blue-collar	3.18	3.70	9.04	35.73	51.65
	Other	6.55	.28	.47	2.21	9.51
	TOTAL	11.47	5.89	13.21	69.43	100.00
P_{13}	Farm	1.31	2.26	2.73	10.51	16.81
	White-collar	2.31	2.06	1.90	6.26	12.53
	Blue-collar	8.08	9.06	11.85	21.51	50.50
	Other	17.37	.77	.75	1.27	20.16
	TOTAL	29.07	14.15	17.23	39.55	100.00
P_{33}	Farm	1.59	2.34	2.34	8.46	14.73
	White-collar	3.71	2.42	1.56	4.77	12.46
	Blue-collar	10.36	9.83	8.68	13.13	42.00
	Other	27.79	1.14	.85	1.03	30.81
	TOTAL	43.45	15.73	13.43	27.39	100.00
P_{53}	Farm	1.60	2.63	2.91	11.40	18.54
	White-collar	3.04	1.95	1.54	6.35	12.88
	Blue-collar	9.08	8.92	9.68	17.30	44.98
	Other	20.93	.83	.73	1.11	23.60
	TOTAL	34.65	14.33	14.86	36.16	100.00

Table 5: Distribution of Poverty by Region and Urbanization

Type of Measure	P_{11}	P_{31}	P_{51}	P_{13}	P_{33}	P_{53}	ρ
Region:							
North East	15.6	16.2	15.3	21.9	27.2	20.9	24.9
North Central	23.3	23.2	23.6	23.7	23.8	24.0	28.9
South	49.9	49.5	50.2	42.0	35.9	43.1	30.6
West	11.2	11.1	10.9	12.4	13.1	12.0	15.6
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Urbanization:							
Rural	45.9	43.9	47.6	42.6	38.0	44.8	30.2
Non-Metrop.-Urban	19.6	20.4	18.6	18.6	17.9	17.6	19.0
Fringe of Met. Area	8.9	8.9	8.9	8.9	8.8	9.0	29.9
Center City of Met. Area	25.6	26.8	24.9	29.9	35.3	28.6	20.9
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0
North-East Only							
Rural	24.0	22.3	23.5	19.8	17.3	20.8	19.3
Non-Metrop.-Urban	15.2	15.0	14.8	13.4	11.3	12.9	14.0
Fringe of Met. Area	16.9	17.2	18.3	14.8	12.8	16.1	30.2
Center City of Met. Area	43.9	45.5	43.4	52.0	58.6	50.2	36.5
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0
South Only							
Rural	55.6	55.0	57.9	55.5	53.9	58.0	41.5
Non-Metrop.-Urban	21.1	22.0	19.7	21.7	22.8	19.9	22.3
Fringe of Met. Area	5.0	4.6	4.6	4.5	4.3	4.2	11.7
Center City of Met. Area	18.3	18.4	17.8	18.3	19.0	17.9	24.5
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Footnotes

[1]See Harold W. Watts, "An Economic Definition of Poverty" (Discussion Paper No. 5, Institute for Research on Poverty, 1967), and Burton A. Weisbrod and W. Lee Hansen, "Assessing Economic Welfare of Consumer Units Through a Merging of Income and Net Worth" (University of Wisconsin, unpublished, July, 1967).

[2]Morton S. Baratz and William G. Grigsby, in association with Martin Rein, Conceptualization and Measurement of Poverty (Institute for Environmental Studies, University of Pennsylvania, July, 1966).

[3]B.L.S. Bulletin, 1966, 1570-2.

[4]See "Estimating Equivalent Incomes or Budget Costs by Family Type," Monthly Labor Review, November 1960, Reprint No. 2357; and Harold W. Watts, "The Iso-Prop Index: An Approach to the Determination of Differential Poverty Income Thresholds," The Journal of Human Resources, Vol. II, No. 1, Winter, 1967.

[5]See Watts, loc. cit.

Note: This research was supported by funds granted to the Institute for Research on Poverty, pursuant to the provisions of the Economic Opportunity Act of 1964.

AREA DIFFERENCES IN LIVING COSTS

Jean C. Brackett and Helen H. Lamale, Bureau of Labor Statistics

Since the advent of the "war on poverty" in the current decade, the fact that poverty is a concept not easily defined has been well documented. Even where agreement can be reached in general terms on what constitutes a minimum standard of living--below which families may be defined as "poor"--there remains the problem of translating such a generalized concept into a specific list of commodities and services that can be priced, as a basis for estimating minimum living costs. Furthermore, there are two other facets of the problem, which add to the complexity of defining poverty or counting the number of families who are poor. (1) Living costs vary with the size of family and with the age and other characteristics of family members; and (2) Living costs vary from place to place--even for the same or an equivalent living standard, for the same family type, and at the same point in time.

With respect to the first of these problems, extensive analyses of consumption data dating back over more than a century have provided a variety of measures of general welfare, e.g., the relative adequacy of diets, the proportion of income spent for various categories of goods, or the proportion of income saved. These measures, either singly or in combination, have been used as the basis for determining scales of equivalent income for families of different size, age, and type. Hence, global estimates of equivalent costs of consumption for different family types can be obtained, if base cost estimates are available for at least one family type. 1/

Studies pertinent to the second problem--measuring the impact of locality differences--have been much more limited. It is frequently assumed that the BLS Consumer Price Indexes for 23 metropolitan areas can be used for this purpose, but this is not the case. These individual area indexes are not based on a uniform "market basket" of goods and services, but on the particular "market basket," or pattern of expenditures of wage-and clerical-worker families, in each area. Like the U.S. urban CPI, each city index is designed to measure changes in price levels over time; and the index weights for the city remain constant over time, except for major revision periods. In the absence of a common set of weights, however, the Consumer Price Indexes for individual cities cannot be used to measure differences in price levels among the cities.

If measurement of intercity differences in price levels were our objective, it would be a relatively simple task to compile such an index using a common set of weights, based, for example, on the U.S. urban average pattern of expenditures for wage-and clerical-worker families. For most purposes, however, interest centers on the question, "How much more does it cost to live in one community than in another?", not simply "How much lower or higher are prices in one area than in another for a theoretical market basket of goods?" Of course, where the cities included in such a comparison are homogeneous with respect to their average level of living, distribution of expenditures, and preference patterns, an intercity index of consumer price levels based on their average expenditure pattern would provide a reasonably good estimate of differences in living costs. Where the cities are heterogeneous, however, their average expenditure pattern would provide a less realistic basis for an intercity price index and, depending upon the degree of heterogenity, an increasingly poor estimate of differences in living costs.

BLS Approach to Measurement Problem

There is no single measure of intercity differences in living costs which will serve all purposes. In our judgment, however, the standard budget approach offers the best solution to a general purpose intercity index for 2 reasons. In the first place, this method of comparison makes it possible to hold constant the age, size, and composition of the family. Thus, variations in requirements associated with family needs are not confused with locality differences. Secondly, the level and manner of living represented by the standard can be held constant for each city in the comparison, even though the cities may be quite different with respect to their actual average levels of living, expenditure distributions, or preference patterns. At the same time, differences in the conditions of living in each locality over which individual families have no control, e.g., climate, transportation facilities, taxes, etc. can be reflected in the comparisons. Hence indexes based on a standard budget measure differences in living costs and not differences in prices only.

New Standard Budgets

In October of this year, BLS published the first of a series of new standard budgets which, when completed, will provide indexes of locality differences in living costs for 3 different living standards and for 2 different family types. The budget now available is for a moderate living standard for a family of 4 persons--an employed husband, age 38; a wife not employed outside the home; and two children, a girl age 8 and a boy 13. 2/ Cost estimates and intercity indexes based on autumn 1966 prices have been compiled for 39 metropolitan areas and for nonmetropolitan areas with populations from 2,500 to 50,000 in 4 regions. The U.S. urban average cost estimate has been used as the base of the intercity indexes.

A similar budget for a retired couple will be published early next year. Also underway are spring 1967 cost estimates and indexes for a lower and higher standard for both family types. Estimates of the cost of the moderate standard will also be made again as of spring 1967. Hence the 6 sets of cost estimates and indexes will be available for the same time period. It is expected that the budgets will be a continuing series, with costs and indexes for the spring of the year published periodically for the same 39 metropolitan areas and 4 regional classes of smaller cities as those included in the first study. Currently, there are no plans to extend the standard budget program to include other places or other types of families.

Locality Differences in Living Cost Components

Comparative living cost indexes based on the new City Worker's Family Budget for a Moderate Living Standard are shown in Table 1. Indexes have been shown separately for each of the budget components in which the comparison is for an equivalent, but not an identical, level of living-in other words, for those components in which both budget quantities (or weights) and prices may vary from city to city. Indexes are also shown separately for federal, and for State and local taxes. This component of the budget reflects not only variations in tax laws in different jurisdictions but also differences in the cost of all other budget components, since these costs constitute the base on which the tax allowances are calculated.

Indexes for the food-at-home component reflect both variation in prices and differences in regional food preference patterns used to calculate the cost of the nutritional standard within regions. Costs varied by as much as 30 percentage points and $530 between Honolulu, based on the regional preference patterns for the West, and the smaller cities in the South. Omitting Honolulu from the comparison, the range in costs was still sizable, amounting to 18 points between Hartford, reflecting Northeastern preference patterns, and the small cities in the South. 3/

Cost differences among cities within the same region reflect only differences in prices. In the West, the range in food prices was 12 percentage points, in the Northeast and North Central regions 7 points, and among cities in the South only 3 points. A special calculation of the cost of food at home using the U.S. food preference pattern in all cities indicated that food prices were highest in Seattle and lower by 12 percentage points in Green Bay, Wisconsin. Food prices in the Southern cities were very close to the U.S. urban average level. Hence it is the use of the Southern regional food preference pattern-and not the level of food prices--which is responsible for the generally lower costs of the food budget in cities in that area.

Indexes for shelter are based on a weighted cost for homes that are rented (25 percent), and homes which families are buying with mortgages contracted for in 1960 (75 percent). These weights were held constant for all areas in the comparison sicne both types of living arrangements are available in each community. However, separate costs were calculated for homes located in the central city and the suburban portions of each community, and the weighted area averages for shelter reflect these locality distributions. On this basis, Champaign-Urbana and San Francisco ranked highest in rental housing costs, Boston and New York in homeowner costs. Hartford and Chicago were among the 5 most expensive cities for both types of shelter arrangements.

Homeowner costs include principal and interest payments, taxes, insurance, fuel, and utilities. Fuel costs reflect variations in requirements resulting from differences in climate, as well as differences in price levels. However, no single one of these components of homeowner costs is responsible for the relative status of the area. For example, relatively high fuel requirements, coupled with high taxes, were responsible for Boston's status as the most expensive city for homeowners. In Chicago, and Cleveland, on the other hand, (6th and 7th ranking cities), costs were high because of the initial purchase price of the house and the subsequent principal and interest payments. The relative level of shelter costs for homeowners depends on a unique combination of the costs in each area for the various items included in the component.

With respect to transportation, Chicago, Philadelphia, and New York had lower costs--by 5 to 10 percentage points--than other cities because it was assumed that 1 in 5 families use public transportation exclusively in these areas. In other words, the weights for auto ownership were adjusted to reflect the greater accessibility of a mass transit system in these than in other areas. Although the same assumption was made for Boston, costs in that area were as high as U.S. urban average costs as a result of relatively higher price levels.

Indexes for clothing also reflect variations in requirements associated with difference in climate. Nevertheless, clothing costs were below the U.S. average in 3 of the 5 coldest areas, and above the U.S. average in 1 of the 5 warmest areas, as the factor of price combined with requirements to determine the level of costs in each area.

For the remaining components of family consumption--food away from home, housefurnishings, household operations, personal care, medical care, clothing materials and services, reading, recreation, education, tobacco, and alcoholic beverages, variations in costs reflect differences in price levels only. Indexes based on the sum of these components (shown in Column 7 of Table 1) indicate that price levels vary by less than 2 percent from the U.S. urban average in half (22) of the 43 areas studied. In cities on the West Coast, prices were from 5 to 10 percent higher, and the regional averages for small cities were from 5 to 10 percent lower, than the U.S. average.

In summary, then, the moderate living standard represented in the CWFB was adjusted in a number of ways to reflect an equivalent level of living in the areas in which the budget was priced. In consequence, intercity comparisons based on the total cost of the budget reflect differences in living costs, and not simply differences in prices. It should also be noted that the relative cost levels of the budget are for established families. The indexes do not reflect differences in living costs associated with moving from one area to another, or costs for recent in-migrants.

Indexes Based on Total Budget Costs

Intercity indexes based on the budget for a lower standard, when they become available, will be more appropriate for use in relation to public assistance and income maintenance programs than the indexes based on the moderate living standard. Nevertheless, analysis of total budget costs for the new CWFB provides some insights into current differentials in living costs for urban areas in all size classes.

Indexes of relative costs for the total budget (U.S. urban average cost = 100) ranged from 85 in the smaller cities in the South to 122 in Honolulu--a spread of 37 points (Table 1). However, 27 of the 43 areas fall within a range of plus or minus 5 percent, or approximately $500, of the U.S. urban average cost of the budget ($9,191).

Among the 7 areas in which total budget costs exceeded the U.S. averages by more than this amount, 5 were large metropolitan areas, each with a population of a million or more in 1960: the New York-Northeastern New Jersey, San Francisco-Oakland, Boston, Buffalo, and Milwaukee areas. Indexes for 14 other areas in this same size class, however, fell within the 5 percent range; and in 3 large cities with populations of a million or more (Atlanta, Dallas, and Houston) costs were lower than the U.S. average by more than 5 percent. Hence differences in living costs are not a function of area size alone. This is also confirmed by the presence of 2 medium-sized cities--Hartford and Honolulu--among the 7 "high cost" areas in the country.

As with the large cities, living costs in the majority of medium-sized cities were concentrated in a relatively narrow range. Also, the 4 cities, of the 17 in this size class (with 50,000 to 1 million population), in which costs were more than 5 percent below the U.S. average were all located in the South--Nashville, Baton Rouge, Orlando, and Austin.

Regional averages for small cities (with populations from 2,500 to 50,000), conceal substantial variations in costs for shelter and smaller variations in food costs. For other budget components, only regional average--not individual city--prices were calculated. Hence nothing is known about the variability of prices or costs among the cities in this size class. On a regional basis, costs in the small cities were below U.S. urban average costs, by 15 and 7 percent in the South and North Central Regions respectively, but only 2-3 percent in the Northeast and West.

"High-and Low-Cost" Living Areas

Why are living costs higher or lower in some cities than in others? An examination of the 7 areas in which costs exceed the U.S. average by more than 5 percent reveals that no single component of family living is responsible for the relative status of the area. Honolulu is the highest ranking city because costs are higher there than in other areas for all major budget components except clothing, personal care, and medical care. But Honolulu is an exception. In other areas, the relative level of total costs results from a unique combination of component costs in each area.

For example, transportation and shelter--and particularly rental shelter costs were higher in Boston than in New York. But for the majority of budget components in which the locality differential was affected by price alone, costs were higher in the New York area than in Boston. This difference, coupled with somewhat higher State and local taxes, made New York the second-, and Boston the third-ranking area, based on total budget costs.

Hartford's price levels were also above Boston's; and food, transportation, and clothing costs were higher in the smaller than in the larger city. But lower shelter costs and the absence of State and local income taxes made total costs in Hartford lower than in Boston. The cost of food at home plus shelter in San Francisco was 12 percentage points below comparable costs in Boston, but higher price levels in the West Coast city for all other components narrowed the overall differential to 2 percentage points.

In Milwaukee, the sixth-ranking city, the costs of all components of family consumption except transportation were lower than in Chicago; but the cost of the total budget was higher in the Wisconsin city, where State and local taxes were third highest among all the areas in the study. Compared with Buffalo, the higher shelter costs in Milwaukee were more than offset by lower food costs in that mid-West city; but higher shelter in combination with higher taxes make Milwaukee slightly more expensive than Buffalo.

Among the 7 cities in which costs were below the U.S. average by more than 5 percent--all located in the South--costs were generally lower than in other areas for food at home, shelter, clothing, and State and local taxes. Transportation costs were not among the lowest in these areas, however. And for those components in which price level was the only factor affecting intercity differences, costs in 3 of the 7 cities--Houston, Dallas, and Atlanta--were approximately the same as the U.S. urban average.

Implications for the Definition of Poverty

What are the implications of these findings for the definition of poverty? Certainly the most obvious one is that a single dollar cost estimate of need, even for a narrowly defined family type, will not be equally representative of requirements in all urban places. However, there is no easy solution to the problem of reflecting actual requirements, short of adjusting the cost estimates on a city-by-city basis.

Furthermore, even if this were administratively feasible, some thought should be given to the circularity implicit in this approach. To some extent any system of living cost differentials will tend to perpetuate the relative standing of different communities. That is, so-called "high-cost" areas, which receive higher allowances, will tend to remain relatively "high-cost"; similarly "low-cost" areas receiving lower allowances will remain relatively "lower-cost". Difference in price levels is only one of a host of factors responsible for differences in living costs in different places. Among others are the long-term average income level in a community, its ethnic background, educational level, age distribution, geographical location, type of industrial development, etc. These factors, which determine the "cost of living" in a community, may also be causal in relation to the problem of poverty. Hence a family living below the poverty line established for a "low-cost" area may need relatively more--rather than less--

than a family living below the poverty line in a "high-cost" area, if the conditions that breed poverty in the "low cost" city are to be eliminated.

1/ One such scale, published by BLS in November 1960, is described in a Technical Note: Estimating Equivalent Incomes of Budget Costs by Family Type (see Technical Reference No. 8). The scale is based on the assumption that families spending the same proportion of income on food have attained equal levels of living. While the scale is useful in estimating equivalent costs of goods and services, or net income requirements after income taxes and occupational expenses, it cannot be applied to individual items or major components of budget costs. A revised equivalence scale, based on information from the Bureau's Survey of Consumer Expenditures, 1960-61, was issued in October 1967.

2/ USDL, BLS, "City Worker's Family Budget for a Moderate Living Standard, Autumn 1966," Bulletin No. 1570-1, USGPO, Washington, D.C. (40 pp.).

3/ Since Honolulu costs were significantly higher than those in the mainland cities for most categories of the budget, comparisons in the remainder of the paper have been limited to the 42 mainland areas.

TABLE 1. INDEXES OF COMPARATIVE LIVING COSTS BASED ON THE CITY WORKER'S FAMILY BUDGET 1/
Autumn 1966

(U.S Urban Average Cost = 100)

Area 2/	Total budget 3/	COST OF FAMILY CONSUMPTION: Total 4/	Food at home	Shelter 5/	Transportation 6/	Clothing	All other 7/	PERSONAL TAXES: Total	Federal	State and local
Honolulu, Hawaii	122	118	121	130	122	97	110	162	137	404
New York-Northeastern New Jersey	111	110	109	126	90	105	106	125	118	194
Boston, Mass.	110	110	110	130	100	100	100	120	117	153
Hartford, Conn.	109	110	110	120	112	103	104	103	114	---
San Francisco-Oakland, Calif.	108	107	100	111	110	106	110	108	112	68
Milwaukee, Wis.	106	103	95	118	102	99	100	131	109	341
Buffalo, N.Y.	106	104	103	109	108	105	100	117	109	201
Seattle-Everett, Wash.	105	107	104	105	113	108	108	98	108	---
Chicago, Ill.-Northwestern Indiana	103	105	101	120	95	101	102	96	105	9
Minneapolis-St. Paul, Minn.	103	100	97	105	102	99	98	129	105	363
Los Angeles-Long Beach, Calif.	103	103	95	98	107	104	110	100	104	60
Cedar Rapids, Iowa	103	102	97	105	103	102	101	111	104	184
Indianapolis, Ind.	102	102	98	106	109	103	100	103	103	103
Washington, D.C.-Md.-Va.	102	101	100	106	101	97	100	110	103	178
Champaign-Urbana, Ill.	102	103	99	116	97	101	100	93	102	---
San Diego, Calif.	101	101	92	100	110	102	105	98	102	57
Cleveland, Ohio	101	103	96	115	101	103	99	92	102	---
Portland, Maine	101	102	108	98	101	108	99	92	101	3
St. Louis, Mo.-Ill.	101	101	102	99	103	100	100	101	101	107
Denver, Colo.	100	100	99	99	106	104	101	102	100	116
Philadelphia, Pa.-N.J.	100	100	107	96	91	102	100	102	100	127
Kansas City, Mo.-Kans.	100	99	100	91	107	101	101	106	104	131
Green Bay, Wis.	99	96	94	94	101	100	97	117	98	301
Wichita, Kans.	98	98	101	92	104	99	99	102	97	145
Northeast, Nonmetropolitan 8/	98	98	104	95	101	97	93	96	97	95
Detroit, Mich.	98	99	98	93	100	102	102	90	96	32
Cincinnati, Ohio-Ky.-Ind.	98	98	98	98	102	100	95	96	96	98
West, Nonmetropolitan 8/	97	96	98	87	104	103	95	107	96	221
Bakersfield, Calif.	97	97	97	83	110	101	102	91	95	50
Pittsburgh, Pa.	97	97	103	87	97	100	99	96	95	109
Lancaster, Pa.	97	97	107	87	95	99	97	95	94	98
Baltimore, Md.	96	94	93	86	99	96	100	104	93	208
Dayton, Ohio	95	96	97	92	101	101	94	87	91	42
Durham, N.C.	95	93	92	89	99	95	95	103	91	221
Nashville, Tenn.	93	95	92	88	102	99	98	80	88	1
Baton Rouge, La.	93	94	95	83	110	91	97	85	88	56
North Central, Nonmetropolitan 8/	93	93	97	90	97	96	89	89	89	94
Dallas, Tex.	92	94	93	82	101	93	101	79	87	3
Atlanta, Ga.	92	92	94	76	101	95	100	84	86	61
Orlando, Fla.	92	93	92	85	102	92	97	78	86	---
Houston, Tex.	91	93	94	76	106	91	101	78	86	3
Austin, Tex.	87	89	93	70	99	92	95	72	79	3
South, Nonmetropolitan 8/	85	86	92	69	99	89	89	75	77	57

1/ The family consists of an employed husband, aged 38, a wife not employed outside the home, an 8-year-old girl, and a 13-year-old boy.
2/ Areas are ranked by the total budget cost level.
3/ The total includes the following components not shown separately: Gifts and contributions, life insurance, occupational expenses, social security, and disability payments.
4/ Includes costs of miscellaneous items not shown separately.
5/ The average costs of shelter were weighted by the following proportions: 25 percent for families living in rented dwellings, 75 percent for families living in owned homes.
6/ The average costs of automobile owners and nonowners were weighted by the following proportion of families: Boston, Chicago, New York, and Philadelphia, 80 percent for automobile owners, 20 percent for nonowners; Baltimore, Cleveland, Detroit, Los Angeles, Pittsburgh, San Francisco, St. Louis, and Washington, D.C., with populations of 1.4 million or more in 1960, 95 percent for automobile owners and 5 percent for nonowners; all other areas, 100 percent for automobile owners.
7/ Includes food away from home, housefurnishings, household operations, personal care, medical care, clothing materials and services, reading, recreation, education, tobacco, and alcoholic beverages. Variations in costs of these components reflect differences in price levels only.
8/ Places with populations of 2,500 to 50,000.

FORMULAS FOR INCOME MAINTENANCE: THEIR DISTRIBUTIONAL IMPACT*

Martin H. David, University of Wisconsin
Jane H. Leuthold, University of Illinois

1. INTRODUCTION

In 1965, $93 billion was spent by public and private organizations to provide income transfers to individuals and households in the United States. Of that amount $5.5 billion was transferred in the form of public assistance; $30.2 billion was transferred through social insurance programs; and $43.1 billion was transferred through other governmental programs. Private direct income payments from welfare agencies amounted to approximately $14.2 billion.[1] In spite of these transfers, which comprise 17.7 percent of personal income, an estimated 35 million individuals were poor according to the standards established by the Social Security Administration and the President's Office of Economic Opportunity.

During the past year there has been considerable discussion of programs seeking to fill the poverty-income gap of the poor, i.e., the difference between the actual income of poor families and what is required for a decent level of living. Among the programs proposed for accomplishing this end are negative rates taxation, guaranteed minimum incomes, and family allowances. All of these programs have certain features in common. They consist of a mathematical and impersonally administered formula of income transfer. The payment is determined by a rate of transfer applied against the income deficiency of the family. Eligibility is conditioned only on an income and/or asset test. Because of these common characteristics, programs of this type are known as formula-based income transfers.

One of the first formulas for income maintenance was proposed by Milton Friedman.[2] Under the Friedman Plan, the income grant is half of the unused Federal family tax exemptions and deductions. A family of four with no income would receive $1,500 (half of (a) four times the Federal exemption of $600 and (b) the minimum standard deduction of $200 plus four times $100 for each exemption). This plan is referred to below as the EX-MSD Plan.

A similar plan, but one not tied to the Federal tax system, is the Lampman-Green Plan.[3] Formula income maintenance is accomplished by a rate applied to the amount by which a poverty standard exceeds income for the family. Lampman and Green assumed that the poverty standard could be reasonably well approximated by $1,500 for the family head plus $500 for each dependent. A family of four with no income would receive $1,500 (1/2 . $3,000). This plan is referred to subsequently as the Income Gap Plan.

Several other formula-based transfer programs have been proposed, notably plans by Tobin[4] and by Schwartz and Theobald.[5] They differ primarily in the level proposed for the poverty standard and the rate structure. Another alternative would be a demogrant which would provide payments to both the poor and the rich. None of these programs will be treated separately in this paper since they are in principle equivalent to the plans discussed above.

Many questions concerning the potential success of formula income transfers remain unanswered because such programs have never been put into practice in the United States. This paper is an attempt to provide a method for analyzing the effectiveness of alternative formula-based programs and of judging their comparative costs. It consists of a series of simulation experiments in which a variety of formula income transfer programs were extended to a sample of poor families.

The sample data used in the simulation are a 3,396 unit cross-section sample compiled by the Survey Research Center at the University of Michigan.[6] The sample includes observations on a number of demographic and income variables for non-institutional households in the conterminous United States in 1959. Low-income families are represented about twice as frequently as non-low income families in this sample. In order to prevent bias in statements about the entire population, the sample is weighted, with low income families receiving the lower weights.

2. FUNDAMENTAL ISSUES RELATING TO FORMULA INCOME MAINTENANCE PLANS

The technique of this simulation involves computing the amounts of the formula income payments for each eligible unit in the sample. This is accomplished by evaluating the formulas shown in the Appendix of this paper for alternative parameter values. The amounts of the payments are sensitive to the parameters of the program: the <u>resource base</u>, the <u>standard of poverty</u>, the <u>receiving unit</u>, and the <u>rate structure</u>.

Under a formula transfer program, if resources of the unit are less than the poverty standard, the unit receives a formula payment. The payment may be proportional to the income deficiency of the unit or it may be graduated to the size of the income deficiency. The poverty standard, the resources, the rate of payment, and the unit to which the payment is made are critical dimensions of formula income maintenance plans. By altering these parameters and examining the distribution results, we are able to analyze the short-run effectiveness of various formulations of formula income transfer programs.

Chart 1 presents an overview of the plans discussed in succeeding sections.

We will discuss each of these dimensions briefly. Comments of others to date have focused on the level of a flat rate and the standard of poverty. We will show that the measure of resources, the receiving unit, and gradation of rates are also important policy issues.

Measure of Resources

The measure of resources used in determining eligibility and the amount of a formula income payment should reflect the capacity of the family to meet its subsistence needs. Among the measures suggested for purposes of formula transfer are adjusted gross income (i.e., income excluding transfers and similar to the Federal tax concept) and total money income (i.e., income including transfers but excluding income in kind). Both these measures of resources were used in this simulation.

The results of the simulation show that total money income is to be preferred to adjusted gross income as a measure of the resources of a unit. Adjusted gross income was first proposed as a measure of resources by Friedman in an attempt to link formula income payments to the Federal income tax. The plan that results is clearly undesirable since payments are distributed to many whose total money income exceeds adjusted gross income by substantial amounts of transfer income. Unless transfer payments are reduced dollar-for-dollar for the amount of formula payment, substantial spillover of payment to high total money income levels occurs (Table 1). Conversely, benefits are less concentrated on the extremely poor. A plan that uses total money income as a measure of family resources and has the same aggregative cost offers substantially larger payments to families whose total money resources are less than $1,500.

The Standard of Poverty

The standard of poverty is a parameter of the transfer system, just as personal exemptions are a parameter of our current tax system. In the following discussion the standard is based on family size. It is called the "poverty standard," although, we recognize that the standard is not identical with any poverty line or true measure of subsistence costs. It would be desirable for the standard to be correlated with the level of subsistence income, with allowances for departures from a "poverty line" where appropriate. Local variations in subsistence costs and economies of scale might imply a poverty standard that would be administratively awkward or would be an incentive to family actions directed solely toward obtaining maximum transfers.[7]

Substitution of total money income for adjusted gross income as a measure of resources does not eliminate difficulties with a plan that uses Federal tax definitions of exemptions and minimum standard deductions as the poverty standard (EX-MSD Plan simulated in Table 1). Table 2 shows that EX-MSD benefits families that are not poor according to a poverty standard proposed by Lampman and Green. That standard is remarkably close to the Orshansky poverty standards, considering its simplicity.[8] Whether spillover to the non-poor is a serious policy matter depends on whether it is considered important that about 3 percent of the aggregate cost would be paid to the non-poor. This payment would go largely to families barely out of poverty (Table 2).

The Lampman-Green poverty standard for the one-person family with no income equals $1,500. This exceeds the value of unused exemptions and deductions. However, each additional family member increases unused exemptions and deductions by $700, while Lampman and Green assume additional subsistence cost at $500. For families of five or more persons unused exemptions and deductions exceed the Lampman-Green standard.

Table 3 contrasts the mean formula income maintenance payment for equal cost plans based on these two standards. The plan based on unused exemptions and deductions (EX-MSD) is based on a 25 percent transfer rate. The plan based on the Lampman-Green standard (Income Gap Plan) is based on an equal cost, 28.5 percent flat rate. As would be expected, the mean payment under EX-MSD exceeds the mean payment under the Income Gap Plan for families of five or more persons.

Payments from the EX-MSD Plan exceed the Income Gap payments for families whose head is over 65 years as a result of the additional income tax exemption currently available to any individual of that age (Table 4).

The Receiving Unit

To minimize the cost of a formula transfer program it would be wise to take into account the income in kind that is received by poor persons who live with relatives "doubled up" in the same household. For this reason it would be natural to apply a formula transfer to the aggregate income of all persons in a family.

Inequities and administrative difficulties could result. Families that undertake to support ailing and indigent relatives in their own homes might not obtain a formula payment, while families that support a relative in another household might still be able to obtain formula transfers for the relative.[9] In addition, the administrators might be plagued by frequent changes in family composition, with the resulting changes in the level of allowable formula transfers.

Another major problem associated with a family unit plan is that it may lead to family fragmentation. If benefits paid to small families are based on a higher per capita transfer than those granted to large families, a family unit plan may be an incentive for families to break up and file several applications for formula income maintenance.[10]

These problems could be solved by using an adult unit (a person eighteen or over, his spouse if he is married, and any children under eighteen who live with him and for whom he is responsible) as the basis for computing formula transfers. However, a plan based on the adult unit as the receiving unit is considerably more expensive than a comparable plan based on the family unit. In other words, a family unit plan entitles recipients to a higher rate of transfer than a comparable adult unit plan of equal cost (Table 5).

The Rate Structure

Figure 1 illustrates three possible patterns of gradation of the rate structure for a family of four members whose poverty standard is $3,000. All plans lead to the same payment to units with no resources. Plans B and C are graduated rate structures. Plan B pays greater benefits to the marginal poor than the flat rate plan. Plan C approaches the benefit level of the flat rate plan only for the "poorest" poor.

Of the three plans, Plan B is the most expensive, Plan C is the least expensive. The schedule of graduated rates used with Plan B is a function of the ratio of the poverty-income gap to the poverty standard. If the family's poverty-income gap is less than one-third of its poverty standard, any increment in resources reduces the formula payment by .75 of the increment. If the ratio is greater than 1/3 but less than 2/3, the formula payment is reduced by half of the movement. Finally if the poverty-income gap is more than 2/3 of the standard, the formula payment is reduced by .25 of any increment resources. As a result, the family with no resources receives 50 percent of the poverty-income gap as a formula transfer under Plan B.

Plan C is the mirror image of Plan B. Increases in resources under this plan lead to a reduction in the formula transfer at rates of .25 and .75 as the ratio of the poverty-income gap to the poverty standard increases from less than one-third to more than two-thirds. Like Plan B, a unit with no resources receives 50 percent of the poverty-income gap as a formula transfer.

Plans B and C are compared in Table 6. Each is also compared to a flat rate plan of equal cost. For families with incomes of less than $1,000, the mean income gap payment from the graduated rate (Plan B) is less than the payment from an equal cost, flat rate plan (Plan A). This situation is reversed for families with income greater than $1,000. Comparison of Plan C and its equal cost, flat rate equivalent (Plan D), shows the reverse situation.

Table 7 compares mean payments under the four plans for <u>adult units</u> of different sizes. The aggregate cost is more than in Table 6, in spite of the fact that adult units contain fewer persons than family units. This finding reflects the fact that many poor adult units live with a unit that is not poor. The resources of the family as a whole are adequate, while those of the dependent are not.[11]

The distribution of formula payments both by adult unit size and by life cycle indicate that large units benefit the most in absolute dollar amounts from the gradation proposed in Plan B. The least benefits go to the older couple and single person (see Table 8).

It is likely that any work effort changes resulting from formula payments will depend on the effective marginal income that an individual can obtain from additional work. The marginal income can be expressed as

$$w\ (1 - p - r)$$

where p is the rate of payroll taxes and r is the transfer rate. The larger r, the smaller the return to additional effort and the greater the probability that the plan will cause a shift in the labor supply function.

To the extent that changes in work effort arise from high rates r the three rate structures differ substantially. Plan B creates the greatest incentive to alter work effort for the marginal poor. Whether such incentives prove a serious problem depends on the degree of labor force attachment of such persons and the latitude for absenteeism, short hours, and discretionary overtime in their plan of employment. At the same time, Plan B offers the least incentive to change work habits to those with no income. Whether that is desirable depends on the likelihood that persons with no income from other sources could be pulled into employment under any circumstances. By graduating the rate structure, changes in work effort can be concentrated on those who are already earning income (as in Plan B) or on those who are not in the labor market at all (as in Plan C).

Some insight into the disincentive issue can be obtained by examining the reported labor force status of the poor (Table 9). Among the poor, 41 percent are employed and 10 percent are unemployed. More than a third of the poor do not consider themselves in a position to work even when no formula income maintenance plan is available. These non-labor force poor include two disparate populations - persons who subsist on their own resources, a small minority who receive assistance payments. For the former, introduction of a formula income maintenance program may reduce the incentive to search for work, an incentive that is already too blunt to bring these adult unit heads into the labor market. For those on assistance, introduction of formula maintenance will provide a positive force to seek work. The effect of such incentives on the labor force participation of these non-labor force groups remains an open question.

For those in the labor force it is unclear how much latitude for the expression of such incentives to change work habits exists under present employment practices. To what extent working habits and desires can be modified by a promise of support at less than the margin of subsistence is again an open question.

Preliminary work by one of the authors using a work-leisure choice model indicates that changes in work effort resulting from a formula transfer program would be minimal. For certain workers in large families or with low wage rates, however, the change in work effort could be substantial. Heads of adult units and spouses with fewer than two children tend to increase hours worked, while spouses with two or more children tend to decrease hours worked for a change in the rate of formula transfer.[12]

3. INTERPRETATION AND SUMMARY

The simulations show clearly that a formula-based income maintenance plan can provide aid to groups that are difficult to locate through categorical programs. The employed poor, the educated poor, and poor with large families and little earning power will all receive benefits.

Any deviation from a constant per capita standard produces a concentration of formula-based income maintenance payments in that direction. For that reason the Lampman-Green poverty line formula gives greater benefits to small families than does the EX-MSD Plan, while the EX-MSD Plan provides greater benefits to the aged.

Second, any plan that places no ceiling on the poverty standard results in extremely high payments to a few large families. This may be desirable, but only if the poverty standard is an acceptable gauge of the need of those large families and if the measure of resources truly reflects their inability to purchase subsistence. If the poverty standard departs from a subsistence level, the resulting formula payment will be a windfall to the large family. This appears to be the case when the poverty standard is based on exemptions and minimum standard deductions. Similarly, if adjusted gross income is used as the measure of the family's resources there will be a few who benefit by large formula-based payments in spite of the fact that their total resources exceed the poverty lines (see Table 1).

Any plan that provides benefits on a standard that deviates from a true subsistence line will give some aid to the near-poor. However, such spillover of benefits may be associated with greater administrative simplicity, reduction of disincentives, and greater acceptability of the formula-based plan. In addition, if the rate of transfer is low, the aggregate amounts paid to the non-poor may not be large (see Table 2).

We anticipated that gradation of transfers could be used to concentrate benefits at various levels of poverty. A plan that focuses on the extreme poor will cost less than a flat rate plan that provides the same benefit at a zero level of income. Conversely, a plan that provides the greatest benefits to the marginal poor costs more than the flat rate plan that provides equal benefits to those with no resources. This latter plan has some interesting anticipated consequences, however. Large families with spouse and young children appear to benefit most. This may be socially desirable. Unfortunately, the plan does imply substantial discontinuities in the rate of taxation of additional income just above and just below the poverty line. Those in extreme poverty are taxed at a low rate on any increments to their earnings.

The simulation indicates clearly that substantial additional costs are associated with use of the adult unit as the unit over which benefits are calculated (see Table 5). The cost could possibly be reduced by imputing income to those who share living arrangements with others. The simulation results presented show true costs only if families do not respond to the value of "transfer splitting" that results from large initial payments to the first member of a household and smaller payments to succeeding members. To the extent that families do respond to that incentive, costs will move to the same level as was simulated for adult units. As we have not incorporated available evidence on undoubling of families in response to income, policy makers will need to judge whether the savings in costs are worth the inequity that results from some families receiving greater benefits than others merely because they are willing and able to rearrange their housing.[13]

The cost and inequity spillover to the non-poor of a program based on adjusted gross income must also be weighed subjectively against the likely effect of alternative rates of transfer on work effort. This simulation provides only either a dollar measure of the difference in cost between two programs using the same rate and different measures of resources, or, alternatively, the difference in rates required for equal cost programs.

Lastly, the results constitute food for thought on the desirability of graduating rates. Arguments can be adduced for either lower-than-average rates to the extreme poor or higher-than-average rates to the extreme poor. The likely work effort effects of grants at different levels of poverty would appear to be an important consideration in the choice of gradations; again we can offer no solution but can illustrate the distributional impact of benefits under whatever program is desired.

The results of the simulation are crude for several reasons. No allowance is made for the response of the poor and the near-poor to a large increase in transfers. No incentives to increase or decrease work effort or family size are incorporated. No effort is made to forecast the response of the state welfare administrations to an income that would be paid directly to the poor by the Federal Government. We view the inclusion of such responses as an important sequel to the present computations. Incentive effects and the accommodation in the existing public transfer programs to formula-based income maintenance cannot be quantitatively appraised on the basis of the present study. When more is known, behavior of poor families and administrators could be added to the present simulation to give better insight into the reactions that may be triggered by a new program of income maintenance.[14]

Summary of Distributional Effects

Table 10 summarizes several aspects of the formula income maintenance payments simulated. The distribution of such payments according to the extent of the income deficiency of the adult unit is shown separately for units headed by an employed person and for all others. Differences in the distribution of payments among the poor and the spillover to the non-poor are indicated in columns 3 and 4 of the table. Columns 5 and 6 provide estimates of the Federal taxes paid by the poor. (Income taxes were simulated without a minimum standard deduction option, per 1959 law, which partially accounts for the positive tax liabilities for units with a poverty-income gap.) The mean social security benefits reported by adult units give some indication of the extent to which social insurance aids the poor, while the last column shows the amount of money income to which formula income payments would be added.

Among the employed one can infer that a poverty gap beyond \$500 results from increasing requirements rather than from decreasing resources. Among the non-employed a somewhat greater drop in income occurs as the poverty gap rises to \$2,000, suggesting a combination of more mouths to feed and fewer resources. Clearly social security plays a major role in maintaining income levels for the small non-employed family. Equally clear, a program of modest cost and low rates of transfer will not eliminate income deficiencies, nor will it obviate the need for support from existing transfer programs.

APPENDIX

Mathematics of the Formula Payment Program

Notation:

- N = amount of formula payment
- t = transfer rate
- Y = resources
- B = poverty standard
- S = family size
- E = earnings
- R = transfer income
- A = annuity value of assets
- X = tax liability
- D = disposable income
- a, b, c are constant

Identities:

$$Y_1 = E + R + A, \quad Y_2 = E + R, \quad Y_3 = F \qquad (1)$$

$$D_{ijk} = Y_2 + N_{ijk} - X \qquad (2)$$

For all programs

$$N_{ijk} = \begin{cases} t_k (B_j - Y_i) & \text{if } B > Y \\ 0 & \text{if } B < Y \end{cases} \qquad (3)$$

where $i = 1, 2, 3$; $j = 1, 2$; $k = 1, 2$.

The subscript i refers to alternative income concepts; j refers to alternative poverty standards; k refers to alternative rate schedules for the income maintenance payment.

For both the EX-MSD and the income gap plan

$$B_j = B_j (S) = a_j + b_j S \qquad j = 1, 2 \qquad (4)$$

For a plan with graduated rates

$$t_1 = t_1 (B_j - Y_i) \qquad (5)$$

where $j = 1, 2$; $i = 1, 2, 3$.

Otherwise a flat rate plan

$$t_2 = c \qquad (6)$$

Some insight into disincentives can be obtained by taking derivatives of N_{ijk} with respect to Y_i and differences with respect to family size S.

For example,

$$\frac{\partial D_{2j2}}{\partial E} = \frac{\partial D_{2j2}}{\partial R} = 1 - t$$

or disposable income increases by only a fraction of earnings or categorical assistance payments.

Given the form of B_j, if $a_j \neq 0$, then it is clear that dissolution of a family of S members into two sub-families sizes S_1 and $S - S_1$ will be advantageous. The family payment will be

$$N^{(f)}_{ijk} = t_k (2a_j + b_j S - Y_i)$$

If a_j is sufficiently large the difference between $N^{(s)}$ and $N^{(f)}$ may induce family dissolution. However, if the formula transfer formula recognizes S as the appropriate administrative unit the form of living arrangement will not affect the amount of the payment. $N^{(s)}$ will be paid in any case.[19]

NOTES

* The research reported here was supported by funds from the National Science Foundation and by funds granted to the Institute for Research on Poverty at the University of Wisconsin by the Office of Economic Opportunity pursuant to the provisions of the Economic Opportunity Act of 1964. The facilities of the Social Systems Research Institute and the helpful comments of H. W. Watts, G. H. Orcutt, and R. J. Lampman have assisted in the conception and execution of this study. Thae Soo Park assisted ably in the compilation of the tables. Errors, value judgments, and conclusions are the authors' and do not reflect views of the sponsors or our colleagues.

[1]"Current Operating Statistics," Social Security Bulletin, Vol. 29, No. 3 (March 1966), pp. 279-301.

[2]Milton Friedman, Capitalism and Freedom (Chicago: University of Chicago Press, 1962), pp. 191-194.

[3]Robert J. Lampman, "Prognosis for Poverty," National Tax Association, Proceedings of 57th Annual Conference (Pittsburgh, September, 1964), pp. 71-81; Christopher Green, Transfer-by-Taxation: An Approach to Improved Income Maintenance, a conference monograph prepared for the Brookings Institution Studies in Government Finance (June 9-10, 1966).

[4]James Tobin, "Improving the Economic Status of the Negro," Daedalus, Vol. 94 (Fall, 1965), pp. 889-895.

[5]Edward Schwartz, "A Way to End the Means Test," Social Work, Vol. 9 (July 1964), pp. 3-12; and Robert Theobald, Free Men and Free Markets (New York: C. N. Potter, 1963), pp. 192-197.

[6]The sample is thoroughly analyzed in J. N. Morgan, M. H. David, W. J. Cohen, and H. E. Brazer, Income and Welfare in the United States (New York: McGraw-Hill, 1962).

[7]See Harold W. Watts, "The Iso-Prop Index: An Approach to the Determination of Differential Poverty Income Thresholds," The Journal of Human Resources, Vol. II, No. 1 (Winter, 1967), pp. 3-18.

[8]Mollie Orshansky, "Counting the Poor: Another Look at the Poverty Profile," Social Security Bulletin, Vol. 28, No. 1 (January, 1965), pp. 3-26.

[9]This would not be technically possible if support payments and income-in-kind received by the dependent were fully reported and included in total money income. It is unlikely that such reporting could be easily enforced.

[10]The problem is identical to the income-splitting problem under current Federal tax law. For an excellent discussion see L. Johansen, Public Economics (Chicago: Rand McNally, 1966), pp. 281-282.

[11]Some qualifications are in order. The dependent and the supporting units may prefer doubling up to other housing arrangements. For example, unmarried sisters may prefer living together. One assumes responsibilities for keeping house and the other earns income. In that case lack of income earned by the "dependent" sister does not necessarily imply inability to maintain herself out of poverty. In the data used here some imputed income from food and housing has been assigned to the dependent adult unit. However, the amount assigned is not included in the total money income measure as it could not easily be used as the basis for a formula transfer.

[12]Possibilities arise for increases as well as decreases in work effort in response to changes in the transfer rate. One of the authors is in fact attempting estimates of potential changes in work effort in the context of a work-leisure choice model. See Jane H. Leuthold, "Formula Income Transfers and the Work Decision of the Poor: An Application of a Work-Leisure Choice Model," unpublished (forthcoming) Ph.D. Diss., University of Wisconsin, 1967.

[13]Some estimates of these disincentives are included in Morgan et al., Income and Welfare in the United States, Chapter 14.

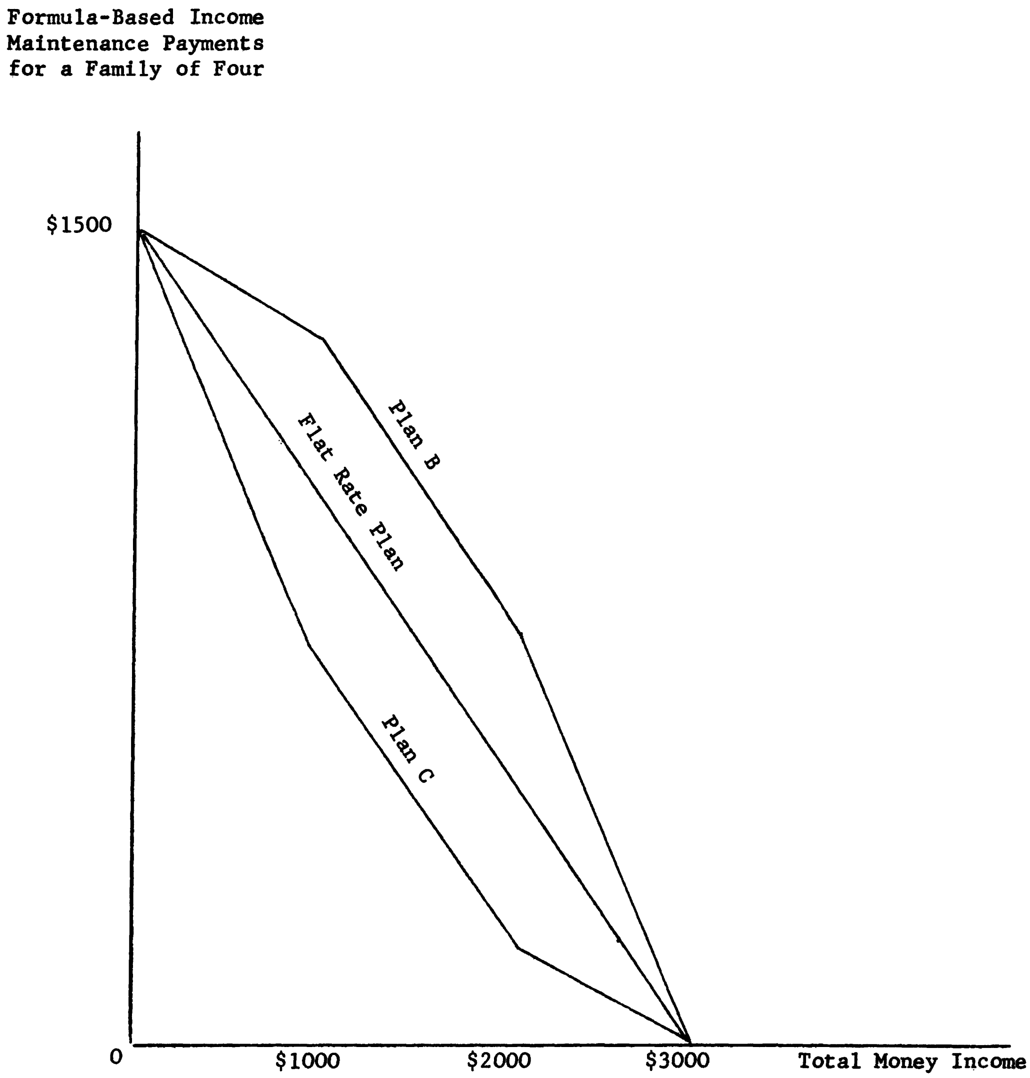
[14]See, for example, M. David, "Design of Simulation Models of the Household Sector," Madison, Wisconsin: Social Systems Research Institute, Household and Labor Market Workshop Paper 6503, August, 1965 (presented before the First World Congress of the Econometric Society, September, 1965).

Chart 1

Formula Income Maintenance Plans

Dimension of the Formula	Options Simulated
A. Resources	A 1 Adjusted Gross Income (excluding transfers and similar to the Federal tax concept)
	A 2 Total Money Income (including transfers, excluding income in kind)
B. Standard of Poverty	B 1 EX-MSD (Friedman-type)
	B 2 Poverty Income Gap (Lampman-type)
C. Receiving Unit	C 1 Families (related individuals occupying a dwelling unit)
	C 2 Adult Units (individuals 18 years of age or older, their spouse, and children under 18)
D. Rate	D 1 Flat rate
	D 2 Graduated rate, decreasing with increases in the income deficiency (Plan B below)
	D 3 Graduated rate, increasing with increases in the income deficiency (Plan C below)

Figure 1



Alternative Income Gap Plans
for Formula Income Maintenance

Table 1

Simulated Formula Income Maintenance Payment to Families Under an EX-MSD Plan: Adjusted Gross Income Compared to Total Money Income as a Measure of Resources within Total Money Income, 1959

Total Money Income	Percent of Family Units	Mean Amount of Payment TMI-Based Plan 43% rate	AGI-Based Plan 25% rate
Negative, zero	1%	$648	$377
1 - 600	3	583	397
601 - 1000	5	530	433
1001 - 1500	6	437	406
1501 - 2000	6	323	361
2001 - 2500	5	342	340
2501 - 3000	5	230	284
3001 - 3500	4	165	170
3501 - 4000	5	88	125
4001 - 4500	5	57	82
4501 - 5000	5	24	36
5001 - 6000	11	12	34
6001 - 7000	9	2	9
7001 - 8000	8	0	2
8001 - 9000	6	0	4
9001 - 10000	4	0	0
Over 10,000	13	0	2
Total, Average	100%	$139	$138
Number of families	2800**		
Aggregate Cost (billions)		$7.4***	$7.4

** The payments are computed according to formulas shown in the Appendix and are applied to a representative stratified sample of U.S. families taken in 1960.

*** Aggregate cost computed by multiplying mean payment by total number of families ($139 x 53.4 mil).

Table 2

Simulated Formula Income Maintenance Payments under an EX-MSD Plan with Resources Measured by Total Money Income: Comparison of Payments to Poor and Non-Poor Families within Total Money Income 1959

Total Money Income	Mean Amount of Payment* Poor**	Non-poor
Negative, zero	$377	$ 0
1 - 600	339	0
601 - 1000	308	0
1001 - 1500	258	0
1501 - 2000	274	***
2001 - 2500	357	36
2501 - 3000	348	17
3001 - 3500	342	3
3501 - 4000	275	1
4001 - 4500	526	2
4501 - 5000	424	2
5001 - 6000	300	6
6001 - 7000	0	2
Over $7,000	0	0
Average	$308	$ 4
Percent of all families	25%	75%

* EX-MSD Plan, total money income base, family unit, 25% rate.

** Poor in the sense that $1,000 + 500S > Y_2$ where S is family size, Y_2 a measure of its resources. See Appendix.

*** Less than $1

Table 3

Simulated Formula Income Maintenance to Poor Families
EX-MSD Plan Compared to a Poverty Income Gap Plan
within Family Size Classes

Size of Family	Mean Amount of Payment: EX-MSD Plan, 25% Rate b/	Mean Amount of Payment: Income Gap Plan, 28.5% Rate b/	Percent of Poor Families	Incidence of Poverty a/
1	$ 131	$207	28%	43%
2	227	208	22	21
3	248	277	10	15
4	324	353	12	17
5	385	365	8	20
6	461	405	9	39
7	718	618	4	35
8	649	481	3	53
9	862	708	3	63
10 or more	1177	686	1	77
Average, Total	$308	$308	100%	25%
Aggregate cost (billions)	$ 4.1 c/	$ 4.1 c/		

a/ Ratio of the number of poor families to the total number of families with this characteristic.

b/ Resources were measured by total money income under both plans.

c/ Aggregate cost computed by multiplying mean payment by total number of poor families ($308 x 19.35 mil).

Table 4

Simulated Formula Income Payments to Poor Families:
EX-MSD Plan Compared to a Poverty Income Gap Plan
within Classes Based on Age of Head

Age of Family Head	Mean Amount of Payment		Percent of Poor Families	Incidence of Poverty
	EX-MSD Plan 25% Rate a/	Income Gap Plan, 28.5% Rate a/		
0 - 24	$166	$258	7	28%
24 - 34	398	392	14	19
35 - 44	448	431	17	18
45 - 54	309	330	19	23
55 - 64	175	257	18	27
65 - 74	273	194	15	39
74 - over	$329	$243	11%	65%
All	$308	$308	100%	25%
Aggregate cost (billions)	$4.1	$4.1		

a/ Resources were measured by total money income under both plans.

Table 5

Aggregate Expenditures and Rates of Transfer for Various Income Maintenance Plans, 1959 1/

Plan Description	Form of Plan Comparison: Equal Costs, Rate of Transfer	Equal Costs, Amount (billion)	Equal Rates, Rate of Transfer	Equal Rates, Amount (billion)	Equal Payments 2/, Rate of Transfer	Equal Payments 2/, Amount (billion)
EX-MSD Plan						
Adult Unit	19%	$4.3	25%	$5.6	25%	$5.6
Family Unit	25	4.3	25	4.3	25	4.3
Poverty Income Gap Plan						
Adult Unit	18%	$4.3	25%	$5.9	23%	$5.5
Family Unit	29.5	4.3	25	3.7	28.5	4.1

1/ The aggregate base to which these rates apply varies according to the unit to which the plan is administered. Thus a 25 percent rate applied to unused exemptions and deductions of family units results in a different payment than 25 percent applied to the corresponding unused exemptions and deductions of adult units (see the Appendix for the formulas used).

2/ The differences in amounts under equal payments are due to the fact that under EX-MSD Plan some non-poor are also eligible to receive income payment.

Table 6

Simulated Formula Income Payments to Poor Families under an Income Gap Plan: A Comparison of Flat and Graduated Rate Plans of Equal Revenue Cost within Total Money Income, 1959*

	Mean Amount of Payment					
Total Money Income	Flat 65% Rate Plan A	Graduated Rate Plan B	Graduated Rate Plan C	Flat 35% Rate Plan D	Percent of Poor Families	Incidence of Poverty
Less than 0	$1163	$895	$895	$626	2%	100%
0 - 600	982	851	660	529	12	100
601 - 1000	817	782	476	440	19	100
1001 - 1500	591	608	302	318	23	98
1501 - 2000	554	586	267	298	16	68
2001 - 2500	761	816	355	410	11	51
2501 - 3000	668	736	292	360	7	35
3001 - 3500	585	663	237	315	5	27
3501 - 4000	342	394	132	184	3	18
4001 - 6000	399	460	153	215	2	10
Over 6000	0	0	0	0	0	0
Ave., Total	$ 701	$698	$381	$378	100%	25%
Aggregate cost ** (billions)	$ 9.4	$9.3	$5.0	$5.1		

* Resources were measured by total money income.

** Discrepancies due to rounding transfer rate.

Table 7

Simulated Formula Income Payments to Poor Adult Units under an Income Gap Plan: A Comparison of Flat and Graduated Rate Plans of Equal Cost within Adult Unit Size 1959*

	Mean Amount of Payment					
Size of Unit	Flat 60%** Rate Plan A	Graduated Rate Plan B	Graduated Rate Plan C	Flat 40% Rate Plan D	Percent of Poor Adult Units	Incidence of Poverty
1	$ 583	$ 554	$ 419	$ 389	54%	54%
2	507	536	309	338	15	20
3	658	679	418	439	8	19
4	779	836	462	519	8	21
5	822	890	480	548	6	23
6	951	1023	561	634	4	36
7	1138	1244	653	758	2	38
8	1182	1312	658	788	1	46
9	1785	1855	1120	1190	2	72
10	***	***	***	***	***	***
Ave., Total	$ 664	$ 671	$ 436	$ 442	100%	33%
Aggregate cost (billions)	$ 14.2	$ 14.3	$ 9.3	$ 9.4		

* Resources were measured by total money income.

** Rate attached to this plan is lower than that illustrated in the previous table because this plan applies to the adult rather than family unit.

*** Insufficient observations for a reliable estimate.

Table 8

Simulated Formula Income Payments to Poor Adult Units under an Income Gap Plan: A Comparison of Flat and Graduated Rate Plans of Equal Cost within Life Cycle, 1959*

Life Cycle	Plan A Flat 60% Rate	Plan B Graduated Rate	Percent of Adult Units	Incidence of Poverty
1. No spouse present, no children, under 45	$650	$599	22%	50%
2. Married, spouse present, no children, wife under 45	441	475	2	11
3. Married, spouse present, children, some under 6, wife under 45	885	961	14	21
4. Married, spouse present, children, none under 6, wife under 45	788	860	4	13
5. Married, spouse present, children, some under 6, wife 45 or older 6. Married, spouse present, children, none under 6, wife 45 or older	842	892	6	30
7. Married, spouse present, no children, wife 45 or older	477	517	10	20
8. No spouse present, no children, 45 or older	537	523	32	57
9. No spouse present, but children	870	884	10	60
Average, Total	$664	$671	100%	33%
Aggregate cost (billion)	$14.2	$14.3		

* Resources were measured by total money income.

Table 9

Distribution and Incidence of Poverty among Adult Units
by Labor Force Status of the Head

Labor Force Status of the Adult Unit Head	Percent of Adult Units		Incidence of Poverty
	Poor	All	
Employed	41%	71%	19%
Unemployed	10	6	60
Retired	14	10	49
Student	9	4	32
Housewife	20	8	79
Other	6	1	--
Total, average	100%	100%	33%

Table 10

Simulated Formula Income Payment to Poor Adult Units: Mean Poverty Income Gap, Mean Payment under Income Gap Plan, Social Security Tax and Benefit Federal Income Tax Liability, and Disposable Income within Labor Force Status and Size of Poverty Income Gap 1959

Labor Force Status of Adult Unit Head 2/	Poverty Income Gap	Distribution of Income Gap (percent)	Amount of Payments: EX-MSD Plan (25% Rate)	Amount of Payments: Income Gap Plan (23% Rate)	Average Social Security Tax	Average Federal Tax Liability	Average Social Security Benefit	Total Money Income less Estimated Federal Taxes 3/
Employed	$ 0	81%	$ 1	$ 0	$104	$783	$22	$6315
	1- 500	5	47	57	38	8	68	2029
	501-1000	5	139	185	26	1	20	1430
	1001-2000	6	342	331	22	0	16	1292
	over 2000	3	838	664	20	0	11	1325
Mean (Employed)	$219	100%	$ 53	$ 50	$ 89	$632	$24	$5389
All Others 1/	$ 0	33%	$ 11	$ 0	$ 35	$249	$523	$3988
	1- 500	12	93	39	7	4	463	1494
	501-1000	18	182	177	4	1	245	927
	1001-2000	35	287	323	3	0	49	325
	over 2000	3	691	593	8	0	14	770
Mean (all others)	$730	100%	$167	$168	$ 15	$ 82	$288	$1790
Aggregate cost			$5.6	$5.5				

1/ Include unemployed, retired, student, housewife, never worked, disabled and not working, and status not ascertained.

2/ At time of interview in March and April, 1960.

3/ Total money income less estimated Federal income and payroll taxes.

DISCUSSION

Benjamin A. Okner*

I think it would be most fruitful to bring these three interesting analyses under the common umbrella of the session title; however, I would like to amend it slightly to "More Light on the Definition of Poverty: What Policy Implications Emerge?" This will give me the opportunity to keep my remarks general and to express some of my own thoughts on this subject. In addition, although this session is sponsored by the American Statistical Association, I assume that most of us are not interested in the poverty definition merely for the sake of "counting heads," but rather because we want to know what policy implications emerge from using different definitions. Although the papers presented did include some discussion on the effect of such changes on policies for alleviating poverty, I feel that this is an area that deserves continual emphasis.

There has certainly been no lack of prior discussion concerning the HEW-OEO poverty definition and the various ways in which it might be modified or refined. Some of the most commonly heard criticisms of this present definition include the omission of any consideration of the family's assets, the use of annual money income rather than income measured over some longer period of time, and geographical differences which influence the cost of achieving some minimum subsistence level of living. Interestingly, the Lamale-Brackett paper which deals solely with the last point indicates that there is relatively little geographic variation in basic living costs. If their data had permitted intra-city analyses of maintaining a given budget level as well as the inter-city comparisons, I would not be surprised to find as much variation (or even more) within the cities they investigated as there is between the cities.

Again, in the paper by Watts, the results did indicate differences in the number of persons defined as poor depending on the definition used; but, the differences found do not seem significantly different from the number as defined under the present HEW-OEO poverty thresholds. I suspect there would be even less variation if the comparison were made with the actual poverty threshold levels rather than with the "$1,500 plus $500 per additional family member" approximation that was used by Watts. However, it is important to note that the composition of the poverty population varies as the number of factors in the definition changes. Such variation could support quite different programs for fighting poverty (e.g., greater emphasis on Head Start rather than on adult job training) and such ambiguity might lead to increased confusion among policymakers.

*The views expressed in this paper are those of the author and do not purport to represent the views of the other staff members, officers, or trustees of The Brookings Institution.

Regardless of how <u>any</u> definition might be refined or extended, there can be little argument concerning the fact that the poor are lacking in income -- or more generally, they lack the means to command sufficient resources to maintain a minimum subsistence level of living. And, regardless of how defined, we can't deny the fact that in our "affluent society" there are a large -- too large -- number of such persons.

For policy purposes, it is not sufficient to define a person or family as "poor" if they lack sufficient income (despite the fact that this is a definition of poverty with which it would be difficult to disagree). Rather than expending large amounts of resources in attempting to refine, extend, or further clarify the poverty definition, I think it would be more useful to accept the fact that there are some 25-30 million persons in this country who are poor -- and would still be classified as poor under any alternative definition -- and devote more time to the questions of why we have such a large poverty population and the most effective means to reduce it.

I think it is fruitful to consider various means of alleviating poverty in the context of both short-run and long-run policies. For the short run, we must develop programs which will provide the poor with sufficient income to maintain at least minimum living levels. Although we already have a large number of income maintenance programs, those specifically directed to the poor consist primarily of the various categorical public assistance programs authorized under the Social Security Act. Although their reasons differ, there is almost unanimous agreement -- both among liberals and conservatives -- that our present programs are seriously deficient and need substantial overhauling. The disagreement, of course, concerns the direction of overhaul and the specifics of what is to be done either to or for our present welfare system. Any critic of public assistance could give you a long list of present deficiencies, but somewhere near the top of the list there would undoubtedly be strong sentiment against the present categorical nature of these plans. Essentially, we now single out certain groups or categories of the poor -- the aged, the blind, the permanently and totally disabled, and families with dependent children -- as being worthy of financial aid from society. In most areas of the country there is very limited assistance available for a poor family which doesn't fit into one of these pigeonholes.

I am somewhat disturbed about adopting the kinds of refinements in the poverty definition that have been discussed because of the implication that it would be desirable to incorporate such adjustments in any new income maintenance plans which might be developed. In light of the evidence presented, the adjustments suggested would be picayune and any <u>possible</u> gains in equity would be overwhelmed by increases in administrative inefficiency. Hopefully, we should learn from our past categorical assistance experience and avoid making the same kinds of mistakes in new programs. It would be easy to predict other unpleasant consequences of such action but rather than pursuing these, let us consider the other side of the coin. What would be involved if such adjustments were not made, for example by city, and there actually are substantial differences in the cost of maintaining a minimum living level? Since none of the income maintenance plans currently being discussed (for example, as in the David and Leuthold paper) is overly generous with the taxpayers' funds, the "dire consequence" of using a uniform rather than an adjusted basis for making payments might be to give a poor person living in a southern city \$1,600 or \$1,700 rather than only \$1,500 per year!

Further, it would seem difficult to justify the logic of regional or city differentials for what I will term "negative tax payments" -- which might be any of the growing number of different non-public assistance income maintenance plans being discussed -- when we do not use such adjustments in the "positive" personal income tax. I can't think of any reason to adopt a social policy that imposes stricter rules of equity to a southern sharecropper than we now apply to a wealthy oil producer!

Although money can be used to help alleviate the immediate problem -- i.e., the lack of it -- for the poor, long-run policy must be aimed at the multi-faceted basic causes of poverty. For many of the presently poor families, this means developing programs through which they will be afforded the opportunity to acquire either education or vocational training so they can earn an adequate income through their own efforts. Such policy, of course, would probably be most effective if aimed at the approximately 15 million children now classified as poor.

It would be foolish, however, to think that any policies -- either short run or long run -- can be devised which will make "taxpayers" out of all the so-called "taxeaters." For those who are aged, or unable to work because of a serious disability or chronic illness, we should expect to support them through new and adequate income maintenance programs for their lifetimes. For families headed by females, I think there is need for extremely careful analysis. As is the case among the non-poor, no doubt many of these mothers would prefer to be out of the house and engaged in some form of employment. But, I doubt that we would want to adopt any kind of national social policy which <u>forces</u> a mother to accept training or employment as a condition for receiving aid (as may be the case under the newly enacted Social Security amendments).

While I have separated the long-run and short-run policy questions for purposes of this discussion, it is obvious that if we are going to win the war on poverty we must proceed on both fronts simultaneously. Just as we would think little of a physician who treated only symptoms and whose patient died because the doctor neglected to concern himself with the underlying causes of some ailment, policymakers must also treat the "whole patient." Providing current income without the needed training programs is insufficient. But, providing only training for people who are hungry or who are ill-clothed because they can't afford such "amenities" is also insufficient.

The admittedly preliminary evidence presented at this session concerning various definitions of poverty do not seem to indicate that definitional refinements will make a very important difference in the poverty profile. While I would not want to disparage such efforts or hinder their progress, in terms of current priorities, I would be for "less light" on the definition of poverty and "more light" on the alleviation of poverty. The need for action is obvious; and the time for action is now. Unless we want to perpetuate the "long, hot summers" and other social disorders of the past few years, it's imperative that we stop talking about ways to fight poverty and inaugurate some meaningful programs. Based on our past experience, these will have to be better programs; they will have to reach more of the poor; and they will have to involve a far greater allocation of our resources than we have been willing to use in the past. There is no cheap way to alleviate poverty. But we should remember that the full social cost of doing nothing or doing too little will probably exceed the money outlays needed to wage an effective battle. From this broader perspective, it's obvious to me that whatever sums are required is money well spent!

DISCUSSION

William G. Grigsby, University of Pennsylvania

Any operational definition of poverty, whether it is based on income or some other variable, should fulfill, as nearly as possible, four requirements: (1) It should divide the population in such a way that all of those who are classified as in poverty are in worse condition, somehow defined, than all of those who are not so classified. (2) It should divide the population in such a way that those who are classified as in poverty are not only relatively deprived but severely deprived. The word poverty implies a serious situation. (3) It should be invariant over time. (4) It should reflect the underlying problems which have occasioned the definition. For convenience these requirments may be termed the criteria of accuracy, severity, invariance, and relevance. The question that I would like to explore briefly is the extent to which the criteria are satisfied by current, commonly-used definitions.

Accuracy

A considerable amount of valuable research, some of which is described in the three papers just presented, has been concerned with this aspect of the measurement problem. Crude income data have been refined to reflect differences in: wealth holdings, family size and composition, the cost of living among cities and regions, spending patterns, and income in kind. Some attention has also been given to the fact that because of the inability of the poor to take advantage of bulk-purchase opportunities, supermarket prices, and reasonable credit terms, a dollar of income does not buy as much in the ghetto as it does elsewhere. Still further work has been devoted to the problem of adjusting income data to include the governmental goods and services which are received by the poor and others.

In effect, all of these efforts have focused on the single task of more accurately measuring per capita real income, so that we may distinguish low-income families from moderate- and upper-income households with some precision. Implicitly, therefore, they assume either that low-income is synonymous with poverty or that it is a very good proxy for it. Given the large number of low-income families in the United States who do not seem otherwise severely deprived and the equally large number of moderate-income families who do have serious problems, there is some reason to doubt this assumption and to raise the question of what it is we are trying to measure. Is it really just low-income, or is it much more than that? Are we more accurately measuring the wrong thing? These possibilities lead us to a consideration of the fourth criterion.

Relevancy

It is now fairly well recognized that every year large numbers of families "move" across income-poverty lines, either in one direction or the other. Which families make such moves and how soon, if ever, they cross back again is not known. Consequently, cross-sectional income data, which are presently used to estimate the extent of poverty in the United States, may not accurately measure how many and which families have low incomes over an extended period of time. This fact has not bothered as many analysts of poverty as it should. Apparently, it is felt that the length of time a family spends below a given income line has no bearing on whether it is in poverty or how severely it is deprived. If the family is below the line, it is assumed to be in poverty, whether it has been there six months or six years. This view may be acceptable as a first approximation. However, the number and severity of problems associated with low income would be very different if every family experienced low income one-fifth of the time than if one-fifth of the families were in this situation perpetually. Clearly, there is a time dimension to poverty which present definitions ignore, with unfortunate results.

Consideration of this dimension of the problem leads to what I believe is a more realistic conception of poverty. Poverty is not low income per se, but the collection of problems that tend to result from being deprived of adequate income over a sustained period of time. More specifically, poverty consists first of the deprivation of physical well-being -- physical comfort, physical health, and safety -- that is experienced by low-income groups to a greater extent than by others; and, second, the deprivation of mental well-being that is occasioned by enforced idleness, being forced to rely on a stigmatizing dole, lack of opportunity to improve, and alienation from the mainstream of society.

The underlying goal of anti-poverty programs is not nearly so much to raise incomes above any particular line, but to reduce the above deprivations and distribute them more nearly at random among income groups. The income distribution which achieves this objective cannot be determined *ex ante*, but only after analyzing the incidence of the deprivations themselves. It is significant, therefore, that all of the current poverty lines carry with them the prior assumptions that: (1) the goods and services which can be purchased with an income that falls just at the line will prevent the deprivations of concern to us; and (2) nearly all families who have the minimum income will consume in the "proper" manner. The first assumption has never been subjected to empirical inquiry. The second is known to be false. Thus, income proxies for poverty are, at the moment, at least two steps removed from reality, and, like GNP, only crudely measure degree of well-being.

Severity

The issues which have just been discussed emerge again with respect to the problem of where to draw the income-poverty line. This problem essentially involves reducing the range within which an arbitrary decision must be made. It is obvious that the lower the line is drawn, the higher will be the proportion of the low-income group who are seriously deprived; but the higher, also, will be the proportion of seriously deprived persons who are not classified as poverty-stricken. The converse is obviously true if the line is drawn high. All of those who are seriously deprived are likely to be classified as in poverty, along with many persons who are not deprived at all. No way around this difficulty has been found, and "objectively" determined poverty lines therefore vary significantly. A major reason for this situation would appear again to be reliance upon erroneous or untested assumptions about the relationship between income and consumption patterns, and in turn between these patterns and income-related problems. The determination of "seriously low" income can only be made with reference to independent measures of the outcomes associated with various income levels. We do not yet have these measures. We have not even specified the relevant outcomes.

Invariance

If, as many persons believe, poverty will always be with us, one reason may be that it is constantly being redefined. We are continually reminded that using the standards of 40 years ago, few of today's population would be in poverty. Similarly, using 1968 standards, most of the population of the roaring and prosperous 1920's would appear to have been destitute. Allowing for changes in definition, the extent of poverty today is not substantially less than it was four decades ago. Because this conclusion seems so clearly implausible, something must be wrong. That something is the failure to specify the various ways in which low-income families are deprived relative to others in society and to seek measurable goals along each of these dimensions. This procedure would not result in a definition that is totally invariant over time, since new goals can constantly be added to any list. It would, nevertheless, introduce much more constancy than is now present.

[1/] For an elaboration of these remarks, see Morton S. Baratz, William G. Grigsby, Conceptualization and Measurement of Poverty, 1966.

VIII

SOCIAL INDICATORS

Chairman, ELEANOR BERNERT SHELDON, Russell Sage Foundation

MEASURING THE TREND IN SOCIAL STRATIFICATION
Abstract

Otis Dudley Duncan, University of Michigan

As K. Svalastoga has stated, "The degree to which positions are filled without respect to social origin ... is the most important clue to [the] system of social stratification" of society. The most appropriate indicator of stratification in this sense appears to be the regression of measures of achieved status on measures of origin status. Some historical data on the correlation of son's with father's occupational status, secured from community studies, serve to illustrate the more salient methodological problems in trend measurement. There is a reasonably convincing case that the measurement of status can be effected in such a way that intertemporal comparisons are conceptually justified, but defects in comparability of data preclude a firm judgment on trend. For what they are worth, the historical comparisons as well as certain inter-cohort comparisons based on retrospective information suggest that there has been no pronounced and continuous lessening or increase in the degree of stratification in the United States over the past several decades. The most important step in monitoring future trends would be a periodic replication of the 1962 survey of Occupational Changes in a Generation.

THREE ASPECTS OF LABOR SUPPLY SINCE 1900

Stanley Lebergott, Wesleyan University

I. PATTERNS OF LABOR SUPPLY: The Role of the Family

Individuals do the work of the world. But are they the decision units for labor supply? In American society decisions made within, and by, the family establish who will seek work, for how long, and where. It is not the children alone who decide whether they will work. The wife does not, alone, choose between a part-time and full-time job, or when to quit the labor force. Family pressures for income enter into all these decisions. Of course these choices are constrainėd by the larger society: fathers in most social classes seeem inevitably to work, children to go to school full time, mothers to keep house, and so on. But options keep changing. Families shift social status. Variation can be observed within any social category of which one can think. And labor force participation can be usefully considered in its family context. We shall see how the sharply different trends in labor supply since 1900 are pulled together into a wildly improbable, but apparently correct, explanatory frame by that context.

The simple facts of labor supply are conveniently summarized in about 50,000 time series for the years since 1940 - plus a hundred or so key monographic studies. From these we refer to only two facts:[1]

> About half the U.S. population of working age (14 and older) is in the labor force at any time: 56% in 1900 and 58% in 1966.
>
> The proportion for males is about double that for females - 78% to 38%.

Behind these simple figures lies the infinitely complex, the continuously changing, pattern of labor supply and its determinants. The rise in output is absolutely limited by the extent to which people can be induced to give up more tantalizing activities in order to begin work earlier in life, continue on in later years, take on a second job as well as a first one, etc. And it is further bounded by their changing preferences in work itself - for factory work rather than domestic service, for jobs above ground instead of the dark of the mine, for well paid work compared to slow-paced work, and so on. The entrepreneur bidding for labor must take most of these preferences as given. For they are fixed by the mute movings of instinct, by ancient tradition, by the subtle interplay of family relationships.

A.

In recent years we have seen two titanic forces shaping labor supply. One to restrict, the other to expand it.

(1) Actions by the government constricted the supply of labor - cutting immigration from overseas to a trickle; insisting on education requirements that result in today's workers entering the labor force close on age 19 rather than - as once - ages 10 to 14; holding out old age pensions that induce men to quit close to age 65 rather than try to hang on to 70.

(2) The expansionist force, unconsciously striving against this restrictionism, was that of advertising, constantly persuading the consumer to new goods and to iterations of old ones. But the consumer lives inside the same skin as the worker. And the new commodity can typically be acquired only by work.[2] Hence it turns out that advertisers have been busily increasing the national labor supply.

B.

Given these massive pressures what were trends of labor input since 1900 for the three major groups - men, children and women?

<u>MEN</u> Two dominant forces characterize the trend in labor input by men.

1. Most obvious is the decline in hours worked. Some 95% to 98% of men in the central ages have always worked. But the hours they put in each week have fallen rapidly since 1900:

Average Weekly Hours in Factories[3]

1900	59
1929	44
1940	38
1966	41

The swoop downward from 1900 to 1929 stands out. But so does the sudden halt, and the grand stability since the 1930's. Union organization during the thirties and Federal wage hour legislation together did little to cut prevailing factory hours. The tendency to exchange more income for more leisure was apparently checked by the end of the mid 1930's. Workers thenceforth took productivity gains almost wholly in money rather than partly in leisure.[4]

2. The other major decline in male labor input was the fall in the percent of older workers in the labor force - from 1900 to the late 1920's, and continued on till today. The mild decline for the 55-64 age group, the drastic one for the

over-65 group, report the impact of rising income and of the Social Security Act:

Worker Rates for Males,[5]
by age

	60-64	65-69	55-64	65+
1900			90.0%	68.4%
1930	86.8%	75.7%	87.1	58.0
1965	78	43	62.6	21.9

<u>CHILDREN</u> For children, too, labor input declined markedly to the late 1920's:

Worker Rates for Children,[6]
by Age

	10-15	14-15
1900	18.2%	
1920	11.3	17.5%
1930	4.7	9.2
1940	2.9	5.2
1950		20.8
1960		17.5
1965		16.0

But instead of that decline merely ceasing at the end of the depression, as that for hours worked, the trend was reversed. Work by children was welcomed at a fast and increasing rate since 1940.

<u>WOMEN</u> If we turn now to the other volatile component of the labor force, women, we find no such reversal, but persistent trends.[6a] (Table 1) The only apparent decline from 1890 to 1940 was for single nonwhite women. Since then rates for both white and nonwhite single women have dropped, and those for married women have risen both persistently and markedly. As a result, rates for nonwhite married women reached equality with those for single ones during the 1950's, while those for white single and married appear likely to cross around 1970.

Table 1

Female Worker Rates, by Marital Status

Year	White			Nonwhite		
	Total	Single	Married	Total	Single	Married
1890	12.1	35.2	2.5	39.5	56.4	22.5
1940*	26.9	47.9	14.6	43.2	45.1	33.5
1951*	31.5	50.5	24.3	41.1	41.3	36.0
1960*	34.1	45.5	29.6	41.2	33.6	40.8
1966*	36.8	42.0	35.0	44.1	32.7	49.4

*March

C.

In the midst of capitalist society there exists an island of primitive communism - the family. Its members share labor and income among themselves with little reference to private market allocation procedures. Yet its decisions and desires help create the varied pattern of labor inputs to the market noted above. Now most American families live above a coarse subsistence level. Hence a significant share of labor is provided by men and women busily working to add optional items to the family expenditure level. Faster junking of automobiles, homes with greater privacy for family members, added TV sets, money for retirement years - all of these are options. But they are options that can be achieved only by extra work on the part of family members. Hence labor supply is ultimately fixed by family expenditure horizons, or, if you will, the associated income horizons.

It is easy to explain why most adult males work. But to explain work patterns in a society well above subsistence levels we must look to the role of income horizons. Suppose that we confront the usual history of increasingly delayed entrance, and earlier retirement, of males from the U. S. labor force, plus the increasingly shorter scheduled work weeks, with the present complex patterns of labor force participation.

The traditional findings imply a flight from labor to leisure as the years have passed. But if so what are we to make of recent data for males in the central age group, 25-44? For these report that in 1965:[7]

40.1% worked overtime (on one job)
Another 8.2% had two or more jobs (averaging 39 hours on their primary job, 13 more on their secondary job).
Some 37.2% had working wives.
Another 11.8% had other family members in the labor force (but not working wives)

These figures may be added together, with only limited error, to reach one conclusion: nearly every family in the central, 25-44, age group was making some extra effort to earn income above and beyond a regular job by the husband. Half did so by additional work of the husband; half by wives or children working. (Some tried both.)

Existing data do not permit us to say whether the search for income has intensified since 1900. About all one can add is that some of these endeavors substitute for taking in boarders and lodgers. (About one-quarter of urban families did so in 1900, although less than 5% do so today.[8]) Work has thus not merely helped to adorn the home with artifacts: it has made it more private.

These differing trends - worker rates rising for married females; declining for males; declining then rising for children; shortening of the scheduled work week but probably expansion of overtime and second jobs - seem susceptible to only one explanation. It is not an explanation to be found in the demand side of the labor market per se, but in the family's labor supply schedule. We have asserted that the schedule is fixed by family income horizons. Data on sources of family income, put together from surveys scattered over two-thirds of a century, tell us something about that horizon. They report an unsuspected constancy:

Proportion of Family Income
Contributed by Husband
(urban families)[9]

1901	80
1935-1936	82
1965	81

Apparently families set their income horizon at 25% more than the head of the family earns. So it was in 1901, when unemployment, and real incomes were low. So it was a third of a century later in the midst of our most extended depression. And so it was in the piping warlike days of 1965. As real hourly wages of husbands rose from 1901 to 1935-1936 other family members adjusted their own participation so that they continued to add 25% to the husband's income. As wages continued to rise over the next thirty years, and unemployment to fall, there was again an adjustment to 25%. Now since wage rates for women and teen agers had also been rising, it is clear that the historic increase in female labor force participation was not required to yield the same real income. Instead, income horizons had shifted. And the family continued to seek 25% more income than the husband provided, despite his higher wage rates and despite his overtime work. Perhaps a moralist or philosopher could explain why the U. S. family always seeks more than the husband's income. But it would require a more protean analyst to explain why the family income horizon apparently kept moving throughout two-thirds of a century to a point 25% beyond the income that the husband could provide.

II. The Decline of the Entrepreneur

A.

Let us turn now to the transition since 1900 from a labor force with a major independent, entrepreneurial component to one with an employee, a worker, orientation. Europe had entered the twentieth century trailing a history that ranged from Feudalism (at its harshest) to sharp class distinctions (at its mildest.) Not so in the United States, which still retained the upward, sanguine outlook of an open society. Our past had been a different one. Moreover, the role of the self-employed remained distinctly greater than in Europe, and workers' prospects of gaining that station were more assured.

In 1900 (as Table 2 reports), about 13 million persons were closely involved in running family businesses. About the same number of persons were employees. The one-to-one ratio between the independent, profit-oriented, group and the hired worker group helped define an open society, helped generate tantalizing worker hopes for personal advance. It was neither happenstance nor a testimony to the power of legal injunction and employer dissuasion that in 1900 less than 4% of the labor force was unionized.[10] Self-employment reached high tide - perhaps not surprisingly - in 1925-26, just midway in the Coolidge years. But its absolute rise from 1900 had been small and its ratio to the number of employees had run almost steadily downward.

The fading of independent entrepreneurship synchronized with the deadly accumulation of mass unemployment in the early 1930s, and the swift rise in the percent of the labor force organized, from 8% in the beginning of the Roosevelt-Truman era (1932) to 26% at its end (in 1952). By 1967 declines in farm employment had decimated the self-employed group, to a figure of not much over 10%.[11] (And few of these self-employed were employers.) That percentage may not exceed the corresponding figure of the Union of Soviet Socialist Republics. Most of the 71 million labor force took orders from other employees - corporate officers, foremen and others who possessed power, but not the power (nor the perspective) of the classic independent entrepreneur. The "employing class" was on its way to extinction.

Table 2

Labor Force and Employment
(in millions)

	1900	1941	1960
Civilian Labor Force	28.4	55.9	70.6
Self Employed and Unpaid	12.7	12.8	10.9
Unpaid Family Workers	3.0	2.0	1.7
Self Employed	9.7	10.8	9.2
Farm	5.8	5.2	2.8
Service	1.1	1.6	2.2
Trade	1.3	2.3	2.4
Construction	.5	.5	.8
Manufacturers	.4	.3	.4
Other	.4	.5	.6
Employees	12.5	35.5	53.4
Domestic Service	1.8	2.1	2.5
Unemployed	1.4	5.6	3.9
Armed Forces	.1	1.6	2.5

Sources: Tables A3, A4, A7 in Manpower in Economic Growth

B.

The swift rise of the employee share in the labor force had its deepest impact in that dynamic, critical sector of the economy where goods were made. Factory production had been somewhat equitably shared between independent proprietors and corporations in 1900. But after 1900 output (and employment) centered in corporations:[12]

Manufacturing Sales
(billions)

	1900	1964
Sole Proprietors	$ 3	$ 7
Corporations	8	453

Not fate nor malignant endeavor accounts for this shift. Increasing returns to scale had appeared throughout the economy: more goods were produced with the same effort and investment as the scale of production increased. But to achieve the larger economy larger business units were apparently required. And to assemble financing for such larger units the corporate form proved more feasible (largely because safer) than the sole proprietorship. Hence the growth of the corporation, hence its taking over of the vast bulk of U. S. output and employment.

Only in part did greater efficiency derive from larger plants, larger production units. True, plant size did rise after 1900. But most factory workers, even today, are not employed in those giant plants with over 1,000 workers. Changes in plant size proved trivial compared to changes in the size of business firms _per se_. And increasingly the larger firms began to dominate the economy. The advance in concentration of employment can be noted most reliably for manufacturing. Between 1900 and 1963, as Table 3 indicates, the proportion of all factory employees in the 185 largest firms rose from 8% to 27%:

Table 3

Concentration of Employment
in factories[13]

	1900	1947	1963
Total	5,308	11,918	12,232
In the 185 largest firms	400	(2,035)	(3,300)
Percent in largest firms	8%	(17%)	27%

Such advancing concentration probably signalled with broad accuracy the increasing centralization of employment in the entire economy.[14] (Although such small-employer sectors as trade and service grew at faster rates than did total employment, that most centralized of employers, government, grew still faster.) Today we reside in neither the early world of small master and humble journeymen, nor in that caricature, "the apogee of monopoly capitalism". Yet Table 3 does suggest a speeding up in concentration after 1947. And one might speculate that by the year 2,000 half our labor force will work for the 250 largest businesses and government units.

III. NEW CONTROL MECHANISMS

Given the growing concentration of employment, the greater role of larger business units, differences in the mode of production appear. Corresponding differences in wages and working conditions may spring up.

There is no evidence that the competition of 400,000 factory enterprises in 1963 produced a different kind of wage determination than the competition of 450,000 in 1900, even ignoring the labor market competition across industry sectors.[15] But whether the growing concentration of employment in larger firms makes a major difference for wage determination (via pattern bargaining etc.) is uncertain. The analytic warfare between those who argue for the impact of market power, union importance and oligopoly on wages as against those who argue for demand forces and the power of competition has not yet been settled - if ever it will be.

There are some changes, however, with respect to the labor force to which one can point with more confidence.

A.

To control workers in larger plants, in greater firms, employers put together control mechanisms that were weak (or superfluous) when workers and owners labored together. One obvious and standard component was direct, hired, supervision. It is probably impossible to report the number of persons engaged in supervision. But Table 4 may tell us something about that trend. As the number of self-employed, who oversaw their own businesses, drifted downward over the years, the number of foremen - a fairly clear-cut and indicative category - increased. The ratio of foremen per employee changed little from 1910 (perhaps from 1900) to 1940. But when unions began to spread in the late 1930's, and wage rates spiralled, a sharper cost consciousness was apparently generated. Supervision became markedly closer after 1940, the ratio of foremen to employee rising at 5 times the rate it did 1910-40.

B.

But there were further alternatives to the urgency that a master could communicate, or the drive that a foreman could command. Some were embodied in another control mechanism: the

the incentive pay system. That system did not rely on personality to shape worker responses, nor on command and status. Instead it trusted to direct financial stimuli.

Incentive pay schemes are, of course, very old. The men who chased Moby Dick were paid a percentage of the sales value of the catch. And sharecroppers have long been classic examples of workers paid by results.

For manufacturing we have trend data of real reliability.[16] Summarized in Table 5, what do they report?

Table 4

Employment in Manufacturing
(000)

	1910	1940	1960
Self-Employed[1]	504	324	383
Foremen[2]	175	293	742
Employees[2]	7,280	10,601	17,530
Foremen per 100 Employees	2.40	2.76	4.23

Sources: (1) Manpower in Economic Growth, Table A-7
(2) 1910: Employees - Gladys Palmer & Ann Ratner, Industrial and Occupational Trends in National Employment (1949), App. III. These data are based on the Population Census, hence comparable with the others shown here. Foremen - 1910 Census, Occupations, p. 91.
1940: 1940 Census, Occupation Characteristics, Table 9.
1960: 1960 Census, Occupation by Industry, pp. 17, 19.

First, and most obviously, the tide was running toward incentive plans: the proportion of all factory workers under incentive plans rose from 18% to 27%. More than two thirds of the sectors shown report increases. Spectacular ones showed up in many of the durable goods industries, such as primary metals from 10 to 46%; stone, clay and glass from 8 to 25%; electrical machinery from 16 to 40%.

Secondly, and no less significant were spectacular declines in some nondurable industries - tobacco falling from 64 to 31%, printing from 15 to 4%, furniture from 54 to 25%.

What are we to make of this mixed pattern - particularly if we note the fairly trivial changes for a few sectors (e.g. food from 10 to 12%)? Most sectors, particularly those with a small proportion using incentive plans originally, found increasing scope for the plans. They presumably proved an effective tactic of cost control by transport equipment and other industries that were increasingly under pressure for higher wage rates.

On the other hand if we made a list of industries who were relatively sluggish in their productivity advance, who generally bombarded the Congress with pleas for tariff protection, we would include many of those that had high incentive percentages to begin with - and then increased them further: leather, from 44 to 63%, apparel from 51 to 59%. And that list would include such other high ratio industries as knitting (65%) apparel (59%) steel (60%) glass (45%). All of which simply suggests that incentive pay schemes proved to be no panacea: like good whiskey they may be effective medicine or, denatured in vegetable compounds, constitute a sorry excuse for genteel toping and evasion of responsibility.

Table 5

Incentive Pay:
Percent of Factory Workers
on Piece Work (1890) or Incentive Pay (1958)

	1890	1958
All Manufacturing	17.9%	27.0%
Tobacco	64.1	30.9
Furniture	54.2	25.0
Apparel	51.3	59.3
Leather	44.0	63.2
Paper Boxes*	45.3	20.0
Printing	14.6	3.8
Textiles	13.4	39.7
Food	10.4	11.5
Chemicals	5.6	8.8
Lumber	4.1	6.3
Instruments	45.2	29.2
Toys, Sporting Goods*	24.5	24.0
Nonelectrical Machinery	21.2	25.9
Fabricated Metals	18.6	23.1
Electrical Machinery	15.6	40.3
Toys and Games	12.7	24.5
Primary Metals	10.1	46.4
Jewelry, Silverware*	10.0	35.0
Stone, Clay and Glass	8.4	25.1
Transport Equipment	4.6	10.4

Sources: 1890: Computed from data in 1890 Census, Manufacturing, Part 1, Table 4.
1958: Monthly Labor Review (May 1960), P. 461.

*BLS coverage too limited to permit showing 2-digit totals for paper or miscellaneous.

III

The third major technique adopted creating spirit in the labor force and controlling labor costs can be inferred from the fact that tobacco and furniture cut their high 1890 incentive percentages markedly by 1958. Bumping up against the limits of the incentive technique they shifted the burden of payroll control on to machine pacing. By 1958 only 3% of cigarette workers (and 13% of motor vehicle workers) were under incentive schemes. The other 97% (87%) were controlled more effectively by the scheduling inherent in the production line, the pacing of the conveyor belt.

In sum, no industry found a single, simple control to replace the eye of the master in the small workshop. Some expanded supervision by foremen. Some put in piecework schemes. Some instituted machine pacing of work. Some combined techniques, their mixture depending on the rate of technical advance, the constraints of materials, the quality of management.

APPENDIX

A. In 1965 21.4% of the male noninstitutional population aged 14-15 was in the labor force, 44.6 of the 16-17 group, 70% of the 18-19 year olds, and so on. (BLS Special Labor Force Report No. 69, Labor Force and Employment in 1965, Table B-1.) Assuming a zero participation rate below age 14, we infer that 21.4% entered the labor force beginning at age 14, that 44.6 minus 21.4 (or 23.2%) entered at ages 16-17, and so on, till the 98% rate for age 30-34, taking that as the effective maximum. At the other end of the age spectrum a problem arises because rates for a given year necessarily do not reflect the nonparticipation of those in our beginning period population who have already died or entered institutions. We pick an arbitrary 65 in the light of participation rates and death rates. Hence the average male enters at 18.5 leaves at 65, based on rates in 1965. (Rates for 1960-64 are much the same.)

B. In January 1966 a labor force survey indicated that males then aged 65-69 had spent 14.6 years on their current job; those aged 60-64 had spent 16.5 years, and so on. Taking 15.5 as the average for men 65 years of age, we then estimate the duration figure for men aged 49.5 (i.e., 65 minus 15.5), and so on. Summing these intervals indicates 12 jobs from age 18 1/2 to 65. (Data on job tenure as of January,1966 from BLS Special Labor Force Report No. 77, Job Tenure of Workers, January 1966, Table A.) Tenure surveys for 1951 and 1963 suggest shorter job duractions were probable during peak hiring periods (e.g., World War II and the Korean War), and longer during slow growth. Using the 1966 report for projections assumes something like the 4.5% to 5.5% average unemployment that prevailed in the past half dozen years. (I here use more recent survey evidence, and a somewhat better procedure than that adopted for a similar estimate in my Men Without Work.)

C. For 1900 Rees gives a daily average of 9.89 hours in manufacturing. (Albert Rees, Real Wages in Manufacturing, 1890-1914(1961), p. 33.) For 1929, 1966: Economic Report of the President, January 1967, p. 244.

D. 1900: Estimated from data for 1901 in U. S. Commissioner of Labor, Annual Report (1904). Pp. 362, 366. 1935-1936: Ratios for particular sampling areas, as shown in the BLS-BHNHE Consumer Purchases Volumes, weighted by the appropriate distribution of nonfarm families from National Resources Committee, Consumer Incomes in the United States (1938), Table 24B. 1965: Bureau of the Census, Current Population Reports, P-60, no. 51, Income in 1965 of Families and Persons in the United States. For nonfarm families, male head, married wife present, the median family income was $7,436, of which the head's income was $6,026 (p. 32), or 81%. Of 39,419 wives 19,816 had a median income of $1,789 or an average of $872 per family, leaving $538 contributed by children and other family members.

E. Data from Manpower and Economic Growth, p. 53 and the writer's "Labor Force and Employment, 1800-1960" in Conference on Research in Income and Wealth, Output, Employment and Productivity in the United States after 1800 (1966), p. 148. Participation rates for ages 10-13, 14-15 in 1940 from the latter source were weighted by population data from 1940 Census, Characteristics by Age, p. 8.

F. BLS, SLF 81, Overtime Hours and Premium Pay, Table J reports 6,949,000 male wage and salary workers working overtime as of May 1966, or 40.1%. BLS, SLF No. 63, Multiple Jobholders in May 1965 reports 1,676,000 males in May 1965 with 2 or more jobs, or 8.2% of those employed in the group. Given the broad stability in the dual jobholder percentage we assume that 8.2% applies to May 1966 as well. Now since about 60% of the self-employed worked longer than 40 hours (BLS, SLF 69, Table D-2), the 40.1% for wage earners plus 8.2% should run to over 50% for all males 25-44. BLS, SLF 80, Marital and Family Characteristics of Workers, March 1966, Table B, indicates that 37.2% of all females, married husband present, aged 25-44 were in the labor force. Since women tend to be married to men slightly older than themselves, the rate for wives of men 25-44 might be 0.5% or so lower.

SLF No. 80, P. A-6 reports 37.2% of females husband present age 25-44 in the labor force. To this we apply a ratio of 30.4% to derive 11.8% for other family members, but not working wives. We derive the 30.4% as the ratio of F-A-21 data indicating 35.5% of all wives in the labor force and 10.8% of husband-wife families with no wife but other members in the labor force.

G. 1900: 1900 Census, Manufac tures, Part 1, pp. lxxxii, 3. We take the 185 organizations for which data are tabulated as equivalent to the 185 largest. Such an assumption is mildly in error: the Carnegie firm was omitted, as well as a few others who might supersede some of the smaller of the 185. 1947: 1963 Census of Manufactures, Vol. I, p.44. 1963: For 1963 the U. S. Census, Concentration Ratios in Manufacturing Industry, 1963, p. 2 gives figures of 26% for the top 150 firms, and 28% for the top 200, from which we interpolate. The same source gives data on concentration of value added, data which show a ten point rise from 1947 to 1963 for the 150 largest firms - and for the 200 largest as well. We reduce the employment ratio for 1963 by ten points to get a 1947 estimate.

FOOTNOTES

1 cf. the writer's Manpower in Economic Growth, p. 512, BLS,SLF 69, p. A-7.

2 Theft,is, of course, one alternative, but mores, internalized values and police all preclude much resort to that solution. Drawing down savings or borrowing is another alternative. However, since families have savings goals, savings must be replenished and debts paid; hence the net answer proves to be: additional work.

3 Estimated in Appendix C. Although estimates for all workers exist, we prefer here to rely on the more reliable data for manufacturing industry - likely to be indicative as well as reliable.

4 It is irrelevant to our present purpose to note that coffee breaks, etc. have probably proliferated. The worker's 40-hour presence on factory premises is required. Additional hours have not (as in earlier decades) been traded for less income.

5 1900-1940: estimates made for developing the annual series in Manpower in Economic Growth, Pp. 393, 402. Cf. 1900 Census, Occupations, P. cxviii; 1930 Census, Occupations, p. 115. 1960ff: BLS, SLF No. 69, Labor Force and Employment in 1965, p. A-10.

6 Estimated in Appendix E.

6a 1890-1960: Data from Manpower in Economic Growth, p. 519. 1966: BLS, Special Labor Force Report No. 80, Marital and Family Characteristics of Workers, P. A-5.

7 Estimates derived in Appendix F.

8 U. S. Commissioner of Labor, Cost of Living and Retail Prices of Food (1904), p. 362 reports for a large (but in some respects biassed) sample of urban families in 1901, that 23% had income from boarders and lodgers; 9% from working wives; 22% from working children. 1960 Census, Families, p. 195 reports just under 4% of all families in urbanized areas as having lodgers.

9 Estimated in Appendix D.

10 Gregg Lewis, Unionism and Relative Wages in the United States (1963), p. 244.

11 BLS, SLF 69, Labor Force and Employment in 1965, Table C-4.

12 1900 Census, Manufactures, Part 1, p. 503, U. S. Internal Revenue Service, Business Tax Returns, 1964-1965, p. 7. Data for "partnerships and firms" in 1900 totalled a further $2.6 billion.

13 Estimated in Appendix G.

14 Such a conclusion appears to be at variance with the well known study by Nutter, and later work built upon that. However that study is questioned by the writer in the Review of Economics and Statistics, November, 1953. Moreover it deals with national income. When we come to employment the role of government is necessarily greater - the contribution of government capital being omitted in the income estimate.

15 For 1900 we roughly adjust the 512,254 establishment figure (1900 Census, Manufactures, Part 1, p. 503). For 1963 we use the IRS total Business Tax Returns, 1964-65, p. 7.

16 1890 data computed from figures in Eleventh Census, Report on Manufacturing Industries of the U. S. (1895), Part 1. Other data from Monthly Labor Review, November 1947, p. 535 and May 1960, p. 461.

CONCEPTS AND MEASURES OF WELFARE

Ida C. Merriam, Social Security Administration

The concept of welfare, in its most general form, underlies any approach to the construction of social indicators. Welfare (and how to achieve it) is, indeed, what social indicators--and economic indicators--are all about.

An important aspect of the measurement of welfare must be fractional or sectorial, growing out of the answers to the questions of "how much," "what quality," "for whom," that should be asked with respect to each separate component of social life and of social and economic change. But this cannot be the entire answer. For these identifiable parts are interacting and potentially conflicting. And the human mind--no matter how sophisticated--persists in wanting to know whether matters in general are better or worse, while the policymaker must ask which emphasis or which choice will contribute most to general welfare.

Whether welfare in this general sense can ever be measured is highly doubtful. General welfare, like positive health, may be impossible to define in operational or policy terms. This is not to say that the idea of welfare has no meaning. But because the meaning is value-laden and complex, all that the social scientist can hope to discover is symbolic indicators and proxy measures. Theoretical welfare economics--which this paper is by-passing--has been built on the recognition that the value or utility of specific goods, services or forms of leisure is not the same for all individuals, and that the presumed end product--satisfaction and equivalent levels of satisfaction--is not directly measurable. We measure observable states and reactions and our best policy goals may be those which stress access to rather than use of particular goods and services. Let me add quickly that "access to" means more than formal availability; it includes the information, the education and the social structure that make possible a genuine choice.

As I observed at the outset, there are many partial welfare indicators--relating to health, housing, civil rights, social disorganization and so on. This paper is drawn from the final chapter of a volume now in preparation that deals with social indicators and the measurement of social change in thirteen specific areas. Presumably the final chapter can assume the importance of all the individual measures.

In searching for a more over-all measure or frame in which to view the specifics, I have come up with nothing startling or new. What I have found myself pushed back to is the level and distribution of income as the most significant general determinants of welfare. Productivity and abundance on the one hand, and equity and social justice on the other become the general frame of welfare.

Abundance--the total quantity of goods and services available--is the more easily measured, and I shall not deal here with any of the problems or issues that are involved, particularly if one attempts to take into account some of the disutilities of technology and economic growth.

Equity is an abstraction almost as difficult as welfare. Notions as to what is fair and just differ among individuals in different circumstances, among societies in different stages of development and over time. But the preference for an equitable society is deeply rooted.

Equity, of course, does not necessarily mean equality. Some degree of inequality would appear to be necessary to provide economic incentives to produce, although our knowledge as to the effects of different types of incentive or disincentive (earnings differentials, progressive taxation, power relationships, status and symbolic rewards) is limited.

Fortunately, it is possible to measure distributional elements without complete agreement on a definition of equity. And if we are not sure as to the desirable shape of the total distribution curve, we can perhaps come closer to agreement on certain cut-off points or negative welfare indicators. The current interest in poverty measures is an example of this approach. The idea of a social minimum has a long history, both in philosophical discourse and in reform movements. The relation of this concept to the general concept of equity may be particularly relevant to currently important issues of social policy.

So long as necessities mean largely subsistence type needs, the distinction between a social minimum concept and a measure of the over-all degree of inequality can be fairly clear-cut. As the general wealth of a society and the goods and services which are perceived as necessities increase, the distinction becomes somewhat blurred. The higher the poverty level in absolute terms, the more critical becomes the question of the distribution of real income both below and just above the level. The social minimum which gains solid public and political support at any given time will certainly be related to and largely bounded by the existing distributional patterns in the middle-income levels.

<u>Defining What is Distributed</u>

Thus far I have intentionally slurred over the question of what is meant by income and to whom it is distributed (the individual or a defined family group) as well as the reference period. The relevant literature is so voluminous--and the unsolved questions so numerous--that I cannot hope to discuss even the major points in a twenty-minute paper. I shall resort, therefore, to some fairly didactic statements and propositions.

If the basic purpose is to measure change in over-all welfare, an indicator that lends itself to repeated use and to the provision of meaningful trend data is what we must look for. An indicator in this context is not intended to mean one series of figures. It does mean an interrelated set of measures, not just a congeries of statistical data. Several different methods of summarizing the distribution of income may be appropriate and desirable. The basic measure should be usable as a classificatory scheme for the analysis of the distributional effects of all the components of economic and social change. Supplementary measures will be needed for specific analytic purposes. They may lead to refinements in the basic measure or simply to better understanding of its limitations and its meaning.

From this perspective, perhaps one can turn to the obvious starting point--current money income--and ask how useful for an analysis of welfare is the distribution of money income alone. I would argue that it is considerably more useful than much current discussion implies, primarily because of the degree to which it is related to other factors that affect the way people are able to live and their control over their life situation.

Recent attention to the problems of the poor has highlighted the fact that ability to use income may in some circumstances be almost as important for the level of living as the amount of income itself. There appears, however, to be a high correlation between the amount of money income and the ability to use it effectively. Efforts to improve the purchasing power of the poor through better access to the market and lessened discrimination of all kinds are highly desirable. In the end, the amount of money income may prove to be the critical factor.

Considerable attention has been devoted to the definition of nonmoney equivalents of money income. An important factor in any comparison of farm and nonfarm incomes is the value of home-grown food or home-produced clothing. International comparisons and trends over time are also affected by the treatment of such items. Of a different character is the imputed value of the use of owner-occupied dwellings or of consumer durables. The latter concept approaches a wealth-income measure. It is important for certain kinds of economic analysis, but may not be essential for a welfare-distribution indicator because of the relation between previous and current money income levels and the quantity and value of durables possessed.

Certain forms of private compensation that may be of growing importance also fall outside the usual measures of current money income. Expense accounts, help in the purchase of a home, stock bonuses, advanced education and training, private pension plans--all benefit disproportionately the higher-income groups. Our knowledge as to the aggregate value and distributional effect of such forms of compensation is very limited. This is an area where much more work is needed.

The growing importance of public expenditures and their close tie to social welfare policy give special importance to their effect on distribution. Total social welfare expenditures alone topped the $100 billion mark in fiscal 1967 and amounted to 13 percent of the Gross National Product, and 43 percent of all Federal, State and local spending. A sizeable part of these expenditures takes the form of transfer payments, most of which have a specific distributional purpose--either a more appropriate distribution of income throughout the life cycle than can be accomplished by the wage system, or a meeting of recognized income deficiencies. In the case of transfers, society is distributing money income, which enters directly into any income count.

But beyond this, publicly-provided services are increasing in scope and importance. With growth in population and pressure into cities, more and more goods--some once free, like clean air and privacy--can be generally available or equitably rationed or distributed only through mechanisms other than those of the marketplace.

The distributional effect of public expenditures involves both tax and benefit impacts. Neither the data nor the theoretical basis for analysis of these effects are as satisfactory as one might wish. Some of the conceptual problems can be illustrated with reference to public education. Public education through high school is intended to be available for all. The value (cost) of the services may, however, vary greatly from school to school, with the lesser per capita expenditures too often associated with poor neighborhoods and families. On the other hand, the imputed addition to income will be proportionately greater the lower the family income. Recent findings that educational success depends more on family and cultural background than on the quantity or quality of schooling provided raise a question as to whether a "correction" of the money income distribution to take account of the redistributive effect (downward) of public education really improves the value of the measure as a welfare indicator. This is not to say that society and social scientists should not be vitally concerned with the relative expenditures on education for different groups in the population. But at least at the lower end of the income scale, what is needed may be not the correction of a money-income distribution through the addition of assumed values for nonmoney income, so much as interpretation through reference to other analytic data.

One general observation may be worth making with respect to all attempts to develop an inclusive income figure. All such attempts involve imputations that rest on more or less arbitrary assumptions as to dollar values and as to incidence. The inclusion of items which are fairly equally distributed among money income classes automatically makes the distribution of total income less unequal.

Furthermore, as Dorothy Brady pointed out in one of the early analyses of income size distribution, an individual or family whose income is 80 percent nonmoney and a family with the same total income all in the form of money are simply not in the same situation or welfare category. To put the point another way, the school child who can get a school lunch if he declares himself needy may consume the same food as the child who buys his lunch, but he does not feel himself to be in the same position.

Assets and Wealth

The importance of assets and wealth as a corrective to the distribution of current money income has also received a great deal of attention and study.

Looking only at the current value of assets, there is a high correlation between levels of current income and asset holding. The poor have little in the way of accumulated wealth. For the near poor and middle-income groups, the value of an owned home is likely to be the major asset. In general, large wealth is associated with large current income. A measure of income plus wealth would show more dispersion than would a measure of current income (including, of course, current yield of assets).

The more important aspect of wealth may be the sense of security and freedom of action and the power and control of the future well beyond the lifetime of the owner that it can give. Ownership of wealth and particular forms of wealth also has different meanings in a dynamic growth economy than in a more static society. Small--or even fairly large--amounts of assets can lose their value as a result of inflation or economic depression. Physical assets--such as an owned home--can appreciate in value or be destroyed in the path of urban renewal or highway construction.

The importance of social insurance derives not only from the transfer of income from the economically active to the economically inactive years (or generations), but from the ability of the system to assure a continuing income with a stable or increasing purchasing power. It represents for the middle- and lower-income groups some of the security of wealth. While social security benefit levels are as low as they are today, the analogy may seem painfully strained. In terms of the potential of the institutional mechanism, it is valid.

Other Measurement Problems

Both conceptually and statistically the unit of reference or of count can significantly affect the apparent distribution of income. In terms of welfare, the individual is the ultimate referent. But the well-being of the individual is dependent on a network of social relations. Most importantly for present purposes, for about half his life, the ordinary individual must look to his family or society for his current consumption needs.

In analyzing the distribution of income and of welfare, the structure of individual earned income is an important subsystem. For more general analysis the unit of count is usually the family, but attention must then be given to demographic changes and family patterns. Increasing employment of women--both cause and effect of other changes--longer years of schooling, earlier retirement, undoubling of families when rising income or the assured income of social security for the aged permits--all have their effects on the distribution of family income. The relationships are neither simple nor static. In the space of this paper, I shall not attempt even to summarize the more important ones.

The conventional time period for measurement of income is a year. Income, however, may fluctuate both within a year or between years. The time period chosen to define "current" income smoothes out or averages fluctuations within the period, accentuates fluctuations between periods. There is no way of avoiding this effect. An important task of analysis is to measure the variability of income and to assess its effects. It is easier to recognize the significance of the distinction between reasonably "permanent" and variable income than it is to separate the two, particularly in trend data. We need more longitudinal family income data to increase our understanding of this aspect of distribution as related to welfare.

Statistical Measures of Inequality

There are a number of ways of measuring income inequality, whatever the definition of income or recipient unit. The distinction that comes first to mind in the current situation is the difference between the budget or poverty-line approach and the income-shares approach. As I suggested at the outset, the former focuses on the social minimum aspect of welfare, the latter on the more general equity aspects. The definition of a social minimum is inherently normative. The analysis of income shares--the proportion of aggregate income received by each fifth or tenth or other portion of the population--does not in itself depend on a concensus as to how equal or unequal the shares should be, although it may lead to conclusions and convictions on the subject.

The fact that the lowest fifth of all families (measured by family income) receive only five percent of aggregate family income may take on a somewhat different coloring depending on the presumed adequacy of the income they get to meet minimum needs. On the other hand, the repeated revisions of measures of minimum need as the general level of living goes up is evidence of the way in which the concept of fair shares permeates the budget approach.

In the few minutes remaining to me, I want to point up some of the other differences in these two types of equity measures.

Poverty Line Measures

It is obvious that any given family income has a different significance depending on the number of people it supports. Most budgets have been constructed for a selected typical family--such as the BLS wage earner's family of four with an employed husband age 38, a wife not employed outside the home, a son aged 13 and a daughter aged 8. To compare the budget cost with the median or mean income for all families of this general type does not tell us very much about the income distribution as a whole.

The usefulness of the poverty-line index and low-income index developed by the Social Security Administration derives from the fact that they are based on approximations of equivalent levels of living for all family types (taking account of farm-nonfarm residence) 1/ and that the cut-off points are applicable to Census current population income data and do provide a measure of how the income of the entire population relates to the defined social minimum. The poverty index cuts into the income distribution at a series of points representing equivalent positions of income adequacy. It thus makes possible not only a count of those above and below the defined level, but also a listing of their characteristics. It tells directly and in easily summarized fashion how many and who are "poor." It can be a stable measure (for short periods at least) and thus provide some indication of changes in the proportion of the population and of different groups affected by the general course of the economy or by specific social policies.

Analyses of the complete income distribution reveal some of the same relationships but from a different perspective. An example may illustrate this point. Analysis of income shares shows full-year rather than part-year employment to be one of the main determinants of personal income distribution, and hence of family income. The poverty analysis highlights a different fact--that a large number of full-year workers are nevertheless poor. Some are poor because they work in very low-paying occupations. More are poor--as defined--because of the size of their families.

An acceptable social minimum is obviously related to the general level of affluence of a society. In a dynamic economy it must therefore change over time. It is easy to reach agreement that what was an appropriate poverty measure in 1900 or 1933 is no longer relevant. It is also possible to get agreement that an acceptable social minimum in 1985 will be higher than today. It is difficult to find a satisfactory method of gradually moving the level up from its present to a hypothetical future position. The difficulty is not simply that of measuring progress in a war with an ever-escalating goal. There are other conceptual problems involved.

The issues involved go in part to the question of optimum allocation of resources at different levels of GNP. They relate to the relative growth of public expenditures for services and for particular types of services. They may involve the distinction between disposable and total money income. The weights for equivalencies may change with demographic changes and as the absolute level of per capita income rises.

Thus far the poverty index has been adjusted only for price changes. The need for a more substantial upward adjustment of the index level has been noted by many commentators. One solution that has much to recommend it would be two concurrent indexes. The SSA poverty and low-income indexes could be continued--adjusted only for changes in purchasing power--through say 1969. A second set of indexes could be adjusted to reflect productivity as well as price changes. The second set could start from the 1959 level, as do the present SSA indexes, or the divergence could start in 1963 or later. The result would be to provide a range of estimates of the number in poverty as well as a forceful demonstration of the way in which the composition of the poor varies depending on the measure of poverty used. The additional cost of the annual tabulations and the confusion in public debate that could result from two sets of figures would be among the disadvantages. The analysis of the two sets of data should provide a better basis for a decision in 1970 as to a new starting level and perhaps a different basis for adjustment over the following decade.

Income Shares

A poverty or low-income line can be misleading if it is treated as though it had some independent reality. It inevitably gives excessive weight to a single position on the income scale. Some of those classified as "poor" are very poor, others close to the line. Some of those whom we say with a sense of relief are above the minimum are hardly enough above to make any real difference. The near poor may be as important for social policy as the poor, particularly with relation to policies that determine income after taxes. Inequality cannot be left out of account as a concern of social policy.

Measures of the total income distribution remain an essential element in the set of interrelated measures that we need to understand changes in welfare. The over-all distribution of income shares is a gross measure which may exhibit great stability in the face of considerable change in the income position of components. For an understanding of trends and of the factors that affect social change, analyses of changes in income distribution within major segments of the population is also needed. Income redistribution analysis adds another dimension.

In a fuller discussion, one would examine the statistical measures that have been used to summarize the income distribution--and particularly the Lorenz Curve and the Gini Index of concentration. Suffice it to say that the usefulness of any such single measure of concentration or dispersion lies primarily in the possibility it offers of testing the effect on over-all inequality of a great variety of actual or hypothetical variations in income patterns. The potential contribution to income theory is greater than any direct use in measurement of welfare.

The major thrust of what I have been trying to say can perhaps be summarized as follows: the measurement of changes in welfare calls for a renewed emphasis on distributive justice and on income distribution; current money income itself is a better measure than our increasingly sophisticated and proper concern with other dimensions of control over resources might suggest; aggregate income distribution measures should be supplemented by a whole series of subsystem measures--for different income units, for age and sex cohorts, for place of residence, for different time periods and for different classifications of income, including the poverty-line type of classification; more important than "correcting" the money-income distribution is its interpretation in the light of as much detailed understanding of interrelationships as economic and social analysis can provide. However the task is formulated, there is much to do.

1/ For a description of the methodology, see articles by Mollie Orshansky in the Social Security Bulletin, January and July, 1965.

Discussion

Charles B. Nam, Florida State University

Within the past two or three years, we have witnessed considerable development of efforts to measure social change as a basis for public and private decision-making. Both through scholarly research and writings and through proposed federal legislation, we have observed a drive to develop a system of national social accounting comparable to that of economic accounting of the past two decades or so.

Where and how this concerted effort to establish a system of social indicators got new impetus is not entirely clear, but several recent additions to the research literature have focused on the topic. The 1966 volume entitled Social Indicators edited by Raymond Bauer, with contributions by Bertram Gross and Al Biderman, among others, provided frameworks for measuring the relative progress which the society was making toward the achievement of national goals and values. Two special issues of the Annals edited by Gross, with contributions by a number of persons in various areas of social life, helped to focus on measurement problems in specific social fields. A Russell Sage Foundation program, initiated by Wilbert Moore and Eleanor Sheldon and reported on at an earlier ASA meeting, will lead to another published volume on monitoring social change.

These scholarly approaches have been complemented by the legislative approach of Senator Mondale of Minnesota and ten of his senatorial colleagues who proposed a "Full Opportunity and Social Accounting Act", which provides for (1) an annual Social Report of the President, (2) a Council of Social Advisers, and (3) a Joint Congressional Committee on the Social Report (consisting of 8 Senators and 8 Representatives) to review the Report and transmit the findings to the Congress, each of these to parallel the present system of economic advice and reporting. Congressional hearings have been held on this proposal.

It is probable that both scholarly and legislative discussions of social accounting found root in the published series of social trends and indicators developed by the Department of Health, Education, and Welfare several years ago, and in the Social Indicators Panel of experts set up by that agency to advise it. Additional impetus for the movement has come from the National Commission on Technology, Automation, and Economic Progress which found that "our ability to measure social change has lagged behind our ability to measure strictly economic change," and from the President's call to HEW to "develop the necessary social statistics and indicators to supplement those prepared by the Bureau of Labor Statistics and the Council of Economic Advisers."

These several developments suggest that the time is perhaps ripe for a new era of broad social accounting similar to that of purely economic accounting in earlier years, and that it is only a matter of how quickly the necessary federal legislation will be passed and the machinery created to provide the basis for charting the direction and pace of social change in our society.

If, in fact, this new era is upon us, social scientists have little time in which to help decision-makers formulate a sound basis for indicating changes in the society. The term "social indicators" itself lacks common definition. On the one hand, it is used to refer to rates, ratios, and other indexes, singly or in combination, which provide signs of increasing or decreasing trends and stability or fluctuation of social phenomena. On the other hand, it is sometimes used to refer to the more complex multivariate models of change, providing not only signs of change but analysis of interrelationships of variables and of factors producing the change.

Both types of approaches have proved useful in tracking the state of the economy. The unemployment rate, GNP, interest rates, consumer price indexes, and other single measures have aided us in detecting when the economy was prospering or had weaknesses, and more detailed economic analyses of the labor force, the market place, and the monetary system suggested alternative avenues of action, if action was needed. With regard to other social phenomena, similar accounting would seem appropriate.

Several steps appear necessary in developing and employing social indicators. Enumeration of these steps and examination of the phenomena of social stratification and walfare in these terms provide one way of discussing two of the papers presented here.

1. Conceptualization of the phenomenon we want to measure. What is the phenomenon in question, and what do we want to know about it?

2. Operational indicators of the phenomenon. What specific items of information can we obtain to operationally measure the phenomenon? Are they valid measures of the phenomenon?

3. Collection of data on the operational indicators. Are they reliable? Are they available frequently enough? Are they collected at critical points in time?

4. Standardization of operational indicators. What are the appropriate statistics for comparing the indicators over time and among groups?

5. Analysis and interpretation of operational indicators. What amount of statistical

change or difference is required to speak of a "real" change or difference? What can we conclude about the operational indicators?

6. How should the operational indicators lead us to interpret the condition of the phenomenon in question, and what action, if any, should be taken on the basis of this?

Dudley Duncan's paper on the trend in social stratification does not come to grips with all of these questions. He first limits his discussion to certain aspects of the general area of social stratification, distinguishing between the static aspects of stratification--the range and distribution of social ranks, which he chooses not to deal with--and the dynamic aspects. On the dynamic side, he distinguishes between measuring the extent of social mobility in the society and the dependence of an individual's achievement on his social origins, the latter which he takes as his central analytical problem.

Professor Duncan would certainly find disagreement from some sociologists concerning his conceptual formulation, since there is still considerable sociological debate regarding the theoretical framework for viewing social stratification. Whether one wishes or not to call the hierarchy of social positions in a society "social stratification" is a matter of semantics. It is clear that the nature of inequalities in social ranks is a matter of social importance, as is the intergenerational transmission of status. In the one case, we are more concerned with the structure of the society; in the latter case, we are more concerned with individual opportunities to achieve. Recognizing this distinction, I would be quick to point out that the author is entitled to single out that aspect of the more general problem which he will treat in a relatively short paper.

With regard to operational indicators of status, Duncan does not attempt to make a case for choosing the variable "occupation." There is a vast literature which supports the choice of occupation as the best single indicator of social status, and yet there are many aspects of status which are imperfectly measured this way. For example, the very topical national issue about the distribution of power and opportunity to achieve it is certainly an element of the social stratification system but is not adequately indexed by occupational rank alone. In terms of availability and reliability of data, the selection of occupation as an operational indicator of one aspect of social stratification seems highly justified. What weaknesses there are in the data are specifically recognized by the author, and his evaluation of these data and attempt to standardize data from different sources are to be admired and emulated.

Duncan's analysis and interpretation of the data he is working with reveals high levels of sophistication and objectivity. His use of correlation and regression methods, and particularly of path analysis, permit him to reach reliable and meaningful conclusions about the trend in social stratification, at least of changes in the intergenerational transmission of occupational status.

In summary, Duncan's careful analysis of occupational data for fathers and sons enables him to reliably conclude that there has been no significant trends in the intergenerational transmission of occupational status. Trends in other aspects of social stratification remain open to question and need to be examined. More attention needs to be paid to problems in the development of operational indicators of social status, both of a substantive and methodological nature. It is no criticism of the author to point out these apparent weaknesses in his paper, since those who know his broader work in this area realize that he has been concerned with all of the issues mentioned. As a student of social stratification, I sleep better at night knowing that Dudley Duncan is devoting much of his energies to this subject.

As I read Dr. Merriam's paper, it seemed that "welfare," as she defined it, was not unrelated to "social stratification," in its broader context. Assuming this, it was interesting that she chose income, the economist's favorite, as the best indicator, in contrast to occupation, the sociologist's favorite.

Dr. Merriam was considerably vague in her conceptualization of general welfare and did not attempt to be convincing in her choice of income as an indicator of welfare. She does not deal with non-economic theoretical views of welfare, reflected in the recent statement of S. M. Miller and associates that "...income is only <u>one</u> of the dimensions of poverty and inequality today" and that there are "...six dimensions of well-being--income, assets, basic services, education and social mobility, political position, and status and self-respect."

Dr. Merriam does evaluate income data, including its relation to assets, but satisfies herself that the high correlations involved justify dealing with current money income alone as an indicator of welfare. She distinguishes between two major approaches to measuring income inequalities--the poverty-line approach and the income-shares approach--and properly identifies these as independently useful measures of separate but equally relevant national problems--of poverty and of the general distribution of income. Her paper does not specifically explore the matter of interpretation of data, nor does she, any more than Professor Duncan, discuss the transition from analysis of operational indicators to understanding of the general phenomenon in question.

In viewing systems of social indicators generally, several points, some of them made by other persons at other times, need emphasis.

1. We need to know what the indicators are for. Biderman's framework of assessing achievement of stated national goals appears to be a logical one and, if agreed upon as a framework, may engender more explicit statements about goals.

2. Distinction between relatively simple indicators of goal achievement and more complex understanding of change needs to be made and both need doing. The latter is crucial to guiding policy-makers to proper social action.

3. The development of new data sources, where information about assessment of goals is lacking, and greater flexibility in data-collection techniques, where some information is available but some is not, should be a concern of government leaders.

4. Problems of definition, analysis, and interpretation of critical variables require greater attention. Utility should take priority over tradition. Such data illnesses as "multiple seriosis" (or inconsistencies among different data sources in the purported measure of the same phenomenon), what Gross has called "hardening of the categories" (or rigidity in the operational measurement of phenomena regardless of utility), and "withdrawal symptoms" (or reluctance of data-gathering agencies to collect data in areas where politicians fear to tread) must be remedied.

Raymond Bauer has written that "For many of the important topics on which social critics blithely pass judgment, and on which policies are made, there are no yardsticks by which to know if things are getting better or worse." If policy-makers are now to adopt and institutionalize a system of social indicators, it is important that knowledgeable social statisticians make recommendations at the outset and continue to advise on modifications and additions. A good start has been made, but we must step up our activity and present our case lest policy-makers go it alone.

DISCUSSION

Sar A. Levitan
Center for Manpower Policy Studies, The George Washington University

Since this is a meeting of the American Statistical Association, it appears proper to discuss the papers in order of the authors' reliance on quantification. This arrangement also permits me to organize my comments in the proper order, ladies first.

Dr. Ida Merriam has established some kind of a record for an American Statistical Association session. There isn't a single number in her paper. She deals rather, with the broad concepts underlying social indicators. These measurements have vital social implications. Time permits comment only on one of her recommendations.

Few would quarrel with Dr. Merriam's observation that poverty indices need continued review and updating to reflect changes in costs of living as well as economic growth and productivity. Nonetheless, I find her proposal to raise the poverty indices developed by Mollie Orshansky and adapted as the "official" government measure denoting poverty disturbing. My objections stem from tactical and operational considerations. Rarely have a series of social indicators, as those developed by the Social Security Administration dealing with poverty, been as rapidly and widely accepted. The present data denoting poverty are based on cost of living measurements in 1959 and have been adjusted for increases in the CPI. Dr. Merriam's proposal would add increased productivity as a factor for adjustment in poverty income criteria and would raise the current income level denoting poverty by some 30 percent depending, of course, on the rate of productivity that would be used.

Strong pressures have already been exerted to raise poverty income criteria and to qualify more persons for participation in antipoverty programs. Thus far the antipoverty warriors have largely succeeded in withstanding these pressures and in restricting the limited resources allocated to these efforts to the 30 million who are now classified as poor. Raising the income criteria would qualify additional millions for these programs, limiting and possibly excluding large numbers of the "hard core" poor. Our social legislation is replete with measures which have been intended to help the poor but have benefited the more affluent. Raising the poverty income criteria would repeat this age old trend.

Aside from this very real operational problem, it also appears to me that it would be tactically wrong to confuse the public at present with a new set of criteria denoting poverty. Using two separate indices would also add to the confusion and to the pressures to qualify more persons for the antipoverty program.

If American Statistical Association papers are ever read, then Professor Stanley Lebergott's discerning paper should provide sufficient subject to dozens of new dissertations and monographs. It seems to me, however, that Lebergott takes his numbers too seriously. In reading his paper, it would be well to remember Henry Clay's admonition--I mean the economist, not the statesman--that "statistics are no substitute for judgment."

I am therefore not as disturbed as Professor Lebergott by the fact that labor force participation of 14 and 15 year-olds has not changed between 1920 and 1960. Assuming that the numbers are right, it is likely that "labor force participation" in 1920 was not the same as it was in 1960. In 1920 a farm boy "participating" in the labor force may have worked from sun-up to sundown. Forty years later the teen-age son of this same farmer was more likely to live in a city and his "participation" in the labor force may have been limited to a few hours of mowing lawns.

I find it more difficult to deal with Lebergott's findings that throughout this century, male family heads have earned 80 percent of family income. My difficulty comes from the fact that Lebergott challanges philosophers or moralists to explain this phenomenon. Being neither a philosopher nor a moralist, I would just suggest the possibility that if the data are correct, the results may be due to a statistical fluke, and leave it at that.

In the final part of his paper, Professor Lebergott treads new ground and raises some interesting questions about the rise of a supervisor "class" which he attributes to the increasing size of business units. He divides manufacturing employment into three groups: self-employed, foremen and employees. Ignoring the first group, he finds that the number of foremen per one hundred employees increased from 2.4 in 1910 to 4.2 in 1960. However, if Lebergott had included self-employed with foremen--both perform supervisory work--

then the trend in the rise of foremen which he stresses would not hold true. In 1910 there were 9.3 self-employed and foremen per 100 employees and in 1960 the proportion of self-employed and foremen to employees declined to 6.4 per hundred.

Professor Otis Dudley Duncan's paper is an elegant illustration of the manipulation of statistics and an attempt to quantify complex social phenomena. It is definitely an "in" paper--all numerology with a regression coefficient as a solution to all problems. The technique is skillful but it is not at all clear whether the results have any relation to reality.

IX

EVALUATION OF FAMILY PLANNING SERVICES IN THE U. S.

Chairman, ARTHUR A. CAMPBELL, National Center for Health Statistics

METHODS OF EVALUATING FAMILY PLANNING PROGRAMS--WITH SPECIAL REFERENCE TO NORTH CAROLINA (A Preliminary Report)

Elizabeth J. Coulter and B. G. Greenberg
University of North Carolina

The recent growth of family planning programs and the medical advances in contraceptive methods bring current needs for evaluating these programs in the community. Related needs for appropriate tools to evaluate family planning programs also arise from the general current interest in measuring effectiveness of health and other government programs in relation to costs to assist in allocation of scarce manpower and other resources.

The purpose of this paper is to discuss some methods of evaluating family planning programs particularly from the standpoint of approaches applicable in the United States. The evaluation process will be illustrated with data for the state of North Carolina which has a history of approximately thirty years of contraceptive services in public health programs.

Goals for Evaluation

The aim of evaluation, as considered in this paper, is to determine effectiveness of programs in meeting goals or objectives. The specific goals to be evaluated are likely to vary with the nature of individual programs and the communities they are established to serve.

The purposes of a family planning (contraceptive) service provided under the auspices of a health department were discussed in 1959 by a state health director and his staff in the article by Norton et al (1) with particular reference to North Carolina. They said that the "main purposes for contraceptive service are spacing or prevention of pregnancy for medical and/or socioeconomic considerations and control of excessive population growth." The authors pointed out that the main factors in establishing contraceptive service in North Carolina were medical ones aimed at improving maternal and child health including mortality and morbidity rates. In accord with this some countries (e.g. Chile) have established a national family planning service to counteract the rising mortality resulting from complications of induced abortions.

The Evaluation Process

Essentially the process of evaluation involves determination of program effectiveness in fulfilling goals or objectives. The use of statistical methods in the evaluation process should be discussed, however, in the framework of the whole process related to public programming. An isolated study of only those aspects of programming related to evaluation can give a distorted picture. It would, for example, be of little value to the community to measure the fulfillment of goals if the goals themselves were improperly chosen.

This paper will focus attention on four phases that can be followed in the use of statistics in continuous programming. These phases include diagnosis, measurement of services, evaluation, and cost-benefit analysis. The first phase, of diagnosis, is considered here as including program design and goal setting, which might also be discussed as a separate phase of the programming.

Phase 1: Diagnosis

An important first step in the initiation of a family planning program should be the development of a program design which can be used in allocation of personnel, facilities, and other program resources. This implies measurement of community needs as carefully and completely as circumstances warrant. Specific tools which can be used in measurement and diagnosis of community needs include population censuses and studies, vital statistics rates and ratios, and special surveys of knowledge, attitudes, and practices of the population. The statistical data developed by such means can help in instituting an appropriate program with realistic goals, and in planning various subsequent stages of program operation with appropriate guide-lines and indices to measure progress and development.

Phase 2: Measurement of Services

Periodic counts of services provided during the operation of a family planning program are needed to determine the extent to which targets are being met and to provide continuing information on the nature of the services furnished. Program service statistics can be prepared, for example, to show characteristics of persons receiving service, types of service provided, and the time and place of service. Such service statistics plus data on financial costs are useful in themselves for program field staff, administrators, and policy-makers; the service statistics also provide important denominators of inputs needed in the evaluation process discussed in Phase 4, cost-benefit analysis.

Phase 3: Evaluation

Evaluation should start in the early stages of the program by examination of program details likely to affect results. Early quasi-evaluation steps may involve, for example, consideration of staffing patterns, policy decisions, administrative organization, budgets, community interest, and other factors known to affect program outcomes. The crucial, ultimate questions that must be answered in evaluation relate, however, to the extent to which goals are accomplished, and whether such accomplishments are due solely to the program or to other community forces that cannot be controlled.

Phase 4: Cost-Benefit Analysis

The fourth phase of programming considers relationships of program gains (or losses) and benefits in accomplishing goals and program costs in terms of personnel, facilities, and other resources. Analyses of such cost-benefit relationships, often called input-output studies, essentially provide ratios of certain measures of accomplishments in phase 3 to selected indices or measures of service chosen from phase 2.

Diagnosis

Numerous types of information about the community are of potential use in setting goals of family planning programs and determining allocation of resources to meet these goals. Potentially useful data include, for example, statistics on personal and socioeconomic characteristics of the population; fertility patterns; knowledge, attitudes, and practices with respect to family planning and means of implementing it; morbidity and mortality rates, and geographic differences in the population characteristics. Some of the required information is usually readily available from such sources as population censuses, vital statistics, and records of health and other government agencies. Special surveys are likely to be necessary to obtain other types of needed information, especially on knowledge, attitudes, and practices with respect to family planning and means of implementing it.

Several examples relating to the population of North Carolina will illustrate uses and sources of various types of community data in setting goals of family planning programs and guiding their subsequent development. The total population of the state used for illustrative purposes was approximately four and a half million in 1960. Approximately one fourth of the population was nonwhite. (Table 1)

Data on personal characteristics and marital patterns of the population of North Carolina show important color differences in marital status of women in the child-bearing ages. Statistics from the U. S. Census bring out the fact that in 1960 relatively larger proportions of the nonwhite than of the white women of child-bearing age were single, e.g. 42.2 per cent of the nonwhite women and 22.6 per cent of the white women at ages 20-24 years. (Table 1) Such data suggest possible contraceptive needs for single as well as married women. These needs will appear in subsequent statistics on fertility patterns.

Major differences in socioeconomic characteristics of white and nonwhite persons in North Carolina also have important program implications. The 1960 Census statistics show, for example, median family income of $1992 for nonwhite and $4588 for white families, and higher unemployment rates among nonwhites than among whites in the labor force. (Table 1) Such data bring out the fact that considerably higher proportions of nonwhite than of white women can be expected to have need of public family planning services designed to provide free or low cost care for those of limited income.

Statistics on natality and fertility patterns in the population are of fundamental importance in development of family planning programs--as well as subsequent evaluation of their effectiveness. Data from 1960 vital statistics for North Carolina show, for example, that birth rates were relatively high among women 20-29 years of age. Color comparisons show that the birth rates were higher among nonwhite than among white women at each of the age groups from 15 through 44, and the excess among the nonwhites was relatively large at ages 15-19 and especially over 30 years. It is particularly interesting to note that the nonwhite women started their child-bearing at relatively younger ages than the white women even though their first marriage tended to occur at older ages, and that the nonwhites had shorter intervals than whites between births of lower birth orders. Illegitimacy rates were in turn considerably higher among nonwhites than among whites. Comparisons by birth order show that in 1960 the excess fertility among the nonwhites compared with the whites was concentrated at birth orders of three or more, and increased considerably in relative magnitude at higher birth orders. (Table 2)

The marriage, natality, and fertility patterns noted for North Carolina suggest several groups in the female population of child-bearing age with potential special needs for family planning service. Specific groups include, for example, mothers in ages of high fertility from 20 to 29 years who may wish help in spacing children, single nonwhite women, and women with several children.

Statistics on knowledge, attitudes, and practice with reference to family size and contraceptive use provide further indication of needs for family planning service, e.g. by giving information about population desires and attitudes with reference to changes in fertility patterns. Recent interview surveys of adult males and females in four low income areas in two North Carolina cities with active family planning (contraceptive) programs and nine low income areas located in four North Carolina counties without large cities but with public contraceptive programs a/ bring out the interesting fact that the mean ideal number of children for each color group and each area ranged from 2.7 to 3.7. The ideal number of children was in some instances lower for nonwhites than for whites even though a reverse fertility pattern has been noted by color. The proportion of respondents who approved of birth control was somewhat higher than the proportion disapproving in the urban study areas and in five of the nine more rural study areas - with a considerable number of respondents in unknown or "it depends" categories with reference to approval or disapproval. The proportion of respondents specifying approval of birth control generally increased with education. (2)

High overall rates of infant, perinatal and maternal mortality in earlier decades of this century as well as excessive rates of such mortality in nonwhite and low socioeconomic population groups have been important factors in the development of public family planning (contraceptive) services in North Carolina. (1) Recent data on infant mortality rates in North Carolina give reasons for continued concern. Statistics for the State for the period from 1960 through 1966 (Figure 1) show, for example, the following:

(1) Since 1960 infant mortality (both neonatal and postneonatal) has shown little decline in either the white or nonwhite races. This "may" suggest that family planning is needed to help raise the level of home environment conducive to survival.

(2) Among nonwhites the postneonatal contribution is almost as great as the neonatal portion. This definitely suggests that in the later period when the mother may be pregnant with her subsequent child she has limited time, energy, and resources to devote to the infant under one year of age. The question arises: can family planning help reduce the postneonatal mortality rates?

Additional statistics on infant and neonatal death rates by birth order, legitimacy status, and color for North Carolina for the years 1959-1961 show considerable excess in such mortality among high birth order, illegitimate, and nonwhite births. (Table 3) These data give further indications of special groups with potential needs for family planning services, such as mothers with several prior children.

Further classification of population and vital statistics by geographic area helps in more specific identification of population groups in need of family planning services. Classifications for North Carolina show, for example, higher income levels but lower fertility rates for metropolitan counties than for less urbanized areas of the State. (3), (Table 6) Such data also show that the illegitimacy problem among the whites is more concentrated in large cities than in other areas of the State while the illegitimacy problem for nonwhites is similar for the cities and the remainder of North Carolina. (4) Census tract statistics for individual cities, such as Charlotte and Raleigh, have also helped to identify low income problem areas of potential program concern.

Measurement of Services

Service statistics collected throughout the development of a family planning program are of considerable use in determination of the extent to which the program is reaching various groups in the community and in subsequent evaluation of program impact on fertility patterns and the general health and well-being of the population. Such service statistics can also be used to consider the most effective combination and location of personnel and material resources and in studying costs in relation to service patterns.

There are several types of potentially useful information about each person served in a family planning program. Specific examples include data on personal and socioeconomic characteristics of persons served, source of referral, prior contraceptive use, method of contraception prescribed, and dates of clinic visits or other contacts with patients. Additional data on personnel, material, and other resource requirements and their costs are also useful.

Several examples of service statistics obtained in public family planning (contraceptive) programs in North Carolina will illustrate some potential uses of such statistics. Historic statistics show that the annual reported numbers of patients receiving contraceptive service through health departments in the State fluctuated from approximately two to four thousand over the period from 1940 through 1963 and then increased sharply to reach 16,516 in 1966. The reported number of persons served has recently shown considerable concentration in several counties, notably Mecklenburg (which includes the city of Charlotte) and to a lesser extent Cumberland (including the city of Fayetteville), Durham, and Forsyth (with the city of Winston-Salem). (1,5)

The relatively old, large Mecklenburg program had a sharp increase in new admissions following the introduction of pills in late 1960 and the intra-uterine devices (IUDs) in 1964. (Table 4) A peak for new admissions was, however, reached in 1965 and a marked decline in new admissions has since occurred. The recent decline in new admissions raises questions about numbers, location, and characteristics of persons interested in family planning who have not already been reached. It is of interest to note in this connection that the Mecklenburg Welfare Department has a program of using homemakers, commonly called "doorknockers", who "have been active since July 1, 1964 in canvassing low socioeconomic status neighborhoods and in following individual leads to women who might benefit from family planning". (6)

Recent service statistics for new admissions to family planning programs operated by health departments in three urban counties in North Carolina show several indications that the programs have reached population groups of special need. The data are for the Mecklenburg program (for the period November 1960 to July 1966), the

Cumberland program (for the period August 1963 to July 1967) and a relatively new program in Wake County, which includes the city of Raleigh (for the period March 1966 to June 1967). The data for the three programs (Table 5) show specifically that:

(1) All three programs had high proportions of nonwhite patients--ranging from 80 per cent for Mecklenburg to about two-thirds for Cumberland.

(2) Each of the programs admitted primarily young women in the ages of relatively high fertility, i.e. under 30 years.

(3) All three programs admitted considerable numbers of single women--following the birth of the first child. Thirty-eight per cent of the nonwhite admissions to the Wake County program were single while 27 per cent of the nonwhites admitted to the Mecklenburg program were single.

It is also of interest to note that one-third of the admissions to the Wake County program were on welfare funds while most of the admissions to the Mecklenburg program were from low income census tracts in the city. (7, 8)

Data by age and parity for white and nonwhite admissions to the Mecklenburg and Wake County programs also suggest some differences in program emphasis in reaching population groups of particular need. (Table 5) Comparisons of the nonwhite admissions to the two programs show, for example, relatively high proportions of the 15-19 year old and single parity women in the Wake County program. The Mecklenburg program included relatively high proportions of nonwhites in the ages from 20 through 29 years and at several of the higher parities. It is also relevant to point out that there was a marked decline in age and parity of new admissions to the Mecklenburg program over the study period from November 1960 to July 1966 (6,8). Hence the two programs are currently more similar than would appear from an analysis only of Table 5.

Comparisons of the Mecklenburg, Wake, and Cumberland programs also show a difference in the method of contraception used, which may affect the length of time patients continue contraceptive usage. The Mecklenburg program has relied heavily on the pills, prescribed for 65 per cent of the white and 80 per cent of the nonwhite new admissions, for whom pills or IUDs were prescribed, over the period from November 1960 to July 1966. (8) The IUDs were used almost four times more frequently than the pills for new admissions to the Wake County program from March 1966 through May 1967 (7), and the Cumberland program has recently used the IUDs almost exclusively. The higher usage of pills in Mecklenburg is not one of time although the pill was available earlier (Table 4).

Follow-up data showing the extent to which persons admitted to a family planning program subsequently drop out for such reasons as pregnancy, discontinued interest, death, or migration from the community are also important for administrative purposes as well as subsequent evaluation of program impact on fertility rates. Illustrative statistics for 458 women fitted with IUDs in the Mecklenburg IUD program in 1964 and followed to mid-1966 show, for example, that 88.9 per cent were still active in the IUD program six months after enrollment in it and 79.5 per cent were active at the end of 12 months. The subsequent six months showed a smaller rate of decline and about 74.7 per cent were still active users after 18 months. b/ (9)

Evaluation

Three general types of measures can be used to evaluate effectiveness of family planning programs at various lengths of time after the initiation of services. Specifically it is of interest to consider:

(1) Early detectable changes in knowledge, attitudes, and practices with reference to contraception and family planning service.

(2) Subsequent effects on natality and fertility patterns, by such classifications as color, age, and parity.

(3) Possible long-range benefits of improved physical or mental health of the mother; lower divorce rates, reduced prematurity and illegitimacy rates; lower maternal, infant, neonatal or fetal mortality rates; higher levels of income and employment, and lower rates of juvenile delinquency.

Changes in knowledge, attitudes, or practices may occur within the first year of the program. Changes in fertility patterns may take up to several years to appear, while some of the long-range benefits of improved health and well-being may not appear for a number of years.

Community surveys at two or more points in time provide a potential means of determining changes in knowledge, attitudes, and practices arising from the initiation of family planning programs. Selected data from clinic records of family planning programs also provide some clues to changes in knowledge, attitude, or practice. Data from the Mecklenburg program clinic records for newly admitted patients over the period from November 1960 to July 1966 show, for example, that only two per cent of the nonwhite and nine per cent of the white patients who accepted pills as a method of contraception through the program had previously used pills. About 30 per cent of the white and 23 per cent of the nonwhite new admissions for whom pills or IUDs were prescribed by the program over the period from November 1960 to July 1966 reported previous use of condoms. (8)

Two different types of approaches can be taken in considering effects of family planning programs on fertility. One approach is to consider use effectiveness of contraceptives among program participants themselves. (10, 11) This involves following women accepting a contraceptive over time to determine periods of protection or exposure and rates of pregnancy or discontinued use for this reason. Another approach to the effect on births is to study various measures of natality and fertility in the community. The latter is, of course, a less sensitive indicator since the denominator includes many nonparticipants. There is a related need in both types of approaches to consider changes in fertility that might have occurred in the absence of the family planning program. For this reason simultaneous control groups or communities are frequently used in such studies.

Fertility rates in the state used for illustrative purposes, North Carolina, declined over the period 1961 through 1966. The decline has been greater in high than in low birth orders, especially among the nonwhites. (12) (Tables 2 and 6) Further study is needed with reference to the possible role of public family planning programs, privately obtained contraceptives, and other factors in the communities of the State in the fertility decline.

Table 6 provides fertility rates for the years 1963-1966 for three metropolitan counties in the State: Mecklenburg and Wake, with family planning programs already discussed, and Guilford, which includes the cities of Greensboro and High Point and has little public family planning service. A considerable drop occurred in the nonwhite fertility rate among residents of Mecklenburg County in 1965 and was followed by a somewhat smaller decline in fertility for this population group in 1966. Neither of the other two counties showed the drop in nonwhite fertility rates in 1965 of the same magnitude as found for Mecklenburg. The white fertility rate for Mecklenburg also failed to show as much decline as the nonwhite rate in 1965 although an interesting earlier decline in the white rate occurred in 1964. As previously noted, there was a sharp increase in new admissions, predominantly nonwhite, to the Mecklenburg family planning program in 1964 and some further increase in new admissions to the family planning program occurred in 1965. (Table 4)

Nonwhite birth rates for Mecklenburg County show that age groups from 20 to 29 years contributed particularly to the decline in nonwhite fertility rates in the county from 1964 to 1965. The decline in birth rate was approximately 22 per cent in each of the fertile age groups of 20-24 and 25-29 years. Smaller declines occurred at ages 30-34 and 15-19 years (18 and 12 per cent, respectively), while little change occurred in birth rates at ages 35-44 years. (Table 7) It is especially interesting to note in studying the data on birth rates by age group that over half of the nonwhite new admissions to the Mecklenburg family planning program in each of the years 1961 through 1965 were from age groups of 20-29 years. (6, 8)

Readily available illegitimacy statistics for North Carolina provide one means of evaluating possible long-term effects of family planning services on the health and well-being of the population. Dr. Elizabeth Corkey noted in a recent article (13) about the Mecklenburg program that a small relative decline in illegitimate births beyond the first occurred for Mecklenburg County during the period 1955-1964--from 49.7 to 44.6 per cent of all illegitimate births. Data on nonwhite illegitimate births for Mecklenburg County residents over the years 1963 to 1966 also show a recent decline in illegitimate births of higher order, especially of five or more, in contrast to an increase of illegitimate first births in 1964 and again in 1966. (Table 8) It is of special interest to note here that the Mecklenburg family planning program has not accepted single women unless they have had at least one birth. It is also relevant that the health and welfare agencies in Mecklenburg County have arrangements to try to reach and discuss possible family planning or other needs of mothers who have had over three illegitimate births, under a state law requiring special consideration of these mothers from the standpoint of the well-being of the children.

Perinatal and late infant mortality rates have also been studied in North Carolina with reference to possible benefits of family planning services. Norton et al, in 1959, noted (1) considerable reduction in perinatal and late infant mortality rates in the State, as well as sharp reductions in maternal mortality rates, in studying the period of the first twenty-one years of experience with a public contraceptive service in North Carolina. They pointed to difficulties in determining the precise role which the contraceptive service of itself exercised in the decline but suggested that such service was nevertheless contributory.

Recent perinatal and postneonatal mortality rates are available in Table 9 for the three urban North Carolina counties for which fertility rates were previously considered: Mecklenburg, with the large family planning program serving particularly nonwhites; Wake, with a relatively newer and smaller program, and Guilford, which does not have a large public family planning program. It is particularly interesting to note that the nonwhite perinatal mortality rate for Mecklenburg County showed a marked decline in 1965 (the same year that the nonwhite fertility rate dropped considerably for the county) and continued to be relatively low in 1966. The white perinatal mortality rate of the county, on the other hand, showed little change over the years 1963-1966. The nonwhite perinatal mortality rate of Wake County also generally showed little change over the period while the nonwhite rate for Guilford has shown some recent decline. Therefore, further study of perinatal and postneonatal death rates within census tracts or neighborhoods of urban counties would be of potential interest, particularly if these can be related to the areas in which the program participants reside. Previously noted high infant death rates among illegitimate and higher order births (Table 3) also suggest that it would be especially interesting to consider possible effects of reduced fertility rates that may occur in these high risk groups on infant mortality rates.

Cost-Benefit Analysis

Analyses of relationships of costs and benefits of family planning programs provide evaluation tools which are particularly important in view of scarcities of human and physical resources to meet the many types of needs in our society. Such analyses include studies of program outputs in fulfilling goals relative to program inputs of personnel, facilities, and other resources. Some of the costs and benefits can be measured in monetary terms; others may require alternative measures, such as number of specified services.

Numerous types of analyses can be made to relate costs and benefits of family planning programs. Administrative studies can, for example, be made to determine best or least expensive combinations of program resources in producing specified results, such as months of protection. Related consideration can also be given to costs of various types of care to the patient, e.g. for transportation to a clinic. There also are many possibilities for study of program benefits or outputs, e.g. in improved health of the mother and child or reduced social problems in the community, in relation to the program costs.

Administrative studies can be made to consider both clinic and other program inputs in relation to specific benefits for the patients. Comparisons of clinic costs per months of protection can, for example, be made by individual clinic, single vs. dispersed clinics, year of program operation, type of contraception prescribed, various combinations of professional and nonprofessional staff, and different combinations of personnel time and educational materials. Other costs to be considered include personnel time required to inform potentially interested persons of available services, and services needed to follow up persons who drop out of the program.

Several types of costs which the patient may incur in coming to the clinic are also relevant. These costs include expenses of transportation, time required to come to the clinic, and associated expenses of child care or hours lost from work. Such costs need to be considered in relation to outlays of funds for the program itself, and to attrition rates. Extra program costs for initial or early clinic visits may, for example, be justifiable if they help reduce costs to the patients and subsequently reduce drop-outs from the program.

The scope of the potential benefits of family planning programs that have already been discussed present many challenges for determination and measurement of program outputs in relation to given inputs or costs. Indices of changes in fertility, improved health or reduced disability, higher levels of education, reduced juvenile delinquency, and other social and community benefits provide possible measures of a non-monetary nature that can be studied. Some of these benefits can also be related to potential economic savings or returns for the population and the community. Reduced disability provides, for example, potential for more productive activities, such as household duties of the mother, and lower medical expenses; improved general health and education levels of the family members give potentials for better employment and in turn higher income; and reduced infant or maternal mortality help save infants and mothers for subsequent years of productive life. Approaches taken by Dublin and Lotka in their book on The Money Value of a Man and more recently by writers such as Fein, Weisbrod and Rice (14-17) present possibilities for measuring economic costs of premature death prior to productive years, lost productivity or income due to disability, and expenses of medical care which are of potential use in measuring economic benefits of family planning, as well as other health services.

This discussion of cost-benefit approaches is intended only to suggest some of the possibilities for applications in family planning programs. The authors are pursuing further work on cost-benefit analyses and expect to publish the results subsequently. An example from the Mecklenburg County family planning program will illustrate, in a preliminary way, a possibility for studying service required in relation to months of protection from pregnancy. Data for 447 women fitted with IUDs in the program in 1964 and remaining active for at least six months c/ show that they had an average of 4.7 contacts with the program, by clinic visit, home visit, or phone call, during the first six months in the IUD program. The average number of program contacts was higher than the regularly scheduled three clinic visits in the first six months for patients for whom IUDs were inserted in 1964 d/, and is higher than one would expect in subsequent months after insertion of the IUD. Additional data for the group of 447 women fitted with IUDs in 1964 and remaining in the program for at least six months show that those who became inactive before July, 1966, had more program visits in the first six months than those who remained active to July, 1966. Specifically, one fourth of the 447 women who became inactive had an average of 5.3 program contacts in the first six months compared with an average of 4.5 program contacts for the three-fourths who remained active. (9) Such data suggest possibilities for further study of visits and costs for family planning program participants, who can in turn be classified by characteristics, type of contraceptive, and length of participation in the program.

Conclusion

The potential relationship of family planning programs to many facets of the complex, ever changing American society present many challenges for continuing development of techniques for evaluating such programs. The challenges for appropriate methodologic approaches relate to the various steps in the programming process from collection of information on characteristics of the community population through study of program services and subsequent consideration of program accomplishments--as well as their relation to program costs. It will be especially useful, at this stage, to have further experience with applications of the evaluation techniques in various types of communities in the United States and careful analyses of results of such experience.

Footnotes

[a/] The survey population included half of the adult household members of both sexes who were 18 years of age or older. The adults to be interviewed were selected at random in each household in a sample consisting of at least 400 households in most of the study areas.

[b/] The group of 458 women consisted of all those fitted with intra-uterine devices during the first year (1964) these devices were provided by the program, except for those served in a small clinic in an outlying area of the county and a few for whom limited information was available. Approximately 80 per cent of the 458 women were new admissions to the Mecklenburg program in 1964. The group of women studied had slightly higher drop-out rates, due to accidental pregnancies, than one might generally expect recently because of the use of small IUDs later discontinued by the program.

[c/] The group of 447 women consisted of the 458 women to whom reference is made in footnote b with the exception of 11 women lost to follow-up by the program in less than six months.

[d/] Subsequently the program has reduced the number of scheduled visits in the first six months to two--an initial and one follow-up visit.

References

1. Norton, J. W. R., Donnelly, J. F., and Lamb, A., Twenty-One Years Experience with A Public Health Contraceptive Service, American Journal of Public Health, 49 (1959), 993-1000.

2. North Carolina Fund, The North Carolina Fund Survey of Low Income Families in North Carolina, Reports on Characteristics of Individuals, 1967, 3b, pp. 37-39; 3c, pp. 43-45; 3e, pp. 41-42; 3f, pp. 41-42; 3g, pp. 46-48.

3. U. S. Department of Commerce, Bureau of the Census, U. S. Census of Population: 1960, General Social and Economic Characteristics, North Carolina, Final Report PC (1) - 35C, pp. 216-220 and 276-284, U. S. Government Printing Office, Washington, D. C.

4. North Carolina State Board of Health, Epidemiology Division, Annual Report of Public Health Statistics Section, 1960, Part 2, pp. 97, 99, 101, 103, and 105, Raleigh, North Carolina.

5. Special tabulations provided by the North Carolina State Board of Health.

6. Siegel, E., Tuthill, R., Coulter, E., Chipman, S., and Corkey, E., A Longitudinal Assessment of A Community Family Planning Program, accepted for publication in the American Journal of Public Health.

7. Omran, A. R., Arnold, C. B., Wells, H. B., and Bethel, M. B., Selected Demographic Data, Wake County, North Carolina. (Mimeographed)

8. Special tabulations prepared as part of a study by the Department of Maternal and Child Health and the Carolina Population Center of the University of North Carolina, with the cooperation of the Mecklenburg County Health Department.

9. Talwar, P. P., Family Planning through IUD Clinics in the Mecklenburg County Health Department in the Year 1964, summer field training report prepared in work with Dr. Elizabeth C. Corkey and Mrs. Marian Haslam of the Mecklenburg County Health Department.

10. Potter, R. G., Jr., Additional Measures of Use Effectiveness in Contraception, Milbank Memorial Fund Quarterly, XLI (1963), 400-418.

11. Tietze, C., Intra-Uterine Contraception: Recommended Procedures for Data Analysis, Studies in Family Planning 18 (Supplement), (1967), 1-6.

12. (a) North Carolina State Board of Health, Epidemiology Division, Annual Report of Public Health Statistics Section, 1960, Part 2, p. 15; 1961, Part 2, p. 15; 1962, Part 2, p. 15; 1963, Part 2, p. 15; and 1964, Part 2, p. 15, Raleigh, North Carolina.

    (b) North Carolina State Board of Health, Division of Epidemiology, Public Health Statistics Section, North Carolina Vital Statistics, 1965, Part 2, p. B-7; 1966, p. 50, Raleigh, North Carolina.

13. Corkey, E. C., The Birth Control Program in the Mecklenburg County Health Department, American Journal of Public Health, 56, Part II, (1966), 40-47.

14. Dublin, L. S. and Lotka, A. J., The Money Value of A Man, The Ronald Press Co., New York, 1946.

15. Fein, R., Economics of Mental Illness, Basic Books, New York, 1958.

16. Weisbrod, B. A., Economics of Public Health, University of Pennsylvania Press, Philadelphia, 1961.

17. Rice, D. P. and Cooper, B. S., The Economic Value of Human Life, American Journal of Public Health, 57 (1967), 1954-1966.

TABLE 1. - Selected Characteristics of the White and Nonwhite Population of North Carolina, 1960

Population Characteristics	White	Nonwhite
Total Population a/	3,399,285	1,156,870
Distribution by Color a/	74.6	25.4
Proportion in Ages Under 18 Years a/	35.0	46.1
Proportion of Women Who Were Single (Not Previously Married) at Ages b/		
15-19	78.4	87.6
20-24	22.6	42.2
25-29	7.6	19.8
30-34	5.0	12.2
35-39	5.1	9.1
40-44	5.1	8.0
Median Family Income c/	$4588	$1992
Per Cent of Civilian Labor Force Unemployed d/		
Male	2.9	6.0
Female	4.9	9.6

Sources:

a/ Based on data in U.S. Department of Commerce, Bureau of the Census, U. S. Census of Population: 1960, General Population Characteristics, North Carolina, Final Report PC (1) - 35 B, pages 38 and 39. U. S. Government Printing Office, Washington, D. C.

b/ Based on data in U. S. Department of Commerce, Bureau of the Census, U. S. Census of Population: 1960, Detailed Characteristics, North Carolina, Final Report PC (1) - 35 D, pages 363-365, U. S. Government Printing Office, Washington, D. C.

c/ U. S. Department of Commerce, Bureau of the Census, U. S. Census of Population: 1960, General Social and Economic Characteristics, North Carolina, Final Report PC (1) - 35 C, page 181, U. S. Government Printing Office, Washington, D. C.

d/ Ibid., page 172.

TABLE 2. - Selected Natality and Fertility Rates Among White and Nonwhite Female North Carolina Residents 15-44 Years of Age, 1960

Population Characteristics	White	Nonwhite
Age Group	Age Specific Birth Rates, Ages 15-44 a/	
Total Ages		
15-44	101.8	148.6
15-19	90.4	140.7
20-24	220.0	281.4
25-29	161.4	215.3
30-34	85.8	142.0
35-39	41.8	80.8
40-44	11.4	23.9
Live Birth Order	Fertility Rates by Live Birth Order 1/ a/	
All Birth Orders	101.8	148.6
1	30.6	29.8
2	26.9	24.9
3	18.8	20.3
4	11.5	17.1
5	6.4	14.4
6	3.4	11.3
7 and over	4.3	31.0
Live Birth Order	Mean Age of Mother at Specified Live Birth Order b/	
1	21.6	19.5
2	24.6	21.8
3	27.1	24.0
4	29.0	25.8
5	30.5	27.5
6	31.6	29.2
Total Births	Percentage of Live Births That Were Illegitimate c/	
Total	2.3	24.5

1/ Statistics include fetal deaths in determining birth order. Fetal deaths accounted for approximately 1.5 per cent of the white deliveries and about 3 per cent of the nonwhite deliveries.

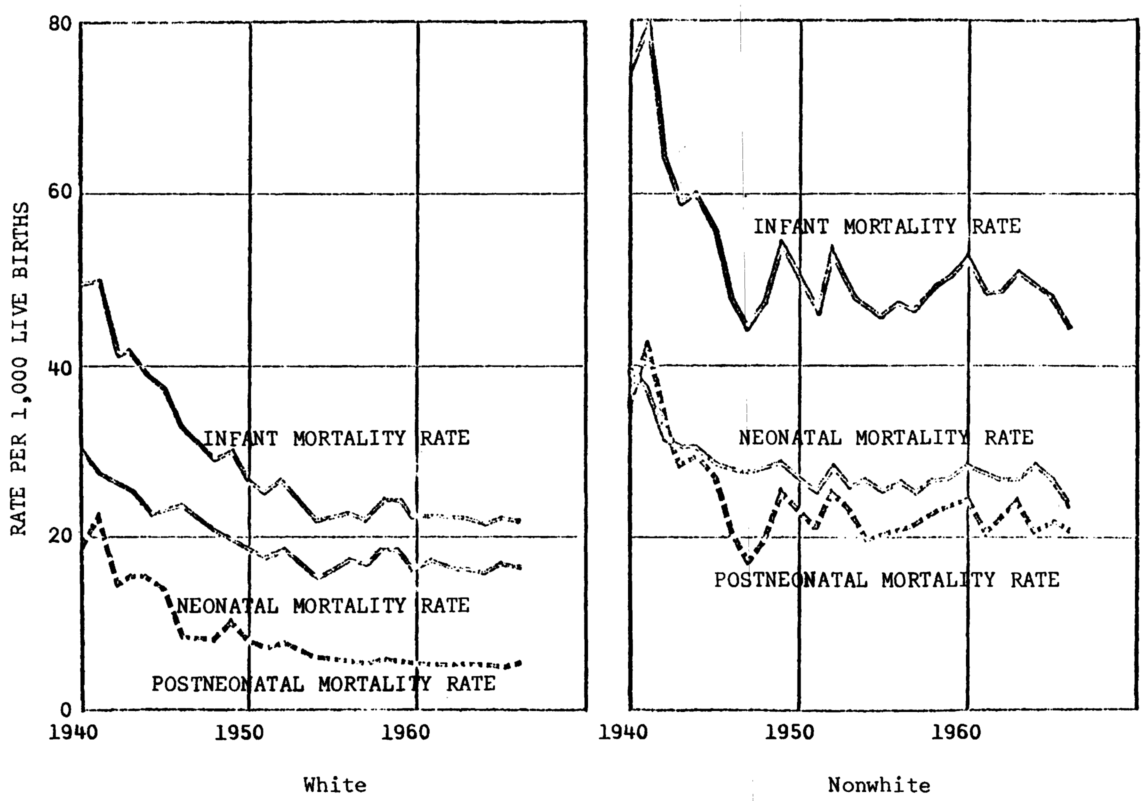
Sources:

a/ Based on natality data in North Carolina State Board of Health, Epidemiology Division, Annual Report of Public Health Statistics Section, 1960, Part 2, page 15, Raleigh, North Carolina, and population data in U. S. Department of Commerce, Bureau of the Census, U. S. Census of Population: 1960, General Population Characteristics, North Carolina, Final Report PC (1) - 35 B, pages 39 and 40, U. S. Government Printing Office, Washington, D. C.

b/ Based on data in North Carolina State Board of Health, Epidemiology Division, Annual Report of the Public Health Statistics Section, 1960, Part 2, page 15, Raleigh, North Carolina.

c/ Based on data in North Carolina State Board of Health, Epidemiology Division, Annual Report of Public Health Statistics Section, 1960, Part 2, page 97, Raleigh, North Carolina.

Figure 1. Infant, Neonatal, and Postneonatal Death Rates per 1,000 White and Nonwhite Live Births, North Carolina, 1940-1966.



Source: Reproduced from North Carolina State Board of Health, Division of Epidemiology Public Health Statistics Section, North Carolina Vital Statistics, 1966, page 4, Raleigh, North Carolina.

TABLE 3. - Infant and Neonatal Death Rates per 1000 Live Births by Birth Order, Color, and Legitimacy Status, North Carolina Residents, 1959-1961 1/

Birth Order	Infant Deaths		Neonatal Deaths	
	Legitimate	Illegitimate	Legitimate	Illegitimate
	White			
1	18.4	43.8	14.5	28.4
2	21.6	69.5	16.9	43.9
3	21.2	88.1	16.6	51.1*
4-5	25.0	76.2	18.3	24.4*
6-7	32.4	98.4*	24.6	49.2*
8-9	34.9	155.6*	24.2	111.1*
10+	40.5	200.0*	26.0	133.3*
	Nonwhite			
1	38.1	51.7	24.9	28.6
2	45.0	64.6	25.1	32.1
3	44.5	74.5	24.1	36.7
4-5	45.1	71.6	23.6	35.8
6-7	49.1	80.8	27.9	33.1
8-9	52.3	91.2	31.6	44.8
10+	56.0	104.3	34.0	55.2*

1/ Fetal deaths were included in determining birth order. The fetal deaths accounted for approximately 1.5 per cent of the white deliveries and approximately 3 per cent of the nonwhite deliveries.

* Rates are based on less than 20 deaths.

Source:

Based on data in: Suksawasdi, R.; The Study of Probability of Survival of Legitimate and Illegitimate Infants Born to Residents of North Carolina, 1959-1961, summer field training report prepared in work with Mr. Glenn A. Flinchum and Mr. Bradford W. Johnson of the North Carolina State Board of Health.

TABLE 4. - Total Number of New Admissions to the Mecklenburg County Family Planning Clinic for Whom Pills and Intra-Uterine Devices Were Prescribed, November 1960 through December 1967 1/

Year	Total	Pills	Intra-Uterine Devices
Total	4514	3493	1021
1960 (October and November) and 1961	89	89	--
1962	212	212	--
1963	314	314	--
1964	1054	689	365
1965	1250	912	338
1966	754	612	142
1967	841	665	176

1/ The data exclude admissions to a small clinic in an outlying area of the county. Services in this clinic are omitted throughout this paper.

Source:

Data are from the following article: Siegel, E., Tuthill, R., Coulter, E., Chipman, S., and Corkey, E., A Longitudinal Assessment of A Community Family Planning Program, accepted for publication in the American Journal of Public Health.

TABLE 5. - Age, Parity, and Marital Status Distribution of White and Nonwhite New Admissions to Family Planning Programs in Mecklenburg, Wake, and Cumberland Counties, North Carolina, Specified Years in the 1960's 1/

Characteristics of New Admissions	Mecklenburg Program, November 1960-June 1966 2/ White	Mecklenburg Program, November 1960-June 1966 2/ Nonwhite	Wake Program, March 1966-May 1967 White	Wake Program, March 1966-May 1967 Nonwhite	Cumberland Program, August 1963-June 1967 3/ White	Cumberland Program, August 1963-June 1967 3/ Nonwhite
	Number of New Admissions					
Total	663	2677	85	230	537	1036
	Per Cent					
Age Group						
Total	100.0	100.0	100.0	100.0	100.0	100.0
Under 15	0.0	0.6	0.0	2.2	0.4	1.3
15-19	22.2	22.2	17.6	30.9	10.9	22.2
20-24	37.0	36.2	33.0	32.6	34.7	37.7
25-29	23.5	21.8	28.2	14.3	27.9	21.6
30-34	9.6	10.8	9.4	11.3	15.7	11.6
35-39	5.3	6.5	5.9	7.0	7.0	4.3
40 and over	2.4	1.9	5.9	1.7	3.4	1.3
Mean Age	24.8	24.9	26.0	24.0	26.6	24.4
Parity 4/					All Colors Combined	
Total	100.0	100.0	100.0	100.0	100.0	
0	3.6	1.2	2.4	1.3	1.0	
1	20.5	19.5	23.5	30.3	17.0	
2	25.4	20.8	24.6	18.3	20.2	
3	26.9	18.1	22.4	16.1	19.1	
4	12.3	13.9	15.3	9.6	14.5	
5	6.6	9.3	4.7	8.3	11.1	
6	2.6	6.4	1.2	4.8	6.0	
7	0.9	4.2	3.5	2.6	3.9	
8 or more	1.2	6.6	2.4	8.7	7.2	
Marital Status					All Colors Combined	
Total	100.0	100.0	100.0	100.0	100.0	
Single	2.0	27.0	5.9	37.8	16.8	
Married	93.1	57.5	91.8	54.3	82.3	
Widowed	0.2	1.3	0	4.8	0.1	
Divorced or Separated	4.7	14.2	2.3	3.1	0.8	

1/ Small numbers of patients of unknown age, parity, or marital status were excluded in calculating percentage distributions for specific characteristics for which unknowns appeared.

2/ Data are limited to patients for whom pills or IUDs were prescribed, which results in exclusion of a very small number of new admissions to the program.

3/ The nonwhite admissions included 43 Indians. There were 21 admissions of unknown color which were excluded in calculating the distribution by age; tabulations by color were not available for the parity and marital status distributions.

4/ Somewhat different procedures were used in determining "parity" for the individual programs. A total of 163 premature births as well as the abortions and stillbirths were excluded for the Mecklenburg program; data for the Wake County program are for living children, and statistics for the Cumberland program include prior fetal deaths as well as live births in determining parity.

Sources:

Statistics for the Mecklenburg program are from special tabulations prepared as part of a study by the Department of Maternal and Child Health and the Carolina Population Center of the University of North Carolina with the cooperation of the Mecklenburg County Health Department. Data for the Wake County program are from a mimeographed report: Omran, A. R., Arnold, C. B., Wells, H. B., and Bethel, M. B., Selected Demographic Data, Wake County, North Carolina. Statistics for the Cumberland program are from special tabulations provided by the North Carolina State Board of Health.

TABLE 6. - Fertility Rates per 1000 White and Nonwhite Females 15-44 Years of Age, North Carolina and Selected Counties, 1963-1966 1/

Year	Geographic Area: North Carolina	Mecklenburg County	Guilford County	Wake County
	White			
1963	97.0	91.8	85.7	89.0
1964	94.8	84.1	81.5	91.0
1965	86.1	76.2	75.5	78.4
1966	82.7	73.7	71.6	78.0
	Nonwhite			
1963	143.7	139.2	125.9	138.6
1964	142.9	139.2	126.0	130.1
1965	133.4	115.9	120.2	129.6
1966	120.9	107.1	108.6	118.3

1/ The small numbers of live births to mothers under 15 and over 44 years of age were included in calculating the fertility rates.

Sources:

a/ Natality data are from the following sources:

(1) North Carolina State Board of Health, Epidemiology Division, Annual Report of Public Health Statistics Section, 1963, Part 2, pages 95, 99, 101 and 105, Raleigh, North Carolina.

(2) North Carolina State Board of Health, Epidemiology Division, Annual Report of Public Health Statistics Section, 1964, Part 2, pages 97, 101, 103, and 107, Raleigh, North Carolina

(3) North Carolina State Board of Health, Division of Epidemiology, Public Health Statistics Section, North Carolina Vital Statistics, 1965, Part 2, pages B-4 and B-5, Raleigh, North Carolina.

(4) North Carolina State Board of Health, Division of Epidemiology, Public Health Statistics Section, North Carolina Vital Statistics, 1966, pages 42, 44, 45, and 47, Raleigh, North Carolina.

b/ Population estimates used in computing rates were made by linear interpolation from numbers of women by color and age as given in the 1950 and 1960 United States Censuses. Formulas used for linear interpolation were provided by Dr. C. Horace Hamilton and are discussed in the article: Hamilton, C. H. and Perry, J., "A Short Method for Projecting Populations by Age from One Decennial Census to Another," Social Forces, 41 (1962), 164-170.

TABLE 7. - Age Specific Birth Rates per 1000 Nonwhite Women 15-44 Years of Age, Mecklenburg County, North Carolina, 1963-1966

Year	Total 15-44	Age Group 15-19	20-24	25-29	30-34	35-39	40-44
1963	136.8	147.5	253.3	190.4	106.9	67.6	16.6
1964	136.8	166.1	260.7	165.9	102.9	52.8	15.1
1965	115.3	145.5	202.6	130.4	84.8	49.0	16.0
1966	105.1	141.2	187.2	113.4	73.6	37.8	8.1

Source:

Natality data are from special tabulations made available by the Mecklenburg County Health Department and the North Carolina State Board of Health. Rates were computed on the basis of population estimates prepared as described in source reference b of table 6.

TABLE 8. - Birth Order Distribution of Nonwhite Illegitimate Live Births to Mothers 15-44 Years of Age, Mecklenburg County, North Carolina 1963-1966

Year	Total	Birth Order 1	2	3	4	5	6	7	8 or more
	Per Cent								
1963	100.0	42.6	24.5	11.9	7.4	4.2	3.8	2.0	3.6
1964	100.0	48.7	23.6	11.9	6.6	3.3	2.0	1.2	2.7
1965	100.0	50.5	22.1	12.2	6.7	3.3	1.8	1.1	2.3
1966	100.0	55.0	24.6	10.7	5.2	2.2	1.1	0.6	0.6
	Number								
1963	554	236	136	66	41	23	21	11	20
1964	664	323	157	79	44	22	13	8	18
1965	612 1/	309	135	75	41	20	11	7	14
1966	635	349	156	68	33	14	7	4	4

1/ Excludes one birth of unknown birth order.

Source:

Natality data are from special tabulations made available by the Mecklenburg County Health Department and the North Carolina State Board of Health.

TABLE 9. - Perinatal 1/ and Postneonatal 2/ Death Rates by Color, North Carolina and Selected Counties in the State, 1963-1966

Year	Perinatal Mortality				Postneonatal Mortality			
	North Carolina	Mecklenburg County	Guilford County	Wake County	North Carolina	Mecklenburg County	Guilford County	Wake County
	White							
1963	29.9	28.9	34.1	21.4	5.4	2.9	4.7	2.6
1964	29.3	26.9	29.9	24.1	5.4	5.3	4.6	3.5
1965	29.7	24.6	23.5	27.9	5.2	4.0	7.2	1.6
1966	29.8	27.6	28.0	35.7	5.6	4.9	5.0	4.2
	Nonwhite							
1963	52.8	50.0	55.7	50.4	24.0	24.1	14.0	18.7
1964	54.7	59.7	68.4	52.0	21.1	16.9	12.9	16.7
1965	51.9	42.5	63.2	59.3	21.6	23.4	15.0	22.3
1966	50.7	41.8	50.6	50.1	21.0	15.1	19.9	20.8

1/ Perinatal deaths include all deaths in the first 27 days of life plus fetal deaths; the perinatal death rates are per 1000 deliveries (including live births and fetal deaths).

2/ Postneonatal deaths include all deaths of infants 28 days to one year of age; the postenonatal death rates are per 1000 live births.

Sources:

The mortality rates are based on natality and mortality data in the following publications:

(a) North Carolina State Board of Health, Epidemiology Division, Annual Report of Public Health Statistics Section, 1963, Part 2, pages 95-96, 99-102, and 105-106, Raleigh, North Carolina.

(b) North Carolina State Board of Health, Epidemiology Division, Annual Report of Public Health Statistics Section, 1964, Part 2, pages 97-98, 101-104 and 107-108, Raleigh, North Carolina.

(c) North Carolina State Board of Health, Division of Epidemiology, Public Health Statistics Section, North Carolina Vital Statistics, 1965, Part 2, pages B-4 and B-5, Raleigh, North Carolina.

(d) North Carolina State Board of Health, Division of Epidemiology, Public Health Statistics Section, North Carolina Vital Statistics, 1966, pages 42, 44, 45, and 47, Raleigh, North Carolina.

THE USE OF MATCHED PAIRS IN THE EVALUATION OF THE DISTRICT OF COLUMBIA DEPARTMENT OF PUBLIC HEALTH BIRTH CONTROL PROGRAM

Louise M. Okada

The D. C. Department of Public Health Birth Control Program began April 1964 at the municipal hospital and six outlying maternal and child health clinics. Because of limited facilities, the Program was directed primarily at medically indigent women who had deliveries within the past three months at the D. C. General Hospital and a very small number of women who were referred by the Department of Welfare. With additional clinics, the Program has expanded to include all women who had a previous delivery.

The time period covered by this study is related to the population utilizing D. C. General Hospital for deliveries between November 1964 through December 1965. During this period, the Birth Control Program was essentially a postpartum program. At the time of discharge after delivery, mothers were invited to attend a session where a birth control film was presented, followed by demonstration and discussion of the various types of contraception. For the interim between delivery and postpartum examination, a two-months supply of foam was distributed. Mothers were told that birth control service is available when they returned for their postpartum examination and were given an appointment prior to leaving the hospital. Two months after registration in a birth control clinic, mothers were given a return appointment to the clinic, at which time they received identification cards to pick up supplies for one year, renewable after a physical examination.

No charge was made for supplies or service and mothers were given two to three months supply per pickup visit. A choice of six methods were offered--pill, foam, diaphragm, rhythm, jelly, and IUD; the use of IUD, however, had barely begun by November 1965. The pills were by far the most popular method used during the period under study.

Description of the population. The population under study offers unique opportunities for research in family planning. Whereas studies in family planning in the United States have been made among white, middle-class, married couples, living together, the respondents in this study are low-income, urban Negro women with high indices of social and personal disorganization.

Family instability results not only in illegitimate births among the very young girls, but repeated history of illegitimate births is a common occurrence among a large proportion of these women in the reproductive ages. Data collected from the live birth certificates from D. C. General Hospital deliveries showed that in 1965, 52 percent of the live births were reported as illegitimate, and among first births this proportion was 76 percent.

One-third of the women delivering live births at this hospital had no prenatal care despite a concerted and well-established program of maternal and child health in this city. This population is characterized by excess fertility resulting from their inability to practice family planning. There is a wide discrepancy between the number of children they would like to have and the number of children they have already borne. Women in our study reported that 60 percent of their live births had been unwanted pregnancies.

Purpose of the study. This study was made possible by a grant from the Population Council for evaluating a postpartum birth control program. The three main objectives of this study were: (1) to measure the reduction in pregnancies as a result of participation in the Birth Control Program; (2) to determine whether certain demographic factors may be related to differential participation in the Birth Control Program; and (3) to measure the "use-effectiveness" or continuation rate in the Program.

Design of the study. We were met with the problem of how much the long-term fertility decline in the United States was affecting the secular trend of the birth rate for our specific population. More importantly, how much effect had the recent introduction of pills in a mass media society on the birth rate over time in this population? Rather than a secular decline, was this population maintaining a stable but high fertility level? We felt that the matched pair design would give us the data needed to answer two of our three objectives: to measure the reduction in pregnancies and to determine which factors, if any, were related to differential participation.

The study universe consisted of fecund Negro women who had a live birth at D. C. General Hospital between November 1964 through December 1965. Excluded from the study were about 130 women who had a fetal death delivery at this hospital during the same period. Excluded also were 640 cases of abortions because of the difficulty of incorporating these cases into the sampling frame since information on their characteristics was not readily available.

Approximately 6,000 live birth deliveries were listed in certificate number order and a systematic 20-percent sample was drawn. Among about 1,200 mothers in the 20-percent sample, 680 who had registered in the Birth Control Program within four months after the month of delivery were identified as members of the Study Group.

Each member of the Study Group was matched by a nonparticipant who also delivered a live birth at the same hospital during the same month, and was within the same age, parity and marital status categories. (The latter was inferred from the legitimacy item on the certificate). These matched nonparticipants comprise our Control Group.

Representativeness of the Study Group was checked by comparing this sample with the total birth control registrants in the Program from January 1965 through February 1966 in terms of age and live-birth order and was found to be fairly representative. On the other hand, the Control Group members are not representative of the nonwhite mothers who had a live birth delivery at D. C. General Hospital during the same time as the

Study Group members, but did not choose to participate in the Program since they are matched controls. On the average, they are somewhat younger in age and lower in parity than nonparticipants in general.

Interviews were conducted in private homes using prepared questionnaires. At first, approximately 20 public health nurses were trained for interviewing on an overtime basis. Subsequently health aides and social service workers were trained and used as interviewers.

The 20-percent sample of approximately 6,000 nonwhite mothers indicated that 56 percent of these women who delivered a live birth at the municipal hospital registered in the Birth Control Program. In contrast to the usual relationship of increased use of birth control with increase in age and parity, among this population, registration in the Program increased only through third births, dropping with fourth and higher births; there was a consistent decline in registration with increase in age, but this relationship failed to hold up in each parity group:

PERCENT REGISTERED IN BIRTH CONTROL PROGRAM AMONG NONWHITE MOTHERS WHO HAD A LIVE BIRTH DELIVERY AT D. C. GENERAL HOSPITAL BY AGE OF MOTHER AND LIVE-BIRTH ORDER, NOVEMBER 1964 THROUGH DECEMBER 1965

Age of Mother	Total	Live-Birth Order					
		1	2	3	4	5	6+
Total	56.0	57.4	60.0	62.6	54.0	53.5	46.4
19 years & under	60.4	58.4	68.5	64.1			
20-24 years	57.5	54.4	50.5	64.4	58.2	62.1	59.1
25-29 years	51.3		66.7	58.8	56.4	50.0	41.6
30 years or more	50.0			57.9	55.0	51.9	47.5

Fifty-five and fifty-seven percent of mothers who had legitimate and illegitimate live births, respectively, registered in the Program.

Findings. For the purposes of estimating the reduction in births and investigating variables which may be related to differential participation, this paper will present data from the first segment of the total 680 pairs in the study--the first 200 pairs. The participants among these first 200 pairs represent essentially a random 20-percent sample of the nonwhite women who had a live birth delivery at the municipal hospital during November 1964 through February 1965 and registered in the Birth Control Program in the first four months of 1965. One hundred sixty one matched pairs out of the first 200 pairs have been interviewed and 14 pairs were dropped where 10 participants moved from the city, 1 died, and 3 were misclassified. The percent followup was therefore 87 percent for this segment.

A comparison of some demographic characteristics between the Study and Control Groups showed that, having been matched by age and live-birth order, the two groups were quite similar on certain demographic variables related to differential fertility or fertility control. These variables were State of birth, residence on a farm and religion:

STATE OF BIRTH

	Study Group	Control Group
Total	161	161
District of Columbia	87	80
Virginia	14	15
Maryland	9	9
Lower South Atlantic 1/	44	48
All other South	1	2
Other	4	6
Not stated	1	1

1/ Includes Florida, Georgia, North Carolina and South Carolina.

EVER LIVE ON A FARM

	Study Group	Control Group
Total	161	161
None 1/	120	121
1 year	1	4
2-5 years	10	4
6-10 years	4	5
11-15 years	8	8
More than 15 years	17	16

1/ Includes a very small number of women who lived on a farm less than 6 months.

RELIGION

	Study Group	Control Group
Total	161	161
Baptist	91	104
Methodist	13	11
Catholic	27	24
Holiness	10	11
All other	18	6
None	1	4
Not stated	1	1

It is interesting to note that among the migrants, the origin of the northward migration to the District was almost solely confined to the eastern seaboard States and, more specifically, to North Carolina and South Carolina. These tables on State of birth and farm residence also show that roughly half the population we are studying are native Washingtonians and that 75 percent never lived on a farm. Thus, much of the high fertility of this population can be attributed to the native slum dwellers.

There was only a minor difference in educational attainment between the Study and Control Groups:

HIGHEST GRADE COMPLETED

	Study Group	Control Group
Total	161	161
0-7 grades	14	5
8th grade	14	20
9-11 grades	96	99
12th grade	32	33
13th & higher	4	4
Not stated	1	0

Some other variables related to fertility behavior are the norms or definitions regarding desired family size and expected family size. In spite of the fact that members of the Control Group already had, on the average, more children at the time of interview, there was no difference between the two groups on these variables:

MEAN DESIRED & EXPECTED FAMILY SIZE

	Study Group	Control Group
Desired family size	2.9	2.9
Expected family size	4.5	4.6
Additional number of children wanted	0.7	0.5

The Study Group had a higher proportion of women who had ever worked and were working at the time of interview:

EMPLOYMENT

	Study Group	Control Group
Total	161	161
Never worked	28	42
Working now full time	52	46
Working now part time	12	5
Working now time not stated	1	0
Not working now worked before	66	66
Not stated	2	2

The chi-square test indicates that the difference in the number of women who never worked between these two groups is significant at the .05 level. It appears that in this population, employment is related to differential fertility. When age is controlled, women who were working at time of interview preferred and achieved a smaller family size than those who had never worked by the time of interview among both Study and Control Group members:

MEAN LIVE-BIRTH ORDER & DESIRED FAMILY SIZE BY EMPLOYMENT STATUS

	Age of Mother at Interview			
	Under 20 yrs	20-24 yrs	25-29 yrs	30+ yrs
Live-Birth Order1/				
Working now (N=116)	1.7	2.6	4.0	6.0
Never worked (N=70)	1.8	4.8	5.8	6.6
Desired Family Size				
Working now	2.7	2.5	2.8	2.6
Never worked	3.3	2.7	2.3	4.5

1/ Includes pregnancies at time of interview. (Data on women who worked before not shown.)

The employment variable may perhaps be an indicator of some qualitative difference between the Study and Control Groups which help to account for the fact that one group participated in the Birth Control Program and the other group did not. This qualitative difference may be related to a type of attitude makeup which makes one group more employable than another. It may also be related to the fact that the Study Group members have a more positive attitude toward preventive health care as evidenced from their past performance in obtaining prenatal care and postpartum examination for their pregnancies and after their deliveries.

The question may be raised as to whether the birth rates between the Study and Control Groups are comparable and whether the reduction in pregnancies due to participation in the Program can be measured in such a manner. The fact that few differences in characteristics could be found between the two groups strengthens the reasonableness of a comparison. Furthermore, pregnancy rates for both groups were roughly equal in the 18-month period prior to the specified pregnancy by which they became part of the sampling frame (hereafter referred to as the "Before" period). That is, excluding one-parity women, the Study Group had an annual pregnancy rate of 49 per 100 women compared to 52 for the Control Group.

By use of the matched pair data, we estimated a reduction in the pregnancy rate of 57 percent due to the Study Group's participation in the Birth Control Program. In the 12 months following their specified deliveries (hereafter referred to as the "After" period), there were 32 pregnancies among the 161 members of the Study Group compared to 75 pregnancies for the same number of Control Group members.

A comparison of pregnancy rates "Before" and "After" their specified deliveries for the Control Group members would indicate whether the pregnancy rate had also declined among those who did not participate in the Program. Because of changes in age and marital status composition between these two periods, in spite of small numbers an attempt was made to adjust for these changes. For the most part, however, these changes were compensated by the inclusion of the younger one-parity women in the "After" period. The "Expected" rates for both groups showed that changes in the age-marital status composition could account for only a small decline in the rates:

ANNUAL PREGNANCY RATES PER 100 WOMEN1/

	"Before" (N=125)	"Expected"	"After" (N=161)	Change "Exp." vs "After"
Study Group	49.1	47.3	19.9	-59%
Control Grp	51.7	50.3	46.6	- 7%

1/ The "Before" rates exclude 36 one-parity women at specified delivery for each group.

From the difference in the "Expected" and "After" pregnancy rates for the Control Group, it can be seen that some decline had occurred which could not be attributed to direct participation in the Birth Control Program. If the pregnancy rate for the Study Group also declined by a like amount without the benefit of participation in the Program, then the difference in the "Before" and "After" rates for the Study Group slightly exaggerated the actual decline due to direct participation.

This is not to say that the small decline for the Control Group occurred completely independent of the Birth Control Program. At least part of the decline can be attributed to indirect effects of the Program; that is, dissemination of information about birth control and the distribution of foam at time of discharge after delivery. Interpregnancy use of contraception, without reference to regularity and length of use, showed that among the Control Group members pill users increased from 2 to 13 before and after their specified deliveries; foam users increased from 16 to 35 during the same time; IUD users increased from 0 to 3 and 4 were sterilized.

Although not directly related to the use of matched pair data, the continuation rate of the participants in the Program will be briefly presented for sake of comparison with other program evaluations of family planning services among the poor urban Negroes. In order to achieve a magnitude of a 57 percent reduction in pregnancies in a 12-month period, what kinds of continuation rate and accidental pregnancy rate were found among the participants in the Program?

The monthly continuation rates for all women using any method from the Birth Control Program and the same rates separately for women on pills only are shown in Tables 1 and 2 in the Appendix. These rates are based on the completed interviews among the first 300 Study Group members representing essentially a 20-percent sample of nonwhite women who delivered a live birth at the municipal hospital during November 1964 through April 1965 and registered in the Birth Control Program during the first six months of 1965. Of these 300 Study Group members, 261 interviews had been completed and 21 had been dropped where 10 had moved from the city, 1 died, 3 misclassified, and 1 institutionalized. The followup rate was thus 94 percent.

The continuation rates in Tables 1 and 2 are about the same when adjusted for the fact that Table 2 (pill only) excludes any woman who did not report use of the pill for at least one month. At six, twelve, and eighteen months after registration percents surviving in the Program and on the pill (unadjusted) were as follows:

PERCENT SURVIVING

	Birth Control Program (N=261)	Pill (N=199)
6 months	65%	72%
12 months	53	58
18 months	44	48

The accidental pregnancy rate at 12 months is almost double for all women participating in the Program compared with those who used pills only:

ACCIDENTAL PREGNANCY RATE PER 100 WOMEN

Birth Control Program	Pill
8.7	4.8

The survival rate by age over time for all women in the Birth Control Program showed that the youngest women, 19 years and under, had the highest dropout rates:

PERCENT SURVIVING BY MONTH AND AGE

	1-6 months	7-12 months	13-18 months
Under 20 yrs (N=80)	59%	41%	31%
20-24 yrs (N=90)	70	63	56
25-29 yrs (N=48)	60	50	38
30+ yrs (N=43)	70	56	54

The relation of continuation in the Program to age, however, is quite irregular.

Summary. In the planning stages of this study the question was raised whether the nonparticipants in the Program would be using sources for contraception other than the Birth Control Program, thereby resulting in somewhat comparable pregnancy rates between those who participated and those who did not. The comparison of the pregnancy rates between the matched Study and Control Groups indicates that the participation in the Program was highly effective in reducing the number of pregnancies among the Study Group members. Comparisons of pregnancy rates "Before" and "After" for both the Study and Control Groups tend to confirm this conclusion. That is, both groups had comparable rates in the 18 months prior to their specified pregnancies. In the 12 months following their specified deliveries, the pregnancy rates fell 60 and 10 percents for the Study and Control Groups, respectively.

It is unlikely that the majority of successful contraceptors among the Study Group would have had a readily accessible source of effective contraception without the aid of the Program. The results of this study lead us to believe that a carefully operated program using pills or IUD will meet with considerable success in reducing unwanted pregnancies among the low-income, urban Negro women.

Table 1

CONTINUATION RATE FOR 261 WOMEN USING ANY METHOD OBTAINED FROM THE BIRTH CONTROL PROGRAM

Months Use	Reason for Discontinuing Use of Any Method from Birth Control Program[1]										No. of women surviving to end of month	Percent surviving	Cumulative exposure months
	TOTAL	Discomfort (pill)	Medical	Harmful	Accidental pregnancies	Planning baby	No need; separated or sterilization	Getting method elsewhere	All other	Not stated			
1	25				2		1	7	15[2]		236	90%	236
2	17	7	1		3		2	2	2		219	84	455
3	19	7	1	1	1	1		3	5		200	77	655
4	16	4	1	1	4			1	5		184	70	839
5	8	1						2	3	2	176	67	1015
6	7	3			1		1		2		169	65	1184
7	7	1	1		1			1	3		162	62	1346
8	9	4		1		2		1	1		153	59	1499
9	5	1	1		1				2		148	57	1647
10	3								3		145	56	1792
11	2	1			1						143	55	1935
12	5				1		1	1	2		138	53	2073
13	5		1	1	2	1					133	51	2206
14	3	1		1	1						130	50	2336
15	5		1				1		3		125	48	2461
16	5	1					2		2		120	46	2581
17	2	1							1		118	45	2699
18	2					1	1				116	44	2815
	145	32	7	5	18	5	9	18	49	2			

[1] Discontinuation means not using a method from Birth Control Program.

[2] Fifteen women did not use a method within 6 months after registration.

Table 2

CONTINUATION RATE FOR 199 WOMEN WHO USED PILLS

Months Use	Reason for Discontinuing Use of Pills									No. of women surviving to end of month	Percent surviving	Cumulative exposure months
	TOTAL	Discomfort	Medical	Harmful	Accidental pregnancies	Planning baby	No need; separated or sterilization	All other	Not stated			
1										199	100%	199
2	15	11	2	0	1		1			184	92	383
3	14	7	1	1	1	1		3		170	85	553
4	14	5	1	1	3			4		156	78	709
5	4	2						1	1	152	76	861
6	9	3		1	1		1	3		143	72	1004
7	5	2	1					2		138	69	1142
8	8	4		1		2	1			130	65	1272
9	4	2	1					1		126	63	1398
10	4	1						3		121	61	1520
11	2			1				1		120	60	1640
12	5				1		1	3		115	58	1755
13	4		1	1	1	1				111	56	1866
14	3	2		1						108	54	1974
15	4		1				1	2		104	52	2078
16	6	2					2	2		98	49	2176
17	1							1		97	49	2273
18	2					1	1			95	48	2368
	104	41	8	7	8	5	8	26	1			

THE PREGNANCY HISTORY APPROACH TO MEASUREMENT OF FERTILITY CHANGE

by

Donald J. Bogue and Elizabeth J. Bogue
Community and Family Study Center
University of Chicago

I
The Problem of Measuring Fertility Change

All over the world there are underway massive programs to reduce high birth rates. Unhappily, as yet no adequate statistical procedure has been devised to evaluate whether these programs are succeeding or failing, and if they are succeeding by how much they are reducing the birth rate each year. The present paper proposes a system which, it is believed, begins to fill this need. Despite the fact that it also has serious limitations, it seems to yield measures which are more reliable, more valid, and capable of more detailed explanatory analysis than any other system yet proposed. It is called the "pregnancy history" approach to fertility study. The central idea is to collect complete pregnancy histories for samples of women in the subpopulations where fertility and fertility change are to be studied, to use techniques of formal demography to adjust these data for known deficiencies and biases, and to devise computer programs that convert these data into conventional demographic measures. This procedure not only is correct in terms of demographic theory, but also seems to be practicable when put to use under the conditions that exist in the developing countries where the "crash" programs for fertility reduction are especially in need of a technique to measure fertility change. This technique has still another interesting virtue: because it is a longitudinal measure, it generates exactly the data needed to pursue some of the newer and more challenging theoretical problems in fertility analysis: fecundability, pregnancy intervals, conception rates under various conditions of use of contraception, and the development of mathematical models of reproduction.

The exposition which follows:

(a) Reviews the "ideal" demographic system for measuring fertility change.
(b) Describes the pregnancy history as a substitute for vital registration.
(c) Spells out the steps for converting the pregnancy history into vital rates.
(d) Lists the biases of pregnancy history data and presents techniques for adjusting for each type of bias.
(e) Describes the procedures to be followed in measuring fertility change from pregnancy history data.
(f) Presents two examples of use of the pregnancy history approach.
(g) Makes a summary evaluation of the technique in comparison with other methods.

II
The "Ideal" System for Measuring Fertility Change

The "ideal" procedures for measuring the level of fertility in a population and short-run changes in fertility are a well established part of demographic methodology. Three sets of measures need to be calculated for two dates:

(a) Age-specific fertility rates (ASFR).
(b) Total fertility rate (TFR)--the sum of ASFR for all ages.
(c) General Fertility Rate (GFR)--ratio of births to women of childbearing age.

The absolute and relative differences in these measures between the first and the second date are universally accepted by demographers as valid measures of fertility change. Table 1 is an example of these procedures, using data for the U.S.

The ASFR, GFR and TFR are superior to the crude birth rate because they exclude the population not exposed to childbearing and maintain a rigorous control over age composition. Moreover, they have a clear and unambiguous meaning. The ASFR is the probability (number of chances in 1,000) that a woman of a given age selected at random from a specified population will bear a child within the next year. The GFR is a similar probability for a woman of childbearing age, without respect to any particular age. In other words, it is a weighted average value of the ASFR probabilities. The TFR is a statement of the average size of completed family (at end of childbearing) that will result if a particular schedule of ASFR were to be in effect for the complete duration of a reproductive span.

In nations with reliable systems of vital registration the data needed to calculate these measures are readily available. For this reason, the nations of Europe, North America, Australia, Japan, Argentina and a few others are able to know precisely at any point in time what their fertility level is and how it is changing. However, in all but a few of the developing nations of Asia, Latin America and Africa births are so incompletely registered that the official vital statistics cannot be trusted, and the calculations of Table 1 cannot be performed. There is little hope that this situation can be remedied within the next fifteen years. There is urgent need to measure fertility change now. This leads to the question,

> "How can fertility levels, fertility changes, and fertility differences between subgroups of the population be measured in the absence of reliable vital statistics?"

III
The Pregnancy History as a Substitute for Vital Registration

A pregnancy history is a chronological record for each woman of childbearing age in a population or a sample, of each pregnancy she has experienced. In addition to the fact of the pregnancy itself, the following information is obtained for each pregnancy.

(a) Outcome of the pregnancy:
   (1) Live birth
   (2) Spontaneous abortion
   (3) Induced abortion
   (4) Stillbirth or other pregnancy loss
(b) Date at which the pregnancy terminated (month and year)
(c) Months of gestation (especially important for all pregnancy losses)
(d) Sex of live-born children
(e) Type of pregnancy (single or multiple birth)
(f) Survival of live born children--is child still living at time of survey?
(g) Age at death of deceased children: Month and year of death or estimated age at death.

A battery of interview questions and a special reporting form have been developed for obtaining these data reliably. These are reproduced in Appendix A. The interviewers are given special intensive training on this phase of the interview. They are required to probe and reprobe to help the respondent recall each and every pregnancy and the pertinent facts about it. Every interval of two or more years without a pregnancy is brought to the attention of the respondent, to try to remind her of any pregnancies that have been overlooked. When the pregnancy history is properly completed, the only unreported pregnancies are due to (a) failures of memory despite the best efforts of especially trained interviewers to assist recall and (b) willful misreporting to hide illegitimate births or induced abortions.

For purposes of analysis we accept this set of data as a substitute for vital registration. In fact, when completed the pregnancy history may be looked upon as a set of certificates for the pregnancies of a particular group of women. In this register has been included not only live births but pregnancy losses as well. By a few calculations it is possible to transform these data into bonafide birth statistics which can then be used to follow the "ideal" system of fertility measurement described above.

IV
Transformation of Pregnancy Histories into Fertility Rates

For a correct interpretation of the fertility measures that are computed from pregnancy histories, it is essential to begin with an appreciation that the data are longitudinal, and pertain retrospectively to the fertility experience of a set of real cohorts, each of which has arrived at a particular stage in its reproductive span at the time of the pregnancy history survey. These data must be manipulated in such a way as to provide cross-sectional data that refer to particular calendar years. (This is the reverse of the cohort fertility problem as it is usually encountered.) The system developed here permits a study of fertility both in the cross-sectional (calendar year) and the longitudinal (real cohort) contexts.

Each fertility rate has two parts: a numerator and a denominator and takes the prototype form

$$N_i = \frac{B_i}{F_i} \cdot K \qquad (1)$$

where i = age of woman

N_i = fertility (natality) rate specific for age i

B_i = number of births to mothers who were age i at time of delivery

F_i = female population who were age i during the interval of time to which the rate refers

K = base of the rates, usually 1,000

The transformation of pregnancy history information into fertility rates requires three related but distinct procedures: one to obtain raw data for the numerators, one to obtain raw data for the denominators, and one to divide the former by the latter.

(a) Numerators. Each pregnancy must be simultaneously classified according to the calendar year in which it occurred and the age of the mother at time of occurrence. This is accomplished by establishing a large matrix in which each row represents one calendar year and each column represents one year of age of mother. For each pregnancy we cumulatively add "1" to the appropriate cell of this matrix according to the age of the mother and calendar year. The calendar year of occurrence is given directly by the pregnancy history. Age of mother at time of occurrence is easily derived from the relation.

$$F_i = D_B - D_F \qquad (2)$$

where F_i = age of mother at date of birth of child

D_F = year and month of birth of mother

D_B = year and month of birth of child

(b) Denominators. Inasmuch as birthdays occur more or less evenly throughout each calendar year, there is no simple way to simultaneously classify women according to age and calendar year. Demographers conventionally resolve this problem by resorting to the concept of person years, and this is used here. We establish a large matrix in which each row represents one calendar year and each column represents one year of age of mother, identical to the one set up for numerators. We then calculate the number of months each woman spends in each age in each calendar year, cumulate these monthss months for all women in the sample, divide by 12 and label the result "person years" spent in age i during calendar year z. The relationship is given by the equations:

$$P_i^z = D_z - D_F \qquad (3)$$

$$P_z^{z+1} = (12 - P_i^z) \qquad (3a)$$

where P_i = number of person-months spent in age i in calendar year z

D_z = December of calendar year z

P_i^{z+1} = number of person-months spent in age i in calendar year $z+1$

D_F = birthmonth of woman

The sum of P_i^z and P_i^{z+1} for any woman is always 12 months, and each of these two values can vary from 0 to 12.[2]

(c) Calculation of rates. By dividing the numerator matrix of step 1 by the denominator matrix of step 2 and multiplying by 1,000 we obtain a rates matrix in which each cell is an age-specific rate for a particular age for a particular calendar year. In other words, the operation conforms exactly to the basic prototype form of equation (1). These rates are not estimates; they are an attempt to make an actual reconstruction of the past and are identical with rates that would have been obtained by a vital registration system, if vital registration and the pregnancy histories were equally complete and accurate. Differences between registered vital rates and pregnancy history vital rates differ only by the degree of completeness between the two and the representativeness of the sample of women for which pregnancy histories are obrained.[3]

Table 2 illustrates the rates matrix that is produced by this procedure. It is one part of the output of a computer program "Pregnancy History Analysis" written by E. J. Bogue.

The longitudinal nature of the data are readily apparent from this table; each diagonal line represents one real cohort. Since the data refer only to women currently in the childbearing years, there are no rates above the principal diagonal.

For all except the very largest samples, data for single years of age and single calendar years are too detailed for practical use. They may be abridged in two ways; (a) by combining ages into 5-year intervals and (b) by combining years into any desired grouping. With the ASFR thus produced it is a simple matter to calculate TFR. GFR is generated directly by the program. Thus, the procedure extracts from pregnancy histories the three basic measures needed for the "ideal" demographic procedure for measuring fertility.

In most developing countries it is possible to collect by direct interview from representative samples of women pregnancy history data that are far more complete and reliable than the data obtainable from the vital registration system. Moreover, an attempt will be made below to demonstrate that the deficiencies of the pregnancy histories can be largely corrected by demographic techniques. It is these two facts which recommend the pregnancy history approach for measuring changes in fertility to evaluate the progress of family planning programs.

Table 1 is only one of several possible tabulations from the pregnancy history data. Before considering the other outputs of the "Pregnancy History Analysis" system, it is necessary to discuss the problem of bias and its correction.

V

Biases in Pregnancy History Data and Techniques of Adjustment

Pregnancy history data are subject to three unique biases:

(a) Misreporting the date at which a pregnancy occurred.

(b) Misreporting the age (date of birth) of the mother.

(c) Failure to report all pregnancies.

By an elaborate editing procedure, involving three separate edits on the electronic computer, gross errors of misreporting dates of events and ages of mothers are detected and adjusted by non-biasing corrections. (Impossibly short intervals between births, impossibly young ages at bearing children, improbably long intervals between marriage and birth of first child are indicators of error). It is believed that by careful interviewing and this editing the ages of mothers at the birth of each child can be fixed within two or three years, even in low-literacy populations, and year of birth of children can be ascertained even more precisely.

When ages of mothers are grouped into 5-year intervals in accordance with usual practice, and the experience of two consecutive calendar years are combined to obtain a two-year average set of rates, it is believed that the first two of the above three biases have been reduced to an acceptably low magnitude.[5]

The third bias, failure to report all events, is inherent in the data and must be corrected by an upward adjustment, according to the presumed nature and extent of the error. A plausible adjustment, and the one recommended is as follows:

(a) Assume that failure to report a pregnancy is strongly concentrated among infants who died during their first year of life. Therefore, <u>discard the pregnancy history data for infants who died during their first year of life and substitute a demographic estimate.</u> This calls for setting up an events matrix which cumulates only live births which survived at least one year.

(b) By independent research estimate what the <u>true</u> infant mortality rate was in the population under consideration during the years for which a fertility measurement is to be made.

(c) Adjust the data for births-that-survived-one year or more for infant mortality by the following equation

$$B^z = \frac{B_s^z}{(1.0 - q_o^z)} = \frac{B_s^z}{S_o^z} \qquad (4)$$

where B^z is the estimated true number of births in year z

B_s^z is the number of births that occurred in year z which survived to the first birthday

q_o^z = the estimated true infant mortality rate in year z

S_o^z = survival factor for year z = $(1.0 - q_o^z)$

(d) Use the adjusted births to compute the ASR and other fertility measures. The procedure outlined above is algebraically equivalent to calculating the ASFR, TFR and GFR first, from the surviving births, and then adjust the rates upward simply by dividing by S_o^z.

Very often in developing countries there will be no estimate of infant mortality (q_o^z) for the years under consideration and for the particular population being surveyed. As will be explained below, the computer program "Pregnancy History Analysis" itself produces a tabulation of this rate. If no alternative source of information is available, it could be assumed that this calculation is only 75 percent of the true value (which will be approximately correct in most situations). As will be shown below, (a) the error that can result from this procedure can affect the estimate of the birth rate by only a negligible amount and (b) tends to be cancelled out when making estimates of change in fertility. For this reason, the technique is a superior one for evaluating the effectiveness of family planning programs.

An alternative strategy is to follow the procedure outlined above, except to use only live-born children still living, and use a life table to reverse-survive each year's births to estimate all deaths to live born children. Model life tables of the United Nations or those prepared by the Princeton Office of population Research may be used for this step. If one distrusts the memory of the respondents (especially where illiteracy is almost complete) this procedure may be preferable to the adjustment for infant mortality. Experience thus far has shown that this procedure yields results that are almost identical with the infant mortality adjustment procedure. Two facts account for this: (a) most child mortality occurs during the first year of life and (b) women appear to remember and report children who survived one entire year with a reasonably high degree of completeness.[6]

A recommended practice is to prepare a "high" "medium" and "low" estimate of fertility. The computed rates adjusted for infant mortality or for child mortality as described above, may be accepted as the "medium" (most plausible) estimate. The unadjusted rates, as they come from the computer, may be accepted as the "low" estimates, for the biases are in the direction of understating fertility. A "high" estimate may be made by assuming that the women failed to report all of their live births who survived by x percent, and to inflate the rates as calculated by a factor of (1.0 - x). Experience thus far suggests that a factor of 5 percent would be a moderate upward adjustment and a factor of 10 percent would be near maximum.

VI
An Example of the Use of the Pregnancy History Procedure to Measure the Level of Fertility

In 1964, the United Nations Demographic Center in Santiago, Chile, sponsored the collection of fertility data from representative samples of females of childbearing age in seven Latin American capital cities. Included in the interview were questions that contained the major ingredients for the fertility history. These data have been processed with the "Pregnancy History Analysis" program, in an effort to establish the level of fertility in each of these places. Table 3 summarizes data for Mexico City, together with other estimates of fertility for Mexico that may be relevant.

Some explanatory comments should be made of these estimates before they are analyzed. The value of the infant mortality rate obtained from the pregnancy history tabulations was 72 per 1,000 live births. This was divided by .75 to obtain an estimated "true" infant mortality rate; the estimated rate is therefore 96 infant deaths per 1,000 live births. The "official" infant mortality for all of Mexico, for 1960, published in the United Nations Demographic Yearbook, and ranked as one of the rates that may be accepted as reasonably correct, was only 69.9 in 1962, the midpoint of the 5-year span of time to which the pregnancy history rates refer. Thus, in this case the pregnancy history approach found a higher level of infant mortality than the official published data; when adjusted upward by 25 percent to obtain a corrected estimate of fertility, the correction for estimated error of memory for deceased infants should be regarded as fully corrected, if not over-corrected.

A close examination of Table 3 reveals the following:

(a) The "medium" estimate of TFR for Mexico City is 92 percent of Dr. Lee Jay Cho's estimate for all of Mexico, based upon census materials and 94 percent of the official estimates published in the Demographic Yearbook of the United Nations. These results seem highly plausible, inasmuch as one would expect a somewhat lower birth rate for the capital city than for the entire nation including rural areas. In fact, if one were to suspect the data of bias, it could be that the pregnancy history estimates for Mexico City are too high.

(b) The pattern of ASFR derived by the pregnancy histories are amazingly close to the estimates of Dr. Cho for ages under 35. The lower rates for Mexico City at ages above 35 are the pattern one would expect for a population just beginning to control its fertility. The pregnancy history ASFR are reasonably close to the United Nations Yearbook reports, for ages between 20-29 years; at ages 15-19 and 40-49 the UN. estimates seem to be less reasonable than the Cho or the pregnancy history estimates.

(c) Even the range between the "low" and the "high" estimate is quite small; the high estimate is only 8 percent greater than the low estimate. The range between the "medium" estimate and the "low" estimate is impressively small; the demographic adjustment for infant mortality increased the level of fertility above the uncorrected estimates by only 2.5 percent. The fact that a high estimate of infant mortality was used to obtain even this difference suggests that the "medium" estimate may lie very close to the true value indeed. In summary, it seems plausible to conclude that the pregnancy history estimates do indeed "bracket" the true level of fertility within a reasonably narrow margin and that the "medium" estimate is an unbiased estimate of the true schedule of ASF, GFR, and TFR for Mexico City.

These results for Mexico are not unique. The results for the other six nations of Latin America are equally consistent with other estimates of fertility levels and other estimates of infant mortality rate. [7]

VII
Use of the Pregnancy History Approach to Evaluate Family Planning Programs

If the pregnancy history approach is as successful in measuring the level of fertility as the above argument suggests, it is readily apparent that a powerful new device is available for evaluating family planning programs. There are two possible research strategies to the use of the pregnancy history technique to measure changes in fertility rates: the prospective and the retrospective design.

A. The prospective research design.
To evaluate a family planning "action project" prospectively it would be necessary simply to follow the

following straightforward study design:

(a) Collect pregnancy history data for a representative sample of the "treatment" population to be subjected to the fertility control program immediately before that program begins or within one year of its start. Simultaneously take a similar sample for a "control group" of essentially the same characteristics not to be submitted to an action program. Compute "high", "medium", and "low" fertility rates for the "treatment" and the "control" populations.

(b) Launch the "action program" among the "treatment populations." Allow it to run for approximately three years. Absolutely no results upon the birth rate can occur for at least 9 months after the start of an action program, and it takes at least 3 months for an action program to get organized and operating on a wide scale. It is therefore completely unrealistic to expect any effect at all within one year. At least two years should elapse, and preferable three, before an attempt is made to measure impact. This amount of time is required for a change in birth rate sufficiently large to have taken place that it can be detected and measured by a sample.

(c) After three years, conduct a second round of pregnancy history inventory. Again compute birth rates for the "treatment" and "control" groups. Calculate the amount and direction of change, using the model of Table 1. If the decline in fertility level in the "treatment" population is significantly greater than the change in the "control" population, if no alternative hypothesis can be found to explain the change, it may be inferred that the action program accelerated the decline in the birth rate.

Unfortunately, no examples yet exist of this approach to family planning evaluation.

B. <u>The retrospective research design.</u> This approach takes advantage of the longitudinal aspect of the pregnancy history. It does not take a before-and-after measurement, as does the prospective design, but merely waits until <u>after</u> the program has been running for about three years and then makes the evaluation. The steps for conducting a retrospective evaluation are as follows:

(a) Conduct a sample pregnancy history inventory among the population where an intensive family planning action program has been underway for two or three years.

(b) Using the longitudinal aspect of the pregnancy history, compute a set of birth rates for the years that correspond to the span of family planning action.

(c) Compute a set of rates for an equivalent number of years <u>immediately preceding</u> the action program.

(d) Compare the fertility level for the period prior to the family planning action program with the fertility level during the program. <u>A comparison of the fertility level for these two dates provides a measure of fertility change.</u> If fertility has declined, the rates for the later date will be significantly lower than the rates for the earlier date, whereas if there has been no change they will be equal. If there has been a fertility increase, the rates for the later date will be higher than the rates for the earlier date. Thus, <u>it is possible to measure recent fertility change with a single interview, taken after the change has taken place.</u> [8]

The above procedure can be used to measure fertility change in populations where there has been no special family planning program, to ascertain whether a secular trend in fertility exists.

The power of the retrospective procedure can be greatly heightened if a retrospective measurement is made on a control group which has not been subjected to the intensive program. If the decline in the "treatment population" is greater than the decline in the "control group" and no alternative explanation can be adduced to account for the result, it may tentatively be assumed that the family planning action program has had a measurable impact upon the population.

Very often it may not be possible to find a "control population," or there may not be sufficient funds and manpower to collect two sets of pregnancy history data. If one is willing to be content with the simple discovery that birth rates either are (a) remaining the same or (b) falling in the treatment area, then it is not necessary to take a measurement for the control group. Under this less rigorous design, the researcher is forced to assume that birth rates would have remained unchanged if there had been no action program, and that all declines in fertility may be attributed to the actions he has taken. In some cases there can be no alternative to this approach. If an entire nation has been inundated with family planning action, then no "control population" exists.

All over the world there are family planning projects which are candidates for evaluation by this retrospective procedure. They were begun without any baseline measurement of fertility and now are desperately in need of an evaluation to learn whether or not birth rates are falling. Although this retrospective design is not as rigorous as the prospective one, it is nevertheless believed capable of assessing whether there is a change in birth rates and the approximate amount.

VIII
Examples of Retrospective Evaluation of Family Planning Action Programs

In a very poor slum area on Santiago, Chili's, outskirts, the University of Chile has been conducting an intensive family planning action program since late 1964. This program was a major experimental effort to combat induced abortion by offering family planning as a substitute. Mass communication, conferences with women coming to a local clinic for health care for their children and themselves, and home visits were made to inform the residents of the area about family planning and to encourage them to come to the clinic for service. The intrauterine device was the principal method of contraception offered.

The director of this study, Dr. Anibal Faundes-Latham,[9] included the questions of Appendix A in a follow-up interview taken with a representative sample in January, 1967. The data were then brought to the University of Chicago and processed with the "Pregnancy History Analysis" program. The results are shown in Table 4. The statistics represented here are the "medium" estimates, making use of the correction for infant mortality. (An infant mortality rate of 90 for the five-year period preceding the interview was estimated by the pregnancy history. This has been inflated to 120 for purposes of calculating the birth rates. The same value is used for both the "before" and "after" period.)

According to the results of Table 4, the average size of completed family in this barrio bajo of Santigo was 7.1 children. After two years of family planning treatment it had fallen to 6.1 children, or by 14 percent. The decline in the general fertility rate (which is a more reliable measure from the sampling point of view) was 19.5 percent. This represents a decline from a crude birth rate of about 48.4 to 39.0 within a period of two years. The declines appear to have been concentrated among the women under 40 years of age. It thus appears that, unless some other explanation for this significant decline can be produced, the experimenters may assume that their program is promoting extremely rapid fertility decline in this area.

The results of the Santiago experiment

should be contrasted with the results that have been obtained for similar tabulations where no special family planning program has been available. In Table 5 we have divided the 5-year interval for Mexico City's medium estimate into two periods analogous to the intervals in Santiago. The period 1962-64 represents roughly 2 1/2 years preceding the interview (the interview was taken in mid-1964), and the period 1960-61 represents the full effort during these years, and the impact upon the total population was quite small. This is reflected in an estimated decline of only 1 percent in the GFR and of 2.6 percent in the TFR.

A. U.S. Family Planning Experiment. Table 6 summarizes the results of an experiment to reduce birth rates in the Old Plantation Belt of Alabama.[10] The data refer to a sample of women who had attended family planning clinics in 8 rural counties in the vicinity of Selma-Montgomery-Tuskegee in response to a special program offering birth control pills at subsidized prices and with Negro family planning educators doing motivational work at maternal and child health clinics and out in the community. This table illustrates the use of the Pregnancy History Analysis Program to compute nuptial fertility rates. Instead of the denominators referring to all women they here refer to all ever married women. (Because of the unusual marriage patterns of Southern Negroes, the data actually refer to "ever exposed" women; an estimated date at which sex relations began to occur more or less regularly was substituted for the date of marriage.) Because all unmarried women not exposed to pregnancy have been removed from the denominators, these rates are very high.

Notes:

The infant mortality reported by the pregnancy history tabulation was 28 per 1,000 live births. This was inflated to 37 (presumed 75 percent complete). The sample is 480 women. It is not possible to compute a TFR for a nuptial population by simply summing ASFR, because this would presume that all females were married at age 15-19. For this reason it is omitted from Table 6.

Table 6 shows that this highly selected group of women who had attended the clinic reduced fertility rates by 42 percent in comparison with the two years preceding the program. In this instance, however, there is much less reason to attribute this result to the particular family planning experiment being conducted than in the Santiago case, for Negro birth rates were falling rapidly throughout the nation during these same years. However, the measured rate of decline is far greater than the U.S. trend, and we must conclude that by attending the clinics these women were highly successful in curtailing their fertility more drastically than the general Negro population. However, we have no experimental way of knowing what action these women would have taken had there been no accelerated family planning program; it is quite possible that they were a select group of highly motivated persons who would have used some other method had the birth control pills not been available. It would have been highly desirable to have a control group against which to compare this sample, but the entire State of Alabama began a free birth control pill program for all indigent citizens only a few months after this experiment began, so that no comparable program that could be truly considered to be a "control group" existed. (A sample of women in the study area who did not go to the clinics was interviewed, but this is not a genuine control group.)

IX

Other Measures Provided by the Analysis of Pregnancy History Data

The pregnancy history provides information concerning several aspects of fertility that hitherto have been researched insufficiently. It therefore offers some fresh opportunities for expanding our knowledge of human fertility. This information is exploited by two computer programs: the "Pregnancy History Analysis" program, described above, and a "Pregnancy Interval" program which is used as the third and final edit before the data are tabulated to obtain rates. Following is a brief listing of the information provided by these two programs.

A. Pregnancy loss rates and infant mortality rates. Each of the following rates is tabulated by single year of age of mother for each calendar year and for any grouping of ages and years desired.

a. Rate of pregnancy loss--number of pregnancy losses per 1,000 women years
b. Probability of pregnancy loss--number of pregnancy losses per 1,000 pregnancies
c. Spontaneous abortion rate--number of spontaneous abortions per 1,000 pregnancies
d. Induced abortion rate--number of induced abortions per 1,000 pregnancies
e. Infant mortality rate--of the infants born in a particular year, the rate per 1,000 who die before reaching their first birthday

It must be acknowledged that all of these rates are subject to serious understatement, especially for events that happened more than three years preceding the interview. However, under good interviewing conditions, the reporting of these events is surprisingly good and provides useful information, as the work of several studies in Latin America has shown. The procedure developed here converts this information into exactly the measures that best permit its analysis.

B. Dates, events, ages

a. Date of conception of each pregnancy
b. Incidence of premarital pregnancy
c. Age at which first exposure to regular sex relations began, independently of marital status

C. Intervals

a. Interval between marriage (or first exposure) and first pregnancy
b. Interval between all successive pregnancies
c. "Open interval"--interval between last pregnancy and date of the interview
d. Length of exposure to pregnancy, with this interval divided into
   (1) Time spent in a state of pregnancy
   (2) Time spent in a state of non-pregnancy

D. Prevalance measures

a. Percentage of women who currently are pregnant (3 months or more)
b. Percentage of women who have experienced a pregnancy loss, by age
c. Percentage of women at each parity at each age
d. Percentage of women who have had a child die, by age
e. Number of children ever born to women, by age

It should be pointed out that the above items are some of the key elements in model-building and mathematical quantification of the conception and reproduction process. The Pregnancy History approach will provide factual data for testing these models.

X

Pregnancy History Analysis as Part of a Larger Research-Evaluation System

The pregnancy history is not taken as an isolated set of observations, but is included as part of a larger more comprehensive "MAKE-UP" interview where the first letters of the term have the following significance:

M--motives for and against family planning
A--attitudes favoring and resisting adoption of family planning
K--knowledge of methods of contraception and of availability of family planning services
E--explanatory variables that account for differences in fertility behavior and family planning adoption
U--use of contraception and use-effectiveness of each method
P--Pregnancy history

This comprehensive interview provides data for a wide variety of variables which can be correlated with the items derived from the pregnancy history. Thus, the pregnancy history approach not only provides the data for measuring fertility change, but when included in a comprehensive interview it is possible to trace the fertility change to the specific persons who accomplished it, and to learn their motives, attitudes, mode of contraception used, the reasons they adopted family planning, and the degree of effectiveness of the methods they employed. It is even possible to learn whether or not they have had contact with one of the official family planning programs, or whether they have received information via mass media. Thus, it offers unexcelled opportunities for linking family planning action to fertility changes for purposes of evaluation.

XI
Comparison of the Pregnancy History with other Systems of Fertility for Measuring Fertility Change

The following four systems have been advanced for measuring short-run changes in fertility in the absence of reliable vital statistics:

(a) Population Growth Estimation (PGE)--A combination of enumeration by repeated home visit (3 or 4 times per year) of births as they occur, linked to an independently maintained system of vital registration, with matching to include births found by one system but not the other. This system has been used successfully in Pakistan, under a program sponsored by the Population Council.

(b) "Open Interval" Analysis (OIA)--An enumeration, at successive intervals, of the time that has elapsed since women have delivered their last child. Under conditions of high fertility these intervals are short; as fertility declines they become longer. If the average length of the intervals increases it implies that birth rates are falling.

(c) Pregnancy Prevalance Analysis (PPA)--An enumeration of the current pregnancy status of samples of women. Under conditions of high fertility, the proportion of women who are pregnant at any particular moment is high; under conditions of lower fertility the proportion is low. If this percentage declines, it implies that birth rates are falling.

(d) Pregnancy History Analysis (PHA)--the system proposed in this article.

All of these systems are too new and too little studied to permit more than a few comparative comments. The following observations concerning the relationship between the pregnancy history approach and the other approaches are submitted in this spirit.

(1) The Pregnancy History Analysis provides, as routine items of output, both the "open interval" and the Pregnancy prevalance measures. As yet, neither of these measures has been "calibrated," that is, the average length of intervals and the average prevalance of pregnancy associated with given levels of birth rates have not yet been determined. The pregnancy history approach provides a highly feasible procedure for doing this, and the authors are pursuing this problem currently, using the data for Latin America and Alabama for the calibration process.

(2) There are some research projects to use laboratory methods (urine samples) to measure pregnancy prevalance. These procedures are costly; both in terms of equipment and personnel.
The combined difficulties of nonresponse and of inconclusive laboratory results may lead to the find that this procedure is no more valid and reliable than the results obtained simply by asking two questions: "Are you pregnant now?" and (if so), "For how long have you been pregnant?" ("How many weeks or months has it been since you menstruated last?")
By limiting the tabulation to pregnancies beyond the third month, it is believed that the simple interview methods can match or surpass the field laboratory methods in completeness and precision.

(3) There have been some suggestions to "bobtail" the pregnancy history approach, such as taking a pregnancy history for only the past five years. (One such version is now being performed in Paskitan, under the sponsorship of Columbia University.) It is believed that the elements of imprecision introduced by this procedure are so great as to render the results unusable. Tying the rates to a fixed date so far in the past can lead to spurious inclusions and omissions far more serious than simple memory lapse. Only by asking low-literacy respondents to account for the totality of their reproductive experience

beginning with all of their living children and carefully reconciling all of the information pertaining to date at marriage, current age, and probing all long intervals without pregnancy is it possible to obtain maximally valid data. The difference, in terms of interviewing time, is small and the gains in terms of precision and additional information concerning intervals, pregnancy loss rates, etc. more than repays the modest extra effort.

(4) The Population Growth Estimation procedure is a completely different system and is the only clear-cut alternative to the Pregnancy History Analysis system. As yet, a rigorous comparative test has not been made, but it is believed that the estimates of fertility levels yielded by the PHA system are just as reliable as those yielded by PGE, and the estimates of change are more reliable. The PHA has the following advantages in comparison with PGE:

(a) PGE requires a sustained effort, over a prolonged period of time. In a one-shot inventory, it is possible to mobilize a crew of high quality workers, train them to peak efficiency, and then dismantle the organization. In the developing countries, a single one-shot effort of the PHA type can be mounted quite nicely through a medical school, a school of social work, or a demographic center--whereas it has proved difficult to employ, retain, and maintain a high level of enthusiasm for PGE operations.

(b) Poor quality of work in PHA results from failure to reduce memory lapse to the point where the demographic corrections produce valid results. Careless interviewing thus tends to underestimate birth rates for earlier years. This leads to an inference that fertility has risen or stayed the same. Poor quality work in PGE results in an undercount of births after a good start, with the result that birth rates apparently decline. Thus, there is a built-in bias against discovering a fall in birth rates in the PHA approach, while the PGE approach has a built-in bias in favor of discovering a fall in birth rates.

(c) The repeated visits of households for purposes of registering births while maintaining a duplicate registration system creates serious problems for contamination in PGE (the interviewer and the registrar both know of each other's work, and have many months to communicate with each other). Also, there are problems of irritating respondents by repeated interviewing, of low interviewer morale because births, deaths, and migration occur with such infrequency that on quarterly visits the normal response is "no event," and the perennial problems of migration, matching names, and reinterviewing in exactly the same household at repeated periods of time. The PHA system by-passes all of these problems.

(d) PGE requires very large samples of households and a long span of time before results are forthcoming. It is prospective only, and hence yields information only after one year or more of work. PHA provides the same level of precision with smaller samples and (because it can be retrospective as well as prospective) after a very short time.

(e) PHA, by reconstructing the past, can provide a baseline to evaluate family planning programs already in operation. PGE can only begin at the present and work into the future. Thus, PHA can evaluate family planning programs that have been underway for two or three years, whereas PGE cannot evaluate the work done previously, but can measure future programs only.

(f) The demographic adjustments that are made to control the biases and deficiencies for PHA are almost identical for the successive intervals of time, and therefore are of the nature of constants that cancel out when measuring fertility change, because they are present in both the "before" and the "after treatment" intervals. The demographic corrections that are made to PGE data can vary independently at each interval of time. For this reason, it is believed that the PHA measurements of fertility change are substantially more precise in most instances than PGE estimates of fertility change.

(5) The major weakness of the PHA is that it is highly sensitive to errors in following a sampling plan. If interviewers fail to interview unmarried but eligible women,

or women who are married but have born no children, the rates are affected directly and drastically. Utmost care in selection, training, and supervision of interviewing is required. The sampling plan must be near-perfect.

The above comments are not intended to argue that the pregnancy history approach should replace the PGE approach. They are only intended to emphasize that the PHA system merits serious consideration, and that it does have some assets for family planning evaluation that are urgently needed around the world today.

FOOTNOTES

[1] See Mortimer Spiegelman, Introduction to Demography, pp. 153-55, 167-68. The "total fertility rate" is equivalent to the gross reproduction rate" taken for all births instead of female births only. It is superior to the GRR for fertility measurement because differences in sex ratio at birth are not allowed to be confounded with fertility level.

[2] For a complete exposition of the concept of person years and the relationship between calendar years and time-in-age, see H. H. Wolfenden, Population Statistics and their Compilation, University of Chicago Press, 1954, especially Chapter 5.

[3] A minor difference is the fertility of women who die during the childbearing years. The pregnancy history approach includes only the fertility of women who survive to be interviewed at a particular age. The fertility of the women of the various cohorts who have died is left out of the numerator, and the person-years of these women is left out of the denominator. As a result, the fertility rates obtained by the pregnancy history approach may be slightly higher than those calculated from vital registers. To the extent that there is differential survival, the sample of surviving women is not representative of the cohorts at earlier years. These are well-known problems of all modes of longitudinal analysis, either prospective or retrospective.

[4] The authors have several versions of this program to accomplish specific types of tabulations, such as nuptiality-specific rates, spouse-present specific rates. Appendix B is the "standard" version of the program for generating rates based on all women of childbearing age.

[5] It should be emphasized that it is not essential that highly accurate data on month and year of birth be obtained to make this system workable. Season of birth may be accepted instead of month of birth, and current age may be accepted instead of year of birth, where more precise data are lacking. Thus, the system can never be worse than a census; with skilled interviewing it can be considerably better. The precision of the rates will depend upon the precision of the basic data. With even the crudest pregnancy histories (so far as dates are concerned), if the count of events is complete the resulting rates will be useful.

[6] The "Pregnancy Analysis" computer program sets up three matrix for life births (a) total live births, (b) births that survived one year and (c) live births still living. It calculates ASFR and GFR on the basis of each. The raw data and the rates for all three matrixes are printed out and therefore are completely available to the researcher for experimenting with alternative systems of adjustment.

[7] A detailed analysis of these data for the seven Latin American capital cities is contained in a forth coming monograph being published by the United Nations Demographic Center, Santiago. Prof. Carmen Miro, director of the Center, is senior author of this

[8] It should be pointed out that the major errors and biases in the pregnancy history data are of such a nature that they tend to hide a decline in fertility when in fact one has occurred, rather than to give a spurious indication of fertility decline:

(a) Women will have less memory loss for recent pregnancies than for ones more distant in the past, so recent birth rates will tend to be higher than earlier ones.

(b) Infant mortality is declining. Therefore, if an average correction for both periods is introduced, it will tend to undercorrect the earlier period and overcorrect the later period, thereby understating fertility at the earlier period and overstating it at a later period.

(c) Women who are older may have more incentive and ability to successfully misreport or lie about illegitimate children. Thus, this type of misreporting may be much more serious for earlier than for current periods, because for current periods the children are present and visible.

(d) When interviewers make errors in selecting women for the sample, there is a tendency to omit single women without children, and especially at the younger ages. The tendency is to exaggerate the fertility of women of younger ages in recent periods.

Thus, when the pregnancy history is used to measure fertility change, it is conservative evaluative technique. If it finds that fertility has declined, this finding has been arrived at in spite of the major biases of the technique, not because of them.

9 Department of Obstetrics and Gynecology, University of Chile. Dr. Faundes has a report of this experiment and its results to day in preparation. This table has been presented through his kind permission.

10 For a description of this study see Donald J. Bogue, The Rural South Fertility Experiments, Community and Family Study Center, 1966.

See, for example, Helen M. Walker and Joseph Lev, Statistical Inference, New York: Holt, Rinehart and Winston, 1953, pp. 68-76. The practice of improving sample precision by aggregating data for fertility behavior of two or more years violates the principles for combining probability samples on a variety of grounds. The presumption made here that they are approximately additive is only a preliminary judgment. The problem is being pursued in more detail with the assistance of Prof. Leo Goodman.

Table 1.--ILLUSTRATION OF THE "IDEAL" DEMOGRAPHIC PROCEDURE FOR MEASURING FERTILITY CHANGE: DATA FOR THE UNITED STATES: 1960 AND 1965

Age	Age specific fertility rates		Change in rates: 1960 to 1965	
	1965	1960	Absolute	Relative
	(1)	(2)	(3)=(2)-1	(4)=(3)/2
15-19 years...........	71.2	89.9	-18.7	-20.8
20-24 years...........	196.8	258.1	-61.3	-23.8
25-29 years...........	162.5	197.4	-34.9	-17.7
30-34 years...........	95.0	112.7	-17.7	-15.7
35-39 years...........	46.4	56.2	- 9.8	-17.4
40-44 years...........	12.8	15.5	- 2.7	-17.4
45-49 years...........	0.8	0.9	- 0.1	-11.1
General fertility rate[a]...............	131.1	156.3	-25.2	-16.1
Total fertility rate................	2928	3654	-726	-19.9

(a) GFR for U.S. is computed on the basis of women aged 15 to 44; most nations of the world use 15 to 49 or 10 to 49.

Table 2. ILLUSTRATION OF MATRIX USED FOR CUMULATING NUMERATORS AND DENOMINATORS AND CALCULATING BIRTH RATES BY PREGNANCY HISTORY ANALYSIS PROGRAM

PERSON YEARS - TOTAL WOMEN - CLINIC - ALABAMA

RATES-SINGLE YEARS TOTAL L.B. / PERSON YEARS

AGE	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
TOT.	134.	160.	245.	290.	357.	417.	395.	415.	443.	404.	399.	393.	372.	380.	349.	344.	309.	380.
1923	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1924	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1925	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1926	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1927	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1928	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1929	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1930	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1931	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1932	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1933	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1934	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1935	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1936	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1937	281.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1938	92.	281.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1939	92.	92.	375.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1940	283.	183.	92.	375.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1941	75.	283.	275.	92.	469.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1942	59.	75.	94.	275.	550.	188.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1943	123.	59.	75.	283.	92.	458.	281.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1944	111.	0.	178.	151.	378.	275.	366.	375.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1945	60.	111.	246.	0.	75.	472.	92.	550.	94.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1946	56.	241.	111.	185.	297.	226.	189.	183.	550.	188.	0.	0.	0.	0.	0.	0.	0.	0.
1947	53.	224.	181.	276.	246.	297.	226.	378.	366.	458.	281.	0.	0.	0.	0.	0.	0.	0.
1948	188.	212.	168.	422.	55.	246.	653.	302.	472.	366.	275.	281.	0.	0.	0.	0.	0.	0.
1949	249.	125.	159.	280.	362.	608.	431.	297.	604.	189.	458.	458.	188.	0.	0.	0.	0.	0.
1950	106.	187.	250.	372.	393.	482.	166.	308.	416.	226.	283.	550.	275.	188.	0.	0.	0.	0.
1951	41.	0.	311.	188.	425.	336.	543.	498.	308.	356.	302.	189.	550.	458.	375.	0.	0.	0.
1952	214.	162.	159.	249.	625.	372.	224.	302.	553.	246.	416.	377.	189.	275.	366.	0.	0.	0.
1953	79.	171.	122.	372.	373.	250.	372.	449.	422.	387.	615.	416.	377.	283.	458.	550.	188.	0.
1954	42.	199.	214.	284.	319.	560.	438.	265.	336.	663.	221.	431.	475.	377.	0.	275.	183.	375.
1955	44.	125.	199.	300.	365.	531.	746.	500.	372.	393.	603.	553.	308.	238.	604.	378.	550.	824.
1956	122.	177.	292.	437.	429.	365.	531.	435.	563.	319.	505.	543.	166.	554.	119.	377.	189.	183.
1957	254.	122.	221.	125.	278.	386.	405.	531.	435.	375.	531.	280.	482.	166.	492.	356.	377.	378.
1958	182.	109.	367.	354.	333.	517.	343.	324.	372.	435.	438.	425.	280.	543.	608.	431.	238.	453.
1959	286.	45.	363.	204.	399.	375.	437.	386.	446.	531.	249.	313.	265.	449.	543.	221.	369.	416.
1960	90.	229.	318.	218.	490.	531.	250.	636.	300.	365.	478.	497.	375.	531.	224.	663.	221.	246.
1961	329.	134.	343.	364.	290.	449.	399.	458.	517.	557.	405.	212.	187.	438.	212.	168.	362.	332.
1962	85.	247.	134.	229.	318.	689.	571.	443.	417.	477.	343.	486.	425.	497.	375.	478.	280.	422.
1963	97.	338.	329.	358.	629.	273.	399.	367.	664.	667.	397.	300.	527.	478.	560.	375.	425.	336.
1964	0.	387.	761.	658.	582.	457.	500.	580.	612.	399.	333.	318.	600.	568.	319.	311.	313.	478.
1965	0.	0.	387.	507.	247.	358.	400.	409.	435.	286.	531.	417.	278.	171.	162.	372.	187.	250.
1966	0.	0.	0.	0.	0.	424.	324.	379.	98.	303.	88.	0.	651.	94.	195.	103.	0.	145.
1967	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.

Table 2. (continued)

CONT.	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	TOTAL
TOT.	307.	342.	293.	304.	255.	227.	181.	'229.	66.	121.	199.	78.	333.	0.	0.	0.	0.	325.
1923	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1924	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1925	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1926	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1927	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1928	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1929	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1930	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1931	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1932	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1933	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1934	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1935	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
1936	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	360.
1937	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	158.
1938	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	234.
1939	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	196.
1940	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	214.
1941	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	232.
1942	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	209.
1943	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	194.
1944	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	233.
1945	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	181.
1946	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	213.
1947	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	284.
1948	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	292.
1949	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	333.
1950	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	314.
1951	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	302.
1952	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	306.
1953	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	334.
1954	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	312.
1955	188.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	380.
1956	366.	281.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	349.
1957	275.	641.	281.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	337.
1958	378.	366.	366.	281.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	362.
1959	151.	283.	275.	458.	281.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	338.
1960	356.	75.	94.	92.	92.	188.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	331.
1961	492.	297.	453.	378.	458.	275.	375.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	354.
1962	276.	308.	356.	377.	94.	275.	92.	188.	0.	0.	0.	0.	0.	0.	0.	0.	0.	373.
1963	362.	442.	369.	297.	226.	94.	275.	275.	94.	0.	0.	0.	0.	0.	0.	0.	0.	387.
1964	449.	422.	387.	246.	475.	302.	189.	366.	0.	188.	0.	0.	0.	0.	0.	0.	0.	437.
1965	106.	168.	241.	55.	185.	238.	75.	283.	0.	92.	94.	0.	0.	0.	0.	0.	0.	272.
1966	304.	133.	0.	507.	120.	0.	0.	0.	0.	0.	164.	0.	0.	0.	0.	0.	0.	179.
1967	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.

Table 3.--ESTIMATED BIRTH RATE OF MEXICO CITY, DERIVED BY THE PREGNANCY HISTORY TECHNIQUE FROM A SAMPLE SURVEY: 1964; COMPARISON WITH OTHER MEASURES OF FERTILITY FOR MEXICO

Age	Pregnancy history estimates			Estimate for Mexico-Dr. Cho	United Nations Yearbook
	High	Medium	Low		
15-19 years.......	103	98	98	105	48
20-24 years.......	297	282	271	293	299
25-29 years.......	316	300	287	312	314
30-34 years.......	270	257	252	257	271
35-39 years.......	160	152	148	192	200
40-44 years.......	64	61	66	94	49
45-49 years.......	8	8	7	0	
General fertility rate............	217	206	201	196	...
Total fertility rate............	6091	5790	5645	6268	6150

Table 4.--ESTIMATE OF CHANGE IN FERTILITY IN SAN GREGORIO FAMILY PLANNING EXPERIMENTAL AREA, SANTIAGO CHILE, 1962-66

Age	Period of family planning: 1965-66	Period before family planning: 1963-64	Fertility change	
			Absolute	Relative
15-19 years........	133	180	- 47	-26.1
20-24 years........	325	374	- 49	-13.1
25-29 years........	267	339	- 72	-21.2
30-34 years........	240	258	- 18	- 7.0
35-39 years........	163	199	- 36	-18.1
40-44 years........	93	74	+ 19	+25.7
45-49 years........	8	0	+ 8	.....
General fertility rate............	182	226	- 44	-19.5
Total fertility rate............	6145	7120	-975	-13.8

Table 5.--ESTIMATE OF CHANGE IN FERTILITY IN MEXICO CITY: 1960-64

Age	2 1/2 years preceding interview: 1962-64	Preceding two years: 1960-61	Change	
			Absolute	Relative
15-19 years.......	98	98	....	....
20-24 years.......	261	310	- 49	-15.8
25-29 years.......	300	300	0	0
30-34 years.......	270	238	+ 32	+13.4
35-39 years.......	137	171	- 34	-19.9
40-44 years.......	74	43	+ 31	+72.1
45-49 years.......	4	14	- 10	-71.4
General fertility rate............	205	207	- 2	- 1.0
Total fertility rate............	5720	5870	150	- 2.6

APPENDIX A

PREGNANCY HISTORY

1.. Have you ever given birth to a child or ever been pregnant?
No.................0 Yes.............1

RECORD ALL ANSWERS TO QUESTIONS 15-16 ON THE PREGNANCY HISTORY CHART. IF THE ANSWER TO 15 IS "NO" WRITE "NEVER PREGNANT" ON THE FIRST LINE OF THE CHART AND PROCEED TO QUESTION 19. IF "YES" ASK: (REMEMBER THAT THE ANSWERS TO THE FOLLOWING QUESTIONS ARE TO BE RECORDED ON THE CHART AND NOT ON THIS PAGE. DO NOT NUMBER PREGNANCIES UNTIL CHART IS COMPLETELY FILLED OUT.)

A. Please tell me the names of each one of your babies who was alive when born. Begin with the oldest.

PROBE: We want to be sure to include all of your children. Did you have any children by another husband(or boyfriend)that you have not mentioned?

WRITE THE NAME IN COLUMN A OF THE CHART - USE THE HEAVY BLACK LINES TO RECORD THE INFORMATION FOR CHILDREN BORN ALIVE. FILL IN THE OTHER COLS. B--J FOR EACH LIVE BORN CHILD.

B. (Outcome of pregnancy for live births is "LB")

C. What was the sex of the child? (ASK ONLY IF CANNOT TELL FROM FIRST NAME)

D. In what year was the child born? In what month of the year? (IF MONTH UNKNOWN ASK: IN WHAT SEASON?)

E. Were you pregnant the full 9 months with this child? IF NOT ASK: During which month of pregnancy was it born?

F. Was this a single birth pregnancy or was it one of twins? (triplets?)

G. Is the child still living?

IF NOT ASK: H-I-J-. IF "YES": PROBE WITH ITEM K.

H. In what year did death occur? In what month?

I. How old was he (she) when death occurred? (IN MONTHS, IF LESS THAN 1 YEAR)

J. What was the cause of death?

K. PROBE: (1) Are you sure you have mentioned all of your children who are living? Are there any who are living away from home that have been forgotten? We want to include sons and daughters who are married or have left home. (INCLUDE ONLY CHILDREN BORNE BY THE WOMAN HERSELF, NOT ADOPTED CHILDREN OR HUSBAND'S CHILDREN BY ANOTHER WIFE)

(2) Are there any other children who were born alive but have died? We want to include any babies that may have lived only a few hours or any that have died after growing up.

2. Many women have pregnancies that do not produce a live baby. Have you ever had a pregnancy that did not produce a live baby, that is, the baby was born dead, or have you ever been pregnant any other time and lost it, because of miscarriage or abortion?

IF "YES" ASK AND RECORD THE FOLLOWING ON BLANK LINES BETWEEN THE OTHER BIRTHS IN THE ORDER OF OCCURRENCE.

A. Between which of the children (live births) did it occur?

D. Date of pregnancy loss - What year was this? What month of the year? (COL. D OF CHART)

E. How many months pregnant were you when the pregnancy loss happened? During which month of pregnancy did the loss occur? (COL. E OF CHART)

B. What was the cause of this pregnancy loss? Sometimes women who become pregnant when they do not want to have another child do something to interrupt the pregnancy. Did you do anything or have anything done to cause the pregnancy to end (Have an abortion)?
Yes.................IA No..............SA or SB
RECORD TYPE OF PREGNANCY LOSS IN COL. B.

F. Do you know if this would have been a single birth or were you not far enough along to tell? RECORD SEX IN COL. C.

AT THIS MOMENT OF THE INTERVIEW PLEASE REVIEW THE ENTIRE HISTORY OF PREGNANCIES AND NOTE THE LENGTH OF THE INTERVALS BETWEEN PREGNANCIES. IF THERE IS AN INTERVAL OF TWO OR MORE YEARS BETWEEN ANY TWO PREGNANCIES OR PREGNANCY LOSSES, ASK:

How does it happen that there is an interval of ____________years between the births of ________ and __________? This is an unusually long time. Is it possible that you were pregnant again and forgot to mention it? Perhaps you were pregnant for only a few weeks?

ENTER ANY ADDITIONAL PREGNANCIES ON PREGNANCY HISTORY CHART IN PROPER PLACE AND ORDER.

3. Are you pregnant now?

No..................0
Yes.................1*
Uncertain, probably...2*

*IF NOW PREGNANT:

A. Order of pregnancy__________________________
B. In which month of pregnancy are you? ____________________
C. Therefore baby is due (month and year) ______________________

AFTER RECORDING ALL INFORMATION ABOUT EACH PREGNANCY, NUMBER EACH PREGNANCY IN THE CORRECT ORDER OF OCCURRENCE AND FILL OUT THE FOLLOWING SUMMARY:

(a) Number of children still living....................____________
(b) Number of live born children now dead..............____________
(c) Total number of live born children (a+b)............____________
(d) Number of abortions: Spontaneous (less than 5 months gestation)................................____________
(e) Number of abortions: Induced (less than 5 months gestation)..____________
(f) Number of stillbirths and miscarriages (more than 5 months gestation)................................____________
(g) Is the woman pregnant now? If "yes", record "1"....____________
Total number of pregnancies.........................____________
(..m of c through g less any multiple births)

Table 6.--ESTIMATES OF CHANGE IN FERTILITY IN RURAL ALABAMA:

Age	Two years of family planning program	Two years preceding program	Change	
			Absolute	Relative
15-19 years.......	422	612	-190	-31.0
20-24 years.......	352	513	-161	-31.3
25-29 years.......	294	448	-154	-34.3
30-34 years.......	200	390	-190	-48.7
35-39 years.......	152	305	-153	-50.1
40-44 years.......	88	190	-102	-53.6
45-49 years.......	0	0	....	....
General fertility rate............	252	433	-181	-41.8

PREGNANCY HISTORY

Preg- nancy order	Name of the child (if live born)	Outcome of Pregnancy LB=live born SA=spontane- ous abortion IA=induced abortion SB=stillbirth	Sex boy or girl?	Date of birth or pregnancy loss		Gestation for each live birth, abortion, miscarriage, stillbirth. In which month of pregnancy did you lose this child?	Type of birth. How many babies?	Is the child still living?	FOR EACH CHILD BORN ALIVE NOW DECEASED: Date of death		FOR EACH CHILD BORN ALIVE NOW DECEASED: Age when death occurred	FOR EACH CHILD BORN ALIVE NOW DECEASED: Cause of death
				Year	Month				Year	Month or season		
	A	B	C	D		E	F	G	H		I	J

X

TRENDS IN NEGRO-WHITE DIFFERENTIALS

Chairman: DANIEL O. PRICE, University of Texas

THE OCCUPATIONAL DISTRIBUTION OF THE RETURNS TO EDUCATION AND EXPERIENCE FOR WHITES AND NEGROES[1]

Lester C. Thurow, Harvard University

Individuals earn income by employing their skills and knowledge -- their human capital. To earn larger incomes they must increase either the quantity or the price of their human capital. Education and on-the-job training provide the principle means for increasing the quantity of human capital. Migration, improvements in information, and the elimination of other market imperfections, such as discrimination, provide the principle means for raising the price of existing capital.

Although price and quantity effects are theoretically distinguishable, in practice the distinction is blurred by using observed income flows to measure indirectly the value of human capital. Price and quantity effects are lumped together as changes in value. In most cases this is not a serious problem since both the individual and society are interested in raising the value of human capital. Real investment is usually necessary to alter either price or quantity. The basic problem is finding that investment which will earn the greatest return. It may be an investment which will increase the quantity of human capital or it may be an investment which will raise the price of human capital.

When observed income flows are used to measure the value of human capital, efforts to measure the specific effects of any one factor must make explicit allowance for the impacts of all other factors. For example, since innate ability (whatever it is and however it is measured) and education levels are probably linked together, the observed income flows that are associated with higher levels of education are caused partially by education and partially by innate ability. If some correction is not made for ability, observed income flows will overstate the actual returns to education.

A similar problem is presented by on-the-job training. Since training and education are associated together,[2] the returns to more education will be overstated if the effects of training are not considered. Since training programs have costs as well as benefits, the error is compounded in calculations of net returns to education. Training benefits are included in the returns to education, but training costs have not been added to the costs of education. Ability did not have any associated costs which needed to be considered. Strenuous efforts have been made to isolate the returns to education from those caused by ability.[3] Less effort has been made to solve the problems presented by on-the-job training.[4] This is not surprising. Practically no direct information is available on either its amount or its costs. The informal aspects of much of the training mean that there is no practical method to obtain direct information. Much of the training is acquired in the course of work and does not result from deliberate training programs. Costs are involved, but they are difficult to estimate.

The problem presented by on-the-job training goes beyond that presented by ability in another way. Since innate ability cannot be altered by definition, society and individuals want to know the returns to increasing education (a variable which can be altered). On-the-job training, however, is not innate. It can be altered. Society and the individual want to know what combination of education and training yields the greatest net return.

The standard technique for isolating the returns to any one factor, such as education, has been to hold all other explanatory factors constant and then note the remaining differences in observed income flows. Either regression techniques or detailed data are used to hold the other explanatory factors constant. Both adjustment techniques assume that the effects of each of the explanatory factors is independent of all other explanatory factors and that their separate effects are additive. Thus, the amount of training is assumed to have no influence on the returns to education and the returns to increasing both education and training are assumed to be equal to the sum of the separate returns to increasing each variable independently.

In fact, many of the explanatory variables which affect income flows are not independent but complimentary. Returns are not additive but multiplicative. This is clearly seen in on-the-job training and education. The returns from training partially depend on the level of formal education possessed by the trainee. Low education levels make some types of training impossible and other types expensive. As education levels rise, training costs fall and the variety of training which can be given expands. Complimentarities also work in the opposite direction. Without training, education is of little value. Most jobs require some knowledge which is peculiar to the job and which is not or cannot be acquired in school. Without this training, education is of little value. With complimentarities the benefits from both education and training will be larger than the sum of the benefits from education and training separately.[5]

Corresponding to the earlier distinction between price and quantity effects on the value of human capital, there are two main sources of complimentarities. Technological complimentarities occur when the skills and knowledge acquired in school are complimentary with the skills and knowledge acquired in training. Price complimentarities occur when market imperfections are reduced in the process of acquiring education and training. Thus, a Negro might receive price complimentarities if higher levels of education and training allowed him to move into occupations with less discrimination.

The degree of complimentarity obviously differs for different jobs. In some jobs there are few complimentarities; in others education and training are linked rigidly together. A priori reasoning leads to the conclusion that complimentarities are important and the data presented below confirm this conclusion.

Ignoring complimentarities leads to biased estimates of the returns to increasing education. Holding training levels constant while observing the returns to education may provide a valid estimate of the returns to education within each training level, but it provides a distorted view of the general returns to education. Part of the returns to education arise from shifting training levels as well as moving up the income ladder within each training level. To estimate the returns to education while holding the level of training constant is to seriously underestimate the actual returns to education.[6]

The existence of complimentarities means that a precise functional relationship must be specified between the value of human capital and its explanatory factors. This function will be called the human capital function. All of the explanatory factors which are linked together by complimentarities must be considered together and their interactions specified explicitly. With complimentarities the impact of education cannot be estimated by itself. The returns to education must be estimated together with the returns to other factors.

This paper begins the task of specifying a human capital function by considering the joint impact of formal education and on-the-job training. These two factors were chosen since they are the major instruments for altering the quantity of human capital. In addition, the impact of training has typically been ignored in calculations of the returns to education. Price effects are investigated by studying the different returns to education and training across race, occupation, and region.

The empirical work presented below is not definitive since all of the relevant explanatory factors have not been considered.[7] The empirical work does, however, confirm the need for a human capital function and its precise specification. The empirical complimentarities between education and training are large and should not be ignored in evaluating programs to alter the distribution of income by changing the distributions of training or education. For males in the American economy the returns to both education and training are approximately nine times as large as the returns to education and training separately. The observed price differences across race, occupation, and region are equally large. For Negroes these price differences severely reduce the complimentarities between education and training as well as their absolute returns.

THE MODEL

Individuals receive formal and informal training while they are at work. One year of work experience will have different effects on the value of an individual's human capital depending on the amount of training received or the impact of work experience on the price of existing human capital. Everyone has one year of experience after one year of work, but the returns from that experience -- the income flows produced by it -- may be very different. If observed income flows are higher, the value of human capital has increased. This may be caused by either increases in the quantity of human capital or its price.

If price effects do not occur, income flows depend on the amount of training received. Thus, the returns to a year of work experience can be used as a surrogate variable to measure the returns to investment in on-the-job training. If labor training markets are in equilibrium, the rates of return on training will be equal for all training projects. In this case different returns to experience would reflect different amounts of investment in on-the-job training. Equal investments would earn equal returns. If labor training markets are not in equilibrium this conclusion cannot be drawn, but the returns to experience still indicate the pattern of gross benefits from training. If work experience affects the price of human capital and labor training markets are not in equilibrium, the function simply measures the gross returns to a year of work experience.

A function analogous to the production function can be used to measure the returns to experience and education and the complimentarities between them. This function is the human capital function (see equation 1). Just as the real variables and parameters on the right-hand side of a production function determine the annual flow of goods and services, so do the real variables and parameters on the right-hand side of the human capital function determine the annual flow of income which will be produced by the stock of human capital. According to the human capital function, income flows depend on the years of education and experience (Ed and Ex) the income elasticities with respect to education and experience (b and c), and a shift coefficient (A_j). Making the assumptions which were outlined above, the amount of on-the-job training is represented by the years of experience and the income elasticity with respect to experience. The shift coefficient represents the impact of different capital-labor ratios, the level of technical progress, discrimination, unionization, market imperfections, and any other relevant factors except education and experience. The shift coefficient differs among different groups and over time, but for any one group and at any point in time it can be regarded as a constant.

$$(1) \quad I_{jik} = A_j \, Ed_i^{b_j} \, Ex_k^{c_j}$$

where I_{jik} = income in occupation j for an individual with i years of education and k years of experience,

A_j = shift coefficient for occupation j,

Ed_k = i years of education,

Ex_k = k years of experience,

b_j and c_j = income elasticities for occupation j.

The impacts of education and experience on incomes can be found by taking the partial derivatives of the human capital function (see equations 2 and 3). As these derivatives indicate, the marginal product of education depends on the shift coefficient, the years of experience, and the years of education already completed. In the same manner the marginal product of experience depends on the shift coefficient, the years of education, and the years of experience already completed. As both of these functions indicate, the returns to either education or experience depend on the level of the other.

$$(2) \quad \frac{\partial I_{jik}}{\partial Ed_j} = A_j b_j Ed_j^{b_j - 1} \; Ex_k^{c_j}$$

$$(3) \quad \frac{\partial I_{jik}}{\partial Ex_j} = A_j Ed_j^{b_j} \; c_j Ex_k^{c_j - 1}$$

Since the income elasticities (b and c) are not constrained, there may be increasing or decreasing returns to increases in education and experience. If the sum of the two elasticities is greater than one there are increasing returns and if the sum of the two elasticities is less than one there are decreasing returns.

There is also no reason why the elasticities should be constant over all ranges of education and experience. College may produce higher returns than high school. To test for such differences, the function can be disaggregated into different ranges of education and experience (see equation 4).

$$(4) \quad I_{jik} = A_j \prod_{g=1}^{n} Ed_i^{b_{jg}} \prod_{i=1}^{m} Ex_k^{c_{jl}}$$

where n = education classes

m = experience classes

If the human capital function is fit to actual income data, the function can be judged by the standard statistical tests. The usual calculations of the returns to education cannot do this. Observed income differences are adjusted for what is believed to be other relevant factors and then education is assumed responsible for the residual. No statistical tests are possible on this latter assumption.[8] The human capital function, however, has the advantage of providing statistical tests of its own validity as well as providing estimates of the empirical size of the relationships.

THE DATA

The human capital function was fitted to 1960 mean income data for males 18 to 64 years of age.[9] Functions were estimated for whites and Negroes,[10] for ten occupations, and for the North and South. Years of experience were calculated by assuming that each individual begins work at 18 if he has finished school by this age. If not, work begins at the school-leaving age. Thus, a college graduate is assumed to begin work at 22. Eighteen was selected as the starting age for those with 12 or fewer years of education since child labor laws and workmen's compensation laws prevent earlier entry into many jobs. A worker's years of experience were found by subtracting his starting age from his current age. To test the hypothesis that different ranges of education and experience have different elasticities, the education variable was divided into three variables (0-8 years, 9-12 years, and more than 12 years) and experience was divided into four variables (0-5 years, 6-15 years, 16-35 years, and more than 35 years). In a cross sectional analysis of male incomes the human capital function worked very well. In most classifications over ninety-five percent of the variation in incomes was explained (see Appendix A).

THE RESULTS: WHITE VERSUS NEGRO

Income elasticities with respect to education and experience were not constant for all levels of education and experience. For white males the income elasticity was 0.23 for elementary education, 0.61 for high school education, and 2.10 for college education[11] (see Table 1). For Negro males the education elasticities were 0.32, 0.32, and 1.49, respectively. The elasticities for Negroes were slightly larger for elementary education, but much smaller for high school and college education.

White male income elasticities with respect to experience were 0.44 for both the first five years and the next ten years of experience, 0.13 for the 16th to 35th years and -0.28 for anything over 35 years of experience. For Negro males the income elasticities of experience were 0.13 in the first five years, 0.56 in the next ten years, and -0.08 for anything over 15 years of experience. Negroes receive less benefit from experience early in their careers and their human capital begins to depreciate sooner than that for whites.[12]

TABLE 1

Income Elasticities of Education and Experience[13]

	Education			Experience				Shift
	0-8	9-12	12 & up	0-5	6-15	16-35	35 & up	Coefficient
Professional								
Northern White	.05	.43	1.54	.43	.43	.20	-.24	$1619
Southern White	.21	.73	1.57	.32	.60	.16		971
Nonwhite	.94	.94	.94	.14	.59			216
Managerial								
Northern White	.16	.58	1.95	.48	.36	.36		1389
Southern White	.11	.80	1.79	.19	.51	.25		1813
Nonwhite	.30	.66	.66	.42	.42			739
Clerical								
Northern White	.09	.24	.83	.25	.48			1695
Southern White	.19	.19	1.30	.52	.19	.19		1005
Northern Nonwhite	.05	.38	.54	.33	.33	.04		1465
Southern Nonwhite	.14	.59	.59	.37	.37	.01	-.12	925
Sales								
Northern White	.14	.43	1.59	.24	.69			1417
Southern White	.24	.59	1.83	.23	.65			979
Craftsmen								
Northern White	.08	.23	1.11	.21	.40	-.06		2225
Southern White	.16	.43	1.31	.39	.39	-.07	-.18	1196
Northern Nonwhite	.13	.13	.67	-.06	.58	-.11		2143
Southern Nonwhite	.06	.58	.58	.33	.33	-.03	-.44	1025
Operatives								
Northern White	.08	.23	1.11	.21	.40	-.06		2225
Southern White	.17	.44	.91	.62	.26	-.06	-.24	829
Northern Nonwhite	.06	.16	.41	.30	.30			1661
Southern Nonwhite	.17	.25	.64	-.07	.55			1233
Laborers								
Northern White	.09	.26	.26	.42	.42	-.15		1212
Southern White	.22	.65	1.07	.76	.15	-.06	-.39	360
Northern Nonwhite	.03	.21		.45	.26	-.02	.15	1235
Southern Nonwhite	.16	.16	.94	.36	.36	-.06		720
Services								
Northern White	.20	.20	.71	.37	.62	-.19		887
Southern White	.14	.78		.92	.26	-.09	-.65	469
Northern Nonwhite	-.01	.32	.32	.87	.05	.05		732
Southern Nonwhite	.07	.25	.74	.38	.38	-.06		783
Farmers								
Northern White	.08	.54	1.31	.54	.32	-.16	-.41	948
Southern White	.33	1.17	2.84	.23	.55	-.11	-.11	556
Total								
White	.23	.61	2.10	.44	.44	.13	.28	879
Nonwhite	.32	.32	1.49	.13	.56	-.08		797
Northern White	.15	.56	2.07	.44	.44	.12	-.31	1107
Southern White	.25	.83	2.15	.45	.45	.11	-.32	724
Northern Nonwhite	.07	.25	1.42	.21	.44	-.05		1537
Southern Nonwhite	.27	.27	1.81	.13	.55	-.11	-.11	728

The marginal products of education and experience can be seen in Charts 1 and 2. For the average white male with twenty years of experience, the value of education falls from $2,004 per year for the first year to $152 for the eighth year and then rises to $1468 for the sixteenth year. The returns to the average Negro are much smaller. The value of education falls from $1384 per year for the first year to $110 for the twelfth year, and then rises to $561 for the 16th year.

CHART 1

MARGINAL PRODUCT OF EDUCATION FOR MALES WITH TWENTY YEARS OF EXPERIENCE IN 1960

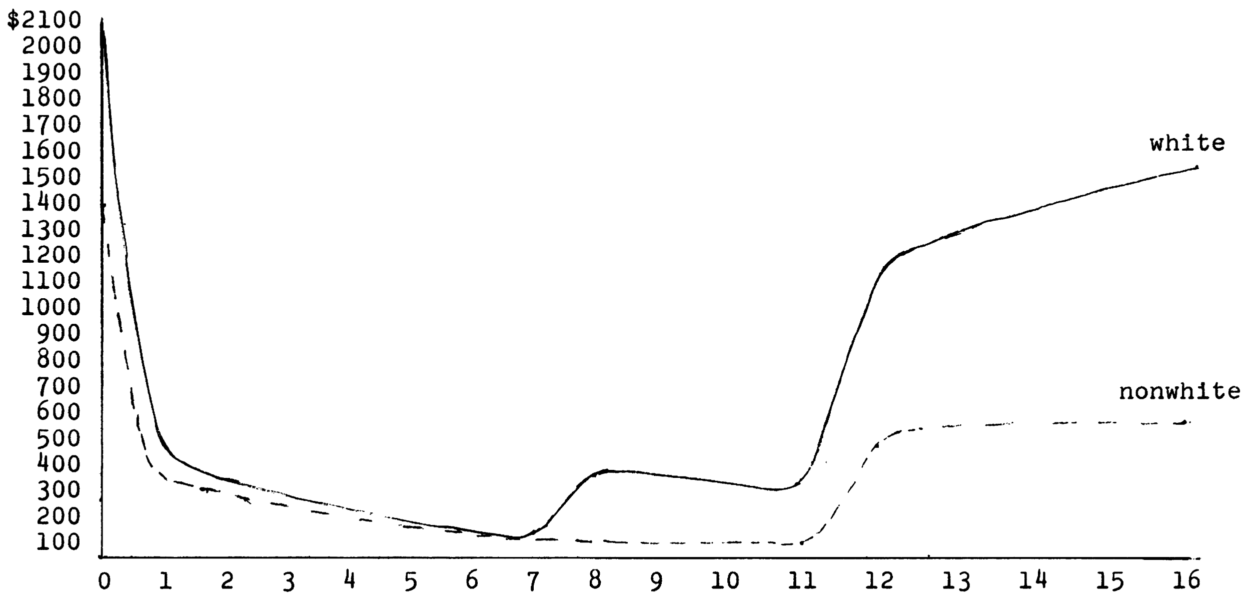
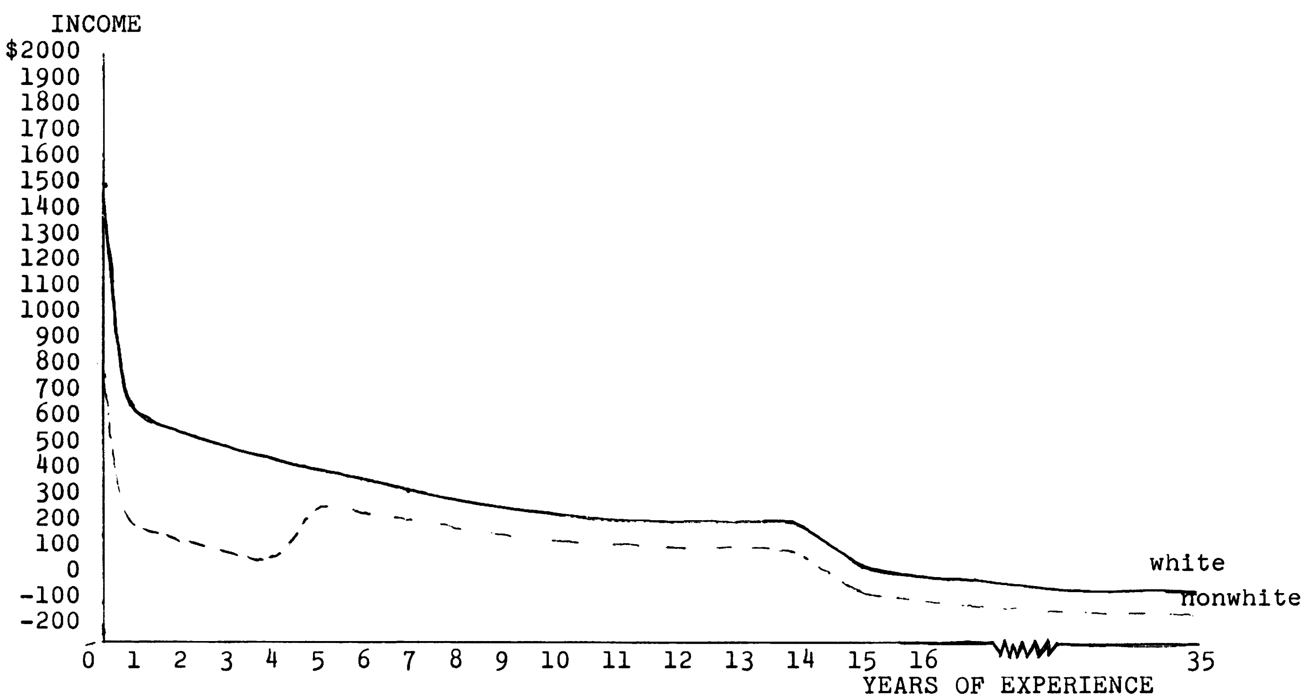


CHART 2

MARGINAL PRODUCT OF EXPERIENCE FOR MALES WITH 10.5 YEARS OF EXPERIENCE IN 1960



As education levels rise, the Negro falls further and further behind. With no education his income is $624 lower than that for a white male with no education (see Table 2). With an eighth grade education he is $1446 behind a comparable white; with a twelfth grade education he is $2356 behind; with a sixteenth grade education he is $5477 behind. The income gap grows as education rises, and grows at an increasing rate. Discrimination hurts the better educated most.

TABLE 2

THE INCOME GAP

Twenty Years of Experience		10.5 Years of Experience	
Years of Education		Years of Experience	
(Whites minus Nonwhites)			
0	$624	0	-$700
8	$1446	5	$1351
12	$2356	15	$1724
16	$5477	35	$2626
(Northern Whites minus Southern Whites)			
0	$1104	0	$38
8	$875	5	$402
12	$467	15	$596
16	$580	35	$712
(Northern Nonwhites minus Southern Nonwhites)			
0	$2030	0	-$32
8	$1241	5	$999
12	$1350	15	$1273
16	$1479	35	$1384
(Northern White minus Northern Nonwhite)			
0	-$605	0	-$489
8	$1180	5	$906
12	$1854	15	$1479
16	$5109	35	$2284
(Southern White minus Southern Nonwhite)			
0	$321	0	-$559
8	$1546	5	$1503
12	$2599	15	$2156
16	$6008	35	$2956
(Southern White minus Northern Nonwhite)			
0	-$1709	0	-$527
8	$305	5	$504
12	$1598	15	$883
16	$4529	35	$1572

Returns to experience do not show the same general U-shaped pattern. The value of experience is very high in early years, but falls off as the amount of experience increases. For the average white male with 10.5 years of education, the marginal product of experience falls from $1471 for the first year of experience to $24 for the 35th year. For Negro males, experience is less valuable and shows a slightly different pattern. The first year of experience is only worth $778, but the returns rise in the sixth to fifteenth year to levels almost equal to those for whites. The biggest gap in the returns to experience comes in the early working years. A Negro male with no years of experience and 10.5 years of education begins with an income $700 higher than that of a white in the same position (see Table 2).[15] After five years of experience, white incomes are $1351 ahead of Negro incomes. At fifteen years the gap is $1724 and at thirty-five years, $2626. Most of the income gap caused by experience is built up during the first five years of a worker's career. During these years of high training investments, Negroes are receiving much less training than whites.

NORTH VERSUS SOUTH

The marginal returns to education are higher in the South than in the North. For a white male in the North with twenty years of experience, a high school education increases his annual income by $4576 above what it would have been with no education. For a similar white male in the South, twelve years of education raise his annual income by $5213, For Negroes the comparable figures are $1906 and $2586. Marginal returns to education are lower for Negroes but the same North-South difference exists.

Although the marginal returns to education are higher in the South, average incomes are higher in the North. Higher shift coefficients and greater returns to experience more than offset the differences in marginal returns to education. A white male with twenty years of experience and twelve years of education earns $467 more in the North than he would in the South, and a Negro male earns $1350 more. The incentives for a Negro to move north to benefit from the higher shift coefficients and more extensive or higher priced training are much greater than those for a white.[16]

As educational attainment rises, the income gap between North and South shrinks, but the gap between whites and Negroes widens (see Table 2). As experience increases, the income gaps both between North and South and between white and nonwhite widen. Southern Negro males are particularly handicapped by a lack of training investments. The first five years of experience increase the income gap between northern and southern whites by $364, but they increase the income gap between northern and southern Negroes by $1031. Northern Negroes receive more investment in on-the-job training, or they have a more favorable price structure than southern Negroes, but they receive much less training or they have a more unfavorable price structure than either northern or southern whites. The first five years of experience increase the income gap between northern whites and Negroes by $1395 and between southern whites and northern Negroes by $1031.

The marginal returns to education are large (see Chart 1 and Table 2). The income differences produced by education, however, are not much larger than those produced by experience (see Chart 2 and Table 2). Differences in the returns to experience and in the training investments that produce these returns explain much of the North-South and white-nonwhite income differences. To close the income gaps would require changes in the distribution of on-the-job training as well as education.

Implications of the above results can be seen in the situation facing a high school student with 10.5 years of education who is trying to decide whether to continue school or begin working. If he is white the first year of experience is worth $1523 in the North and $1343 in the South. Another year of education is worth $12 and $15 respectively. If he is Negro, a year of experience would be worth $1179 in the North and $604 in the South. Another year of education would be worth $16 and $19 respectively. By the time the individual has accumulated twenty years of experience however, the relative values of having previously obtained another year of education or experience are reversed regardless of color. The extra year of education is more valuable than the extra year of experience. For northern whites with twenty years of experience the extra year of education raises incomes by $312 and the extra year of experience raises incomes by $37. A similar reversal exists for the Negro. If time horizons are short and discount rates are high, the individual will drop out of school. With low time horizons and low discount rates the individual will stay in school.

OCCUPATIONAL DIFFERENCES

The same general pattern of returns is visible in almost all occupations. Shift coefficients are higher in the North. Experience is more valuable in the North; education is more valuable in the South. Within each region the returns to both education and experience are lower for Negroes, but the returns to experience for northern Negroes are sometimes higher in specific occupations than the returns to experience for southern whites. Among nonwhites the returns to experience are much higher in the North and the returns to education are slightly higher in the South (see Table 1).

Craftsmen present the principle exception to these rules. The returns to education for northern Negroes are high and the returns to experience are low. Northern Negroes receive little investment in training, but a high level of education allows them to move into higher paying jobs within the crafts. In the South, experience has a much larger return for Negroes, but education a smaller one. More training is done, but formal education is not a route into higher paying jobs.

Although the general pattern of returns across occupations is similar, the magnitudes of the effects differ widely (see Table 3). This is true both across occupations and within occupations. The large differences can be seen within the professional occupations. For the northern white professional worker, 35 years of experience and no education results in an annual income $4685 higher than that achievable with no education and no experience. The same 35 years of experience is worth $1255 to a southern white but only $5 to a Negro.[17] The gains from education are reversed. The increase in income resulting from sixteen years of education is $283 for the northern white, $652 for the southern white, and $1542 for all Negroes.

Incomes rise as education or experience rises. More important is what happens when education and experience increase simultaneously. The impact is much greater than the sum of the two separate impacts. Northern white managers provide a good example (see Table 3). If the effects of having thirty-five years of experience and sixteen years of education were no greater than the sum of the two separate effects, thirty-five years of

TABLE 3.

Annual Increase in Incomes Due to Education and Experience

	Experience 35 years of experience and no education.	Education 16 years of education and no experience.	Both 16 years of education and 35 years of experience
Professional			
Northern White	$4685	$ 283	$12642
Southern White	$1255	$ 692	$11777
Nonwhite	$ 5	$1542	$ 6943
Managerial			
Northern White	$2870	$ 399	$18442
Southern White	$2799	$1754	$14832
Nonwhite	$ 539	$ 294	$ 6795
Clerical			
Northern White	$2473	$ 556	$ 6809
Southern White	$1381	$ 175	$ 5411
Northern Nonwhite	$2731	$ 227	$ 7787
Southern Nonwhite	$1239	$ 252	$ 5061
Sales			
Northern White	$2140	$ 918	$10901
Southern White	$ 838	$1098	$10150
Craftsmen			
Northern White	$2599	$ 918	$ 7736
Southern White	$1470	$ 380	$ 7869
Northern Nonwhite	$ 283	$3233	$ 4111
Southern Nonwhite	$1707	$ 207	$ 4006
Operatives			
Northern White	$1625	$1170	$ 5155
Southern White	$1280	$ 83	$ 6233
Northern Nonwhite	$2471	$ 260	$ 4737
Southern Nonwhite	$ 160	$2436	$ 3113
Laborers			
Northern White	$2079	$ 138	$ 4652
Southern White	$ 676	$ 11	$ 5351
Northern Nonwhite	$2812	$ 42	$ 3735
Southern Nonwhite	$ 778	$ 209	$ 3452
Services			
Northern White	$1039	$ 252	$ 5390
Southern White	$1375	$ 8	$ 4665
Northern Nonwhite	$3353	$ 3	$ 4022
Southern Nonwhite	$1443	$ 112	$ 3169
Farmers			
Northern White	$1889	$ 115	$ 6036
Southern White	$ 248	$1375	$ 9487

experience and sixteen years of education would raise incomes by $3269 per year above that of the individual with no education and no experience. In fact, increasing education and experience simultaneously raises incomes to $18,442 above what they would have been with no education and no experience. The combined effect of education and experience is almost six times as great as the sum of the two individual effects. The complimentarities between education and experience are large but they differ widely by occupation, race, and region.

Factors other than education and experience play an important role in income differences. These other factors are measured by the shift coefficients. They include the impact of capital-labor ratios, unionization, technical progress, discrimination, and other market imperfections. The shift coefficients range from $2225 for northern white craftsmen to $360 for southern white service workers. Shift coefficients are higher in the North than in the South, but there is an interesting racial difference. The shift coefficients for whites and nonwhites are approximately equal in the South. In the North the shift coefficients for Negroes is higher than that for whites. Northern Negro males are in occupations where the returns to education are smaller and where the investment in training is less, but they are occupations and industries, such as durable goods, with high shift coefficients. Southern Negroes receive even less training and they are not located in occupations with high coefficients.

IMPLICATIONS

The existence of strong complimentarities means that the returns from programs designed to improve education, on-the-job training, or shift coefficients are heavily dependent on what is happening simultaneously to each of the other variables. Increasing education will have little effect on incomes if the individuals work in areas with low shift coefficients and little training. Conversely, education will have a large impact on incomes if the individuals work in areas with ample training and high shift coefficients. The same complimentarities affect the returns to programs designed to alter either training or shift coefficients. If the other necessary factors are not present, the observed returns to any one program will be very low. Thus, any program designed to affect one of the variables, such as on-the-job training, will appear to fail unless it is coordinated with other programs to alter simultaneously the structure of education and shift coefficients. This means education programs, training programs, and efforts to move individuals into areas with high shift coefficients must be coordinated. The combination of policies which will produce the greatest income changes at the least cost cannot be determined abstractly. The present positions of the individuals to be aided must be determined. Only then can the marginal benefits and costs be determined.

Given the general characteristics of the poor, large returns could be earned by remedial programs designed to raise everyone in the labor force to at least eighth grade standards of literacy. The social benefits from such a program are large, but the benefits are also large from a narrow economic point of view. The marginal income flows from raising education levels in this range are great and the complimentarities with on-the-job training programs are very important. Unless an individual possesses an eighth grade standard of literacy he is under a very severe competitive handicap and as general education levels rise this handicap will grow. Since most individuals with less than an eighth grade standard of literacy are beyond the normal school age, efforts to bring the working population up to this standard must focus on adult education programs. This is precisely the area where the least effort has been made in educational programs for the poor. Concentrating on children might eliminate poverty in the long-run, but the long-run is intolerably long. Something must be done for those who are going to be in the labor force for the next thirty years.

Education plays a vital role in eliminating the differences between the income distributions for whites and Negroes. Negroes receive less education and part of the observed differences in the economic returns to education for Negroes is caused by differences in the quality of the education that is provided. The principle need, however, is for more on-the-job training. More than 80 percent of the difference between white and Negro incomes is explained by differences in the returns to experience. Negroes receive much less training. Unless this defect can be overcome, education programs will have little impact on the incomes of Negro Americans.

APPENDIX A

Equation (1) is estimated in the following form:

$$\ln I = A + b_1 \ln Ed_1 + b_2 \ln Ed_2 + b_3 \ln Ed_3 + C_2 \ln Ex_2 + C_3 \ln Ex_3 + C_1 \ln Ex_4$$

Where I = mean income level

Ed_1 = years of education possessed by individual up to a maximum of 8

Ed_2 = years of education possessed by individual up to a maximum of 12

Ed_3 = total years of education possessed by individual

Ex_1 = years of experience up to a maximum of 5

Ex_2 = years of experience up to a maximum of 15

Ex_3 = years of experience up to a maximum of 35

Ex_4 = total years of experience.

To calculate the elasticities for different levels of education and experience, the 'b' and 'c' coefficients are added together. Thus, the elasticity for the 0-8 educational range is $b_1+b_2+b_3$, the elasticity for the 9-12 range is b_2+b_3, and the elasticity for the above 12 range is simply b_3. The elasticities of experience are calculated in a similar manner. Since the log of zero is negative infinity, individuals with no education are inserted at the value of 0.01 years of education rather than at their actual value. Variables were dropped from the regressions when they did not exceed their standard errors.

South

	A	b_1	b_2	b_3	c_1	c_2	c_3	c_4	$\bar{R}^2$	S_e
Professional White	6.8783 (.1773)	-.5180 (.2208)	-.8388 (.3392)	1.5712 (.2222)	-.2790 (.2129)	.4396 (.1699)	.1593 (.0689)		.94	.101
Farmers White	6.3208 (.1505)	-.8403 (.1225)	-1.6659 (.2241)	2.8351 (.1619)	-.3205 (.1937)	.6568 (.1115)		-.1111 (.0426)	.99	.061
Managers White	7.5026 (.1215)	-.6933 (.1513)	-.9945 (.2324)	1.7938 (.1523)	-.3231 (.1459)	.2547 (.1164)	.2540 (.0472)		.97	.069
Clerical White	6.9129 (.1136)		-1.1153 (.1187)	1.3032 (.1666)	.3266 (.0836)		.1926 (.0287)		.93	.081
Nonwhite	6.8295 (.0638)	-0.4523 (.0886)		.5936 (.0578)		.3620 (.0356)	.1253 (.1061)	-.1165 (.1138)	.99	.029
Sales White	6.8867 (.1216)	-.3530 (.1520)	-1.2339 (.2442)	1.8282 (.1666)	-.4181 (.1199)	.6476 (.0560)			.97	.076
Crafts White	7.0870 (.0782)	-.2641 (.1064)	-.8830 (.1727)	1.3085 (.1179)		.4586 (.0455)	.1179 (.1117)	-.1829 (.1140)	.97	.053
Nonwhite	6.9329 (.0872)	-.5226 (.1156)		.5819 (.0894)		.3582 (.0600)	.4119 (.1763)	-.4380 (.1681)	.94	.055
Operatives White	6.7201 (.1186)	-.2689 (.0823)	-.4711 (.1533)	.9091 (.1197)	.3618 (.1626)	.3166 (.0912)	.1821 (.1202)	-.2410 (.1076)	.98	.040
Nonwhite	7.1172 (.0849)	-.0746 (.0658)	-.3903 (.1802)	.6379 (.1512)	-.6196 (.1075	.5528 (.0400)			.98	.031
Services White	6.1501 (.2509)	-.6456 (.1639)	.7807 (.1220)		.6635 (.3537)	.3478 (.1947)	.5652 (.2560)	-.6520 (.2354)	.94	.084
Nonwhite	6.6637 (.0835)	-.1844 (.1146)	-.4924 (.2860)	.7420 (.2351)		.4434 (.0526)	-.0625 (.0378)		.95	.056
Labor White	5.8871 (.2323)	-.4290 (.1361)	-.4147 (.4044)	1.0673 (.3727)	.8111 (.3291)	.2144 (.1857)	.3284 (.2436)	-.3915 (.2102)	.96	.064
Nonwhite	6.5797 (.0823)		-.7769 (.3566)	.9400 (.3476)		.4184 (.0633)	-.0627 (.0468)		.93	.064
White	6.5847 (.1150)	-.5755 (.1631)	-1.3252 (.2564)	2.1544 (.1706)		.3359 (.0639)	.4303 (.1703)	-.3164 (.1746)	.97	.081
Nonwhite	6.5898 (.1088)		-1.5479 (.1728)	1.8131 (.1534)	-.4190 (.1418)	.6586 (.1131)		-.1117 (.0511)	.95	.079

FOOTNOTES

1. The research reported herein was performed pursuant to a contract with the Office of Economic Opportunity, Executive Office of the President, Washington, D.C., 20506, and was based in part on data collected by the Equal Employment Opportunity Commission, Washington, D. C., under section 709 (c) of the Civil Rights Act of 1964, in cooperation with the Office of Federal Contract Compliance of the United States Department of Labor.

2. Mincer, Jacob, "On-the-job Training: Costs, Returns, and Some Implications," The Journal of Political Economy. Supplement, October 1962.

According to Mincer, on-the-job training accounted for 55 percent of the total training costs of those with a college education, 46 percent for those with a high school education, and 69 percent for those with an elementary education. See Mincer, op. cit.

3. For the most comprehensive attempt see: Hannoch, Giora, "An Economic Analysis of Earnings and Schooling". Journal of Human Resources, Summer 1967.

4. For the one major article see: Mincer, op. cit.

5. The same kinds of complimentarities would exist between migration and education or training.

6. This is equivalent to holding occupations constant while studying the returns to education. Many of the returns to education occur by moving across occupations rather than within occupations.

7. Most notably innate ability.

8. The additivity assumption can be compared with the human capital function of this paper by testing the equation

$$I = a + b\,Ed + c\,Ex.$$

(cont'd.)

The form of the function given in equation (1) does a better job of fitting the observed data. The additive model leads to negative incomes for groups with little education and experience.

9. All data come from the 1960 Census volume, "Occupation by Earnings and Education." North refers to North and West in census classifications.

10. Negroes and nonwhites will be used interchangeably since 92 percent of nonwhites are Negroes. All data refers to the census category nonwhite.

11. The income elasticities with respect to education are greater than those for any occupation since a large part of the gain is explained by shifts across occupations rather than within occupations. Since those with high education levels are concentrated in occupations with high shift coefficients and high income elasticities with respect to experience, the mean rise in income from low to high education levels is greater for the total population than it is for any individual occupations.

12. Age can cause skills to depreciate and skills can become obsolete. Earlier depreciation for Negroes may be due to lower health standards or less permanent skills.

13. In several occupations there were not enough Negroes to estimate the human capital function.

14. The function is fit in such a way as to provide step discontinuities in the value of the function at eight years of education and twelve years of education. Given market imperfections, completing high school may be much more valuable than dropping out one day before graduation. An alternative hypothesis would be that there are kinks in the marginal product curves at these points but not step discontinuities. This possibility is currently under study.

15. This is caused by an income elasticity with respect to education which is higher for Negroes for the first eight years than for whites. At higher education levels whites would have higher incomes.

16. To gain the benefit of higher returns to experience a Negro would have to move north to get the better training. Experience probably does not become more valuable by simply moving north, but this could happen if discrimination were lowering wages more in the South than in the North.

17. Most of the nonwhite male professionally are in the South.

18. For some numerical examples see: Thurow, Lester C., The Economics of Poverty and Discrimination, The Brookings Institution, 1968.

19. For estimates of shift coefficients by income classes see: Ibid.

TOTAL

	A	b_1	b_2	b_3	c_1	c_2	c_3	c_4	$\bar{R}^2$	S_e
Total White	6.7783	-.3796	-1.4847	2.0986		.3132	.4086	-.2778	.97	.081
	(.1161)	(.1646)	(.2588)	(.1721)		(.0645)	(.1718)	(.1762)		
Nonwhite	6.6807		-1.1696	1.4910	-.4292	.6384	-.0760		.97	.060
	(.0838)		(.1319)	(.1177)	(.1130)	(.0898)	(.0396)			
Professional Nonwhite	5.3753			.9377	-.4535	.5910			.94	.082
	(.4437)			(.1595)	(.1664)	(.0900)				

North and West

	A	b_1	b_2	b_3	c_1	c_2	c_3	c_4	$\bar{R}^2$	S_e
Professional White	7.3896	-.3824	1.1089	1.5436		.2366	.4359	-.2393	.95	.082
	(.1341)	(.1794)	(.2650)	(.1731)		(.0655)	(.1729)	(.1772)		
Farmers White	6.8539	-.4601	-.7650	1.3069	.2193	.4774	.2552	-.4102	.97	.055
	(.1508)	(.1130)	(.1810)	(.1237)	(.2136)	(.1170)	(.1427)	(.1325)		
Managers White	7.2361	-.4268	-1.3716	1.3540	.1163		.3604		.97	.077
	(.1312)	(.1668)	(.2506)	(.1624)	(.0731)		(.0267)			
Nonwhite	7.2895	-.3355	-.1535	.5384		.2878	.0448		.97	.038
	(.0689)	(.0918)	(.1514)	(.1058)		(.0355)	(.0244)			
Clerical White	7.4355	-.1496	-.5919	.3301	-.2349	.4804			.97	.050
	(.0802)	(.1011)	(.1609)	(.1071)	(.0790)	(.0369)				
Nonwhite	7.2895	-.3355	-.1535	.5384		.2878	.0448		.97	.038
	(.0689)	(.0918)	(.1514)	(.1058)		(.0355)	(.0244)			
Sales White	7.2564	-.2894	-1.1652	1.5907	-.4345	.6947			.96	.085
	(.1364)	(.1704)	(.2739)	(.1823)	(.1345)	(.0628)				
Crafts White	7.7073	-.1421	-.8818	1.1069	-.1861	.4605	-.0599		.95	.059
	(.0966)	(.1186)	(.1913)	(.1289)	(.1118)	(.1899)	(.0393)			
Nonwhite	7.6699		-.5345	.6671	-.6415	.6903	-.1094		.93	.052
	(.3725)		(.2445)	(.2339)	(.4949)	(.2234)	(.0567)			
Operatives White	7.3526	-.0977	-.1746	.3777		.4420	-.0749		.97	.047
	(.0662)	(.0933)	(.1481)	(.0974)		(.0351)	(.0258)			
Nonwhite	7.4149	0.0995	-.2512	.4121		.2062			.97	.036
	(.0528)	(.0733)	(.1825)	(.1494)		(.0108)				
Services White	6.7878		-.5176	.7142	-.2475	.8165	-.1942		.93	.097
	(.1361)		(.2143)	(.1911)	(.1835)	(.1458)	(.1644)			
Nonwhite	6.5962	-.3309		.3244	.8269		.0463		.95	.059
	(.1179)	(.0932)		(.0634)	(.0973)		(.0256)			
Laborers White	7.1000	-.1724		.2613		.5659	-.1486		.96	.061
	(.0847)	(.0865)		(.0555)		(.0484)	(.0357)			
Nonwhite	7.1189	-.1771	.2054	.1912	.2808	-.1670	.1478		.97	.038
	(.1171)	(.0821)	(.0640)	(.1675)	(.0965)	(.1256)	(.1147)			
White	7.0092	-.4056	-1.5140	2.0699		.3161	.4333	-.3094	.96	.087
	(.1241)	(.1760)	(.2767)	(.1841)		(.0690)	(.1837)	(.1884)		
Nonwhite	7.3378	-.1772	-1.1721	1.4192	-.2271	.4833	-.0456		.95	.066
	(.1077)	(.1322)	(.2133)	(.1437)	(.1247)	(.0991)	(.0438)			

DIFFERENCES BETWEEN NEGRO AND WHITE WOMEN IN MARITAL STABILITY AND FAMILY STRUCTURE: A MULTIPLE REGRESSION ANALYSIS

Myron J. Lefcowitz, University of Wisconsin *

The story line of this paper is very simple. We start with the known empirical generalization that Negroes and Whites differ with respect to marital stability and the dominance of women within the family.

Starting with Frazier, attempts have been made to explain this differential by examining Negro-White differences in social structural location and culture.[1] The implication, of course, is that these social and cultural factors are direct causes of marital instability and family structure independent of race; that is, if there were no racial differences other than skin pigmentation, marital instability would still be directly related to difference in culture and social location. In brief, race leads to different life experiences and stands as a summary of different life experiences.[2]

Very few attempts have been made to partial out these intervening factors from the relationship between race and marital stability. In general, the Frazier hypothesis that the differential rates are largely a function of the impact of slavery and subsequent emancipation in interaction with the urbanization of the Negro has been accepted.[3]

Persons in other social categories have also migrated to the cities -- although not necessarily at the same rate as Negroes -- and have also been susceptible to the impact of urbanization. Therefore, we should be able to get some maximum estimate of the current relevance of the Negro's historical situation for their greater marital instability and differential family structure, by partialing out the effects of social-ecological factors for which data are available and which affect both Negroes and Whites.

One footnote is in order here. Even if we were to find that all differences between Whites and Negroes disappear, this does not mean that there were not true differences at the time Frazier was doing his analysis, or that these differences were not indeed a function of the unique development of the family among Negroes as compared with Whites in the United States. At the time Frazier was doing his work, Negroes in large numbers were beginning to make their move from southern rural areas to the cities and to the north. The Negro family may now be going through a new historical stage, and is worth looking at for that reason alone.[4]

Recently, moreover, some systematic attempts to examine Negro-White differentials with respect to family stability and structure have been made. Bernard and Udry, both using 1960 Census data for two different population groups, arrive independently at the conclusion that controlling for socio-economic differences between non-Whites and Whites does not significantly reduce the differential in marital stability.[5,6] Udry's analysis, in fact, suggests that the differential might even increase with income.[7] A somewhat earlier unpublished paper of mine suggests that taking family income into account does significantly reduce the differential in the proportion of female-headed families except at the very lowest of income levels -- that is, under $3,000 a year.[8]

To give some idea of what is involved, take a look at Table 1. It shows the proportion of husband-wife families among Whites and non-Whites in 1960 -- by poverty status of family, and age of family head. Poverty status is measured by the Social Security Index developed by Mollie Orshansky, and corrected for 1959 price levels.[9] As can be seen, the largest difference between the races is among those families where the head is less than 25 year of age, and the family is under .7501 of the poverty line. The smallest differential is in those families where the head is aged, and the family income puts them at 25 per cent over the poverty line or better. Since poverty status is in part defined by the number of people in the household, it is possible that using the family as the unit of analysis may disguise the actual difference. As we can see in Table 2, however, the pattern of White/non-White differences remains substantially the same for the proportion of <u>persons</u> in husband-wife families as for proportion of <u>families</u>.

These tables suggest that with increasing affluence the distribution of family types among non-Whites begins to resemble that of Whites; but the differences are not obliterated. There are, however, other differences between Negroes and Whites that are also related to marital stability. Table 3 summarizes some of these differences by Whites and non-Whites for ever-married women. Non-Whites are more likely than Whites: to be

*I am most grateful to Harold Watts for his technical guidance in the preparation of this paper.

younger, to be nearer to the time of first marriage, to have less education, and to live in the South or in urban areas. Moreover, among both Whites and non-Whites, women married more than once differ from the ever-married on the same characteristics (and presumably even more from the once-married) although not necessarily in the same direction.

Thus, we have a set of social-ecological factors which appear to be differentially distributed with respect to race as well as with marital stability.

Method. Since a multivariate analysis involving so many variables is extremely difficult to handle through cross-tabulation, multiple regression analysis has been used here to ascertain what happens to the relationship between race and marital stability when all these dimensions of the differential social-ecological position of Negroes and Whites is taken into account.

In addition to the variables already mentioned, the following variables have also been taken into account: (1) the region in which the respondent was born, (2) whether they moved at all between 1955 and 1960 and (3) the relationship of this migration to their 1955 SMSA status. (See Appendix I for the definition of all independent variables.) The 1/1000 Census tape for 1960 was used. Our sample consists of all ever-married women over the age of 14, plus those women who were heads of families but had never been married. (It is of some interest to note that 2.3 per cent of our total sample fell into this latter category.) Because our interest is in comparing Negroes with Whites, all other non-Whites were excluded from the sample.

All variables, independent and dependent, were entered as continuous or ordinal attributes except race, region (current residence or birth), 1960 residence, and migration between 1955 and 1960. These attributes were treated as dichotomous or dummy variables, and therefore either zero or one. The possible nonlinear relationship between marital stability and age, duration of time since first marriage, education, and poverty status were also taken into account in the construction of the variables. For example, age was broken into three variables so that the slopes for persons less than age 30, 30 to 50, and over 50 could be independently calculated.

The definition of the dependent variables requires somewhat more discussion. (See Appendix I.) Present marital status and whether the sample members had been married once or more than once was used to define marital stability. It is very easy to agree that persons who are currently married, with spouse present, and have been married only once are the most stable; that those persons who have been married more than once without a spouse present are the most unstable; and those married only once, but without a husband present, are in between. There is some problem, however, in assigning widows -- after all, a 65-year-old woman married to the same man for 40 years could hardly be called maritally unstable. On the other hand, what about those persons who are married with spouse present, but have been married more than once -- are they stable or unstable? To ascertain whether any differences would result from varying classifications of widows and the married-more-than-once-but-with-spouse-present, four different indices of marital stability were constructed.

Family stability was measured by dividing the sample into: the married with spouse present and both only married once; the married where either had been married more than once; and female headed-families. The first was scored as most stable and the third as least. One variation was also tried -- to divide those women with spouse present by their own frequency of marriage only.

To examine female dominance in the family, the following indices were constructed: First, women were considered to be least dominant in a husband-wife family where the wife was not the chief income recipient; and most dominant where the family was headed by a woman. The husband-wife family where the wife was the chief income recipient was scored as intermediate to the two extremes. The second index was basically the same, except that the relative earnings of husbands and wives were used to differentiate the husband-wife family -- (1) where the husband's income was greater than the wife's she was considered to be less dominant; (2) where it was equal to or less than the wife's she was considered to be more dominant -- and as before, women who were the sole heads of their families were considered to be most dominant.

I want, at this time, to interject that I completely agree with the objection that none of these definitions of stability and female dominance is adequate. We are all familiar with those households where the husband brings home all the bacon, but the woman wears the pants. What I would claim, however, is that the census data used to operationalize these concepts are the best available. It behooves us who would criticize it to produce more adequate data. In the meantime, let us see what the available data tell us.

Results. The basic strategy used in the analysis started with the relationship between race and the various indicators of marital and family stability and female dominance, and investigated what happened to that relationship as different variables were introduced into the regression. Here, the coefficient of race and partial correlation of race with the dependent variables told our main story. Two sub-plots also were developed -- one was to ascertain the effect of poverty status by introducing it last into the regression; the second was to look at the change in R^2. Table 4 presents the results.

What are they?

1. The coefficient of race and the partial correlation of race with the dependent variables are both reduced by approximately half when fully regressed.
2. About half of that decrease is accounted for by poverty status alone.

To summarize: whatever race means in relation to marital stability and female dominance, half of that meaning is a summary statement of the relationship between location in a social-ecological system and marital stability -- with poverty status being particularly relevant. (It is of interest here that Lee Rainwater, in his comments on the Bernard Paper, states that if the battery of traditional demographic variables were taken into account, "Perhaps then the average difference between homogenized White and Negro categories could be reduced by as much as half.")[10]

3. A mean difference between Negroes and Whites with respect to the dependent variables does remain, given the included conditions. (In all the T-ratio is highly significant.) This difference could result from many factors. Urbanization may indeed have a larger impact on Negroes than on Whites. Movers are more likely than non-movers to be unstable. We are unable to tell, however, whether the differential is larger for Negroes than Whites.) Rainwater has suggested group process variables (e.g. community support of norms with respect to fidelity). Bernard has suggested culture and social psychological variables such as goal-striving and self-esteem. Who knows? Perhaps the mere fact that Negroes are less likely than Whites to be Catholic is a factor.
4. What may be a more important question is whether the mean difference in marital stability between Negroes and Whites, given the social-ecological conditions, is socially relevant. We can see in Table 4 that The R^2 is negligible when race is the only variable in the regression, and increases greatly with the introduction of the other variables. The removal of race would, therefore, have a negligible effect on the explained variance. This effect has been estimated and can be found in Table 4.

Conclusion. When all is said and done, what do we have? The social-ecological position of the races does account for half of the mean difference between the races in marital stability. Once all factors are "controlled", race still remains a statistically significant factor with respect to marital stability. By itself, race at no point helps account for much of the variance in marital stability. What is left of the relationship between race and marital stability, therefore, although statistically significant seems hardly socially relevant. To a large extent, then, race stands as a summation of social-ecological position. This is, of course, without doubt itself a function of the patterns of discrimination and prejudice with respect to the Negro in our society.

Footnotes

[1]E. Franklin Frazier, The Negro Family in the United States (Chicago: University of Chicago Press, 1939).

[2]See Lee Rainwater, "'Marital Stability and Patterns of Status': A Comment," Journal of Marriage and the Family, November, 1966, p.442.

[3]"The widespread and continued family disorganization among Negroes in cities . . . is one of the results of the impact of the urban environment upon the simple and loose family organization of the Negro folk." E. Franklin Frazier, "The Negro Family in Chicago," in Ernest W. Burgess and Donald J. Bogue, Contributions to Urban Sociology (Chicago: University of Chicago Press, 1964), p. 404.

[4]See Jessie Bernard, Marriage and Family Among Negroes (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1966).

[5]Jessie Bernard, "Marital Stability and Patterns of Status Variables," Journal of Marriage and the Family, November, 1966. The analysis was done on men in the 45-to-54-year age bracket.

[6]J. Richard Udry, "Marital Instability by Race, Sex, Education, and Occupation Using 1960 Census Data," The American Journal of Sociology, Vol. 72, No. 2, September 1966. Data for the age group 25-34 were calculated.

[7]J. Richard Udry, "Marital Instability by Race and Income Based on 1960 Census Data," The American Journal of Sociology, Vol. 72, No. 6, May, 1967.

[8]Myron J. Lefcowitz, "Poverty and Negro-White Family Structures," White House Conference "To Fulfil These Rights," November, 1965.

[9]Mollie Orshansky, Social Security Bulletin, January and July, 1965.

[10]Rainwater, op. cit., p. 444.

Table 1

Percentage of Husband-Wife Families,
by Poverty Status, Race, and Age of Head

Age of Head:	Race	Poverty Status: Under .7501	.7501-1.25	Over 1.25
Under 25	White	80	93	96
	Non-White	62	85	87
	W-NW Differences	18	8	9
25 - 64	White	74	86	93
	Non-White	63	79	86
	W-NW Differences	11	7	7
65 and over	White	75	85	79
	Non-White	63	76	78
	W-NW Differences	12	9	1

Table 2

Percentage of Persons in Husband-Wife
Families, by Poverty Status, Race, and Age of Head

Age of Head:	Race	Poverty Status: Under .7501	.7501-1.25	Over 1.25
Under 25	White	82	94	97
	Non-White	65	85	89
	W-NW Differences	17	9	8
25 - 64	White	80	91	95
	Non-White	67	83	87
	N-NW Differences	13	8	8
65 and over	White	75	84	78
	Non-White	64	73	76
	N-NW Differences	11	11	2

Table 3

Social Characteristics of Ever-Married Women
and Women Married More Than Once, by Race

		Women Married	
		Ever	More Than Once
Median Age:			
	White	44.4	48.7
	Non-White	41.6	49.3
Median Years Since First Marriage:			
	White	21.2	27.6
	Non-White	18.8	28.2
Median Education:			
	White	11.3	10.0
	Non-White	8.7	7.8
Percentage in South:			
	White	27.5	28.9
	Non-White	54.6	57.3
Percentage in Urban Residence:			
	White	71.2	74.7
	Non-White	76.8	78.1

Source: U. S. Census of Population: 1960, Marital Status. PC(2)-4E, Tables 1,2, 4.

Table 4. Measure of simple and partial effects of race on marital stability, family stability and female dominance.

Dependent Variable:	For Race Coefficient	Partial Correlation	R^2 With Race	Without Race
1. M_1				
a. Simple regression	-.325 (.012)*	-.119	.014	
b. Multiple regression without poverty index	-.249 (.012)	-.088	.133	.126
c. Full regression	-.162 (.012)	-.057	.153	.15
2. M_2				
a. Simple regression	-.311 (.012)	-.114	.013	
b. Multiple regression without poverty index	-.221 (.012)	-.076	.085	.08
c. Full regression	-.155 (.013)	-.053	.096	.093
3. M_3				
a. Simple regression	-.257 (.009)	-.126	.016	
b. Multiple regression without poverty index	-.187 (.009)	-.089	.149	.142
c. Full regression	-.118 (.009)	-.056	.171	.168
4. M_4				
a. Simple regression	-.296 (.010)	-.130	.017	
b. Multiple regression without poverty index	-.239 (.010)	-.106	.197	.188
c. Full regression	-.141 (.010)	-.063	.234	.231
5. F_1				
a. Simple regression	-.364 (.011)	-.156	.024	
b. Multiple regression without poverty index	-.284 (.012)	-.115	.085	.073
c. Full regression	-.192 (.012)	-.077	.11	.105
6. F_2				
a. Simple regression	-.343 (.011)	-.15	.023	
b. Multiple regression without poverty index	-.272 (.014)	-.112	.083	.071
c. Full regression	-.178 (.012)	-.073	.109	.104

Table 4 cont.

Dependent Variable:	For Race: Coefficient	For Race: Partial Correlation	R^2: With Race	R^2: Without Race
6. Dom_1				
a. Simple regression	-.308 (.010)	-.139	.019	
b. Multiple regression without poverty index	-.255 (.011)	-.108	.076	.066
c. Full regression	-.152 (.011)	-.064	.110	.106
7. Dom_2				
a. Simple regression	-.347 (.011)	-.152	.023	
b. Multiple regression without poverty index	-.283 (.011)	-.118	.101	.088
c. Full regression	-.148 (.011)	-.063	.161	.158

*The numbers in the parenthesis are the standard errors of the estimate of the coefficient.

APPENDIX I

INDEPENDENT VARIABLES

1. Race: Negroe = 1, White = 0

2. Region: South = 1, Other = 0

3. Birthplace: South = 1, Other = 0

4. Present residence: Rural, Non-SMSA Urban, SMSA Fringe, SMSA Center City (Four dummy variables)

5. SMSA residence in 1955 and 1960: Same SMSA in 1955 and 1960, Different SMSA in 1955 and 1960, Non-SMSA in 1955, Other (Four dummy variables)

6. Residential mobility: Same house in 1955 and 1960, Not same house in 1955 and 1960, Other (Three dummy variables)

7. Age:
   a. Age of individual
   b. Number of years over 30 (if any)
   c. Number of years over 50 (if any)

8. Years since first marriage:
   a. Total number of years
   b. Number of years over 10 (if any)
   c. Number of years over 20 (if any)

9. Education:
   a. Total years of education
   b. Twelve years of education = 1
   c. Sixteen years of education = 1
   d. Years of education beyond high school (if any)
   e. Years of education beyond 4 years of college (if any)

10. Poverty Index:
    a. Ratio of total family income to poverty line for family's size and farm-nonfarm residence
    b. Excess of poverty index over .4999 (if any)
    c. Excess of poverty index over .9999 (if any)
    d. Excess of poverty index over 1.4999 (if any)

Dependent Variables

1. Marital Stability

a. M_1 = 1 when married, spouse present, and married only once
= 0 all others married only once
= -1 married more than once or never-married family head

b. M_2 = 1 when married, spouse present, or widowed and married only once
= 0 all others married only once
= -1 all others

c. M_3 = 1 when married, spouse present, or widowed, and married only once
= 0 all others married only once; or other married, spouse present; or other widowed.
= -1 all others

d. M_4 = 1 when married, spouse present, and married only once
= 0 all others married only once; or other married, spouse present
= -1 all others

2. Family Stability

a. F_1 = 1 when in husband-wife family, husband and wife married only once
= 0 in all other husband-wife families
= -1 female family head

b. F_2 = 1 in husband-wife family, wife married only once
= 0 in all other husband-wife families
= -1 female family head

3. Female Dominance

a. Dom_1 = 1 when in husband-wife family, wife not chief income recipient
= 0 all other husband-wife families
= -1 female family head

b. Dom_2 = 1 when in husband-wife family, husband's income is greater than wife's
= 0 all other husband-wife families
= -1 female family head

WHITE-NEGRO OCCUPATIONAL DIFFERENTIALS*

Karl E. Taeuber, University of Wisconsin

Study of the past may be the only rational basis for assessment of the future, but history never repeats itself. Impaled on the horns of this dilemma, I have been contemplating the relevance of study of past census data to assessment of future trends in occupational differentials. I should like to speak briefly to the uses of the past, then marshall evidence for the irrelevance of the past, and finally outline the beginnings of a rather complicated and unsatisfying resolution.

As a demographer who has wandered into the domains of race relations and occupational structures, I carry with me the demographer's predilection for differentiating between prediction and projection. Demographers' predictions in the 1930's of imminent population decline have been confronted with the harsh reality of 200 million Americans. From this bitter experience it has become a matter of professional image-building to maintain that we project rather than predict.

The art of projection flourishes. Past trends in components of population change (fertility, mortality, immigration, emigration) can be measured and plausibly narrow limits can be placed on likely short-term future trends in all except fertility. Even though birth rates fluctuate in pesky fashion, sophisticated analysis of detailed data can remove many irregularities from the series.

Extension of the projection model to sub-populations, such as occupation groups, is feasible conceptually, but processes of entry and exit into a sub-population are much more complicated and past trends are very inadequately measured. Hence reliance tends to be placed on rather straight-forward extrapolation techniques.

The extrapolation approach at its simplest takes trends in white and Negro occupational distributions among major census categories (Table 1) and carries them forward. The implicit model tends to be an occupational structure initially resembling a caste situation (with Negroes at the bottom) but now being transformed in more or less regular fashion by Negroes pushing up into successively higher occupational strata. Each of us can speculatively add to Table 1 a column representing 1970 or 1975. I need not belabor the limitations of this means of projection. It can be improved upon by assembling a more detailed occupational series and calculating appropriate indices of change in racial composition, but a convincing rationale for any specific extrapolation system is difficult to provide.[1]

Another approach is to extrapolate flows or change processes rather than successive cross-sectional distributions. A recent instructive example is provided by Lieberson and Fuguitt's use of an intergenerational occupational mobility tabulation as the basis for a transition matrix in a Markov process.[2] A supplement to the March, 1962, Current Population Survey, "Occupational Changes in a Generation," provided a tabulation of occupation of father by occupation of son.[3] This yielded an intergenerational transition matrix for major occupations. The initial white and Negro occupational distributions (for 1960) were quite dissimilar. Application of the matrix yielded second generation distributions much less divergent one from another (index of dissimilarity declined from 40 for the first generation to 13 for the second.). By the third generation the occupational distributions were very similar (index of dissimilarity of 4). Lieberson and Fuguitt also considered a process in which a father-son educational level transition matrix was applied to whites and Negroes alike. If Negroes were to translate educational attainment into occupational level in the same manner as whites, then two generations of "racially neutral" father-son educational mobility would yield highly similar white-Negro occupational distributions.

In summary, the Lieberson-Fuguitt analysis shows that application of white intergenerational mobility patterns (whether of education or of occupation) to Negroes would greatly reduce white-Negro occupational differences in one generation and would virtually eliminate them in two generations. I have four brief comments on these results:

1) The speed of convergence is more rapid than might have been expected given discussion of a vicious circle of disadvantages and a culture of poverty. The occupational stratification system in the United States is sufficiently open that low origins, in and of themselves, are not an impenetrable barrier to occupational advance. On the other hand, two generations represent about 40 years according to the assumptions of the model, and this assumes instant complete elimination of race as a factor in occupational mobility. Each year is a long time from the perspective of those seeking rapid change.

2) The model has no acceptable rationale as a basis for prediction. The authors freely acknowledge that "occupational structure and changes in structure could not be deduced from intergenerational occupational mobility."

3) The results speak mainly to the character of the basic transition matrix. There are no occupational statuses which are excessively difficult to get into or out of. Hence the character of the initial occupational distribution of Negroes makes little difference. The cube of the matrix is nearly at equilibrium. Hence only two generations suffice to redistribute Negroes and whites nearly to the equilibrium distribution.

4) The fact that white and Negro occupational distributions have been and continue to be divergent probably means that the intergenerational occupational mobility matrix for Negroes differs from that for whites. The instant substitution of a new matrix for Negroes being grossly unrealistic, attention should be directed to the difference.

The separate matrices were unavailable to Lieberson and Fuguitt, but were subsequently assembled and examined by Duncan.[4] The summary Negro intergenerational occupational mobility matrix is radically different from that for whites. Paradoxically, as pointed out by Duncan, its dominant feature is a form of equalitarianism. Among Negroes, for each father's occupation a majority of sons obtains unskilled or semiskilled employment. Among whites, most sons retain or improve upon father's occupational level. Differences between Negro and white intergenerational mobility processes in part represent the educational and other disadvantages experienced by Negroes, but in large part are attributable to racial discrimination in the labor market. Transformation of the Negro matrix into the white matrix, therefore, would require a variety of substantial social changes.

There are a number of technical difficulties with this application of a transition matrix, such as the indefinite time period to which it applies, the disregard of differential fertility, the glossing over of patterns of occupational mobility within a career, the difficulty of bringing to bear other relevant variables such as educational level of father and son, and the assumption that a single transition matrix adequately characterizes a variety of social mobility processes. Many of these difficulties can be overcome to some degree by consideration of more complex sets of mobility data. Yet this would still not solve the extrapolation problem--how rationally to designate the appropriate Markov or other stochastic process for projection into the future.

To be sure it should be possible to concoct a rationale for some particular approach. But my concern is more deep-seated, arising from that side of the original dilemma that says that history never repeats itself. In very large measure the history of occupational transformations cannot repeat itself. The Negro mobility matrix for the first half of this century was dominated by the transformation from farm origins to nonfarm activities. This transformation occurred through mass migration and urbanization and a restructuring of the total labor force. In 1966, three-fourths of all Negroes lived in cities and only 6 per cent of employed Negro males were farmers or farm laborers. This particular transition has largely run its course; it cannot continue. The majority of employed Negro males holds unskilled or semi-skilled jobs in the laborer, operative, and service categories. If the Negro occupational distribution is to converge toward that of whites, the next great transition must be into the skilled and white collar levels. To be sure mobility matrices from the past entail movement from lower manual jobs to higher levels, but the flows were small in magnitude. A radical change in magnitude would necessarily involve considerable change in the structure of flows. At the minimum, the manner in which the total occupational structure can change imposes constraints on the possible Negro mobility patterns and imposes a dependence of white and Negro patterns on one another.

There is an additional difference of future from past, and that is the likely degree of deliberate intervention into the racial aspects of occupational mobility, intervention by private pressure groups as well as by various levels of government. The aggregate rate of economic growth may or may not dominate future employment trends as many would argue it has in the past, but the number, variety, and effectiveness of specific programs to foster Negro occupational mobility seem certain to increase.

The final task set for this discussion is to move toward a resolution of the dilemma. Obviously we must project, and obviously we must rely on the past. The question is not whether the past is a satisfactory basis for projection--it isn't--but how best to use the past. I would like to draw an analogy to the situation confronting demographers in the 1930's. Trends in natural increase seemed subject to linear or logistic extrapolation. The net reproduction rate, a recently derived descriptive measure, seemed to provide a basis for prediction--it was even called an "intrinsic" rate. These simple techniques proved inadequate, however, and it required two decades of developments in data (cohort fertility series) and techniques to reach the current state of the art. Whatever the shortcomings of national population projections there is at least a clear perception of the relevant population dynamics, specific components can be measured currently and assessed against assumed trends, and the assumptions may easily be modified and new projections made. To attempt to predict future white-Negro occupational differences by simple techniques, is, I would assert, to repeat the mistakes of the 1930's. The equilibrium vector produced by a Markov process is really no more than a descriptive measure of the transition matrix, and hence is no more of a predictive device than the intrinsic rates of a stable population model. Until we have data and techniques for describing the underlying mobility processes, we cannot expect to do very well even at short-term projection of occupational distributions of whites and Negroes.

What are the needed data? I could speculate about the full range of flow data one might like from social security records if detailed occupation and race were available, or from the new Equal Employment Opportunities Commission series if age of worker were available. Clearly we need analysis of the occupational histories of successive cohorts of whites and Negroes and of the forces facilitating and inhibiting Negro occupational mobility. But rather than launch into a general discussion, I should merely like to illustrate the prior importance of taking a much more complex and detailed view of the labor force.

The structure of the labor force involves a matching of skills with tasks, job seekers with prospective employers, occupation with industry. To call attention to the relevance of the industrial dimension to racial occupational trends, the time series in Table 2 suffices to document a wide variation among industries in utilization of Negroes in the unskilled and semiskilled labor forces. Contemplation of these data and of preliminary tallies from a special tabulation of per cent Negro for very detailed occupations and industries from the 1960 Census leads me to suggest an alternative model of occupational assimilation. Earlier I mentioned the common implicit notion of steady movement of Negroes upward in the hierarchy of major occupations. If the labor force structure is viewed as a lattice of detailed occupations and industries, the channels of Negro advance will probably prove to be much more varied. In particular, there may be certain industries--government being one obvious example--which provide opportunities at a number of occupational levels. Patterns of labor mobility between industries may then serve to disperse Negroes at selected occupational levels without a corresponding Negro presence at other occupational levels lower in the prestige hierarchy. For example, Negro white collar workers may flow from government into various highly regulated industries even though craft unions remain closed to Negroes. Viewing the labor force as a structure of thousands of discrete types of employment may, in the long run, facilitate our ability to identify channels of Negro advance and to anticipate and guide programs of planned intervention.

The dilemma is not resolved. Negro occupational trends cannot be foreseen in detail. But the core of my argument is that the rationality of projections must be increased by expansions of the relevant data base rather than by application of fancier formulae to existing data. As with population projections, the aim should be not perfect prediction but movement toward an ever-improving accounting system which lets us see where we are in some detail and shows us the processes which got us there and are moving us on.

Notes:

[1] A detailed occupational series for 1940 to 1960 together with indices of change have been assembled by Daniel O. Price for his forthcoming Census Monograph on Negro population. Construction of appropriate indices of change is also addressed in R. David Mustian and C. Horace Hamilton, "Measuring the Extent, Character, and Direction of Occupational Changes," Social Forces, 45 (March, 1967), 440-444.

[2] Stanley Lieberson and Glenn V. Fuguitt, "Negro-White Occupational Differences in the Absence of Discrimination," American Journal of Sociology, 73 (September, 1967), 188-200.

[3] Current Population Reports, Series P-23, No. 11.

[4] Otis Dudley Duncan, "Patterns of Occupational Mobility among Negro Men," paper presented at the 1967 Annual Meetings of the Population Association of America.

* This paper draws on a continuing program of research on Negro occupational assimilation, co-investigator Alma F. Taeuber, supported by funds granted to the Institute for Research on Poverty at the University of Wisconsin by the Office of Economic Opportunity, pursuant to the provisions of the Economic Opportunity Act of 1964. The speculations are the sole responsibility of the author.

TABLE 1.--Occupational Distribution of White and Negro Male Labor Force, 1910-1966

	1910		1920		1930		1940		1950		1960		1966	
	White	Negro	White	Negro	White	Negro	White	Negro	White	Negro	White	Negro	White	Negro
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Professionals	3.3	1.1	3.4	1.1	4.2	1.5	6.0	1.9	7.0	2.2	11.3	4.0	13.2	4.4
Managers	8.6	1.1	8.6	1.0	9.8	1.3	10.5	1.7	11.8	2.3	14.6	2.7	14.5	2.6
Clerical and Sales	10.1	1.0	11.4	1.4	14.0	1.7	14.4	2.3	14.6	4.3	13.8	7.0	14.0	7.9
Craftsmen	15.5	3.6	17.5	4.4	17.5	4.8	15.8	4.5	19.4	8.1	20.0	9.0	20.0	10.3
Operatives, Service, Laborers	30.3	37.6	31.1	43.0	31.2	50.0	31.9	48.4	32.7	59.1	31.0	62.7	31.8	68.3
Farmers	18.9	25.6	17.3	24.9	14.4	21.8	14.3	20.8	10.2	13.3	6.3	5.7	4.5	2.1
Farm Laborers	13.3	30.0	10.7	24.2	8.9	18.9	7.1	20.4	4.3	10.7	3.0	8.9	2.0	4.4

Source: Dale L. Hiestand, Economic Growth and Employment Opportunities for Minorities (New York: Columbia University Press, 1964) Table II. 1966 data are for March, from Employment and Earnings and Monthly Report on the Labor Force, 12 (April, 1966), Table A-23, and Current Population Reports, Series P-20, No. 168, Table 16. Comparability from one date to another is imperfect.

Source Note for Table 2: Compiled from the respective decennial censuses and from unpublished material obtained from the Bureau of the Census on occupational and industrial comparability. This is a preliminary tabulation; fuller documentation will be prepared at a later date.

TABLE 2, -- Per Cent Negro among Employed Male Laborers and Operatives, by Industry, 1910-1960.

	Laborers, n.e.c. 7/				Operatives, n.e.c. 7/			
Industry	1910	1930	1950	1960	1910	1930	1950	1960
Sawmills	33.4	37.2	37.8	31.3	10.4	9.0	23.0	17.8
Furniture	5.9	8.2	20.8	22.2	2.0	2.4	6.5	9.2
Glass	7.0	8.8	5.6	7.5	1.4	2.3	3.6	4.1
Cement 1/	-	-	-	-	5.9	8.4	16.0	16.1
Structural clay	20.4	20.6	21.7	25.6	8.3	8.5	15.1	17.5
Pottery	3.9	6.5	4.3	5.9	0.7	1.2	1.7	3.2
Misc. nonmetallic	10.7	12.4	19.2	18.9	4.0	2.7	5.5	9.8
Motor vehicles	1.2	13.6	25.0	24.4	0.3	2.2	10.3	12.3
Shipbuilding	24.4	26.3	37.4	39.6	3.3	9.3	17.6	17.0
Electrical machinery 2/	1.4	3.8	8.5	12.3	-	-	-	-
Meat products	9.1	22.5	29.2	24.9	6.1	14.6	20.2	19.7
Canning	4.3	11.6	14.1	16.4	2.6	13.0	13.0	19.3
Dairy 1/	-	-	-	-	0.2	0.7	3.6	3.9
Grain mills	11.9	15.3	20.7	20.5	6.1	7.2	12.6	15.1
Bakery	10.0	13.4	22.0	18.0	4.1	5.3	9.4	11.3
Confectionery	6.7	9.2	14.7	17.7	2.8	3.8	10.1	12.0
Beverage	7.4	18.6	14.3	22.5	2.4	6.8	6.5	9.3
Misc. and not spec. food 3/	-	-	-	-	12.0	6.6	16.1	21.7
Tobacco	50.4	62.9	68.3	56.4	10.1	16.1	27.1	29.5
Knitting 4/	-	-	-	-	1.7	0.5	0.9	5.9
Dyeing 4/	-	-	-	-	1.2	1.1	2.7	5.0
Carpets 4/	-	-	-	-	2.4	1.0	5.6	6.6
Yarn	9.6	13.8	23.1	24.4	1.0	1.1	4.0	5.1
Misc. textiles 4/	-	-	-	-	0.7	2.2	7.6	10.9
Apparel 5/	7.4	16.0	18.5	23.2	2.3	5.1	4.8	6.7
Misc. fab. textiles 5/	-	-	-	-	1.1	1.0	7.2	11.2
Pulp mills	2.6	9.5	23.3	25.5	0.6	1.7	6.2	5.8
Misc. paper	1.6	5.3	22.7	23.1	1.0	1.2	5.4	5.8
Paperboard	2.4	5.1	15.4	16.2	0.6	1.3	8.5	9.8
Printing	15.7	10.4	21.4	19.0	1.1	2.0	6.6	7.2
Synthetic fibers	-	14.9	14.0	23.4	-	0.8	0.9	2.9
Prints	4.4	12.3	14.8	16.2	2.1	4.2	8.8	13.8
Drugs and misc. chemicals 3/	28.0	38.9	34.4	30.4	8.6	9.6	14.8	10.3
Petroleum refining	8.1	19.3	22.8	22.8	1.0	2.1	3.4	3.6
Misc. petroleum 2/	-	-	-	-	20.6	20.9	22.0	22.2
Leather 6/	5.7	7.9	12.7	10.7	1.9	2.9	7.1	9.7
Footwear 6/	-	-	-	-	1.9	0.5	0.9	1.2
Leather products 6/	-	-	-	-	1.0	0.9	5.4	10.4
Not spec. mfg.	9.4	15.1	24.8	39.3	2.6	4.4	11.0	11.4
Construction	18.7	21.4	25.0	25.9	3.4	4.0	10.3	12.3
Railroads	15.2	22.2	27.3	27.7	5.2	8.2	11.7	9.9
Transportation and telecom.	16.0	17.3	22.4	25.5	3.3	3.7	10.3	13.0
Business services	4.8	19.7	22.0	17.1	2.7	6.3	11.5	14.2
Public administration	18.8	22.1	29.7	33.6	2.8	5.2	18.5	21.3
Trade 1/	23.9	25.3	20.1	16.9	5.0	6.6	13.6	16.4
Misc. light mfg. 4/	16.9	21.4	13.1	14.9	2.7	3.2	5.3	6.9
Misc. heavy mfg. 2/	6.2	15.4	23.3	22.9	1.6	3.3	6.1	6.9
Personal services and all other	20.8	20.9	34.2	30.5	6.7	8.0	13.5	16.4
TOTAL	17.2	21.0	25.3	24.3	3.2	4.7	8.7	9.7

NOTES:

1. For laborers, cement and dairy are included with trade.
2. For laborers, misc. petroleum is included with misc. heavy mfg. For operatives, electrical machinery is included with misc. heavy mfg.
3. For laborers, misc. and not spec. food are included with drugs and misc. chemicals.
4. For laborers, knitting, dyeing, carpets, and misc. textiles are included with misc. light mfg.
5. For laborers, misc. fab. textiles is included with apparel.
6. For laborers, footwear and leather products are included with leather
7. Some specific laborer and operative jobs are separately identified in census tabulations, e.g., longshoremen; the rest are aggregated together as "n.e.c." -- not elsewhere classified.

DISCUSSION

By Tobia Bressler, Bureau of the Census

The difficulties facing students of demography when trying to measure or project Negro-white differentials are not a result of their fear of treading on thin ice nor their lack of imagination. It may be that they are faced with the same type of problem the microbiologist encounters when he attempts to examine a strain of bacteria which has developed a new structure and pattern of growth because of an increase of new stimulants and irritants in its environment.

Mr. Thurow's paper presents a well documented explanation of a procedure for measuring the returns in income from improvements in education and job related training. He points out early in his paper the need to know what combination of on-the-job training and formal education yields the greatest return to the individual.

In the process of measuring this the author computes a human capital function which has as one of its factors years of work experience. I should like to raise a question about the procedure used for obtaining this item. (Years of experience--e.) The human capital function was estimated separately for the white and nonwhite population and this is as it should have been but the number of years of experience was computed in the same fashion for both the white and nonwhite population...by subtracting the work-starting age from the current age. This technique gives equal weight to all workers, those who may have worked full-time, worked only on a part-time basis, and workers with long periods of unemployment. Should an adjustment be included in the computation of "e" to take account of the white-nonwhite differential in employment patterns?

Mr. Thurow correctly points out that his computation of human capital does not include a factor for innate ability and, therefore, will overstate the returns to education. The degree to which this omission affects the results will also depend on whether innate ability is assumed to vary by race.

The inclusion of age along with race, occupation, and region as variables in the investigation of the returns to education and training would provide an interesting measure of change over time.

XI

CONTRIBUTED PAPERS I

Chairman, JOHN D. DURAND, University of Pennsylvania

MULTIPLE CLASSIFICATION ANALYSIS WITHOUT ASSUMPTION OF INTERVAL MEASUREMENT, LINEARITY, OR ADDITIVITY: A COMPARISON OF TECHNIQUES

James M. Carman, University of California (Berkeley)

The interest of this paper is data analysis, not inference. In survey research, one is commonly faced with the problem of analysis, often exploratory analysis, of data from a relatively large number of subjects on which values of a number of variables have been collected. In social science research, the measures we have of these variables often do not meet the standards that statisticians would like. We are faced with mixed interval, ordinal and nominal data, nonlinearity, nonorthogonality, and interactions. Thus, the restrictions of the common variations of the general linear regression model are not often met. Aided by the computer, the rate of development of operational, often heuristic, schemes for analysis of these kinds of data has increased in recent years. This paper will review some of these newer techniques and empirically compare their relative efficiencies and shortcomings.

In attempting to see the relationship between variables, the analyst is inevitably faced with the problem of having more data than can be comprehended by the human mind at one time. The particular problem discussed here is one where the task is to relate a large set of predictors to some specific dependent variable in such a way as to isolate intervening conditions and discard spurious and irrelevant variables. In this process it is necessary to reduce the quantity of data to a level of rapid comprehension.

Classification Techniques

It should be noted that in some problems of this general type the practice is to perform a data reduction operation prior to analysis of the effects of the predictors on a particular dependent variable. There are two related approaches to data reduction, each of which has developed a variety of models based on whether the data are normally distributed or simple classifications.

The first of these approaches is the taxonometric approach--that is, reduce the number of subjects by placing them into subcategories so that the nature of everyone in the subcategory is more like each of the other n-1 persons than he is like any other person in any other subcategory. For continuous variables, the models of Tryon [11] and Cattell [3] are well known. For nominal and ordinal data, McQuitty has made substantial contributions [8].

The second approach is the factor analysis approach--that is, reducing the number of predictors by collapsing them into construct factors and then constructing measures for the construct based on weighted factor scores.

While a few of the variables used in the example which follows are factor scores, in the main, this approach to data reduction has been avoided for two reasons. First, there is serious question as to whether the measures of our data (or for that matter, most psychological data) meet the requirements of the factor analysis model. Second, the factor analysis model is most appropriate when one has a number of measures of one or several closely-related variables or constructs. In the survey research problem, as opposed to the psychological test problem, one has a number of variables which are intercorrelated, but which relate to characteristics and attitudes stemming from very different question formats and which refer to different time periods in the subject's life. Thus, a priori, it is difficult to judge which variables should be proxies for a single construct. Consequently, both statistical and behavioral theory would suggest that factor analysis solely for the purpose of data reduction should be avoided.

We turn, then, to a search for analysis techniques which will give insight into the effects of a number of predictors on a dependent variable and, at the same time, provide some amount of data reduction. One recent and fresh approach to this problem has been made by James Coleman [4]. Unfortunately, Coleman's technique requires that the dependent variable be dichotomous. In addition, he has not, to the best of my knowledge, solved the interaction problem. Since Coleman's problem can be solved more efficiently by a dummy regression or, more precisely, a two-group discriminant model, it was not included in the empirical comparison to follow.

The one approach which does seem to have more merit than a long series of exploratory dummy regressions is a branching technique. (It is important to note, however, that neither branching or dummy regression will handle the problem of nominal predictors with a nominal criterion variable of more than two levels.) A binary branching schema does provide a kind of data reduction which the dummy regression model does not. In dummy regression, all levels of all nominal predictors must be established as potential predictors. In a branching technique, the algorithm searches for that split in the classification which maximizes the distance between the mean value of the dependent variable in the two subcategories. While other branching techniques are to be found in the literature, the one which has been developed most completely is the Sonquist-Morgan Automatic Interaction Detector Algorithm [10]. This technique is a center of interest in the empirical comparison.

An Example

As a vehicle for comparing the effectiveness and efficiency of some models for social science data analysis, we have chosen a problem from the study of consumer behavior. The data came from the Berkeley Food Panel, a study in which the food shopping habits of panel members were studied over a period of fifteen weeks [1].

The particular problem of interest here was whether characteristics of the respondents would predict the stability through time of their buying patterns with respect to the food chains they patronized. After some collapsing of small, independent stores into groups, it was possible for a respondent to have shopped in twenty-three different chains or independent stores. Each respondent was classified as having stable or unstable buying patterns during the period based on whether her pattern rejected a null hypothesis of temporal symmetry in a test involving the Kruskal-Wallis H-statistic [2]. Thus, the dependent variable was dichotomous, taking a value of 1 for unstable patterns and 0 for stable patterns.

The predictors were social, environmental, economic, demographic, psychological, attitudinal, and behavioral characteristics of the respondents collected during the course of the panel study. These were typical social science data in that a few were true interval measures, some were rank measures, and many were simple classifications. There was considerable correlation between predictors and, for the continuous measures, linearity was not a good assumption. In all, we had about 95 predictors: 25 continuous, 36 ordinal, 7 dichotomous, and 27 nominal, with an average of 6 levels each. There were 235 observations.

How might one approach analysis of these data? Cross-classification analysis is probably the most obvious approach, but consider what is required. First, the interval and ordinal scale would have to be treated as classifications and, as a start, 95 two-way tables produced. Even if one could cope with this many tables, the analysis would be void of any investigation of joint effects. If interactions and intercorrelations were considered, the problem gets completely out of hand. Even with the computer, cross classification requires a great deal of setup for very little data reduction.

Another approach might be to analyze the data as a dummy regression problem. The chief advantages of this approach are the very large amount of data reduction it achieves plus the availability of a variety of convenient computer programs. Unfortunately, there are a number of serious disadvantages. Initially, additivity would be assumed and, for the continuous variables, one would probably assume linearity also. The ordinal variables would have to be converted to dummies. Most serious is the fact that, in this example, there are insufficient degrees of freedom to analyze the data initially as a dummy regression.

Automatic Interaction Detector

A more fruitful approach proved to be analysis of the data with the binary branching techniques of the Morgan-Sonquist Automatic Interaction Detector (A. I. D.) schema. Using the amount of data reduction as a criterion, A. I. D. falls in between the cross-classification approach and the dummy regression approach. While the number of pages of output, amount of brute-force study, and number of reruns necessary to get meaningful results are significantly less than in cross classification, do not expect to get instant answers. The A. I. D. trees presented in Exhibits 1, 2, 3, and 4 each required at least two computer runs, each produced approximately 100 pages of computer printout, and each required this investigator about one-half hour to digest. With the present version of the A. I. D. algorithm, data analysis is still an investigator activity rather than a computer activity.

Using the amount of flexibility and generality as criteria, A. I. D. comes off significantly better than either cross classification or dummy regression. The continuous variables must be treated as categories, but any predictor can be restricted to have a monotonic relationship with the criterion variable. Thus, A. I. D. can discover nonlinear relationships with continuous or ordinal predictors without reporting spurious and meaningless minima and maxima. A. I. D. is well suited to analyze classification data and, of course, it is ideally suited to handle interactions between predictors [9].

The A. I. D. algorithm solves the degrees of freedom problem by calculating the deviation of every observation from its branch mean and making the deviation available as the dependent variable in a subsequent analysis. On any one run it is important not to introduce so many predictors that the degrees of freedom become used up before some important predictors have had a chance to enter the analysis. We have found it useful to follow the practice of cross-classification analysis and enter predictors in time order of occurrence for the respondent. For example, in the first tree the predictors relate to the respondent's childhood experience and her environment; the second tree predictors relate to general personality characteristics which are, in part, a function of background; the third tree predictors are attitudes related specifically to homemaking; the fourth tree predictors are characteristics of shopping behavior which are, themselves, a function of the predictors in the earlier trees.

The results of the A. I. D. analysis are presented in Exhibits 1 through 4. It should be emphasized in passing that the best method for summarizing and presenting A. I. D. results is not obvious or well established.

Dummy Regression

It is useful, for comparison purposes, to see how the A. I. D. results would compare with results from a dummy regression. Regression analysis is possible now because the results of the A. I. D. analysis can be used to collapse some categories and to eliminate variables which the tree analysis showed to be poor predictors. We introduced 44 predictor variables and dummies into a standard stepwise linear regression program. Thirty-two of those entered with alpha risks of less than .30. The results are summarized in Exhibit 5.

Following are eleven hypotheses which might be advanced, based on the regression results:

Proneness toward unstable food store shopping patterns:

1. Increases with income.
2. Decreases with asset accumulation.
3. Increases with cultural status, i.e. education and occupation status.
4. Is greatest among the unmarried under 45 years of age.
5. Is least if shopper's Mother lives nearby.
6. Is inversely related to the degree of training as a child on the value of money and to dissatisfaction with present economic situation.
7. Is greatest among those with high religious commitment.
8. Is greatest among those who are interested homemakers and mothers, but not devoted cooks or shoppers.
9. Those with unstable patterns are liberal in their economic thinking, don't make a special effort to please others, and "have a complete, realistic, practical respect for the facts."
10. Increases with weekly food expenditures.
11. Is least among those with the greatest amount of store choice.

The eleven generalizations leave out some rather disturbing inconsistencies within the regression findings. Only a part of these inconsistencies can be traced to multicolinearity, which was clearly evident. One is also struck by the low fraction of variance explained.

What is even more disturbing is that when we analyze the trees, we find that three of these eleven generalizations do not appear to be correct interpretations of the data.

1. The income and asset factors do not show up in the trees at all and one wonders if the regression results are not related in some way to the social class effect which shows up significantly in both the regression and the trees.
2. The trees show that it is not being unmarried which is related to unstable buying practice. Rather, the relationship is with family structure. The least stable are families with four or more children living at home; second are young families with older children at home; the most stable are older families with no children at home. Again, the regression model is confusing because of a failure to cope with an interaction between life cycle and family structure.
3. While the Yeasay and Personality types agree between the regression and tree analyses, the relationship of the Economic Conservative scale is not as clear. The tree analysis shows this scale to be interacting with the Yeasay scale in a fashion which suggests, on balance, an effect the reverse of that shown by the regression results.

Another way to compare the results is to compare the statistics in Exhibit 5. The R^2 and β^2 statistics for the discriminant (regression) analysis have the usual interpretation. The proportion of variance explained by the trees is simply the between group sum of squares over the total sum of squares. There is no adjustment for loss of degrees of freedom; yet clearly, this statistic is a function of the number of observations and the number of groups.

The reduction in unexplained variance from any one split can be calculated from the program by:

$$D = \frac{TSS_i}{TSS_T} - \left(\frac{TSS_j}{TSS_T} + \frac{TSS_k}{TSS_T}\right)$$

where i is the parent group and j and k the resultant groups. There are other statistics which can be calculated from the A. I. D. output which have intuitive appeal because of their parallel to analysis of variance. However, the critical distinction between them is that the A. I. D. model involves sequential solution with the statistics generated at each branch, while the ANOVA model involves a simultaneous solution. In general, one would expect that in Exhibit 5 the A. I. D. reductions in unexplained variance would overstate β^2. This is not true in many cases, leading to the conclusion that the results given by the two models are different.

To summarize, the regression analysis, even after some initial doctoring of the data based on the tree analysis, explained only 18 percent of the total variance, passed over seven predictors which the tree analysis indicated were important, and yielded results which in many instances mislead the analyst in understanding the information contained in the data.

The A. I. D. analysis, on the other hand, leads to a much better understanding of the data, but can give misleading results when the number of observations in a branch gets small. It is important not to introduce too many predictors in one run. For example, one final A. I. D. run introduced 30 predictors which were shown in earlier runs to be important. Only 13 of the most powerful of these entered the analysis before the degrees of freedom had been exhausted.

Holmes' Substrata Analysis

Another branching scheme which appeared to offer some usefulness to the analysis problem at this point was Holmes' Substrata Analysis [6]. This technique was developed by the late Jack A. Holmes in a project which was trying to identify the factors and mechanism which leads some children to read at an earlier age than others. The technique did help Holmes to gain new insight into the reading process. In this scheme a set of first-level predictors are regressed on the criterion variable. Then a set of second-level predictors are regressed on each significant predictor in the first-level analysis. If desired, a set of third-level predictors may be regressed on each second-level predictor. In this way a tree of regressions is constructed. Each regression is the standard, stepwise, linear, additive algorithm. The user may allow all predictors to be eligible to enter the analysis at any level or may specify the level at which they are to be considered. The user may also specify "fundamental" variables which are not permitted to be criteria in subsequent levels.

In some ways the Substrata Algorithm appears to be similar to A. I. D. Predictors may have a direct influence on the criterion or may only work through a first-level predictor. In many key respects, however, the two techniques are quite different. For one thing, at each level the Substrata Algorithm makes all of the usual linearity, additivity, independence assumptions of the general linear regression model. Therefore, even though it is a branching model, it is not a very general model. On the contrary, it is quite specific and requires the analyst to start with a theory which will justify the substrata model. In our problem, the model looked reasonable, i.e. stability is a function of shopping habits which, in turn, are functions of personal characteristics, personality, and early training. In practice, however, the results from this model did not match up with theory. Predictors entered at wrong levels and individual regressions did not make as much sense as the single equation regression model. The total amount of output was just as great as with A. I. D., but supplied much less information.

Data Reduction and Real Time Analysis

This problem of a large volume of output is a serious one. If the data will not permit the luxury of reduction to a single, simple correlation matrix, then any analysis scheme will not yield the amount of data reduction common in regression analysis. Further, since our problem is one of heuristic data analysis and not inference, the analyst learns more about how to proceed as he goes along. These two characteristics--large volumes of data and a heuristic process--make real time computer analysis the next logical step in the development of branching processes. F. H. Westerfelt developed at the University of Michigan a stepwise, polynomial, regression procedure which maximizes predictability with a minimum number of terms. David Evans developed, at Berkeley, a way to display this and alternative models on an oscilloscope, while the analyst interacts with the computer in real time. The day is not far off when the A. I. D. trees presented here can be generated in real time with visual display output in such a way that a large variety of alternative orders of entry and reentry into the analysis can be accomplished in the time required to study the output from one run in a batch processing system. Thus, it should soon be possible to teach the logic of data analysis developed by Hyman [7] over twelve years ago without having the student and instructor feel the frustration of having no analytical technique for making this logic operational.

REFERENCES

[1] Bucklin, L. P., and Carman, J. M. The Design of Consumer Research Panels: Conception and Administration of the Berkeley Food Panel. Berkeley: Institute of Business and Economic Research, University of California, 1967.

[2] Carman, J. M., and Stromberg, J. L. "A Comparison of Some Measures of Brand Loyalty," Working Paper No. 26. Berkeley: Research Program in Marketing, Institute of Business and Economic Research, University of California, 1967.

[3] Cattell, R. B. Handbook of Multivariate Experimental Psychology. Chicago: Rand McNally and Co., 1966.

[4] Coleman, James S. Introduction to Mathematical Sociology. New York: Free Press, 1964, Ch. 6.

[5] Harmon, H. H. Modern Factor Analysis. Chicago: University of Chicago Press, 1960.

[6] Holmes, J. A. Substrata-Factor Differences Underlying Reading Ability in Known Groups. Report to the U. S. Office of Education under Contract No. 538-8176 (1961).

[7] Hyman, H. Survey Design and Analysis. Glencoe: Free Press, 1955.

[8] McQuitty, L. I. "Typal Analysis," Educational Psychological Measurement, 21 (1961), 677-696.

[9] Morgan, J. N., and Sonquist, J. A. "Problems in the Analysis of Survey Data, and a Proposal," Journal of the American Statistical Association, 58 (1963), 415-435.

[10] Sonquist, J. A., and Morgan, J. N. The Detection of Interaction Effects. Ann Arbor: Survey Research Center, University of Michigan, 1964.

[11] Tryon, R. C., and Bailey, D. E. "Cluster and Factor Analysis." Berkeley: Computer Center, University of California, 1965.

EXHIBIT 1

FIRST A.I.D. TREE
DISCRIMINANT FUNCTION FOR SHIFT IN SHOPPING PATTERN

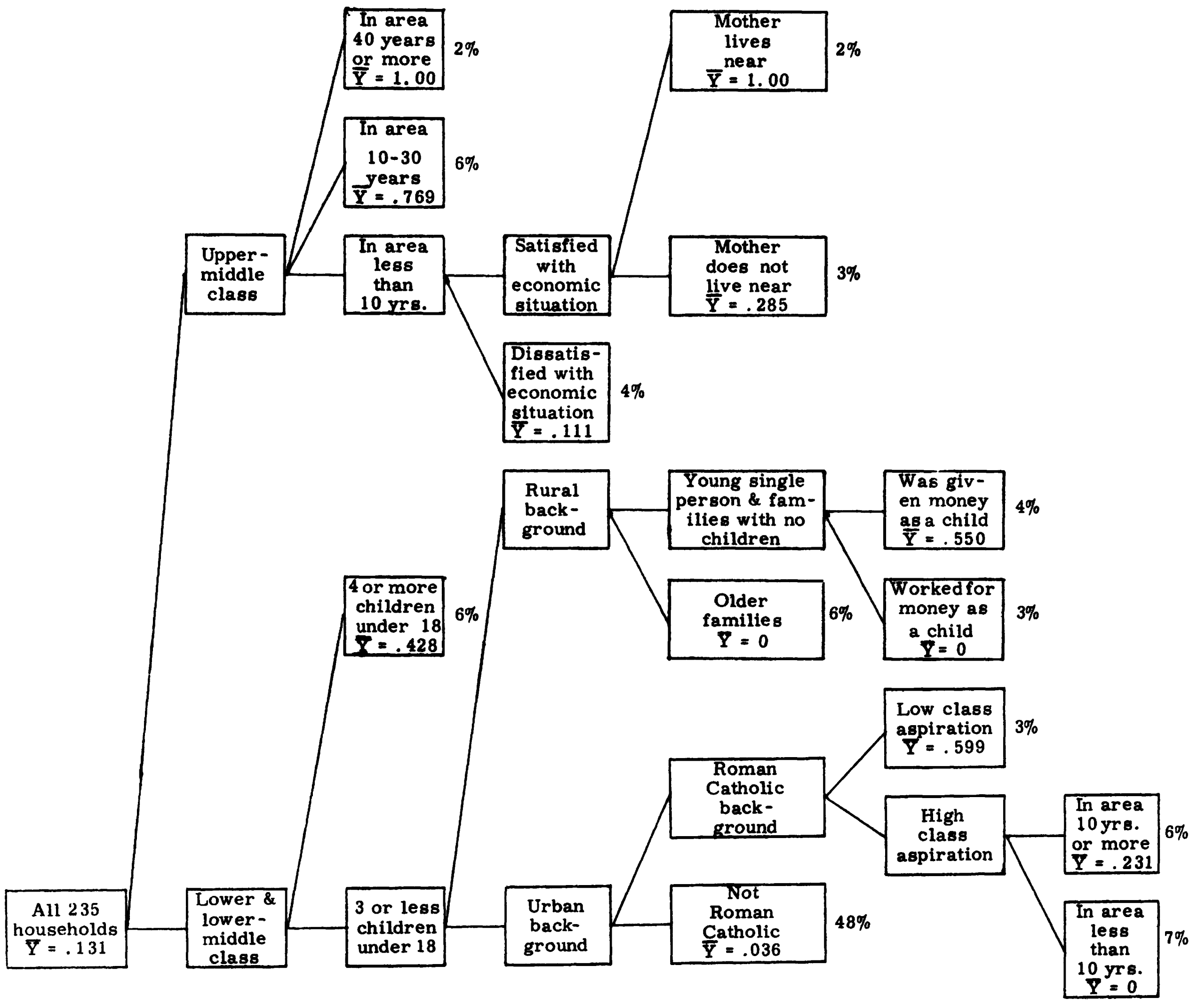


EXHIBIT 2

SECOND A.I.D. TREE
DISCRIMINANT FUNCTION FOR SHIFT IN SHOPPING PATTERN

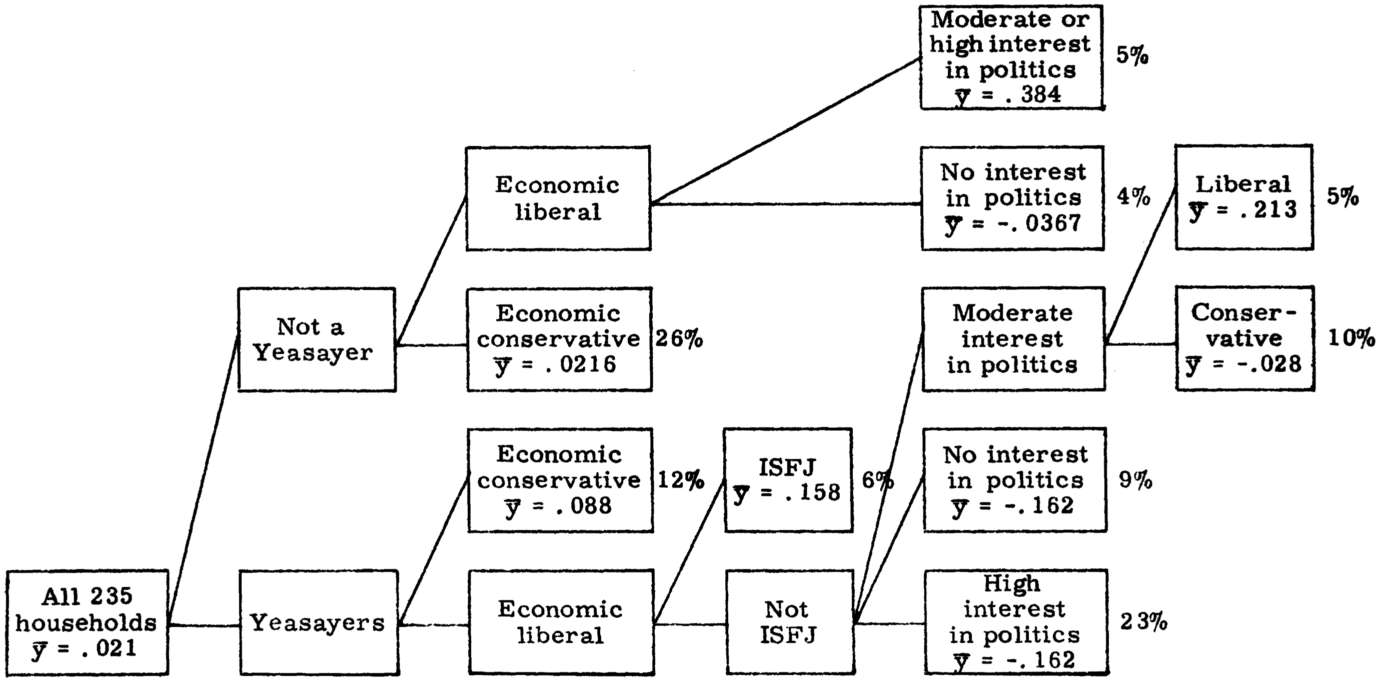


EXHIBIT 3

THIRD A.I.D. TREE
DISCRIMINANT FUNCTION FOR SHIFT IN SHOPPING PATTERN

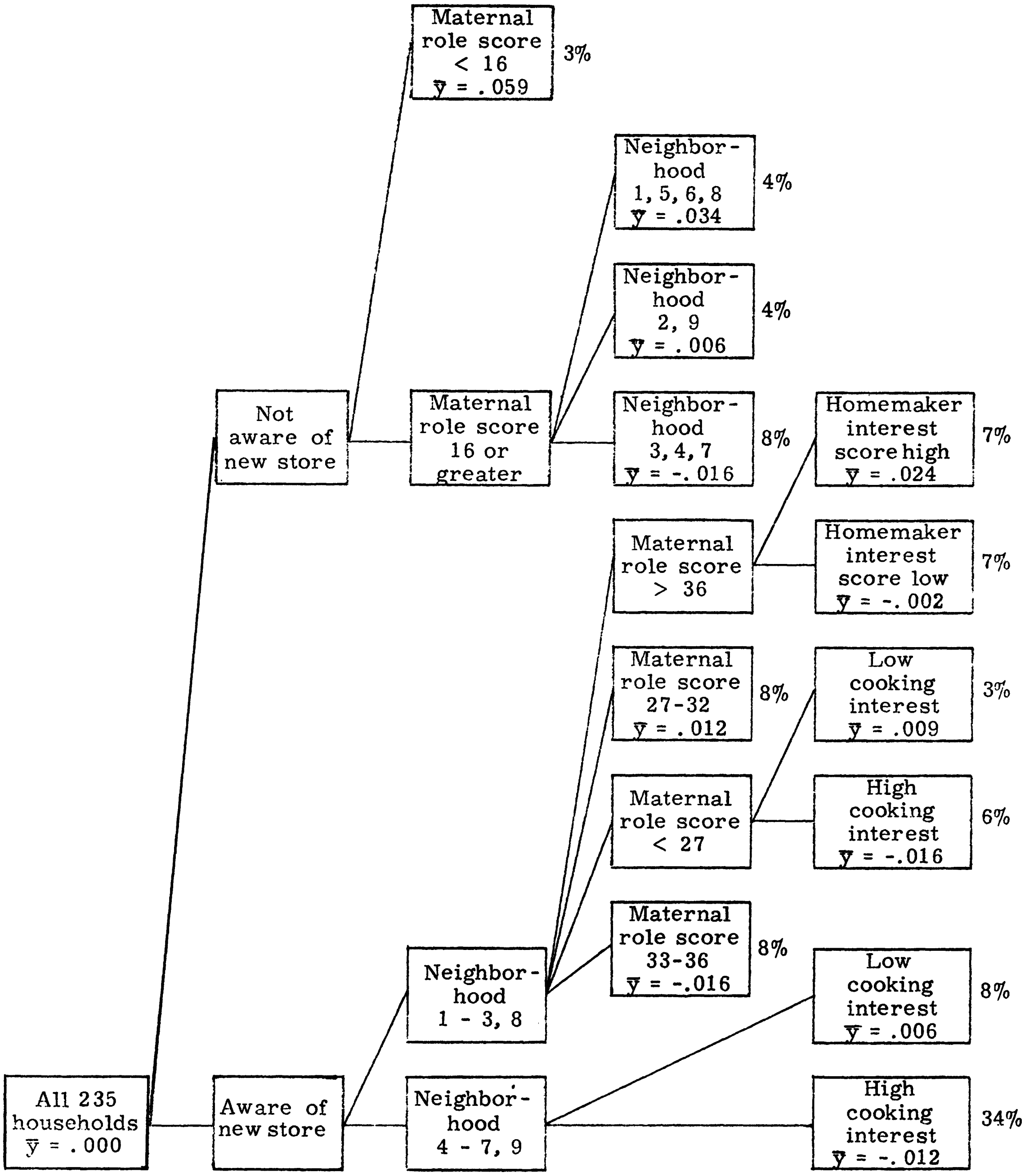


EXHIBIT 4

FOURTH A.I.D. TREE
DISCRIMINANT FUNCTION FOR SHIFT IN SHOPPING PATTERN

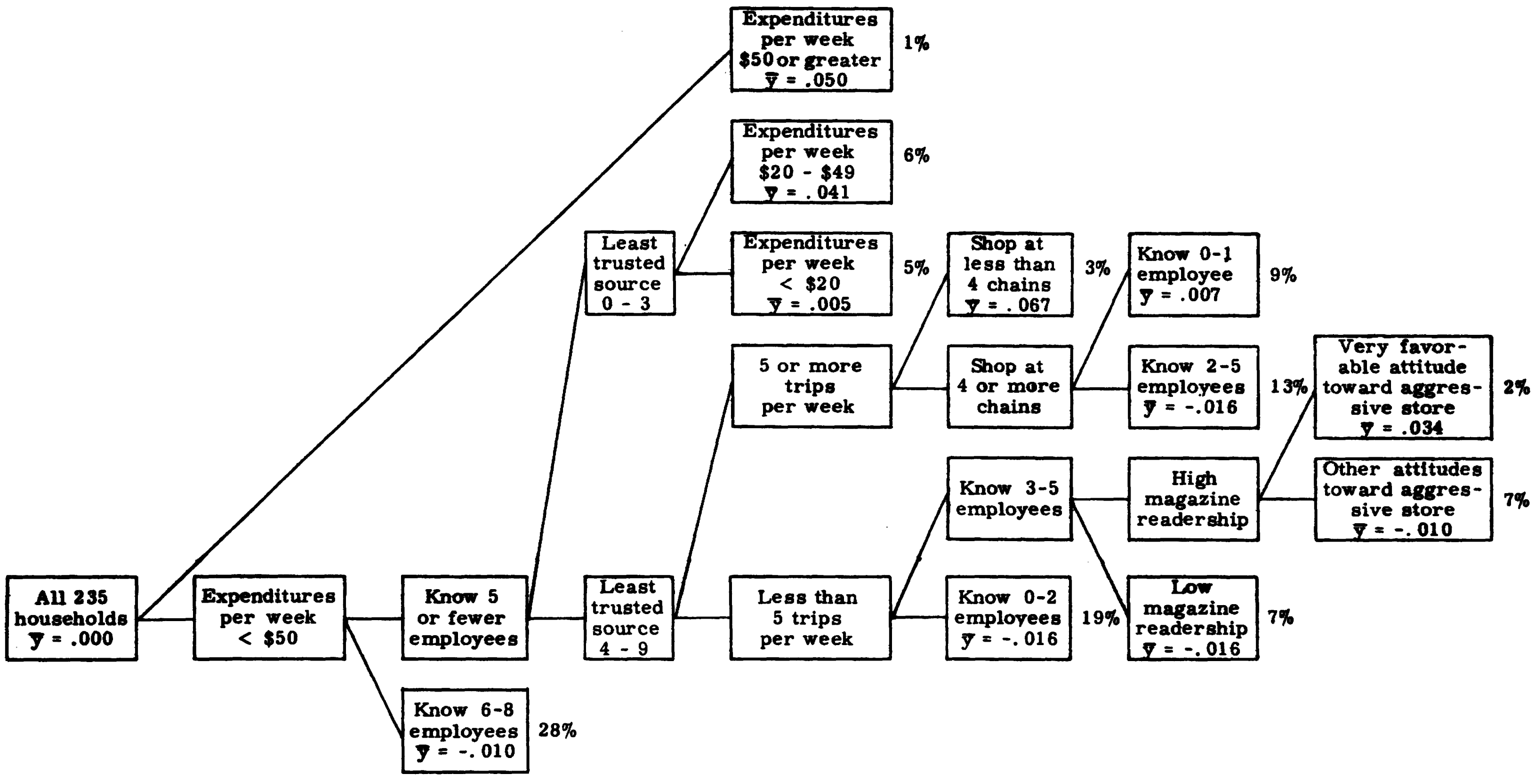


EXHIBIT 5

COMPARISON OF LINEAR DISCRIMINANT AND A. I. D. RESULTS

Proportion of Variance Explained By:	
Discriminant function, adjusted	.18
First tree, no adjustment	.41
Second tree	.17
Third tree	.30
Fourth tree	.30
$1 - (1-R_1^2)(1-R_2^2)(1-R_3^2)(1-R_4^2)$	.72

Predictor	Discriminant Function		A. I. D. Split Reduction in Unexplained Variance	Comments
	Significance	β^2		
Effects of Background, Social, Demographic, and Economic Environment:				
Social class	.01	.0428	.0365	
Income	.05	.0454	-	
Property value	.05	(-).0428	-	Sig. corr. with income
Many investments	.25	(-).0094	-	Sig. corr. with income
Life cycle	.05	.0299	.0270	Splits are not the same
Number of children under 18	ns	-	.0419	
Mother lives near	.15	(-).112	.0468	
Early independence training	.20	(-).083	.0311	
Rural background	ns	-	.0281	
Tenure in area	ns	-	{.0120 .0468	
Roman Catholic index	.05	.0279	.0162	
Religious involvement score	.10	.0310	-	
Dissatisfied with economic situation	.15	.0161	.0360	
Class aspiration	ns	-	.0330	

EXHIBIT 5--Continued

Predictor	Discriminant Function Significance	Discriminant Function β^2	A.I.D. Split Reduction in Unexplained Variance	Comments
Effects of General Personality Characteristics:				
Economic conservative	.10	(-).0172	.0103 .0165	
Yeasaying score	.28	.0061	.0160	
Personality type ISFJ	.20	.0071	.0174	
Politically active	ns	-	.0064 .0973	
General conservative	ns	-	.0095	
Effects of Attitudes Related to Homemaking:				
Maternal role dominant	.05	.0454	.0712 .0578	Significant correlation
Homemaker role dominant	.10	.0182	.0442	between these
Cooking interest score	.15	(-).0174	.0157 .0094	three predictors
Aware of new supermarket	.10	(-).0142	.0283	
Magazine readership	.01	.0502	.0209	
Frequency of entertaining at home	.05	(-).0437	-	
Frequency of entertaining neighbors	.20	.0166	-	Significant correlation with home entertainment
Effects of Shopping Behavior:				
Number of market employees known	.10	(-).0146	.0248 .0173 .0150	
Don't trust home economists	.10	.0372	.0361	
Don't trust friends for food information	.10	.0552		
Don't trust store clerks	.20	(-).0071		
Husband influential in setting food budget	.20	(-).0079	-	Significant correlation with life cycle
Favorable attitude toward aggressive store	ns	-	.0192	

EXHIBIT 5--Continued

Predictor	Discriminant Function		A.I.D. Split Reduction in Unexplained Variance	Comments
	Significance	β^2		
Live in Neighborhood 8	.05	.0279	.0508	
Live in Neighborhood 7	.10	(-).0135	.0230	
Weekly food expenditures	.01	.1239	.0244 .0333	
Number of stores visited per week	.01	(-).0835	-	
Number of different stores visited in 15 weeks	.05	.0339	.0633	Significant correlation with stores per week
Number of shopping trips per week	-	-	.0491	
Mean interval between shopping trips	.20	(-).0204	-	
Mean expenditures per trip	.05	(-).0061	-	Significant correlation with expenditures and interval between trips

CLASSIFICATION ERRORS AND MEASURES OF ASSOCIATION IN CONTINGENCY TABLES*

F. G. Giesbrecht, Research Triangle Institute

It has become a well recognized fact that statisticians must be cognizant of both sampling and non-sampling errors in the analysis and interpretation of data obtained from sample surveys. For the purpose of this study, the term response error is rather loosely defined to include all effects which result in an incorrect classification in the final tabulations. These can be due to such diverse sources as deliberate falsification by the respondent or incorrect recording by the interviewer.

The purpose of the model described in this paper is to provide a means for investigating the effect of response errors on selected measures of association in contingency tables and to aid in the design of special surveys for the purpose of estimating these errors. The model for 2 x 2 contingency tables contains a total of 13 parameters, including three basic probabilities and ten response error parameters. The response error parameters are defined as conditional probabilities. The two characteristics will be referred to as A and B with the respective complements being $\bar{A}$, (not A) and $\bar{B}$ (not B). Consequently an individual is identified as belonging to both A and B, i.e. AB, A and not B i.e. $A\bar{B}$, not A but B i.e. $\bar{A}B$ or finally neither A nor B and denoted by $\overline{AB}$. These four classes are disjoint and exhaustive.

Let P_B be the probability that a randomly selected individual belongs to class B. Let $P_{A|B}$ be the conditional probability that a randomly selected individual from class B also belongs to class A. Similarly $P_{A|\bar{B}}$ is the conditional probability that a randomly selected individual who is not in class B, is in class A. It follows that the probability that a randomly selected individual will belong to A and not to B is equal to $P_{A|\bar{B}}(1-P_B)$. Probabilities for the other three possibilities have equivalent definitions. The special case in which $P_{A|B} = P_{A|\bar{B}}$ is the one in which there is no association between the two factors.

Response Error Parameters

The three basic parameters defined in the previous section would be sufficient if there were only sampling errors. However, the actual classification (abbreviated as ac) will at times differ from the true classification (abbreviated as tc). Now define the response error parameters:

β_1 = Pr (ac is B | tc is B).

β_0 = Pr (ac is $\bar{B}$ | tc is $\bar{B}$).

These two probabilities do not depend on the A classification. A slightly more flexible model can be obtained by introducing four probabilities for errors in the B classification and allowing a dependence on the actual A classification.

α_{11} = Pr (ac is AB | tc is AB and ac is B).

α_{01} = Pr (ac is $\bar{A}B$ | tc is $\bar{A}B$ and ac is B).

α_{10} = Pr (ac is AB | tc is $A\bar{B}$ and ac is B).

α_{00} = Pr (ac is $\bar{A}B$ | tc is $\overline{AB}$ and ac is B).

γ_{11} = Pr (ac is $A\bar{B}$ | tc is AB and ac is $\bar{B}$).

γ_{01} = Pr (ac is $\overline{AB}$ | tc is $\bar{A}B$ and ac is $\bar{B}$).

γ_{10} = Pr (ac is $A\bar{B}$ | tc is $A\bar{B}$ and ac is $\bar{B}$).

γ_{00} = Pr (ac is $\overline{AB}$ | tc is $\overline{AB}$ and ac is $\bar{B}$).

Since a randomly selected individual can belong to any one of four classes and be assigned to any one of four classes, the parameters define the likelihood of the 16 distinct possibilities. The probability that a randomly selected individual will be assigned to class AB is

$$\beta_1\alpha_{11}P_{A|B}P_B + (1-\beta_0)\alpha_{10}P_{A|\bar{B}}(1-P_B)$$
$$+ \beta_1(1-\alpha_{01})(1-P_{A|B})P_B$$
$$+ (1-\beta_0)(1-\alpha_{00})(1-P_{A|\bar{B}})(1-P_B) .$$

Similarly the probability a randomly selected individual will be assigned to class $A\bar{B}$ is

$$(1-\beta_1)\gamma_{11}P_{A|B}P_B + \beta_0\gamma_{10}P_{A|\bar{B}}(1-P_B)$$
$$+ (1-\beta_1)(1-\gamma_{01})(1-P_{A|B})P_B$$
$$+ \beta_0(1-\gamma_{00})(1-P_{A|\bar{B}})(1-P_B) .$$

The probability of being assigned to class $\bar{A}B$ is

$$\beta_1(1-\alpha_{11})P_{A|B}P_B + (1-\beta_0)(1-\alpha_{10})P_{A|\bar{B}}(1-P_B)$$
$$+ \beta_1\alpha_{01}(1-P_{A|B})P_B$$
$$(1-\beta_0)\alpha_{00}(1-P_{A|\bar{B}})(1-P_B) .$$

Finally, the probability of being assigned to class $\overline{AB}$ is

* Research supported by Bureau of the Census, Contract No. Cco-9191.

$$(1-\beta_1)(1-\gamma_{11})P_{A|B}P_B + \beta_0(1-\gamma_{10})P_{A|\bar{B}}(1-P_B)$$

$$+ (1-\beta_1)\gamma_{01}(1-P_{A|B})P_B$$

$$+ \beta_0\gamma_{00}(1-P_{A|B})(1-P_B) .$$

Effect of Response Errors on the Chi-square Statistic

A study of the non-centrality parameter of the χ^2 test statistic when there is no association, that is $P_{A|B} = P_{A|\bar{B}}$ and β_0 and β_1 are the only response error parameters not equal to one, verifies the statement by Bross that errors of this type do not disturb the validity of the χ^2 test. However, these errors do decrease the non-centrality parameter when $P_{A|B} \neq P_{A|\bar{B}}$. This is the phenomenon of loss of power of the χ^2 test in the presence of response errors. However, response errors do not always have these effects. For example, if $\alpha_{11} = .9$, $P_B = .5$, $P_{A|B} = .9$, $P_{A|\bar{B}} = .9$ and all other parameters equal one, then the non-centrality parameter is not zero even though there is no true association and hence an invalid test. For this case the non-centrality parameter turns out to be .016. However if in the above case, $P_{A|B} = .7$, then the non-centrality parameter is .101. This can be compared to .062 when there are no response errors. The implication is an increase in power. Other cases can be examined in a similar manner. Similar techniques can be used to study the effects of combinations of response errors on other measures of association.

Estimation of the Response Error Parameters

Since it is reasonably simple to examine the effects of the response error parameters, the interesting problem is to estimate these parameters. Assume that a second interviewer is assigned the task of reinterviewing a random sample of those individuals already surveyed. Based on the responses to characteristics A and B in both interviews, each individual will be assigned to one of 16 classes. These classes are combinations of the four possible assignments as a result of the first interview and the four following the second interview. If the two interviews are assumed to be independent and that the same response error parameters apply to both interviews, then the expected values of the fraction in each of the 16 categories are as given in Table I. The task now is to determine which functions of the parameters are estimable. An examination of Table I shows immediately that the expected number assigned to the class A and $\bar{B}$ by the first interviewer and to class AB by the second interviewer is equal to the expected number assigned to class AB by the first interviewer and to class A and $\bar{B}$ by the second interviewer. There are five other pairs with matching expected values. Also the sum of all 16 frequencies is equal to unity, implying a maximum of 9 degrees of freedom for purposes of estimation. However, there are 13 parameters in the model. It is obvious that not all parameters are estimable. The problem now is to find which parameters or functions of the parameters are estimable. The method for locating these is an application of the definition of an estimable function, that is, a function is estimable if there exists a function which estimates it. These functions are located by equating the observed relative frequencies to the expected values and then solving the resulting equations for meaningful functions of the parameters. The estimators obtained in this manner may not be optimal in any sense. It is simply a verification that the function can be estimated. Once it has been verified that a set of functions of the parameters is estimable, one can use any of the standard methods, such as maximum likelihood or minimum chi-square to obtain estimates with desirable properties. If one of these methods is chosen, one will need to use one of the iterative, numerical techniques to arrive at the final solutions.

An example of the method for verifying that certain parameters are estimable is as follows:

The expected value of the sum of the four classes for which both interviewers have recorded that the individual belongs to class B is

$$\beta_1^2 P_B + (1-\beta_0)^2 (1-P_B) .$$

Similarly the expected value of the sum of those recorded as $\bar{B}$ by both interviewers is

$$(1-\beta_1)^2 P_B + \beta_0^2 (1-P_B) .$$

This leads to two equations in three unknowns. If one of the three is known and P_B does not have an extreme value then one can solve for the remaining two. Hence two of the three are estimable.

Further applications of this technique lead to other estimable functions. Unfortunately the set of estimable functions obtained in this manner is not unique, but rather is a function of the assumptions one is willing to make. For example, the above derivation illustrates that if one knows β_0 then β_1 and P_B are estimable. Alternatively if one assumes $\beta_0 = \beta_1 = \beta$ then β and P_B are estimable. The appropriate choice for any given situation depends on the supporting information available from other sources.

It can be shown that if it is assumed that if β_1, α_{01}, α_{00}, γ_{01} and γ_{00} are known and β_1 is not equal to an extreme value, then P_B, $P_{A|B}$, $P_{A|\bar{B}}$, β_0, α_{11}, α_{10}, γ_{11} and γ_{10} are estimable. Alternatively, if it is assumed that $\beta_0 = 1$ and that $\alpha_{11} = \alpha_{01}$, $\alpha_{10} = \alpha_{00}$, $\gamma_{11} = \gamma_{10}$, and $\gamma_{01} = \gamma_{00}$, then β_1, P_B, $P_{A|B}$, $P_{A|\bar{B}}$, α_{11}, α_{10}, γ_{11} and γ_{01} are estimable.

ACKNOWLEDGEMENTS

The author acknowledges the benefit derived from several informal discussions at the Bureau of the Census. Participants included Max A. Bershad, William N. Hurwitz, Tom Jabine, Leon Pritzker and Dr. Benjamin Tepping. The author also acknowledges the assistance and encouragement by Dr. D. G. Horvitz of the Research Triangle Institute.

REFERENCES

[1] Assakul, Kwanchai and C. H. Proctor, "Testing Hypotheses with Categorical Data Subject to Misclassification," N. C. Inst. of Statistics, Mimeograph Series No. 448, Raleigh, N. C., (1965).

[2] Bross, Irwin, "Misclassification in 2 x 2 Tables." Biometrics 10, (1954), 478-486.

[3] Mote, V. L., "An Investigation of the Effects of Misclassification on the χ^2 tests in the Analysis of Categorical Data," N. C. Inst. of Statistics, Mimeograph Series No. 182, Raleigh, N. C., (1957).

Table I Expected Values of the Frequencies of Various Types of Classification in Two Independent Interviews

	Classification			
	First Interview	Second Interview	Fraction	Expected Value
1	AB	AB	$f(A_1B_1A_2B_2)$	$\beta_1^2\alpha_{11}^2\ P_{A\|B}P_B + (1-\beta_0)^2\alpha_{10}^2\ P_{A\|\bar{B}}(1-P_B) + \beta_1^2(1-\alpha_{01})^2(1-P_{A\|B})P_B$ $+ (1-\beta_0)^2(1-\alpha_{00})^2(1-P_{A\|\bar{B}})(1-P_B)$
2	AB	$A\bar{B}$	$f(A_1B_1A_2\bar{B}_2)$	$\beta_1(1-\beta_1)\alpha_{11}\gamma_{11}P_{A\|B}P_B + \beta_0(1-\beta_0)\alpha_{10}\gamma_{10}P_{A\|\bar{B}}(1-P_B)$ $+ \beta_1(1-\beta_1)(1-\alpha_{01})(1-\gamma_{01})(1-P_{A\|B})P_B + \beta_0(1-\beta_0)(1-\alpha_{00})(1-\gamma_{00})(1-P_{A\|\bar{B}})(1-P_B)$
3	$A\bar{B}$	AB	$f(A_1\bar{B}_1A_2B_2)$	$\beta_1(1-\beta_1)\alpha_{11}\gamma_{11}P_{A\|B}P_B + \beta_0(1-\beta_0)\alpha_{10}\gamma_{10}P_{A\|\bar{B}}(1-P_B)$ $+ \beta_1(1-\beta_1)(1-\alpha_{01})(1-\gamma_{01})(1-P_{A\|B})P_B + \beta_0(1-\beta_0)(1-\alpha_{00})(1-\gamma_{00})(1-P_{A\|\bar{B}})(1-P_B)$
4	$A\bar{B}$	$A\bar{B}$	$f(A_1\bar{B}_1A_2\bar{B}_2)$	$(1-\beta_1)^2\gamma_{11}^2P_{A\|B}P_B + \beta_0^2\gamma_{10}^2P_{A\|\bar{B}}(1-P_B) + (1-\beta_1)^2(1-\gamma_{01})^2(1-P_{A\|B})P_B$ $+ \beta_0^2(1-\gamma_{00})^2(1-P_{A\|\bar{B}})(1-P_B)$
5	AB	$\bar{A}B$	$f(A_1B_1\bar{A}_2B_2)$	$\beta_1^2\alpha_{11}(1-\alpha_{11})\ P_{A\|B}P_B + (1-\beta_0)^2\alpha_{10}(1-\alpha_{10})\ P_{A\|\bar{B}}(1-P_B)$ $+ \beta_1^2\alpha_{01}(1-\alpha_{01})(1-P_{A\|B})P_B + (1-\beta_0)^2\alpha_{00}(1-\alpha_{00})(1-P_{A\|\bar{B}})(1-P_B)$
6	AB	$\bar{A}\bar{B}$	$f(A_1B_1\bar{A}_2\bar{B}_2)$	$\beta_1(1-\beta_1)\alpha_{11}(1-\gamma_{11})\ P_{A\|B}P_B + \beta_0(1-\beta_0)\alpha_{10}(1-\gamma_{10})\ P_{A\|\bar{B}}(1-P_B)$ $+ \beta_1(1-\beta_1)(1-\alpha_{01})\gamma_{01}(1-P_{A\|B})P_B + \beta_0(1-\beta_0)(1-\alpha_{00})\gamma_{00}(1-P_{A\|\bar{B}})(1-P_B)$

Table I Continued

7	$A\bar{B}$	$\bar{A}B$	$f(A_1\bar{B}_1\bar{A}_2B_2)$	$\beta_1(1-\beta_1)(1-\alpha_{11})\gamma_{11}\, P_{A\mid B}P_B + \beta_0(1-\beta_0)(1-\alpha_{10})\gamma_{10}P_{A\mid\bar{B}}(1-P_B)$
				$+\ \beta_1(1-\beta_1)\alpha_{01}(1-\gamma_{01})(1-P_{A\mid B})P_B + \beta_0(1-\beta_0)\alpha_{00}(1-\gamma_{00})(1-P_{A\mid\bar{B}})(1-P_B)$
8	$A\bar{B}$	$\bar{A}\bar{B}$	$f(A_1\bar{B}_1\bar{A}_2\bar{B}_2)$	$(1-\beta_1)^2\gamma_{11}(1-\gamma_{11})P_{A\mid B}P_B + \beta_0^2\gamma_{10}(1-\gamma_{10})P_{A\mid\bar{B}}(1-P_B)$
				$+\ (1-\beta_1)^2\gamma_{01}(1-\gamma_{01})(1-P_{A\mid B})P_B + \beta_0^2\ \gamma_{00}(1-\gamma_{00})(1-P_{A\mid\bar{B}})(1-P_B)$
9	$\bar{A}B$	AB	$f(\bar{A}_1B_1A_2B_2)$	$\beta_1^2\alpha_{11}(1-\alpha_{11})P_{A\mid B}P_B + (1-\beta_0)^2\alpha_{10}(1-\alpha_{10})P_{A\mid\bar{B}}(1-P_B)$
				$+\ \beta_1^2\alpha_{01}(1-\alpha_{01})(1-P_{A\mid B})P_B + (1-\beta_0)^2\alpha_{00}(1-\alpha_{00})(1-P_{A\mid\bar{B}})(1-P_B)$
10	$\bar{A}\bar{B}$	AB	$f(\bar{A}_1\bar{B}_1A_2B_2)$	$\beta_1(1-\beta_1)\alpha_{11}(1-\gamma_{11})P_{A\mid B}P_B + \beta_0(1-\beta_0)\alpha_{10}(1-\gamma_{10})P_{A\mid\bar{B}}(1-P_B)$
				$+\ \beta_1(1-\beta_1)(1-\alpha_{01})\gamma_{01}(1-P_{A\mid B})P_B + \beta_0(1-\beta_0)(1-\alpha_{00})\gamma_{00}(1-P_{A\mid\bar{B}})(1-P_B)$
11	$\bar{A}B$	$A\bar{B}$	$f(\bar{A}_1B_1A_2\bar{B}_2)$	$\beta_1(1-\beta_1)(1-\alpha_{11})\gamma_{11}P_{A\mid B}P_B + \beta_0(1-\beta_0)(1-\alpha_{10})\gamma_{10}P_{A\mid\bar{B}}(1-P_B)$
				$+\ \beta_1(1-\beta_1)\alpha_{01}(1-\gamma_{01})(1-P_{A\mid B})P_B + \beta_0(1-\beta_0)\alpha_{00}(1-\gamma_{00})(1-P_{A\mid\bar{B}})(1-P_B)$
12	$\bar{A}\bar{B}$	$A\bar{B}$	$f(\bar{A}_1\bar{B}_1A_2\bar{B}_2)$	$(1-\beta_1)^2\gamma_{11}(1-\gamma_{11})\ P_{A\mid B}P_B + \beta_0^2\gamma_{10}(1-\gamma_{10})\ P_{A\mid\bar{B}}(1-P_B)$
				$+\ (1-\beta_1)^2\gamma_{01}(1-\gamma_{01})(1-P_{A\mid B})P_B + \beta_0^2\gamma_{00}(1-\gamma_{00})(1-P_{A\mid\bar{B}})(1-P_B)$
13	$\bar{A}B$	$\bar{A}B$	$f(\bar{A}_1B_1\bar{A}_2B_2)$	$\beta_1^2(1-\alpha_{11})^2\ P_{A\mid B}P_B + (1-\beta_0)^2(1-\alpha_{10})^2P_{A\mid\bar{B}}(1-P_B)$
				$+\ \beta_1^2\alpha_{01}^2(1-P_{A\mid B})P_B + (1-\beta_0)^2\alpha_{00}^2(1-P_{A\mid\bar{B}})(1-P_B)$

Table I Continued

14	$\bar{A}\bar{B}$	$\bar{A}B$	$f(\bar{A}_1\bar{B}_1\bar{A}_2B_2)$	$\beta_1(1-\beta_1)(1-\alpha_{11})(1-\gamma_{11})\ P_{A\|B}P_B + \beta_0(1-\beta_0)(1-\alpha_{10})P_{A\|\bar{B}}(1-P_B)$
				$+\ \beta_1(1-\beta_1)\alpha_{01}\gamma_{01}(1-P_{A\|B})P_B + \beta_0(1-\beta_0)\alpha_{00}\gamma_{00}(1-P_{A\|\bar{B}})(1-P_B)$
15	$\bar{A}B$	$\bar{A}\bar{B}$	$f(\bar{A}_1B_1\bar{A}_2\bar{B}_2)$	$\beta_1(1-\beta_1)(1-\alpha_{11})(1-\gamma_{11})P_{A\|B}P_B + \beta_0(1-\beta_0)(1-\alpha_{10})(1-\gamma_{10})P_{A\|\bar{B}}(1-P_B)$
				$+\ \beta_1(1-\beta_1)\alpha_{01}\gamma_{01}(1-P_{A\|B})P_B + \beta_0(1-\beta_0)\alpha_{00}\gamma_{00}(1-P_{A\|\bar{B}})(1-P_B)$
16	$\bar{A}\bar{B}$	$\bar{A}\bar{B}$	$f(\bar{A}_1\bar{B}_1\bar{A}_2\bar{B}_2)$	$(1-\beta_1)^2(1-\gamma_{11})^2P_{A\|B}P_B + \beta_0^2(1-\gamma_{10})^2P_{A\|\bar{B}}(1-P_B) + (1-\beta_1)^2\gamma_{01}^2(1-P_{A\|B})P_B$
				$+\ \beta_0^2\gamma_{00}^2(1-P_{A\|\bar{B}})(1-P_B)$

OCCUPATIONAL CLASSIFICATION: CURRENT ISSUES AND AN INTERIM SOLUTION*

Jerome B. Gordon, Bureau of Applied Social Research, Columbia University

Introduction

Of late, there has been a ground swell of interest in and criticism of the occupational classification currently used by the U.S. Bureau of the Census.[1] What has distinguished this recent spate of professional discussion from those in the past is the unanimity of a broad cross section of the users of Census occupational data on the fundamental lack of relevance of the current scheme in supplying meaningful data for analytical purposes. Increasingly, both sociologists and economists have become aware of the difficulties in using the broad or detailed occupational data for analysis of, for example, changes in socio-economic status or projections of manpower requirements.

Concurrent with this interest and criticism has been the experimentation with possible "candidates" for replacement of the existing structure. This experimentation has taken two basic directions.

First, development of "Convertibility Lists" between the existing Census occupational classification and the Bureau of Employment Security Dictionary of Occupational Titles, as well as a similar "List" between the D.O.T. and the International Standard Classification of Occupations (I.S.C.O.) of the International Labour Office (I.L.O.).[2] Second, examination of existing occupational taxonomic principles and development of new taxonomic principles resulting in more "homogeneous" detailed as well as broad occupational groupings.[3] Expectations from the work above would be: 1) a reduction in classification errors (e.g., reduction of the proportion of individuals reported as "N.E.C."); 2) the construction of job-families based on job content for projection of manpower supply and demand by, for example, skill; 3) availability of suitable data for analyzing and evaluating alternative manpower policies and programs.[4]

The purposes of this paper, then, are three: first, to briefly review the dimensions of the criticism of the existing Census occupational classification; second, to examine some of the suggested solutions for revising the current occupational classification; third, to display and discuss a major occupational schema devised by the author, meeting some of the objections in the current debate within the constraints of the existing occupational classification.

Dimensions of the current criticism

According to Webster the generally accepted definition of "Occupation" is: "the principal business of one's life; a craft, trade, profession or other means of earning a living."[5] The sociologists and students of occupational data are a bit more rigorous in defining the "pigeonhole" marked "occupation." Reiss, for one, defines "occupation" as follows:

> The social valuations attached to work in a society may be thought of as referring to both the kind of work a person does and the situation in which one works. The specific kind of work a person does in a socially evaluated work situation generally is thought of as a job, while an occupation refers to the characteristics that are transferable among employers.[6]

Some of the dimensions used to define "work" or "job" are:[7] 1) specific tasks or functions performed; 2) the purpose for which the job is done; 3) the materials, tools, and equipment used; 4) the standard to be met in the working environment; 5) the education and training needed by a worker to perform a job. Hence, an "occupation" is the sum total of homogeneous work characteristics that are, theoretically, transferable among employers, industries and classes of workers.

This leads us to the fulcrum of criticism in the literature. That is, that the taxonomic principle of "homogeneity" in defining both broad and detailed occupations is being violated in the current Census classification. Witness this comment:

> "...occupational codes are not formed solely by reference to the similarity in tasks performed by individual incumbents of specific jobs. Instead, work settings, industrial affiliations, (class of worker) and other factors are used to define detailed occupational groups."[8]

A more fundamental criticism is that this "heterogeneity" in the existing classification is fostered by the current standards of determining broad as well as detailed groups: the use of "socio-economic status."[9] This standard, according to one analyst, results in sizable enumeration and classification errors. For example, the residual categories, the n.e.c. groups, of the major occupations contain over one-third of those employed![10] Further, even if socio-economic position is used as a guide to determining major occupational groups, it does not seem to be applied rigorously. Again, this comment from the investigative literature:

> "...despite the obvious use of socio-economic criteria to form them (the major occupations), they do not comprise a grouping of detailed

*This paper is part of a larger study on the "Demography of the Middle Years" under sponsorship of the Russell Sage Foundation.

occupations effected on the basis of status alone....we know...that ratings assigned by respondents to specific jobs or occupations falling in different major occupational groups overlap appreciably."[11]

Lastly, the compounding of the heterogenity in content of the occupational groups, the problems with the magnitude of the residual categories--and the large expected classification and enumeration errors, present the most important problem: the relevance of the data. Among labor force and manpower specialists there is an increasing realization that the current classification is not providing useful information for a host of analytical areas.

Current occupational data are not amenable to analysis of labor resource allocation decisions, e.g., projection of manpower supply and demand requirements.[12] There appears to be too little information on a) the number of jobs available at different combinations of skill and wage levels, and b) the numbers of workers possessing different skill levels at prevailing or expected wage levels. Hence, the current occupational structure falls down at precisely the point where the economist, for one, needs it--that is, supplying data on the range of substitution possibilities between occupation groups.

With this in mind, let us turn to the several schemes suggested as possible substitutes for the existing Census occupational classification, and the associated problems in changing to another system of classification.

Suggested occupational classifications

While there have been innumerable suggested schema put forth for classifying occupations, most of these fall into two "families" of technique. The first, which was the route followed in the development of the current Census schema, is the so-called "direct approach"--that is, the classification of the labor force by an independent criterion--for example, ranking occupations by median level of education or "socioeconomic status" score.[13] The second, currently being used by development planners, is the "distilled technique." This consists of first grouping occupations by the "work performed" or "job content" and then stratifying the groups in terms of "skill prerequisites." Most of the current experiments with occupational classification schemes have been of the latter variety.

Simple moves towards the type of homogeneous occupational structure desired by analysts are those currently involved in matching Census and Dictionary of Occupational Titles data. The distinction between the Census and D.O.T. classifications rests largely on the stratification of the blue-collar occupations by the stage and type of production process. Recent developments leading towards a "convertibility" list between the two classification schemes have attempted to define similar "work content" differences in the lower white-collar occupations, i.e., clericals.[14]

A slightly more advanced schema was that used by Parnes in his examination of manpower and skill requirements in the OECD Mediterranean Region Research Project.[15] Using the International Standard Classification of Occupations (I.S.C.O.) of the International Labour Office (I.L.O.), Parnes coded detailed occupations by their skill prerequisites, i.e., the level of education required. The resulting four basic groups formed the basis for determined skill requirements for individual countries.[16]

By far the most sophisticated suggestion put forth to date is the creation of "job-families."[17] "Job-families" would comprise aggregates of common demand characteristics of detailed occupations. That is, the "families" would have relatively common content--the tasks the individual job-holder performs--and would be stratified by degree of complexity involved in the conduct of work. Job content, of course, would be intimately interwoven with educational and training prerequisites. Further, considerations of wage structure and mobility patterns would also affect the definition of "job-families."[18] Thus, occupational classes or job-families would be relatively homogeneous. In economic terms, there would be a high degree of substitutability of detailed occupations within each class. Similarly, occupational classes or "job-families" would be differentiated from one another by a low degree of substitutability between one another, i.e., low cross-elasticity of demand. On the supply side of the picture, similar ground rules would be employed to define so-called "worker" classes. Advocates of such a classification system contend that the benefits accruing will be "adaptability to changes in technology and educational policy to isolate new jobs and hence new skills which alter substitution possibilities."[19]

Problems of changing the existing occupational classification

While some of the suggestions put forth in the previous section would not result in too great a modification of the existing Census occupational schema, others, particularly those directed towards creation of so-called "job-families," would. The problems to be faced in changing the existing classification system can be viewed as falling into three areas: comparability considerations, measurement and analysis, and costs and benefits.

All of the suggestions discussed earlier generally have the following beneficial characteristics: 1) all are concerned with devising more homogeneous occupational classes; 2) all of them are directed toward better definition of the labor force, e.g., reduction of the proportion of the total labor force "not elsewhere classified" (n.e.c.); 3) all of them expect, as an end product, better and more relevant data for use in manpower and labor force analysis.

Let us now turn to the problems engendered by changing the occupational structure.

The ideal of creating an occupational classification composed of more homogeneous groupings is plagued by the bogey of comparability. A radical departure from the existing system will cause considerable mental anguish in the community of census users, one of whose primary desires is the investigation of historical trends. Further, to recast prior census materials in any contemplated schema would be a costly operation. Lastly, but not an unimportant consideration, is the matter of timing and tooling up for any change in the occupational classification. A substantive revision in the Census occupational schema could not be implemented much before the 1980 Census at the earliest. Rigorous statistical standards and numerous tests of any radical schema would have to be established before any intended classification system could be implemented.

While the comparability issue is a constraint in the consideration of alternative occupational classifications, it is not a rigorous one. Supplemental classification schemas have been devised and used with the existing Census occupational structure. Limited departures from the existing codes, in the directions desired by the current crop of critics, can be handled without much degradation in comparability. The suggested modification presented in this paper is one such revision.

While the development of an occupational code with numerous homogeneous groups may be a desirable goal, it runs right into another constraint: statistical reliability. A very refined structure, which is desired by many of the critics of the existing one, when crossed with other socio-economic variables for analysis purposes, will very quickly run out of expected cases per cell. Thus, the accuracy of parameters at very refined levels of aggregation will become more dubious than they are at present.[20] Depending upon the number of variables used to define an occupational code, too fine a structure may represent a less objective and hence less independent yardstick for statistical analysis.[21]

The foregoing is meant only as food for thought in the evaluation of alternative occupational classifications; it is not meant to throw cold water on thinking about or design of alternatives. To the contrary, what it does do is define the decision space within which alternatives should be considered.

A Suggested Interim Solution

Discussion

The 10 or 12 Census major occupational groups are a mixed bag. The present classification places certain detailed occupations in what would seem *a priori* to be the wrong major occupational group. For example, in a recent study of differential occupational mobility by color, a significant amount of net in-movement was observed in a residual service category for white males.[22] On closer examination, a significant proportion of the net in-movement in this residual service category was explained by the growth of "protective service" occupations. The socio-economic characteristics of this group were not only significantly different from the other "Service" groupings, but also were significantly different from the socio-economic characteristics of the major occupational groupings purportedly above "Service Workers" in the occupational hierarchy, i.e., operatives, craftsmen and clericals.[23] Hence, contrary to Scoville's criticisms, the current Census major occupational groups would appear to be a "weak" ordering of occupations on the basis of socio-economic characteristics. Numerous other examples of mis-classification, using the existing major occupational groupings, can, I am sure, be documented.

Thus, it would appear that a first attempt at deriving more homogeneous major groupings can be to use the existing materials on the socio-economic characteristics of occupations in conjunction with their relationship to, for example, the functioning and development of the economy. Once a definition of the socio-economic characteristics of the detailed occupations was established, the resulting ranking could be broken into an initial major grouping based on the relative importance of the occupational socio-economic characteristics. A final major occupational classification could then be developed by segregating the occupations into groups on the basis of their relationship to technological change. This composite major occupational code would have the benefits of greater homogeneity than is presently available with the Census groups, greater adaptation to change in detailed occupation, as well as greater analytical clarity in considering questions of skill requirements.

These desirable features of such a major occupational classification should not be underestimated. No matter what the form or composition of the index used to compute the socio-economic score for the detailed occupation, the original breaks in the ranking of socio-economic scores on the basis of relative importance will not change.[24] Thus, the introduction of totally new detailed occupations--achieved through better definition of the current residual "n.e.c." categories or through technological change--will not perturb the original structure of major occupational groupings. It will simply expand the number of occupations falling in a specific socio-economic grouping. This will significantly reduce the problem of intercensus comparability, and lessen the likelihood of classification errors due to arbitrary allocation standards. The major differences between Censuses will, however, be very sensitive to the standard used in further subdividing the detailed occupations into groups for analysis of manpower and skill requirements.

These desirable features of such a major occupational classification should not be underestimated. No matter what the form or composition of the index used to compute the socio-economic score for the detailed occupation, the original breaks in the ranking of socio-economic scores on the basis of relative importance will not change.[24] Thus, the introduction of totally new detailed occupations--achieved through better definition of the current residual "n.e.c." categories or through technological change--will not perturb the original structure of major occupational groupings. It will simply expand the number of occupations falling in a specific socio-economic grouping. This will significantly reduce the problem of inter-census comparability, and lessen the likelihood of classification errors due to arbitrary allocation standards. The major differences between Censuses will, however, be very sensitive to the standard used in further subdividing the detailed occupations into groups for analysis of manpower and skill requirements.

With these considerations in mind, an attempt was undertaken to produce: 1) a system of major occupations on the basis of relative importance of socio-economic characteristics of detailed occupations, using 1960 Census materials; 2) a further subdivision of the resulting major occupational code, using the "livelihood code" system suggested by Jaffe.[25]

Data and methods

As mentioned previously, the objective of the exercise was to recast the present Census Occupational classification into a more homogeneous one. This was accomplished by utilizing the existing ordering of the detailed occupations by Socio-Economic Status (SES) Scores and devising new major occupational groupings on the basis of differences in the relative values of the Socio-Economic Status (SES) Scores.[26] Relative values for Socio-Economic Status (SES) Scores were estimated, using the Ordinal-Relative Value Conversion technique developed by the Research Analysis Corporation (RAC).[27]

The technique has the following features.

First, estimates of relative value are derived through successive chains of coordinates of plotted points of half-values from a truncated ordering.[28]

Second, estimates of relative value from each chain of plotted points of half-values are normalized to derive a common scale.

Third, normalized estimates of relative value from each chain of plotted points of half-value are averaged on the common scale to derive a final composite estimate.

Fourth, where relative values are lacking at the extremes of the ordering--for example, in the truncated portion--they are built up from pairs of order numbers which have equivalent values to the order numbers previously dropped during the estimating.

The decision rule used in defining the major occupational groupings was that major breaks in the occupational ranking occurred when second differences between composite relative values of Socio-Economic Status (SES) reached zero.

Suggested Alternative Interim Major Occupational Codes

The following three tables depict the "menu" of possible major occupational groups, using the class intervals in Socio-Economic Status (SES) Scores, derived through the RAC Ordinal-Relative Value technique. These represent different levels of aggregation using combinations of the Current Census Codes, the Socio-Economic Status Score class intervals and the Jaffe "livelihood" codes. Based on a benchmark minimum of 100,000 cases per grouping, the table below summarizes the number of possible major occupational groups for each alternative.

Occupational Classification Scheme	Number of possible major occupational groups having a minimum of 100,000 per group
1. Current Census Occupational Code	12
2. Socio-Economic Status (SES) Score - (Table 1)	16
3. Jaffe Livelihood Codes	19
4. Socio-Economic Status (SES) Score x Major Jaffe Livelihood Occupational Codes - (Table 3)	41
5. Socio-Economic Status (SES) Score x Current Census Occupational Code - (Table 2)	52

Choice of an appropriate alternative from the array suggested here will depend largely on the scope and nature of the investigation. Modification of these suggested alternatives is of course possible. For example, analyses may be undertaken using the alternative occupational codes and new groups formed based on the degree of similarity in mobility patterns, or differences in the wage structure.

The Socio-Economic Status (SES) Score major occupational code, shown in Table 1, represents a moderate improvement towards greater homogeneity over the existing Census major occupational code. By cross-classifying the Socio-Economic Status (SES) Score code with the current Census major occupational code, we can get a better feel for the extent of heterogeneity in the existing Census major occupational classification. For example, the range in Socio-Economic Status (SES) Scores for "Professionals" is approximately 50 points, and in the case of "Operatives" it is over 70 points. (See Table 2.) It would appear, then, that part of the lack of association between Socio-Economic Status (SES) and the current major occupational hierarchy is explained by the extensive variability in Socio-Economic Status within each major occupation. Hence, one improvement, using the current Census code, might simply be to develop a major occupational code stratified by the Socio-Economic Status (SES) Score intervals as shown in Table 1.

A more involved modification of the existing code is the development of a major occupational code based on a cross-classification of the Socio-Economic Status (SES) Score intervals and the Jaffe "livelihood" code. The result would be an identification of occupations by their involvement--or non-involvement--with the development of and use of technology and other economic functions, i.e., "Modern," "Classical" and ancillary occupations; e.g., administration and distributive functions, service functions, agricultural functions and the pool of industrial unskilled. Table 3 shows such a code for the 1960 male labor force. The breakdown is useful in several respects. First, it enables one to identify groups for more refined analysis of changes in the occupational requirements of the economy over time, i.e., during different stages of economic development and growth. Second, it prevents the masking of trends, presently a problem in the use of the existing Census major occupational code for forecasting occupational requirements. Third, the recasting of the existing Census occupational code into this suggested format can be accomplished with a minimum of significant technical problems.[29]

The Usefulness of SES Score Groupings: The Measurement of Socio-Economic Inequality

A variety of tools has been used by sociologists and economists to detect dissimilarities in status and income. The SES Score groupings presented in this report provide another means of estimating the presence and extent of inequalities in both status and income within and between occupational groups and countries. The measure of socio-economic inequality can further be analyzed to detect sources or possible causes of the inequalities, as well as--in a policy sense--the programs that might ameliorate them.

To illustrate this use of the SES Score classification scheme, data on the distribution of socio-economic status for the United States and Panama were used.[30] The extent of socio-economic inequality in both countries and the differences between them can be graphically portrayed through the device of a Lorenz curve.

Socio-economic status Lorenz curves were developed for the United States and Panama by plotting the cumulative proportion of units arrayed in order, from the lowest SES Score group to the highest, against the cumulative proportional share of the aggregate socio-economic status accounted for by these units. Theoretically, if all units had exactly the same amount of socio-economic status, the Lorenz curve would be represented by a 45^{o} line drawn through the origin. For a variety of reasons, curves drawn to actual data will fall below the diagonal. Generally, the greater the inequality in socio-economic status, the greater will be the area between the diagonal and the Lorenz curve. A measure of inequality or concentration is the Gini coefficient, which is the proportion of the total area under the 45^{o} line that is between the diagonal and the Lorenz curve.[31]

Examining the socio-economic status score Lorenz curve drawn for the United States and Panama in Figure 1, one can quite easily see the greater degree of socio-economic inequality in Panama. In the case of Panama 50 per cent of the units have 25 per cent of the aggregate socio-economic status, while in the case of the United States the same proportion of units has 35 per cent of the aggregate socio-economic status. This is further summarized when one examines the Gini coefficients of socio-economic status concentration. For the United States it is .232, while for Panama it is .312. The question remaining, in a speculative sense, is what are the possible causes for socio-economic inequality.

In an a priori sense, a lower limit for the aggregate socio-economic inequality in a country or region is the statistical distribution of the abilities of the basic population. Hence, depending upon the shape of the distribution of abilities, in a very real sense there will never be perfect socio-economic equality in a country or society. That is, the socio-economic Lorenz curve will depart from the diagonal by some minimal area between the diagonal and the Lorenz curve as a function of the distribution of abilities of the population. Anything above this minimal degree of socio-economic inequality is the product of several factors: 1) the mix of institutional arrangements in the country or region, 2) the degree of industrialization, 3) the proportion of the total population of working force age employed in the market sector of the economy, 4) the universality and availability of education on all levels, 5) status consistency, 6) the ease of socio-economic mobility, etc.

In terms of a continuum of economic development--ranging from the less to the more developed countries--it is possible that socio-economic inequality may be significantly related to the extent of underemployment, and the rigidity of societal and government controls on upward socio-economic movement. Thus, a country with substantial urban and rural underemployment, low levels of literacy, and limited social mobility opportunities may have relatively greater socio-economic inequality than a more advanced nation with a majority of its working force employed in non-agricultural pursuits, a substantial number of its younger population enrolled in secondary and higher education institutions, and a government, one of whose major policies is the reduction of social immobility.

Thus, the existence of the socio-economic concentration measure permits us to undertake a whole new range of investigations in a previously large and unmeasured aspect of development. In a policy vein, we can analyse the determinates of socio-economic inequality within or between countries over time and detect those factors which offer relatively greater leverage in the reduction of socio-economic inequality. We now have a socio-economic target or indicator that could tell us how far development programs or governmental regulations have gone in reducing socio-economic inequalities. In the instance of international aid programs such as the Alliance for Progress, through appropriate analysis of the relationships of socio-economic inequality to the stage of economic development and rate of economic growth, we could determine the aggregate amount and type of effort required to achieve national integration. In the case of racial integration in the United States, one could determine how much or how little progress has been made in reducing the socio-economic inequality of the American Negro to the approximate lower relative limit of that of the rest of American society.

In any case, the potential use of socio-economic status grouping, such as the one developed here, can serve a variety of statistical and analytical purposes.

Conclusions

1. While innumerable problems beset the existing Census major occupational classification, several means are currently available for developing new major occupational codes with varying degrees of technical feasibility.

2. A series of alternative major occupational codes can be developed by stratifying occupations by the relative importance of their respective Socio-Economic Status (SES) Scores.

3. The alternative major occupational codes developed, using this scheme, offer greater homogeneity in occupational grouping than is possible with the current code, greater usefulness to investigators involved in manpower analysis and forecasting, as well as a vehicle for future research into occupational classification.

FOOTNOTES

1. See, for example, the collection of papers on the subject in the Proceedings of the Social Statistics Section, American Statistical Association, 1966, pp. 176-208.

2. See L. Lewis, "Development of a Convertibility List Between the DOT and Census Classification Systems," in Proceedings, op. cit., pp. 204-206.

3. R. W. Hodge and P. M. Siegel, "The Classification of Occupations: Some Problems of Sociological Interpretation," in Proceedings, op. cit., pp. 176-192.

4. J. Scoville, "Making Occupational Statistics More Relevant," Proceedings of the Business and Economic Statistics Section, American Statistical Association, 1965, pp. 317-323.

5. Philip Babcock Gove, ed., Webster's 3rd New International Dictionary. Springfield: G. & C. Merriam, 1961, p. 1560.

6. Albert J. Reiss, et al., Occupations and Social Status. New York: The Free Press of Glencoe, 1961, p. 10.

7. G. Alexandrin, "Notes on Occupational Classification," Quarterly Review of Industrial Relations, Vol. 21, No. 4 (1966), p. 531.

8. Hodge and Siegel, op. cit., p. 178.

9. Scoville, op. cit., p. 317.

10. Scoville, ibid., p. 318.

11. Hodge and Siegel, op. cit., p. 185. See also: Otis Dudley Duncan, "A Socio-Economic Index for All Occupations," in Reiss, op. cit., pp. 109-138.

12. See W. Lee Hansen, "Labor Force and Occupational Projections," Proceedings of the Industrial Relations Research Association, 1965, pp. 10-20.

13. See J. Scoville, "The Development and Relevance of U.S. Occupational Data," Industrial and Labor Relations Review, Vol. 19, No. 1 (1965), pp. 70-79; and U.S. Bureau of the Census, "Methodology and Scores of Socioeconomic Status," Working Paper No. 15, Washington, D.C., 1963.

14. See Lewis, op. cit., p. 205.

15. See H. S. Parnes, Forecasting Educational Needs for Economic and Social Development, Organization for Economic Co-operation and Development, Paris, 1962.

16. Parnes, op. cit., p. 26. The four groups were:

Class A: All occupations for which a university education ...is required.

Class B: Occupations for which two or three years of education beyond the secondary level...are required.

Class C: Occupations for which a secondary education...is required.

Class D: All other occupations.

17. Scoville, "Making Occupational Statistics More Relevant," op. cit., p. 320.

18. See Stanley Lebergott's suggestion for stratifying detailed occupations by wage levels in Proceedings of the Social Statistics Section, American Statistical Association, 1966, p. 207.

19. G. G. Cain, W. L. Hansen and B. A. Weisbrod, "Occupational Classification: An Economic Approach," Monthly Labor Review, Vol. 90, No. 2 (1967), pp. 48-52.

20. See Oskar Morgenstern, On the Accuracy of Economic Observations, 2nd ed. (Princeton: Princeton University Press, 1963), especially Chapter XIII, "Employment and Unemployment Statistics," pp. 228-234.

21. See Robert Solow's comment on the problem of level of aggregation of analysis in reply to G. H. Orcutt and A. M. Rivlin, "An Economic and Demographic Model of the Household Sector; A Progress Report," in Demographic and Economic Change in Developed Countries, Universities-National Bureau of Economic Research, Princeton University, 1960, pp. 318-321.

22. See A. J. Jaffe and J. B. Gordon, "Occupational Mobility for White and Nonwhite Males: 1950-1965," The New York Statistician, New York Area Chapter of American Statistical Association, Vol. 18, No. 4 (1966), pp. 1-3.

23. Jaffe and Gordon, op. cit., Table 1.

24. That is, the components of the index may be independent--hence the index would be "additive"--or the index may be interdependent--hence the index would be multiplicative. Weighting of components of a socio-economic index would be a thorny issue, and one which would require more experimentation to test the sensitivity of final composite scores to changes in the weighting schemes.

25. See A. J. Jaffe, "Suggestions for a Supplemental Grouping of the Occupational Classification System," Estadistica, March 1957, pp. 13-23. Jaffe's schema utilized a simple dichotomy of occupations between those classes of occupations that involved the use of technology and those that did not, i.e., "Classical" versus "Modern" occupations. Detailed codes were developed by Jaffe and Froomkin for the 1960 Census and a 5 per cent sample tape of the labor force was created, using the code.

26. The Socio-Economic Status (SES) Scores computed and published by the Census Bureau with the 1960 Census of Population, provide a multiple-item index of socio-economic ranking by occupation. The three components are occupation, educational attainment and family income for the chief income recipient. Scores are obtained for each component on a scale ranging from 1 to 100 and averaged to provide a single over-all figure. For more information on the methodology and use of Socio-Economic Status (SES) Scores, see U.S. Bureau of the Census, "Methodology and Scores of Socioeconomic Status," op. cit.

27. See, for example, Jerome B. Gordon, "Conversion of Ordinal Values to Relative Values for Research and Exploratory Development Project Proposals," The Research Analysis Corporation, McLean, Va., December 1965. Reprinted in The Proceedings of the 15th Military Operations Research Symposia, Office of Naval Research, 1966.

28. "The RAC technique makes the assumption that there is an underlying regularity between rank and relative value....As can be seen the RAC technique rests only on the assumption about the regularity of the universe of relative values implying the ability to interpolate between directly estimated points. A logarithmic scale (is used) because the half-values plotted in tests exhibited general linearity where the scale was logarithmic." Gordon, ibid., pp. 3 and 9.

29. The suggested occupational codes will not, of course, remedy the problems of reporting and definition. These are additional issues: issues that impinge on any occupational classification. However, the occupational codes suggested here offer great flexibility in terms of use with the existing code, as well as adaptability to changes in definition of detailed occupations.

30. Socio-Economic Status Scores for Panama were developed from tabulations of the educational, earnings and employment characteristics of detailed occupations in the Metropolitan areas of the Republic, drawn from the monthly Current Population Survey of the Republic of Panama. A special memorandum listing the detailed occupations falling into the major SES group classification for the Republic of Panama, as well as a similar one prepared for the United States, are available from the author upon request.

31. For a graphical explanation of how the Gini coefficient is derived, see James Morgan, "The Anatomy of Income Distribution," The Review of Economics and Statistics, August 1962, pp. 281-282.

APPENDIX A*

STEPS IN CONVERSION OF ORDINAL TO RELATIVE VALUES

The problem is to estimate relative values for a set of objects or items where the value rankings are given. It is assumed that, in the course of the estimating procedure, usefully accurate judgments can be made to determine which of any two given pairs of objects has a combined value for the pair that more nearly approximates the value of some given, single object.

Presented below are formal steps. The following symbols are used:

n The number of objects to be valued.

p A number in the sequence
$1, (1+1/3), (1+2/3), 2, \ldots, (n-1/3), n$.

$\bar{p}$ A number in the set
$1, 2, \ldots, n$.

V_p The value associated with p. The p may be subscripted. For any $\bar{p}$,

$$V_{\bar{p}} = V_{\bar{p}-1}.$$

p_j The coordinate on the p scale for the jth plotted point of a sequence of plotted points.

$\bar{p}_j$ The $\bar{p}$ closest in magnitude to p_j.

$\bar{p}_{jk}$ $\bar{p} + k$, where k is -1, 0, or +1.

$p'_{(j+1)k}$ An estimate of what p_{j+1} would be if $\bar{p}_{jk}$ were the last previously plotted point.

The steps in the procedure follow.

1. Select a p_1 (the first of the p_j) from among the highest-ranking $\bar{p}$, but low enough to be in a region where it is <u>judged</u> that the ratios between values associated with adjacent $\bar{p}$ are not significantly greater than the ratios between values for adjacent $\bar{p}$ in the region of middle $\bar{p}$ levels. In the absence of ability to make such judgment, a $\bar{p}$ at about the 90th percentile will be satisfactory.

2. The first plotted point is (p_1, V_{p_1}), where V_{p_1} is selected arbitrarily (but preferably near the top of the graph to leave room for later plotted points of lower value). The graph paper has a horizontal arithmetic scale for p and a vertical logarithmic scale for V_p. The units of measurement on the vertical scale are arbitrary.

3. For each of the three possible levels of k, estimate a $p'_{(j+1)k}$ such that, as nearly as possible,

$$V_{\bar{p}_{jk}} = V_{p'_{(j+1)k} + 1} + V_{p'_{(j+1)k} - 1}.$$

4. Calculate

$$p_{j+1} = p_j - {}_{k}(\bar{p}_{jk} - p'_{(j+1)k}\ /3$$

5. Plot the point $(p_{j+1}, V_{p_j}/2$; and then call this point the new base (p_j, V_{p_j}) for additional iterations of steps 2 through 5. Continue iterations until a point is plotted after which no further points can be plotted, because the step 3 condition can no longer be approximated.

6. Connect all plotted points by a straight line between adjacent points. The lines as joined at the plotted points are hereafter referred to as a plotted chain.

7. The plotted points and the straight lines between them on a plotted chain provide estimates of the relative values $(V_{\bar{p}})$ for all $\bar{p}$ within the range of the plotted chain. Additional estimates may be made for later averaging by calculating additional plotted chains, starting from points with coordinates between p_1 and p_2 of the first plotted chain. Calculate as many additional plotted chains as desired, with the starting point of each so determined that the intervals on the p scale between starting points of all plotted chains are approximately equal.

8. Tabulate the $V_{\bar{p}}$ for all plotted chains and for all $\bar{p}$ covered by any of the chains.

9. Retabulate dividing each tabulated $V_{\bar{p}}$ for a plotted chain by the $V_{\bar{p}}$ for some middle level of $\bar{p}$, using the same $\bar{p}$ to determine the divisor value in the case of each of the plotted chains. The result is a table of scaled relative values where, for each plotted chain, the relative value is unity at the same level of $\bar{p}$.

*Bernard Sobin and J. B. Gordon, "Improvement of Army Methods of Determining Research and Exploratory Development Programs," T-482, Research Analysis Corporation, McLean, Va., September, 1966.

10. Calculate an arithmetic mean of scaled relative values across plotted chains for each level of $\bar{p}$. Make sure that the divisor in calculation of an arithmetic mean is equal to the number of plotted chains that have scaled relative values for the $\bar{p}$ concerned; do not consider a missing scaled relative value as a zero. The arithmetic means that are calculated with as many items as there are plotted chains are the final estimates of relative values within the $\bar{p}$ ranges of the plotted chains.

11. Where an arithmetic mean has been calculated with fewer items than the number of plotted chains (because one or more of the plotted chains does not have sufficient range), multiply that mean by a scale factor. The scale factor is calculated as

$$S_p = M_{p^*}/M_{p^*}^{(p)} \text{ , where}$$

S_p is the scale factor for the arithmetic mean at order number p ; M_{p^*} , the mean at the order number nearest to p that is within the range of every one of the plotted chains; and $M_{p^*}^{(p)}$, what the mean at p^* would have been had it been calculated only with items from plotted chains that include p within their ranges.

12. Calculate the value to be associated with the smallest $\bar{p}$ that is above the range of the plotted chains, calculating it as a sum of any two of the arithmetic means calculated in steps 10 and 11 as are estimated to be most nearly equal to it in aggregate value.

13. Calculate the value to be associated with the next largest $\bar{p}$ in the same way, except that the selection of pairs of values can be not only from the adjusted arithmetic means calculated in steps 10 and 11, but also any values calculated in later steps. Continue this procedure iteratively until values have been estimated for all $\bar{p}$ above the range of the plotted chains.

Table 1

MAJOR OCCUPATIONAL GROUPS BASED ON RELATIVE VALUES OF SOCIO-ECONOMIC STATUS (SES) SCORES: EMPLOYED MALES (000), 1960

Group #	SES Score Range		Total No. of Employed Males (000)
1	96 - 100		1,947.4
2	90 - 95		2,585.9
3	81 - 89		4,232.8
4	76 - 80		2,486.5
5	71 - 75		4,197.4
6	64 - 70		2,101.0
7	56 - 63		6,202.5
8	49 - 55		2,067.5
9	40 - 48		3,579.1
10	31 - 39		5,782.5
11	25 - 30		1,130.5
12	20 - 24		490.1
13	14 - 19		2,319.9
14	1 - 13		753.8
15	*		2,387.6
16	**		1,201.8
		TOTAL	43,466.3

*Farmers and Farm Managers

**Farm Laborers and Foremen

NOTE: These groups do not have Socio-Economic Status Scores calculated for them.

Source: U.S. Bureau of the Census, U.S. Census of Population: 1960. Detailed Characteristics. U.S. Summary. Final Report PC(1)-1D. U.S. Government Printing Office, Washington, D.C. Table 202, pp. 1-522-29.

Table 2

CROSS-CLASSIFICATION OF MAJOR OCCUPATIONAL GROUPS BASED RELATIVE VALUES OF SOCIO-ECONOMIC STATUS (SES) SCORES AND THE CENSUS MAJOR OCCUPATIONAL GROUPS: EMPLOYED MALES (000), 1960

Group #	SES Score Range	Census Major Occupational Groups: Professionals	Managers	Clericals	Sales	Craftsmen	Operatives	Service	Laborers	Farmers & Farm Managers	Farm Laborers	Occupation Not Reported
1	96 - 100	1,582.0	365.4	--	--	--	--	--	--	--	--	--
2	90 - 95	915.5	1,440.7	173.7	56.0	--	--	--	--	--	--	--
3	81 - 89	1,217.9	1,219.5	188.9	1,367.8	238.7	--	--	--	--	--	--
4	76 - 80	278.7	572.5	192.5	133.4	1,241.1	68.3	--	--	--	--	--
5	71 - 75	235.9	380.5	1,974.5	--	1,049.6	188.9	368.0	--	--	--	--
6	64 - 70	210.2	174.2	205.1	4.0	1,331.4	138.4	37.7	--	--	--	--
7	56 - 63	29.7	208.3	258.6	1,211.8	1,831.1	2,621.5	28.3	13.2	--	--	--
8	49 - 55	0.8	192.8	12.6	--	1,217.6	627.9	0.2	15.6	--	--	--
9	40 - 48	8.6	33.2	48.2	--	215.3	2,972.0	224.1	77.7	--	--	--
10	31 - 39	--	--	3.8	--	1,320.4	1,323.1	896.5	251.7	--	--	1,987.0
11	25 - 30	--	--	--	--	--	230.8	104.4	795.3	--	--	--
12	20 - 24	--	--	--	182.1	43.6	10.2	--	254.2	--	--	--
13	14 - 19	--	--	--	--	--	267.6	939.4	1,112.9	--	--	--
14	1 - 13	--	--	--	22.7	--	192.4	61.3	477.4	--	--	--
15	Farmers	--	--	--	--	--	--	--	--	2,387.6	--	--
16	Farm Wkers.	--	--	--	--	--	--	--	--	--	1,201.8	
TOTAL		4,479.3	4,629.6	3,015.4	2,977.8	8,488.8	8,641.1	2,659.9	2,998.0	2,387.6	1,201.8	1,987.0

Sources: U.S. Bureau of the Census. U.S. Census of Population: 1960. Detailed Characteristics. U.S. Summary. Final Report PC(1)-1D. U.S. Government Printing Office, Washington, D.C. Table 202, pp. 1-522-29.
See also Table 1, preceding page.

Table 3

CROSS-CLASSIFICATION OF MAJOR OCCUPATIONAL GROUPS BASED ON RELATIVE VALUES OF SOCIO-ECONOMIC STATUS (SES) SCORES AND MAJOR "LIVELIHOOD" CODES
EMPLOYED MALES (000), 1960

		Major Livelihood Code Groups						
Group #	SES Score Range	Modern[1]	Classical[2]	Managerial, Administrative & Distribution[3]	Service[4]	Industrial & Commercial, Unskilled[5]	Forestry, Fisheries & Agriculture[6]	Occupation Not Reported[7]
1	96 - 100	1,238.1	343.9	365.4	--	--	--	--
2	90 - 95	808.1	107.4	1,670.4	--	--	--	--
3	81 - 89	748.9	707.7	2,776.2	--	--	--	--
4	76 - 80	1,588.1	--	898.4	--	--	--	--
5	71 - 75	1,391.8	82.6	2,355.0	368.0	--	--	--
6	64 - 70	1,429.5	250.5	383.3	37.7	--	--	--
7	56 - 63	4,227.9	254.4	1,678.7	28.3	13.2	--	--
8	49 - 55	1,601.6	244.7	205.4	0.2	15.6	--	--
9	40 - 48	2,913.5	282.4	81.4	224.1	77.7	--	--
10	31 - 39	1,244.5	1,399.0	3.8	896.5	251.7	--	1,987.0
11	25 - 30	230.8	--	--	104.4	795.3	--	--
12	20 - 24	10.2	43.6	182.1	--	254.2	--	--
13	14 - 19	260.6	7.0	--	939.4	919.5	193.4	--
14	1 - 13	192.4	--	22.7	61.3	325.5	151.9	--
15	Farmers	--	--	--	--	--	2,387.6	--
16	Farm Wkers.	--	--	--	--	--	1,201.8	--
TOTAL		17,886.0	3,723.2	10,622.8	2,659.9	2,652.7	3,934.7	1,987.0

1. Modern occupations are Census Occupation Codes 100 for Professionals, 400 and 410 for Craftsmen, and 600 for Operatives.
2. Classical occupations are Census Occupation Codes 200 and 210 for Professionals, 500 for Craftsmen, and 610 for Operatives.
3. Managerial, Administrative and Distribution occupations are Census Occupation Codes 230 and 240 for Managers, 300 and 310 for Clericals, and 330 for Sales Workers.
4. Service occupations are Census Occupation Codes 800 for Private Household Service Workers, and 810 for Service Workers.
5. Industrial and Commercial Unskilled occupations are Census Occupation Code 902 for Laborers.
6. Forestry, Fisheries and Agricultural occupations are Census Occupation Codes 220 for Farmers and Farm Managers, 900 for Farm Workers, and 901 for Laborers.
7. Occupations not reported are Census Occupation Code 995.

Sources: U.S. Bureau of the Census, U.S. Census of Population: 1960. Detailed Characteristics. U.S. Summary. Final Report PC(1)-1D. U.S. Government Printing Office, Washington, D.C. Table 202, pp. 1-522-29.
Unpublished Jaffe "livelihood" codes for the 1960 U.S. Census listing of Detailed Occupations.

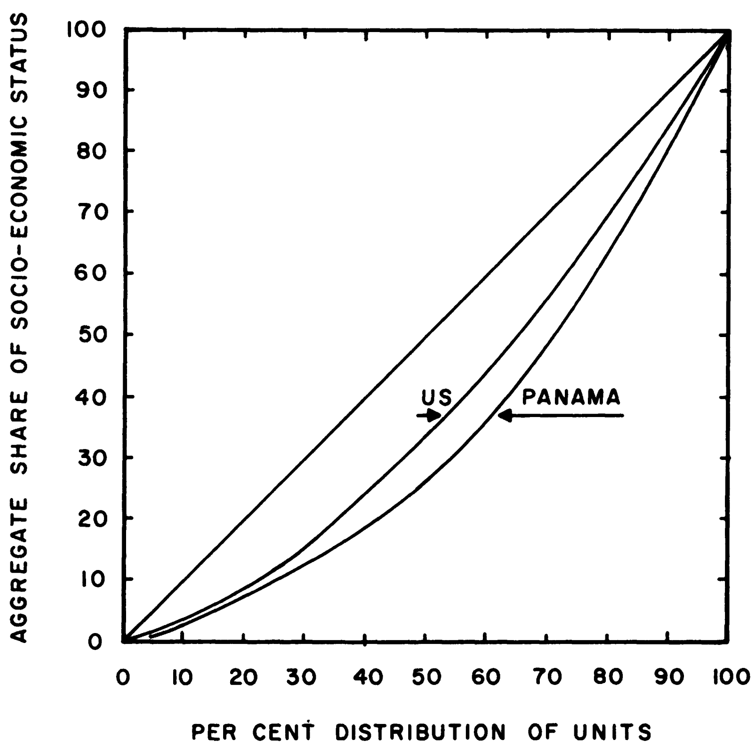


FIG. I. LORENZ CURVES OF SOCIO-ECONOMIC STATUS FOR THE UNITED STATES AND PANAMA

DEFINITION AND MEASUREMENT OF THE NONFARM SELF-EMPLOYED

Irving Leveson*
National Bureau of Economic Research

Self-employment encompasses many different occupations ranging from peddlers to physicians. It can exist under a wide variety of legal (and illegal) and institutional arrangements--sole proprietorships, partnerships, small owner-managed corporations, franchises etc. This diversity makes the definition and measurement of the self-employed difficult, and limits the clarity with which aggregate data on the self-employed can be interpreted. Measurement problems have hampered studies of many aspects of economic activity and limited analysis of the self-employed themselves.

Frequently held reservations about the quality of data on the number of proprietors stem in part from the elusive nature of the activities of persons without an established place of business such as peddlers and door to door salesmen, of persons whose places of business are in their own homes, and of seasonal businesses. Lack of agreement about conceptual problems of measurement is compounded upon differences between sources in treatment of the groups in question. A systematic examination of these problems is attempted here.

Definition of the Self-Employed

Many attributes of the work of the self-employed might be considered as the basis of a definition. The performance of entrepreneurial functions such as risk taking, autonomy in performing one's work, and ultimate authority within the firm are possible criteria. However, these characteristics do not clearly distinguish the independent proprietor from the salaried manager, director, or the stockholder. Such a distinction would be made if we counted as self-employed, persons who are substantially residual income recipients and are active in a business or profession.

This definition is preferred to a more general one which would include all persons "working for themselves." Persons who sell only their own labor are usually engaged in production directed by an economic unit--often a household and sometimes a firm--and not by themselves. The residual income definition is analytically convenient since under it the status of a person in self-employment corresponds with the existence of a firm. While there may be some ambiguity in classifying persons who work with small quantities of other factors of production, the count of the self-employed would probably not be seriously affected by choice of a point of demarcation.

Many persons receive large parts of their incomes in forms such as commissions and tips which, although not residuals, depend directly on the level of business activity and therefore are probably closely tied to profits. The individual can often have an immediate influence on the amount of such incomes. Such groups as independent craftsmen may also have greater incentives than if they worked for someone else. These persons may be termed the quasi-self-employed. It would be useful to know whether the behavior and characteristics of the quasi-self-employed are closely related to those of the self-employed, particularly since for those who desire to be paid on an incentive basis, the two forms of employment may be close substitutes. The definition of this group will not be considered here.

Many persons engaged in criminal activity could in principle be considered as self-employed or quasi-self-employed. The oldest profession is characterized by self-employment, much the same as many of the newer, less notorious ones. The number of self-employed criminals is quantitatively important. In 1965 there were three and a half million crimes against property. In the same year there were 1.2 million part-time self-employed and 577,000 wage and salary workers who were self-employed on secondary jobs according to the Current Population Survey. Criminal activity may be an important substitute for these legitimate activities. Here, however, we confine our attentions to those activities that are normally measured in the GNP accounts and regular employment series.

Current statistical practice is to classify persons according to their class of worker status in the major labor market activity of a single week. All of the number, hours, earnings, and characteristics of many multiple jobholders are attributed to their major employment in this way. This practice affects the classification of some

* The author wishes to thank the Ford Foundation, and the National Bureau of Economic Research, and Columbia University for their support. The findings reported herein have not yet undergone the full critical review accorded the National Bureau's studies and are therefore to be considered solely the views of the author. The comments and suggestions of Victor R. Fuchs and Jacob Mincer are greatly appreciated. Thanks are also due to Charlotte Boschan for her programming efforts.

information tabulated by occupation, industry and other characteristics, in addition to data on class of worker. If we wish to know the contribution to production of the self-employed or how individuals allocate their time (aside from the question of multiple job-holding), we do not wish to count the time of self-employed spent in wage and salary employment, while we would like to include the time spent in self-employment of wage and salary workers who are self-employed on secondary jobs. One solution to this problem is to classify portions of a person's hours, earnings, etc. according to the proportion of his time spent in each activity over a period of time. We would then obtain a number of full-time equivalent proprietors and a total manhours of proprietors which would correspond with entrepreneurial income as it is now being measured.[1]

Another question which arises is the treatment of proprietors of small, closely held corporations. Since they are similar in function and behavior to proprietors of unincorporated businesses, they should be included in the count of the self-employed for many problems of analysis. Interest in the unincorporated sector arises either when we wish to supplement existing information on the corporate sector to obtain a more complete picture of an industry or the economy, or when we are specifically concerned with legal status. For a great many questions such as understanding changes in opportunities in self-employment for disadvantaged groups or the role of small firms in the economy, attention should focus on all self-employed including those in small, closely-held corporations. It is therefore necessary to consider what effects inclusion of corporate proprietors would have on the number of self-employed.

Existing Estimates of the Number of Nonagricultural Self-Employed and Their Definitions

Sources

Before considering the quantitative aspects of the questions raised, let us briefly review the information available. The main features of some of the major sources of data on the number of self-employed are summarized in Table 1. All series exclude proprietors of corporations by definition. The sources vary greatly as to the frequency with which data appears and the amount of detail available. The Census of Population provides detailed information based on enumeration of households. The labor force concept of employment which it adopted beginning in 1940 is used by the other series. Earlier Census data applied to gainful workers.[2] The Current Population Survey (CPS) provides monthly data on a current basis for the postwar period from household surveys. Recently Stanley Lebergott has prepared annual estimates of the number of nonagricultural self-employed since 1900, based largely on Census of Population data, and designed to be consistent with the concepts and coverage of the CPS.[3]

Data on the number of proprietors of unincorporated business is collected approximately every five years for many service industries on employer reports in the Census of Business. Data for other industries appears in the Census of Manufactures, Census of Mineral Industries and the Annual Survey of Manufactures. The Office of Business Economics has estimated the number of self-employed since 1929, as a component of the series on the number of persons engaged in production, relying heavily on these censuses. Because the Census of Business benchmarks are so infrequent, the OBE uses information on the number of unincorporated business tax returns for interpolation and current extrapolation.

Official Definitions

The official definition in the Census of Population would appear to include many persons who are not residual income recipients by defining the self-employed as

> Persons who worked for profit or fees in their own business, profession, or trade, or who operated a farm either as an owner or tenant. Included here are the owner-operators of large stores and manufacturing establishments as well as small merchants, independent craftsmen and professional men, farmers, peddlers, and other persons who conducted enterprises of their own. Persons paid to manage businesses owned by other persons or by corporations, on the other hand, are classified as private wage and salary workers (or, in some few cases, as government workers).[4]

Class of worker status was ascertained by asking "Was this person self-employed in his own business, professional practice of farm?" Many construction craftsmen, visiting nurses and others who provide little in addition to their own labor services are counted as self-employed. The CPS follows the Census of Population criteria. Both classify persons by their major activity at a point in time.

The 1958 Census of Business questionnaire asks for information on the number of "Proprietors or partners who worked 15 hours or more during the one week ended nearest November 15, 1958." The hours limitation is intended to include only those devoting a major portion of their time to the business. This has the effect of reducing double counting of proprietors of more than one establishment. It is not clear which individuals within businesses are being classified as proprietors or partners, although collection of information on legal status does assure exclusion of corporate proprietors.

Corporate Self-Employed

The Number of Corporate Self-Employed

Now let us examine the effects of including the corporate self-employed. Very little direct information is available on the number of small, closely held corporations. In 1960 there were 90,000 Small Business Corporation income tax returns. These originated from operating domestic firms with fewer than ten shareholders which elected to be taxed as individuals. Another half million returns of a total of 1.1 million returns came from corporations which were more than 50 per cent owned by one stockholder.[5]

The size distribution of all corporations may give a better indication of how many small, closely held corporations there are. Table 2 indicates that over two-thirds of all corporations' tax returns had business receipts of less than $100,000 and five-sixths had less than $200,000 in 1962. The percentages that had value added below these are even higher. About a million corporate tax returns indicated business receipts of below $200,000 in 1962. While not all of these firms were closely held and the number of tax returns tends to overstate the number of businesses, some of these firms had more than one proprietor and some larger firms were undoubtedly closely held. This would suggest that the number of corporate proprietors is about 1 million.

Treatment of Corporate Proprietors in the Census of Population and CPS

The Census Bureau established an experimental panel, the Monthly Labor Survey (MLS), about one-fourth the size of the CPS, and the CPS Methods Test, a three area sample, to sharpen measurement of labor force activity. These surveys, which were discussed at the 1965 meetings, found that many corporate proprietors were reporting themselves as self-employed, contrary to the official definition, because information on legal status of their business was not being elicited. The Monthly Labor Survey (MLS), reported on by Robert L. Stein and Daniel B. Levine, showed that when the question on self-employment was followed by a question as to whether the business was incorporated, approximately one million fewer non-farm self-employed were indicated then by the CPS, which had no such question, for the first half of 1965. Similar results from the CPS Methods Test were presented by Joseph Waksberg and Robert B. Pearl. A follow-up mail questionnaire regarding legal form of the business, sent to those reporting initially as self-employed during April 1963-December 1964, indicated that the count of self-employed would be reduced by about 750,000 if corporate proprietors were excluded.[6] A comparison between the CPS and MLS for the full year 1966 based on a MLS sample increased to half the size of the CPS has been published. The difference there is also about 750,000.[7] These figures can be considered as estimates of the number of proprietors of small, closely held corporations. Their level is consistent with the number implied by the tax return data.

No direct information is available on the extent to which corporate proprietors report themselves as self-employed in different groups or over time. Some indication of the relative importance of corporate proprietors in different industries can be obtained, however, by comparing the CPS and MLS estimates. The CPS, which includes proprietors in corporations exceeded the MLS estimate by 39 per cent in mining and manufacturing, 5 per cent in construction, 6 per cent in transportation, communications and public utilities, and 18 per cent in trade and in services for the first half of 1966.

The Effect of Including Corporate Self-Employed on Recent Changes

The shifting of firms into corporate status limits the value of the reported number of self-employed as a measure of the total number including those in corporations in recent years. Table 3 compares the CPS and OBE number of self-employed in manufacturing, trade and services with the number of establishments and employees for 1954-1963. Changes in the number of establishments better reflect changes in the total number of self-employed since establishments of corporations are included. However, they probably tend to somewhat overstate the growth of the number of proprietors since they include increases in the number of establishments of large firms which are not associated with changes in the number of proprietors. Data from the CPS and OBE for the three industries combined show the number of proprietors to be growing much less rapidly than the number of establishments because of the growth of corporations. The similarity of changes in the CPS series to changes in the OBE number of proprietors suggests that the changes in CPS number of self-employed will not be significantly biased as a measure of changes in the number of proprietors of unincorporated businesses. This is helpful for gauging current changes since the CPS series is available monthly on a current basis and is not subject to large revisions like the OBE series.

Changes in the number of establishments are much more similar to changes in the number of employees than are changes in the total number of self-employed. The comparison of establishments and employees suggests that the percentage of employed persons in all nonagricultural industries who are self-employed, including corporate proprietors, declined slightly from 1954-1963 with all of the decline coming since 1958. The decline in the ratio of corporate proprietors to employees appears to have been much greater in trade than in services.

Beginning in January 1967 the CPS has excluded corporate proprietors by making use of information now requested on legal status from persons responding as self-employed. It would be useful if the questions on class of worker and legal status were tabulated so as to provide separate data on the number of corporate and noncorporate proprietors. This would also permit calculation of a current series which is comparable to the earlier one.

Some Problems With Existing Estimates

Differences Between the Census of Population and CPS

A number of other sources of difficulty exist with present estimates. According to the Gordon Committee Report, 970,000 more nonagricultural self-employed were indicated in the April 1960 CPS than in the 1960 Census of Population.[8] Table 4 compares the annual averages of the number of nonagricultural self-employed from the CPS with the April Census of Population figures for 1940, 1950 and 1960. The CPS show higher levels than the Census of Population in all industries and years. While the two sources show 1940-60 percentage changes in the total number of nonagricultural self-employed which are close, 15.6 per cent for the Census of Population compared to 18.1 per cent for the CPS, individual industries showed wide divergence. Differences between sources in the size of changes are not consistent in direction across industries or decades.

In the opinion of Census Bureau personnel, about half of the differences between the sources results from the Census of Population assumption that persons who do not report class of worker status are private wage and salary workers. Furthermore, the CPS elicits better reporting of self-employment in marginal enterprises, such as those conducted part-time in the home.[9]

Differences Between Establishment and Household Series

Substantial differences often exist between data from household series and information collected from employers. Establishment data counts multiple jobholders at each place of work. In recent years there have been about 400,000 wage and salary workers on their primary jobs who are self-employed on secondary jobs.[10] In addition an unknown number of proprietors are counted at more than one establishment. Another difference is that all corporate proprietors are excluded by the Census of Business because information on legal form of organization is made use of. Important differences may exist in the information on industry given by households and employers.[11]

I compared the number of self-employed in the 1960 Census of Population to the average number in the 1958 and 1963 Censuses of Business, with minor adjustment for the exclusion of small establishments in the latter, for 20 trade and service industries which could be easily matched. The unweighted mean ratio of the Census of Population to the Census of Business count was .97 and its standard deviation was .12. The total number of self-employed in the 20 industries was lower in the Census of Population by 5 per cent. The uneven effects of inclusion of corporate proprietors and classifying persons not reporting class of worker status as wage and salary workers in the Census of Population, and the double counting of multiple jobholders in the Census of Business, plus differences in industry classification explain the dispersion of the differences among industries.

Additional problems appear when we examine time series data. Chart 1 compares the number of self-employed as reported by the OBE and the CPS series with Lebergott's extension. Because of the smoothing introduced by interpolation between infrequent benchmarks, the OBE series is less sensitive to cyclical and war related changes. Furthermore, revisions have at times been large. The latest OBE revision cut the number of construction self-employed in half in recent years reducing the estimated number of proprietors in all nonagricultural industries by one-tenth.[12] After this change, the OBE data show about the same rise as the CPS since 1950, compared to the more rapid rise in the unrevised OBE series. The revision also eliminated the declines in the recession years of 1954 and 1958. The remaining difference in level comes largely in services. Another inconsistency is a large decline in the OBE construction estimates from 1948 to 1954 while the CPS showed a slight rise.

A large discrepancy in trend existed between the CPS and OBE for the period 1949 to 1952. Part of it was removed by the revision. Table 5 shows the size of the differences and effect of revisions. Part of the remaining differences may be because the OBE relies on Census of Business data for the month of November which, in the benchmark year of 1954, was about 200,000 above the annual average according to the CPS, while there was no strong seasonal pattern in other years. The impotance of eliminating such inconsistencies between "official" series can be appreciated with reference to current changes. If forces related to the war in Viet Nam produce differences of the kind that existed during the Korean War, our understanding of current developments would be most unsatisfactory.

Estimates of the Number of Self-Employed in 1960

Next I make use of information on the content of available series in order to derive estimates of the number of self-employed according to alternative concepts and sources in 1960. Estimates are presented for self-employed in unincorporated businesses and all self-employed, alternatively derived classifying persons by their major activity at a point in time and dividing their time between activities during the year, and alternatively based on OBE and CPS data. These estimates are presented in Tables 6 and 7.

The estimate of corporate proprietors was based on the findings of the expanded Monthly Labor Survey. Current Population Survey data permitted elimination of multiple jobholders. An estimate was made for the number of proprietors with negligible amounts of residual income. Large differences exist between the published numbers of self-employed and my estimates of the number in unincorporated businesses. The level in the published figures is about the same as the estimates of the total number of self-employed. The reconciliation has successfully eliminated the differences between sources. Estimates based on allocation of time are higher because more wage and salary workers have second jobs as self-employed than self-employed have secondary jobs as wage and salary workers. Either classifying at a point in time or over a period, and whichever source is used, it is estimated that the number of nonagricultural self-employed in unincorporated businesses was about 5 million in 1960 while the total number of self-employed was nearly 6 million.

Available data suggest that if similar estimates were made for 1950, we would observe the 1950-60 trend little affected by changes in the extent of multiple jobholding. The rate of growth in self-employment would be reduced by the exclusion of persons with small amounts of residual income, but the effect of including corporate proprietors would be large enough so that the total number of self-employed would show a greater rise than existing series.

Summary

The heterogeneity of the self-employed makes their definition and measurement difficult. It is maintained that the appropriate definition of a self-employed person for problems of economic analysis is one who is substantially a residual income recipient and is active in a business or profession. The number of self-employed is better defined as the average number over the year in full-time equivalents, than by classifying persons by their major activity and at a point in time. While we may wish to examine proprietors of unincorporated businesses in order to supplement existing information for the corporate sector, for most problems of analysis, proprietors of small, closely-held corporations should be included in the count of self-employed. In 1960 the number of nonagricultural self-employed in unincorporated businesses was about 5 million while the total number was nearly 6 million, according to the preceding criteria. Reductions in self-employment as a per cent of all employment have been greatly overstated in recent years as a result of the growing importance of the corporate form of organization.

Estimates of the number of nonagricultural self-employed differ widely between sources for many reasons. The Current Population Survey and Census of Population have been counting up to a million corporate proprietors as self-employed. However, the CPS began excluding proprietors of corporations from the count of the self-employed in January 1967. The Census of Population seriously undercounts marginal proprietors and classifies persons not reporting class of worker status as wage and salary workers. The OBE series on active proprietors of unincorporated enterprises double counts multiple jobholders and proprietors of more than one establishment since it is based on establishment reports. Before the latest OBE revisions, there were large differences between the CPS and OBE construction estimates which seriously affected the total for all industries. Trends in the CPS and OBE series have at times sharply diverged. Adjustment for major differences results in close reconciliation of the OBE and CPS series in 1960.

It is recommended that the use of a residual income criterion to define the self-employed be investigated, that information for persons with multiple activities be distributed among those activities, that separate data on the number of corporate proprietors be published by the CPS, and that the OBE publish more information on its estimating procedures.

Footnotes

1. Such a measure of the number of self-employed would not be sensitive to the undercount of those working few hours. This is a constant source of difficulty in obtaining consistency over time in the Census of Business on which the OBE series is based and is at least a potential source of error in the CPS. For a comparison of employment both at a point in time and over a period of time with comparable earnings data see Irving Leveson, "Nonfarm Self-Employment in the U.S.," unpublished Ph.D. dissertation, Columbia University, 1967, Chapter IV.

2. Briefly, the 1930 gainful worker group includes all persons who reported a gainful occupation, regardless of whether they were working or seeking work at the time of the census. The group includes all persons who usually worked at gainful labor, regardless of when they worked. The 1940 labor force on the other hand, was determined by activity during a particular period - the last week of March 1940 - and includes only persons who were working, or with a job, or seeking work in that week.

Alba M. Edwards, Sixteenth Census of the United States: 1940, Population, Comparative Occupation Statistics for the United States 1870 to 1940, Washington: U.S. Government Printing Office, 1943, p. 7.

3. Stanley Lebergott, Manpower in Economic Growth, New York: McGraw-Hill Book Company, 1964, Table A-4.

4. U.S. Bureau of the Census, Census of Population: 1960, Vol. 1, Characteristics of the Population, Part 1, United States Summary, Washington: U.S. Government Printing Office, 1964, p. LXXIII.

5. U.S. Internal Revenue Service, Statistics of Income...1960-61, Corporation Income Tax Returns, Tables 17 and 35.

6. Robert L. Stein and Daniel B. Levine, "Research in Labor Force Concepts," Proceedings of the Social Statistics Section of the American Statistical Association, 1965, pp. 218-26 and Joseph Waksberg and Robert B. Pearl, "New Methodological Research on Labor Force Measurements," Proceedings of the Social Statistics Section of the American Statistical Association, 1965, pp. 227-37. The CPS Methods Test applied to all industries, including agriculture. However, since there are relatively few corporations in agriculture, the results are very close to those that would be obtained for nonagricultural industries alone.

7. U.S. Bureau of Labor Statistics, "New Definitions for Employment and Unemployment," reprinted from Employment and Earnings and Monthly Report on the Labor Force, February 1967, Table 11.

8. U.S. President's Committee to Appraise Employment and Unemployment Statistics, Measuring Employment and Unemployment Statistics, Washington, U.S. Government Printing Office, 1962, Table J.4.

9. Robert B. Pearl, Chief, Demographic Surveys Division, U.S. Bureau of the Census, letter to the author, January 13, 1966. The assumption of private wage and salary worker status applies to persons who do not report employment status and are allocated to the labor force on the basis of the procedure used for allocation of unknown cases (about 3 per cent of the working age population) and those reporting in the labor force but with missing entries for class of worker, occupation and industry (about 5 per cent of persons reporting themselves in the labor force). The CPS-Census Match found that in April 1960 the CPS reported 536,000 nonagricultural self-employed who had reported being out of the labor force in the Census of Population compared to 202,000 persons indicating nonagricultural self-employment in the Census of Population who were classified as out of the labor force in the CPS. U.S. Bureau of the Census, Evaluation and Research Program of the U.S. Censuses of Population and Housing, 1960, Accuracy of Data on Population Characteristics as Measured by CPS-Census Match, Series ER 60, No. 5, Washington: U.S. Government Printing Office, 1965.

10. U.S. Bureau of Labor Statistics, Multiple Jobholders in December 1960, Special Labor Force Report, No. 18, Table 1.

11. These differences could arise for self-employed if wives respond in the household interviews but husbands respond in employer reports.

12. This change is supported by data from the 1/1000 sample which shows that approximately the amount of self-employment income in construction according to the OBE estimate was earned by the smaller number of persons. At the very least this demonstrates that the earlier OBE estimates of income of unincorporated enterprises and active proprietors were inconsistent with each other.

Table 1

Major Sources of Data on the Number of Nonagricultural Self-Employed

	Primary Sources			Secondary Sources	
	Census of Population	Current Population Survey	Census of Business	Lebergott	Office of Business Economics
Frequency	decennial	monthly	irregularly	annual	annual
Period covered	up to 1960	1947 to present	1929 to 1963	1900 to 1946	1929 to present
Reporting unit or primary source	households	households	establishments	CPS and Census of Population	Census of Business
Multiple jobholders double counted	no	no	yes	no	yes
Corporate proprietors excluded by definition	yes	yes	yes	yes	yes
Some corporate proprietors included in practice	yes	yes	no	yes	no
Coverage of marginal enterprises	poor	good	good	CPS level, poor for changes	good
Level of industry detail	detailed	major	detailed	major	detailed
Other detail available	detail on labor force and personal characteristics	some data on labor force and personal characteristics	location, firm size	none	none

TABLE 2

Number and Distribution of Tax Returns by Size of Business Receipts and Legal Form of Organization, Nonagricultural Industries, 1962

Size of Business Receipts	Sole Proprietor-ships	Partner-ships	All Un-incorporated Returns	Cor-pora-tions	All Returns
Number (thous.)					
Under $5,000	2,334	167	2,501	181	2,682
$5,000-$10,000	840	83	923	142	1,065
$10,000-$25,000	1,089	137	1,226	148	1,374
$25,000-$50,000	696	126	822	168	990
$50,000-$100,000	425	117	542	168	710
$100,000-$200,000	194	80	274	182	456
$200,000-$500,000	77	48	125	93	218
$500,000-$1,000,000	14	12	36	87	123
$1,000,000-$5,000,000	4	6	10	9	19
$5,000,000 or more	-	1	1	6	7
Total*	5,675	777	6,452	1,188	7,640
Percent Distribution					
Under $5,000	41.1%	21.5%	38.8%	15.2%	35.1%
$5,000-$10,000	14.8	10.7	14.3	12.0	13.9
$10,000-$25,000	19.2	17.6	19.0	12.5	18.0
$25,000-$50,000	12.3	16.2	12.7	14.1	13.0
$50,000-$100,000	7.5	15.1	8.4	14.1	9.3
$100,000-$200,000	3.4	10.3	4.2	15.3	6.0
$200,000-$500,000	1.4	6.2	1.9	7.8	2.9
$500,000-$1,000,000	.2	1.5	.6	7.3	1.6
$1,000,000-$5,000,000	.1	.8	.2	.8	.2
$5,000,000 or more	-	.1	-	.5	.1
Total*	100.0%	100.0%	100.0%	100.0%	100.0%
Percent under $100,000	94.9%	81.1%	93.2%	67.9%	89.3%

Source: U.S. Internal Revenue Service, Statistics of Income...1962, U.S. Business Tax Returns, Tables 6, 16 and 34.

*Excluding receipts not reported.

TABLE 3

Comparisons of Changes in Self-Employment With Changes in the Number of Establishments and Employees, Manufacturing, Trade and Services, 1954-63

		1954	1958 (thousands)	1963	Percentage Change 1954-58	1958-63	1954-63
CPS Self-Employed:	Manufacturing	424	408	362	- 3.8%	-11.3%	-14.6%
	Trade	2,481	2,449	2,203	- 1.3	-10.0	-11.2
	All Services	1,745	1,984	2,319	+13.7	+16.9	+32.9
	Total	4,650	4.841	4,884	+ 4.1%	+ .9%	+ 5.0%
OBE Self Employed:	Manufacturing	410	363	360	-11.5%	- .8%	-12.2%
	Trade (excluding automobile services)	2,299	2,440	2,219	+ 6.1	- 9.1	- 3.5
	Selected Services*	1,000	1,086	1,267	+ 8.6	+16.7	+26.7
	Total	3,709	3,889	3,846	+ 4.9%	- 1.1%	+ 3.7%
Establishments:	Manufacturing	286	298	312	+ 4.2%	+ 4.7%	+ 9.1%
	Trade	1,974	2,075	2,016	+ 5.1	- 2.8	+ 2.1
	Selected Services*	786	979	1,062	+24.6	+ 8.5	+35.1
	Total	3,046	3,352	3,390	+10.0%	+ 1.1%	+11.3%
Employees:	Manufacturing	16,099	16,035	16,962	- .4%	+ 5.8%	+ 5.4%
	Trade	9,679	10,750	11,499	+11.1	+ 7.0	+18.8
	Selected Services*	2,362	2,904	3,262	+22.9	+12.3	+38.1
	Total	28,140	29,689	31,723	+ 5.5%	+ 6.9%	+12.7%

Source: U.S. Bureau of the Census, Censuses of Business and Manufactures and U.S. Office of Business Economics.

*Excludes professional services and nonprofit organizations.

TABLE 4

Comparison of Levels and Changes in the Number of Self-Employed in Nonagricultural Industries in the Census of Population and Current Population Survey, by Industry, 1940-60

	1940	1950 (thousands)	1960	Percentage Change 1940-50	1950-60	1940-60
Census Total*	4,600	5,121	5,317	11.3	3.8	15.6
Manufacturing	269	403	366	49.8	-9.2	36.1
Construction	482	640	733	32.8	14.5	52.1
Trade	1,827	2,115	1,906	15.8	-9.9	4.3
Service	1,573	1,498	1,780	-4.8	18.8	13.2
Other	415	449	472	8.2	5.1	13.7
CPS Total	5,390	6,069	6,367	12.6	4.9	18.1
Manufacturing	324	407	383	25.6	-5.9	18.2
Construction	582	696	758	19.6	8.9	30.2
Trade	2,344	2,562	2,443	9.3	-4.6	4.2
Service	1,589	1,883	2,175	18.5	15.5	36.9
Other	551	521	608	-4.6	16.7	10.3
CPS minus Census:						
Total	790	948	1,050	1.3	2.1	2.5
Manufacturing	55	4	17	-24.2	3.3	17.9
Construction	100	56	25	-13.2	-5.6	-21.9
Trade	517	447	537	-6.5	5.3	-.1
Service	26	385	395	23.3	-3.3	23.7
Other	136	72	136	-12.8	11.6	-3.4

Note: CPS data are averages of monthly figures.

Source: U.S. Bureau of the Census, Census of Population, 1940: Industrial Characteristics, Table 6, 1950: Industrial Characteristics, Table 8, 1960: U.S. Summary, Detailed Characteristics, Table 6, Stanley Lebergott, Manpower in Economic Growth, New York: McGraw-Hill, Inc., 1964, Table A-7.

*Includes industry not reported.

TABLE 5

Comparison of the CPS Number of Self-Employed With the OBE Number Before and After Revision, Nonagricultural Industries, 1949-52
(thousands)

Year	CPS	OBE, Unrevised	OBE, Revised
1949	6,208	5,662	5,654
1950	6,069	5,729	5,721
1951	5,869	5,901	5,782
1952	5,670	6,023	5,822

Source: CPS and revised OBE data from Table A-1. Unrevised data from U.S. Income and Output Tables VI-13 and VI-16.

TABLE 6

Alternative Estimates of the Number of Self-Employed in Unincorporated Businesses and the Total Number of Self-Employed, Measured by Classifying Persons by Their Major Activity at a Point in Time, Nonagricultural Industries, 1960
(thousands)

	Based on OBE	Based on CPS
1. Number of self-employed as published	5,941	6,367
2. less corporate self-employed		-743
3. less wage and salary workers who are self-employed on secondary jobs	-399	
4. less reported self-employed with negligible residual income	-600	-700
5. Number of self-employed in unincorporated businesses	4,942	4,924
6. plus corporate self-employed	743	743
7. Total number of self-employed	5,685	5,667

Sources: Line

1 U.S. Office of Business Economics, The National Income and Product Accounts of the United States, 1929-1965, Statistical Tables, Tables 6.4 and 6.6, and U. S. Bureau of Labor Statistics, Labor Force and Employment in 1960, Special Labor Force Report No. 14, Table C-4.

2 U.S. Bureau of Labor Statistics, "New Definitions of Employment and Unemployment," reprinted from Employment and Earnings and Monthly Report on the Labor Force, February, 1967, Table 11.

3 U.S. Bureau of Labor Statistics, Multiple Jobholders in December, 1960, Special Labor Force Report No. 18, Table 1.

4 Sum of selected occupations including clerical and kindred workers, carpenters, electricians, operatives, clergymen, nurses, waiters, insurance agents and dressmakers outside of factories. A smaller number was used for the OBE data because some part-time proprietors are excluded.

5 Line 1 minus lines 2-4.

6 Same as line 2.

7 Line 5 plus line 6.

TABLE 7

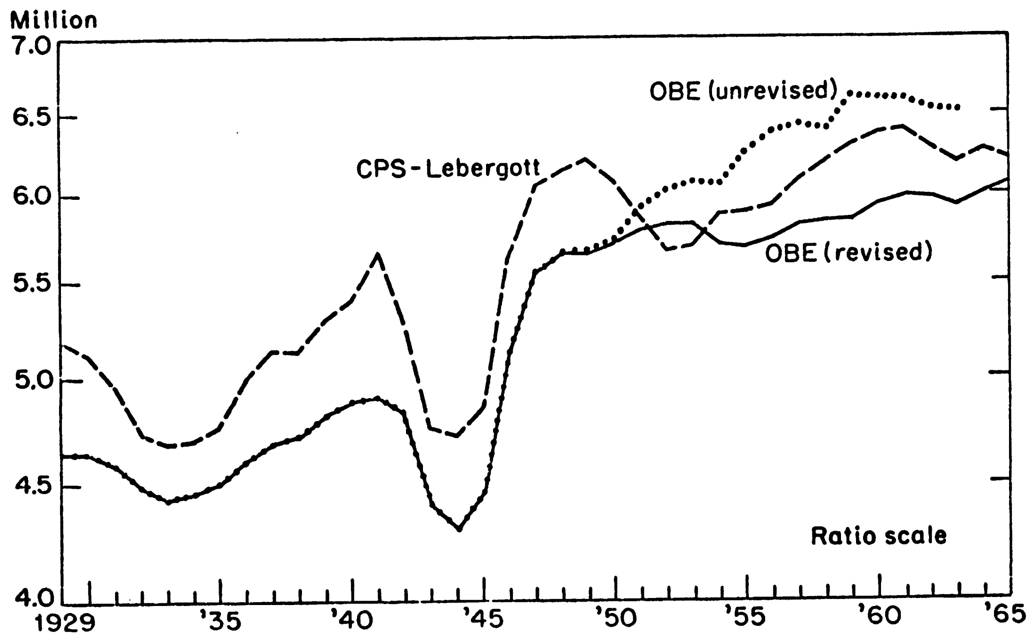
Alternative Estimates of the Number of Self-Employed in Unincorporated Businesses and the Total Number of Self-Employed, Measured by Classifying the Portion of a Person's Time Spent in Each Activity During a Period in that Activity, Nonagricultural Industries, 1960
(thousands)

Line		Based on OBE	Based on CPS
1	Number of self-employed as published	5,941	6,367
2	less corporate self-employed		-743
3	plus adjustment of wage and salary workers with secondary jobs as self-employed to full-time equivalent self-employed	-200	80
4	less adjustment of self-employed with secondary jobs as wage and salary workers to full-time equivalent self-employed	-36	-36
5	less reported self-employed with negligible residual income	-600	-700
6	Number of self-employed in unincorporated businesses	5,105	4,968
7	plus corporate self-employed	-743	743
8	Total number of self-employed	5,848	5,711

Sources: line

1 and 2	Same as Table 6, lines 1 and 2.
3	80 per cent of Table 6, line 3 with an extra adjustment for the OBE series because some part-time proprietors are excluded; 20 per cent for CPS.
4	Same as Table 6, line 3. 80 per cent of the hours of proprietors with secondary jobs were estimated to be in self-employment.
5	Same as Table 6, line 4.
6	Line 1 minus line 2 plus line 3 minus lines 4 and 5.
7	Same as line 2.
8	Line 6 plus line 7.

Chart 1
OBE and CPS-Lebergott Number of Nonagricultural Self-Employed, 1929-65



Source: Stanley Lebergott, <u>Manpower in Economic Growth</u> New York: McGraw-Hill Book Company, 1964, Table A-3, U.S. President, <u>Manpower Report of the President, March, 1966</u>, Table A-9; U.S. Office of Business Economics, <u>U.S. Income and Output</u>, Tables VI-13 and VI-16, and <u>SCB</u>, July 1962 and July 1964; U.S. Office of Business Economics, <u>The National Income and Product Accounts of the United States, 1929-1965, Statistical Tables</u>, Tables 6.4 and 6.6

AGGREGATE INCOME AND INCOME SIZE DISTRIBUTION ESTIMATES FOR SELECTED STANDARD METROPOLITAN STATISTICAL AREAS

Mitsuo Ono, George Patterson, and Emmett Spiers*
U.S. Bureau of the Census

The United States Decennial Population Census provides income data once every 10 years for areas such as counties and Standard Metropolitan Statistical Areas (SMSA's). There has been a steadily growing demand for current income data (Census basis) for these areas. Accordingly, a research program has been initiated to develop estimates of aggregate income and income size distribution using available Administrative Record sources, such as Internal Revenue Service (IRS) tax base information. Preliminary research findings are included in this paper which consists of three parts. The first part analyzes estimation procedures used to obtain aggregate income levels for 17 SMSA's. The second part describes a technique using lognormal probability graph paper to derive income size distribution data. The third part outlines directions for further research.

Derivation of Aggregate Income Estimates

Table I presents estimates of aggregate income (Census basis) for 17 large SMSA's for income years 1963 and 1965, respectively. These SMSA's were limited to areas which do not cross State boundaries and which had no changes in area definition since 1959. Two methods, designated I and II, were used to develop alternative estimates. These procedures are outlined in the Methodological Appendix. In essence, method I relates SMSA Census data with IRS published adjusted gross income (AGI) information and personal income data published by the Office of Business Economics. Method II relates SMSA Census data with personal income data only. Moreover, method II implicitly assumes a constant SMSA/State per capita income ratio over time.

For 1963 and 1965, respectively, 10 and 11 of the 17 SMSA estimates developed by using the two alternative methods showed a net percentage difference of less than 5 percent. In 1965, the following SMSA's had net percentage differences of greater than 5 percent: Atlanta, Georgia; Baltimore, Maryland; Miami, Florida; Newark, New Jersey; Pittsburgh, Pennsylvania; and San Diego, California. An analysis of these differences revealed that in all of these areas, the SMSA/State ratio of per capita AGI indicated decreases (from 6 percent in Newark SMSA to 16 percent in Baltimore SMSA) from 1959 to 1965 (see table II). These net percentage differences are related directly to the assumed consistency over time of the SMSA/State per capita income ratio.1/

As a check on the estimate, aggregate income totals from the Current Population Survey (CPS) for three SMSA's were computed for 1963 and 1965. It should be noted that CPS and Census data are not strictly comparable since the former covers the noninstitutional population only (see the Methodological Appendix for detailed definition). As shown in table III of the six estimates developed by method I for 1963 and 1965, five showed net percentage differences of less than 10 percent (between CPS data and method I estimates).

Derivation of SMSA Income Size Distribution

This procedure uses lognormal probability graph paper. In essence, this method attempts to capitalize on three empirical findings: (1) for larger population areas, income size distribution usually does not change rapidly over time, (2) State and SMSA income size distributions are generally similar and (3) income size distributions of joint tax returns and family units are fairly uniform. Essentially, the following steps are involved in this graphic technique:

1. IRS published information shows biennially AGI distribution by all returns and joint returns for SMSA's. As noted above, joint returns are assumed to follow distribution of families. Individual returns are obtained by subtracting joint returns from all returns. Individual returns are assumed to follow the distribution of unrelated individuals.

2. For income year 1959 the income size distributions of families and joint returns for the State are plotted on lognormal probability paper. The 1959 family income distribution for the SMSA is also plotted. For income year 1965, the income size distribution of joint returns for the State is plotted. These plotted distributions are examined for uniform curve types (including position and shape). The same procedure is followed for SMSA data. (For 1959, SMSA joint returns distribution is not published but 1961 data may be used.)

3. The general consistency of these plotted distributions for the State and for the SMSA, covering income years 1959, 1961, and 1965 are examined. If these distributions show similar curve types, then it is assumed that the 1965 joint return distribution can be used to estimate 1965 family distribution. An analysis is made to determine whether the SMSA and State joint return distributions have shifted proportionately over time. The percentage increases in median income for joint returns between 1961 and 1965 for the SMSA and for the State are obtained. If these rates of change are similar then the percentage increase in the median income of joint returns for the State between 1959 and 1965 are applied to the 1959 SMSA median family income to obtain an extrapolated SMSA median family income for 1965. Based upon this extrapolated median figure, 1965 family distribution data for the SMSA are plotted following that of the joint return distribution for the SMSA.

* Suggestions by Dr. Murray Weitzman are gratefully acknowledged. Comments represent views of the authors and not necessarily those of the Bureau of the Census.

Table I.--COMPARISON OF 1963 AND 1965 TOTAL MONEY INCOME (CENSUS BASIS) FOR SELECTED SMSA'S--METHODS I AND II

SMSA	1963 Method I (million dollars) (1)	1963 Method II (million dollars) (2)	1963 Net percent difference Col. (1)/Col. (2)	1965 Method I (million dollars) (1)	1965 Method II (million dollars) (2)	1965 Net percent difference Col. (1)/Col. (2)
Atlanta, Ga.	2,536.4	2,587.5	-2.0	2,981.5	3,227.5	-7.6
Baltimore, Md.	3,800.7	4,225.1	-10.0	4,303.1	5,144.9	-16.4
Buffalo, N.Y.	2,759.1	2,984.1	-7.5	3,176.5	3,277.3	-3.1
Chicago, Ill.	17,611.7	18,156.7	-3.0	19,872.6	20,708.2	-4.0
Dallas, Texas	2,860.4	2,912.7	-1.8	3,391.5	3,413.2	-0.6
Denver, Colo.	2,558.8	2,635.3	-2.9	2,801.4	2,870.0	-2.4
Detroit, Mich.	9,179.2	9,612.9	-4.5	11,428.3	11,861.0	-3.6
Miami, Fla.	2,103.0	2,338.9	-10.1	2,393.1	2,709.4	-11.7
Minn.-St. Paul, Minn.	4,108.6	4,059.6	+1.2	4,529.7	4,661.2	-2.8
New York, N.Y.	31,106.2	30,327.9	+2.6	34,295.8	33,833.3	+1.4
Newark, N.J.	5,076.8	5,081.9	-0.1	5,381.5	5,754.4	-6.5
Paterson-Clifton-Passaic, N.J.	3,828.6	3,550.7	+7.8	3,988.4	4,049.7	-1.5
Pittsburgh, Pa.	4,843.6	5,104.5	-5.1	4,945.0	5,745.2	-13.9
San Bernardino-Riverside-Ontario, Calif.	2,096.5	2,085.7	+0.5	2,397.7	2,405.3	-0.3
San Diego, Calif.	2,282.9	2,784.8	-18.0	2,658.3	3,068.9	-13.4
Seattle-Everett, Wash.	3,326.0	3,095.6	+7.4	3,452.8	3,471.1	-0.5
Tampa-St. Petersburg, Fla.	1,673.8	1,670.1	+0.2	1,912.0	1,942.4	-1.6

Table II.--RELATIONSHIP OF TOTAL INCOME--NET PERCENT DIFFERENCES TO SMSA/STATE PER CAPITA AGI RATIO

SMSA		Aggregate income: Net percent difference between Methods I and II (1)	SMSA per capita AGI (2)	State per capita AGI (3)	Ratio SMSA/State per capita AGI Col.(2)/Col.(3) (4)	Net percent difference 1965/1959 Col.(4) (5)
Atlanta, Ga.	1959		1,708	1,145	1.49	
	1965	-7.6	2,348	1,703	1.38	-7.4
Baltimore, Md.	1959		1,915	1,982	0.96	
	1965	-16.4	2,134	2,638	0.81	-15.6
Buffalo, N.Y.	1959		1,997	2,207	0.91	
	1965	-3.1	2,328	2,689	0.87	-4.4
Chicago, Ill.	1959		2,349	2,076	1.13	
	1965	-4.0	2,883	2,656	1.09	-3.5
Dallas, Tex.	1959		1,879	1,402	1.34	
	1965	-0.6	2,361	1,772	1.33	-0.7
Denver, Colo.	1959		1,978	1,692	1.17	
	1965	-2.4	2,450	2,148	1.14	-2.6
Detroit, Mich.	1959		2,044	1,817	1.12	
	1965	-3.6	2,699	2,486	1.09	-2.7
Miami, Fla.	1959		1,768	1,424	1.24	
	1965	-11.7	2,072	1,894	1.09	-12.1
Minn.-St. Paul, Minn.	1959		1,989	1,558	1.28	
	1965	-2.8	2,590	2,087	1.24	-3.1
New York, N.Y.	1959		2,423	2,207	1.10	
	1965	+1.4	2,996	2,089	1.11	+0.9
Newark, N.J.	1959		2,370	2,121	1.12	
	1965	-6.5	2,839	2,716	1.05	-6.2
Paterson-Clifton-Passaic, N.J.	1959		2,192	2,121	1.03	
	1965	-1.5	2,761	2,716	1.02	-1.0
Pittsburgh, Pa.	1959		1,926	1,785	1.08	
	1965	-13.9	2,130	2,300	0.93	-13.9
San Bernardino-Riverside, Calif.	1959		1,603	2,104	0.76	
	1965	-0.3	1,944	2,578	0.75	-1.3
San Diego, Calif.	1959		1,876	2,104	0.89	
	1965	-13.4	2,053	2,578	0.80	-10.1
Seattle-Everett, Wash.	1959		2,072	1,837	1.13	
	1965	-0.5	2,616	2,328	1.12	-0.9
Tampa-St. Petersburg, Fla.	1959		1,405	1,424	0.99	
	1965	-1.6	1,836	1,894	0.97	-2.0

Table III.--COMPARISON OF AGGREGATE TOTAL MONEY INCOME OBTAINED FROM CURRENT POPULATION SURVEY, METHODS I AND II (CENSUS BASIS), AND IRS STATISTICS OF INCOME SELECTED SMSA'S--1963/1965

(In million dollars)

SMSA	CPS aggregate income (1)	Aggregate income Method I (2)	Aggregate income Method II (3)	Adjusted gross income (4)	Net percent difference (2)/(1) (5)	Net percent difference (3)/(1) (6)	Net percent difference (4)/(1) (7)
1963							
Chicago, Ill............	16,688.3	17,611.7	18,156.7	16,574.6	+5.5	+8.8	-0.7
Detroit, Mich...........	8,646.3	9,179.2	9,612.9	8,782.3	+6.2	+11.2	+1.6
Pittsburgh, Pa..........	5,325.1	4,843.6	5,104.5	4,807.0	-9.0	-4.1	-9.7
1965							
Chicago, Ill............	18,128.5	19,872.6	20,814.9	19,377.9	+9.6	+14.8	+6.9
Detroit, Mich...........	10,497.2	11,428.3	11,550.9	11,138.4	+8.9	+10.0	+6.1
Pittsburgh, Pa..........	5,609.5	4,945.0	5,718.3	5,064.6	-11.8	+1.9	-9.7

Source: (1) March 1964 and 1966 CPS Consumer Income Tabulations--Families.
(2) Statistics of Income--Individual Income Tax Returns, Internal Revenue Service, 1963 and 1965.

4. The same steps can be followed to obtain estimated 1965 distribution data for unrelated individuals from individual returns. This general procedure was used except for few modifications to obtain data for Denver, Colorado, one of the 11 SMSA's in 1965 which showed a net percentage difference in aggregate income of less than 5 percent using methods I and II. In the graphic analysis for Denver, the 1965 Census family distribution was estimated using the 1965 joint return distribution for Denver after extrapolating the 1959 median Census family income. The rate of increase was based upon changes in median values of joint returns for the State, between 1959 and 1965. This is described in more detail in the Methodological Appendix.

Directions for Further Research

Empirical findings have revealed that distributions of wage and salary income obtained from both IRS and Census data are generally similar. Differences between IRS and Census income size distribution data can be related to three relationships: (1) whether or not the distribution of income types other than wage and salary income changes over time in proportion to distribution of wage and salary income, (2) whether or not the IRS tax return population changes over time in proportion to the Census consumer unit population, and (3) whether or not the area-State data ratio remains consistent or changes over time. A major task is to investigate further the extent of the applicability of these "proportionality" assumptions to areas other than those covered in this paper.

The data presented herewith are considered first estimates. These data are being compared with other available data sources, e.g., published State tax data, in order that the adequacy of these data can be evaluated further.

Other estimation methods are also being developed. For example, under certain conditions, a simple extrapolation of 1959 SMSA Census data using rates of increase developed from AGI data for the SMSA results in similar aggregates obtained from using method I.

The estimation procedure on income size distribution involves a subjective analysis of the position and shape of the curve types. It is planned to formalize these procedures in an estimation model so that information can be analyzed more objectively.2/

Footnotes

1. A detailed explanation of the relationship between methods I and II is shown in Section III of the Methodological Appendix.

2. IRS tends to report fewer low income families than Census in their respective income size distributions. Income tax data normally not cover persons receiving less than $600 annual income and exclude certain transfer payments, e.g., Social Security, unemployment compensation, etc. and "occupational" payments. Consequently, some means must be found to adjust for these differences in coverage. One possible method is to develop mathematical relationships whereby Census data are functionally related to IRS data at decile values.

METHODOLOGICAL APPENDIX

I. DEFINITIONS

Standard Metropolitan Statistical Areas.--Except in New England, a Standard Metropolitan Statistical Area is a county or group of contiguous counties which contains at least one city of 50,000 inhabitants or more, or "twin cities" with a combined population of at least 50,000. In addition to the county, or counties, containing such a city or cities, contiguous counties are included in a SMSA if, according to certain criteria, they are essentially metropolitan in character and are socially and economically integrated with the central city. In New England, SMSA's consist of towns and cities, rather than counties.

Between 1959 and 1965, the range of income years compared in this study, there were amendments to the definition of many SMSA's throughout the United States. These definitional changes were generally additions or deletions of towns and/or counties. A listing of the amended areas is located in Part IV, pages 45-52 in Standard Metropolitan Statistical Areas, 1967, prepared by the Office of Statistical Standards, Bureau of the Budget.

Income Reported in the Census and Current Population Survey (CPS).--In the 1960 Census, total income was the sum of money received by persons 14 years old and over from wages or salaries, net income (or loss) from self-employment, and income other than earnings.

Income from wages or salary was the total money earnings received for work performed as an employee, including wages, salary, pay from Armed Forces, commissions, tips, piece-rate payments, and cash bonuses earned.

Self-employment income was net money income (gross receipts minus operating expenses) obtained from a business, farm, or professional enterprise in which the person was engaged on his own account or as an unincorporated employer. Gross receipts included the value of all goods sold and services rendered. Expenses included the costs of goods purchased, rent, heat, light, power, depreciation charges, wages and salaries paid, business taxes, etc.

Income other than earnings was money income received from sources other than wages or salary and self-employment, such as net income (or loss) from rents or receipts from roomers or boarders; royalties; interest, dividends, and periodic income from estates and trust funds; Social Security benefits; pensions; veterans' payments, military allotments for dependents, unemployment insurance, and public assistance or other governmental payments; and periodic contributions for support from persons who were not members of the household, alimony, and periodic receipts from insurance policies or annuities.

Receipt from the following sources were not included as income: money received from the sale of property, unless the recipient was engaged in the business of selling such property; the value of income "in kind," such as free living quarters or food produced and consumed in the home; withdrawals of bank deposits; money borrowed; tax refunds; gifts and lump-sum inheritances or insurance benefits. Further information is found in Consumer Income, Series P-60, No. 51, Current Population Reports.

Adjusted Gross Income (Less Adjusted Gross Deficit).--Adjusted gross income was gross income from all sources that are subject to income tax minus (1) ordinary and necessary expenses of operating a trade or business, (2) expense deductions attributable to rents and royalties, (3) expenses of outside salesmen attributable to earning salary or other compensation, (4) expenses of travel, meals, and lodging while away from home overnight paid by an employee with respect to services rendered, (5) transportation cost related to the performance of services as an employee, (6) expenses for education required to maintain salary, status, or present employment, (7) expenses paid or incurred in connection with service as an employee under a reimbursed or other expense allowance arrangement with the employer, (8) exclusion of allowable sick pay if the sick pay was included in gross salary, (9) depreciation and depletion allowed life tenants and income beneficiaries of property held in trust, (10) deductible losses from sales of capital assets, and other property, (11) deduction equal to 50 percent of this excess of net long-term gain over net short-term capital loss, (12) net operating loss deduction, (13) contributions to a retirement fund by the self-employed, (14) reasonable expenses incurred in moving from old residence to new residence at new place of employment, and (15) any other deductions or exclusions from gross income.

Deficits in adjusted gross income occur when deductions allowed for the computation of adjusted gross income, as stated above, exceeded the gross income. See Statistics of Income - 1965, Individual Income Tax Returns, Internal Revenue Service, for further details.

The definition of adjusted gross income is subject to amendment through changes in law. Items 13, 14, and 15 of the 1965 definition were not part of the 1959 definition.

OBE Personal Income.--The personal income totals developed by the Office of Business Economics include, among other items, the following types of nonmoney income which are not included in the Census definition: wages received in kind, the value of food and fuel produced and consumed on farms, the net rental value of owner-occupied homes, the property income received by mutual life insurance companies, and the value of the services of banks and other financial intermediaries rendered to persons without the assessment of specific charges. These items of income in kind account for about 5 percent of total personal income. The Census definition of income, on the other hand, includes such items as regular contributions for support received from persons who do not reside in the same living quarters, income received from roomers and boarders residing in households, and employer contributions for social insurance which are not included in the personal income series. These items, however, represent a much smaller income total than the nonmoney items included in OBE personal income. For further information, see pages 49-65, Personal Income by States Since 1929 prepared by the Office of Business Economics, Department of Commerce.

II. POPULATION COVERAGE

Office of Business Economics.--In general, the population data used in computing per capita personal income by States are the midyear (July 1 of income year) estimates published by the Bureau of the Census. These estimates exclude Federal civilian and military personnel stationed outside the continental United States.

Adjusted Gross Income (IRS Returns).--This population universe is composed of all persons reported on all 1965 individual income tax returns and all prior-year delinquent returns for each State and SMSA. The prior-year delinquent returns comprise less than 1 percent of all returns filed.

Population Reported by the Bureau of the Census.--The population data used in the calculation of per capita income in 1959 was the 1960 Census enumeration of total residents in each SMSA and State as of April 1, 1960. This enumeration included all civilian institutional and noninstitutional, as well as the total military population living in the SMSA. For 1963 and 1965, the population of each area was extrapolated for April 1, 1964 and 1966, respectively, from population estimates furnished in the Series P-25 reports. Differences in population coverage between independent Census estimates and CPS totals are explained later in this appendix.

Per Capita Income.--The per capita income is derived by dividing the total income of the SMSA by the population in the SMSA. This description is applicable to both the Census and OBE income series as described above.

III. ESTIMATING METHOD - AGGREGATE INCOME

A. Method I.--This method of SMSA income estimation assumes that a proportional relationship exists in the same period in time and over time between Census income and Internal Revenue Service Adjusted Gross Income (AGI) data. It assumes also that Census per capita income increases over time at the same rate as OBE per capita income. This ratio-estimation method involves a three-stage procedure and is described in terms of estimating SMSA income for 1965. The first step involves an adjustment of the July 1, 1966 State population estimates to conform with the estimate as of the Census date of April 1, 1966. The second step requires obtaining aggregate income (Census basis) for the State in which the SMSA is located. This is accomplished through multiplication of the 1965 per capita income for the State by the estimated State resident population (as of April 1, 1966). The State 1959 per capita income is extrapolated by a ratio of change calculated from OBE per capita income for 1959 and 1965. The final step involves obtaining the proportion of SMSA income to the State total income (as derived above). A ratio-estimation formula relating 1959 and 1965 Internal Revenue Service Adjusted Gross Income (AGI) data with Census data is employed.

B. Method II.--As in method I, three steps are also required to estimate SMSA income by this method. In the first step, the 1965 midyear SMSA resident population is adjusted as of the "Census" date of April 1, 1966. In the second step, the 1965 per capita income (Census basis) is calculated for the State and SMSA. The SMSA per capita income is derived from use of the relationship between 1959 State and SMSA per capita income. This method of income estimation assumes that the State and SMSA per capita income (Census basis) relationship remains stable over time. It assumes also, as in method I, that State per capita income (Census basis) increases over time at the same rate as State per capita (OBE) personal income. The final step involves the estimation of 1965 SMSA aggregate income. The SMSA resident population estimated as of April 1, 1966 is multiplied by the 1965 SMSA per capita income (Census basis).

The above procedures with appropriate population and income data adjustments were employed also in the estimation of corresponding 1963 aggregates shown in tables I and III.

See illustrations of methods I and II on the following pages.

C. Method I vs. Method II.--The elements of information that are required to perform the calculations by each of these methods are compared in the diagram below. All cells that are marked with an "X" indicate that the particular element is a necessary part of the estimation process.

	Method I		Method II	
Element	State	SMSA	State	SMSA
Per capita personal income (OBE), 1959........	X		X	
Per capita personal income (OBE), 1965........	X		X	
Per capita income (Census), 1959...........	X	X		X
Census population, 4/1/60..	X	X		
Census population, 7/1/65..	X	X		X
Census population, 7/1/66..	X	X		X
Adjusted gross income (IRS), 1959...............	X	X		
Adjusted gross income (IRS), 1965...............	X	X		

It is evident from the above scheme that method II estimates can be calculated using data from the Bureau of the Census and the Office of Business Economics. In contrast, method I totals require published data from three governmental sources--Bureau of the Census, Internal Revenue Service (IRS), and the Office of Business Economics (OBE). Moreover, the adjusted gross income data furnished by IRS is only available biennially. This lack of information precludes the possibility of estimating income annually by method I, unless additional techniques such as interpolation are introduced.

A definite relationship exists between method I and II estimates based on the differences in the rate of growth from 1959 to 1965 of population and AGI estimates between the State and SMSA. Specifically, if:

$$R_I = \frac{\text{SMSA AGI,1965}}{\text{SMSA AGI,1959}} \bigg/ \frac{\text{State AGI,1965}}{\text{State AGI,1959}} \quad \text{and}$$

$$R_P = \frac{\text{SMSA population,1966}}{\text{SMSA population,1960}} \bigg/ \frac{\text{State population,1966}}{\text{State population,1960}}$$

then the method I estimate equals $\frac{R_I}{R_P}$ multiplied by the method II estimate.

Thus, if both State population and State AGI change over time at the same rates as for the SMSA, method I will equal method II. The relative ease of calculation and the availability of annual Census population estimates lend support to a preference for using method II. However, method I can be modified so that it requires data obtained only from the Bureau of the Census and the Internal Revenue Service. To find a modified method I estimate (designated as method III) multiply the 1959 SMSA aggregate income (Census basis) by the ratio of increase of the SMSA adjusted gross income (AGI) between 1959 and 1965.

$$\text{Thus: Method III} = \text{1959 SMSA aggregate income (Census basis)} \times \frac{\text{SMSA AGI,1965}}{\text{SMSA AGI,1959}}$$

Method III is related to method I by the following relationship:

$$\text{Method I} = \frac{\text{State personal income,1965}}{\text{State personal income,1959}} \times \frac{1}{\dfrac{\text{State AGI,1965}}{\text{State AGI,1959}}} \times \text{Method III}$$

Method III has the same ease of calculation as method II without having to make the assumption that SMSA income increases at the same rate as the State income.

Current Population Survey (CPS) Aggregate Income.--Income distributions for all families and for all unrelated individuals for selected SMSA's in the United States were tabulated from the March Current Population Survey. Estimates of total money income shown in table III were computed by

ILLUSTRATION OF METHOD I

	State of Colorado	Denver SMSA	SMSA Income / State Income
A. Adjustment of July 1, 1966 State Population Estimates to Conform with Census Base of April 1, 1966			
(1) July 1, 1966	1,955,000		
(2) July 1, 1965	1,949,000		
(a) Difference (Δ)	6,000		
(b) .75Δ	4,500		
(3) July 1, 1965	1,949,000		
(4) April 1, 1966 population (est.)	1,953,500		
B. Estimating State Per Capita Income (Census Basis)--1965			
(1) Per capita income (Census basis)--1959	1,889		
(2) Per capita personal income--1959	2,196		
(3) Per capita personal income--1965	2,706		
(4) Rate of change in per capita personal income--1965/1959	1.2322		
(5) Per capita income (Census basis) estimate--1965 (B1 X B4) (1,889 X 1.2322)	2,328		
C. Estimating Aggregate State Income--1965			
(1) Population (est.) April 1, 1966	1,953,500		
(2) Per capita income (Census basis) (est.)--1965	2,328		
(3) Aggregate income (Census basis) (est.)--1965 ($000)	4,547,748		
D. Estimating SMSA Aggregate Income			
(1) Per capita income (Census basis)--1959	1,889	2,167	
(2) Population, April 1, 1960	1,753,947	929,383	
(3) Aggregate income (Census basis)--1959 ($000)	3,313,206	2,013,973	.608
(4) IRS Adjusted Gross Income--1959 ($000)	2,967,757	1,838,082	.619
(5) IRS Adjusted Gross Income--1965 ($000)	4,196,165	2,632,842	.627
(6) Aggregate income (Census basis) (est.)--1965 ($000)	4,547,748		
(7) SMSA share based on 1959/1965, State/SMSA proportional relationship			.616 1/
(8) Denver SMSA aggregate income (Census basis)--1965 ($000)		2,801,413	

1/ $\frac{619}{627} = \frac{.608}{X}$

$X = .616$

ILLUSTRATION OF METHOD II

	Denver SMSA	State of Colorado	SMSA Income / State Income
A. Adjustment of July 1, 1965 SMSA Population Estimates1/ to Conform with Census Base of April 1, 1966			
(1) July 1, 1965	1,073,000		
(2) July 1, 1964	1,071,000		
(a) Difference (Δ)	2,000		
(b) 1.75 Δ	3,500		
(3) July 1, 1964	1,071,000		
(4) April 1, 1966 population (est.)	1,074,500		
B. Estimating State and SMSA Per Capita Income (Census Basis)			
(1) SMSA/State per capita income ratio--1959	2,167	1,889	1.1472
(2) Per capita personal income--1959		2,196	
(3) Per capita personal income--1965		2,706	
(4) Rate of change in per capita personal income--1965/1959		1.2322	
(5) Per capita income (Census basis) (est.)--1965 (1,889 X 1.2322)		2,328	
(6) Per capita income (Census basis) (est.)--1965 (2,328 X 1.1472)	2,671		
C. Estimating SMSA Aggregate Income (Census Basis)			
(1) Per capita income (Census basis)--1965	2,671		
(2) Population, April 1, 1966	1,074,500		
(3) Denver SMSA aggregate income--1965 / C(1) X C(2) / ($000)	2,869,990		

1/ July 1, 1966 population estimates are currently available for selected SMSA's. For these SMSA's, the resident population is calculated as three-quarters of the difference between July 1, 1966 and July 1, 1965 population totals published in the P-25 series reports (Bureau of the Census).

multiplying the frequencies in each income interval by the corresponding mean estimate and then summing these products. For the upper open-end class interval, the Pareto formula was used. These products were then aggregated into subtotals for families and for unrelated individuals. The two subtotals were combined into a grand total for families and unrelated individuals.

IV. GRAPHICAL ANALYSIS--INCOME SIZE DISTRIBUTION

Chart 1.--Joint Returns vs. Families, Denver SMSA and Colorado State, 1959 and 1965.

Family and joint return income size distribution data for the State in 1959 appears to show a uniform pattern. The shape of the State and SMSA curves also show a uniform pattern. The 1965 distribution was estimated as follows:

A. 1959 median income of joint returns for the State of Colorado was estimated at $5,800. The 1959 median income of families for the State was estimated at $5,800.

B. 1965 median income of joint returns for the State was estimated at $7,400. The net percentage increase between 1959 and 1965 for median income of joint returns was 27.6 percent. This rate of increase was used to extrapolate the 1959 Denver SMSA median family income figure ($6,600) to obtain the estimated 1965 SMSA median family income figure ($8,400). Based upon the 1965 median figure, the 1959 family income distribution curve was shifted upward to obtain the estimated SMSA 1965 family income distribution.

Chart 2.--Individual Returns vs. Unrelated Individuals, Denver SMSA and Colorado State, 1959 and 1965.

State 1959 income distributions of individual returns and unrelated individuals are not uniform. Consequently, the extrapolated 1965 income distribution of unrelated individuals is not as adequate as the family income distribution. It also appears that the relationship between income distributions for Denver SMSA and for the State is not consistent. One method that can be used to obtain preliminary estimates (subject to further review) of income size distribution of unrelated individuals is described below:

A. The rate of increase between 1961 and 1965 of the median income of IRS individual returns in the SMSA was obtained. Using a ratio-estimation formula, the rate of increase (between 1959 and 1965) for median income of individual returns for the SMSA was computed.

B. Using this rate of increase (between 1959 and 1965), the 1959 median income figure of unrelated individuals in the SMSA was extrapolated to 1965. The relationship between the 1959 distribution for unrelated individuals and the 1965 distribution for individual returns was examined. Since the curve types appear generally similar, estimated SMSA 1965 income distribution for unrelated individuals was plotted following the SMSA 1959 distribution for unrelated individuals.

V. LIMITATIONS

The SMSA's examined in this paper are limited to those which do not cross State boundaries. As can be seen in the illustrative computations for methods I and II, the total and per capita income relationships between the SMSA and the State which contains it underly these methods, in addition to the explicit relationship assumed among IRS, Census, and OBE income data.

Differences in Income Concepts and Population Coverages.--In addition to limitations noted previously, consideration must also be given to differences in the income definitions and population coverage employed in the estimation of Census, CPS, IRS, and OBE data in any comparative assessment of totals derived from the various sources. The income concepts used in Census and CPS are basically the same. However, differences do exist in the population coverage. The CPS excludes the institutional population and most members of the Armed Forces living on post. These two groups were included in the Population Census. Secondly, college students are generally enumerated at their own home in the Current Population Survey and classified as family members, but in the Census they were enumerated at their college residence, usually as secondary individuals.

The IRS income data are not directly comparable with those obtained from Census or CPS. Income, as defined for tax purposes, differs from the concept employed by the Bureau of the Census. For example, certain types of receipts such as veterans' payments, Social Security benefits, and relief payments, which constitute the main income source for some families, are excluded from income tax coverage. Moreover, the coverage of income tax statistics is less inclusive because persons receiving less than $600 (less than $1,200, if 65 years old and over) are not required to file returns.

Both the CPS and IRS income totals are based on sample data, and, as such, are subject to sampling variability. For additional information on the CPS sample selection, see Technical Paper No. 7, The Current Population Survey - A Report on Methodology, U.S. Department of Commerce, Bureau of the Census. Comparable background information on the IRS sample of tax returns is available in the Statistics of Income - 1965, Individual Income Tax Returns, Internal Revenue Service.

Chart 1.--JOINT RETURNS VS. FAMILIES

DENVER SMSA AND COLORADO, 1959, 1965

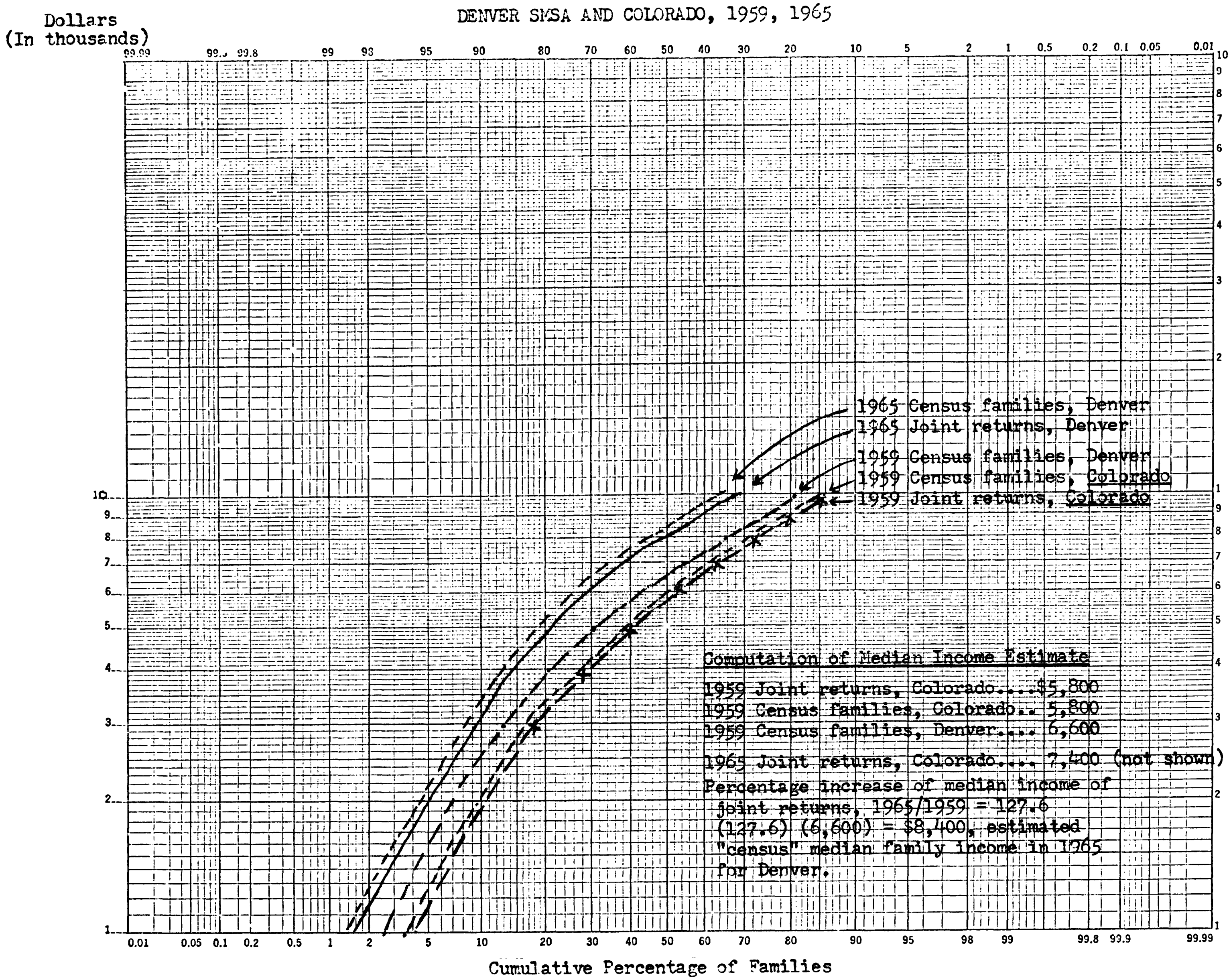
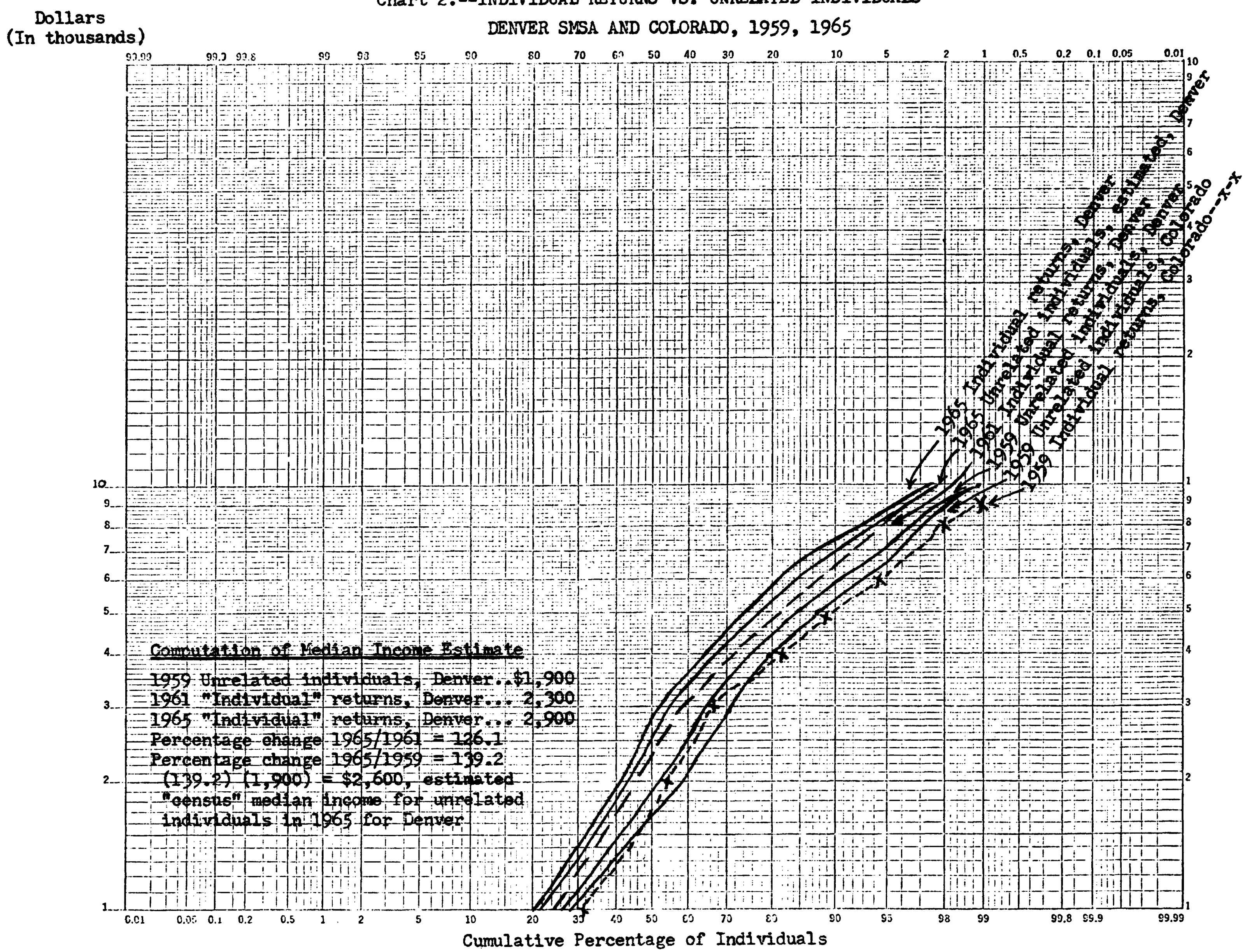


Chart 2.--INDIVIDUAL RETURNS VS. UNRELATED INDIVIDUALS
DENVER SMSA AND COLORADO, 1959, 1965



POVERTY IN THE RURAL POPULATION OF THE UNITED STATES*

Gordon F. Sutton
Office of Research, Plans, Programs and Evaluation
Office of Economic Opportunity

Concentration of Poverty in Rural Areas

Table 1 shows the total population and the poverty population of the United States in 1960 according to color and urban-rural residence. It is clearly apparent that poverty is found disproportionately among the rural population. While the nonwhite population is somewhat more urbanized than white population, the color differential for the poverty population is even more pronounced. In any case the prevalence of poverty among whites in urban areas is only three-fourths that of all whites and the prevalence among the urban nonwhite population is just over four-fifths that of the total nonwhite population.

The differential prevalence of poverty among urban and rural populations is a source of considerable concern among administrators who, in weighing program allocation and program impact, find more effective anti-poverty program alternatives in urban areas. However, it may lessen the burdens of such administrators to consider that students of demography have for some time been interested in alternative descriptions of rural and urban population which may be more sensitive to issues which are likely to plague policy makers, whether dealing with poverty or with a number of other matters.

TABLE 1. TOTAL POPULATION AND POPULATION IN POVERTY BY COLOR, URBAN AND RURAL, UNITED STATES, 1960 (CIVILIAN NONINSTITUTIONAL POPULATION...) (In Thousands)

Poverty status and color	Total	Urban	Rural		
			Total	Nonfarm	Farm
All Classes					
Total	175,035	122,287	52,748	39,206	13,542
White	155,206	107,860	47,346	35,408	11,938
Nonwhite	19,829	14,427	5,402	3,798	1,604
Poverty Class					
Total	38,684	21,294	17,390	12,471	4,919
White	27,719	14,583	13,136	9,519	3,617
Nonwhite	10,965	6,711	4,254	2,952	1,302
Percent Distribution					
All Classes					
Total	100.0	69.9	30.1	22.4	7.7
White	100.0	69.5	30.5	22.8	7.7
Nonwhite	100.0	72.8	27.2	19.1	8.1
Poverty Class					
Total	100.0	55.0	45.0	32.2	12.8
White	100.0	52.6	47.4	34.4	13.0
Nonwhite	100.0	61.2	38.8	26.9	11.9

*Prepared for presentation at the December 1967 meetings of the American Statistical Association.

Deriving Population Geography for Issues Relevant to Program Application

In the administration of anti-poverty programs, the fundamental distinctions between urban and rural program concepts are related to issues such as accessibility to prospective program recipients of employment opportunity and employment alternatives, and such as the local availability of highly skilled technicians and administrators, among other features often favorably associated with large populations in densely settled areas (urban and metropolitan). Here, the concern is with activity pattern and with involvement or potential involvement of populations in essentially urban metropolitan activity, at one extreme, as opposed to farming and other rural-oriented village or open-country lifestyles on the other.

An Alternative to Urban-Rural Residence

This paper reports on an attempt made to describe the total population and the population in poverty according to a scheme which would be more sensitive to the accessibility and activity concerns expressed above. One element introduced in the preliminary search for an alternative reflects an additional concern that some basically rural areas are sustaining conditions of a stable or growing economy which many others are not.

Background

Some remarks by Vincent Whitney in his study of changes in the rural-nonfarm populations of the United States, 1930-1950 are worth noting. Whitney found that 22,000,000 persons, more than one-half of the rural nonfarm population in 1950, were located in counties containing centers of from 10,000 to over 1,000,000 people. He concluded: "Despite the fact that the rural nonfarm population differs noticeably from the urban population in some respects, the majority of the rural-nonfarm people are clearly persons with urban orientations and associations." /4, p. 363/

More recently, Warren Robinson, in a study of rural population by metropolitan status, showed that nonmetropolitan rural population remained nearly constant, at just over 40,000,000 persons, over the period 1900-1960. / 2 /

By contrast, the rural portion of metropolitan areas, using "retrojection" as the technique for reconstructuring metropolitan areas prior to 1950, grew "...by 20 percent or more in every decade since 1900 and has consistently exceeded the national average growth rate." /2, pp. 176-177/

Under the assumption that the new patterns of population growth--particularly in rural areas--and the post-1920 local communication and transportation technologies are associated with radically altered community settlement patterns and patterns of area-to-area accessibility, a modified approach to urban-rural classifications is described below. The classification is intended to serve those concerned with locally-based anti-poverty programs in the context of a changing society.

Method

The 1962 city-county data file as prepared by the Bureau of the Census, corresponding to the 1962 County and City Data Book / 3 /, was used to classify all counties in the U.S. according to the following arrangement:

1. Metro-urban: Metropolitan counties or those with population 50 percent or more urban in 1960.
   a. Metro: Metropolitan counties.
   b. Nonmetro: Nonmetropolitan counties with 50 percent or more urban.
2. Stable, nonmetro-urban: Nonmetropolitan counties with population (1) less than 50 percent urban in 1960, and (2) without absolute loss, 1950-1960.
   a. No migration loss: Without net loss due to migration, 1950-1960.
   b. With migration loss: With loss due to migration.
3. Declining, nonmetro-urban: Nonmetropolitan counties with population (1) less than 50 percent urban in 1960, and (2) displaying absolute decline, 1950-1960.
   a. Level-of-living index 100 or more: Farm operator level-of-living index in 1959 less than 100, regardless of size of farm population in county.
   b. Level of living index less than 100: Index below 100.

Once having identified the counties according to this scheme, we prepared tabulations of a file of unpublished 1960 Census data which provides county summaries of population by poverty status.

In the modified classifications, the basic dichotomy is that of population within metropolitan areas or within counties with 50 percent or more of their respective population's classified as urban, on the one hand, and population located in nonmetropolitan predominantly rural areas on the other. The first class, that which I have termed metro-urban, refers to areas in which populations are clearly within the immediate influence of the urban society such that accessibility would not appear to be a critical problem

[1] Poverty status is defined according to the Social Security Index as described in / 1 /.

to urban or city-oriented institutions (poverty programs). This assumption ignores variations in population size and density by minor civil division which may provide severe problems of another sort.

Functional interrelationships of population in such areas are assumed to be of a common urban kind. Assuming the salience of occupational roles in defining such relations, the ties--either directly or through opportunity--to urban employments are evident and are distinguishing with respect to the objective of the classification.

Outside the metro-urban areas, both absolute population change as well as change due to migration are considered. The predominantly rural "stable" areas were defined as those with no absolute population loss over the decade 1950-1960. The stable class, then, includes both counties with and without losses in population due to migration.

In stable or growing rural areas, while access to urban institutions may be less than in the metro-urban areas, the presumptive needs, in speculative consideration of anti-poverty treatments, for schemes for economic development or, alternatively, for encouraging out-migration are thought to be less severe (opportunities are thought to be more prevalent) than is the case among the less than economically healthy areas--the predominantly rural "declining" areas.

Finally, in nonmetro-urban "declining" areas the farm operator level-of-living index was used, regardless of the size of the farm component, separating counties at levels of 100 or greater as opposed to those of less than 100. This was done on the assumption that a further scaling of declining areas as to their economic health might be accomplished. I must admit to more uncertainty in the utility of this step than of the others. At any rate, I should like to turn to the results of the tabulations.

Results

Tabulations of the county populations according to the rubic described above, which I shall term the urbanism-stability area index, yields the distribution of the total population and the rural population, by poverty status, as shown in table 2. Thus, while the prevalence of poverty is greater among rural persons than urban, the rural poverty population is more heavily concentrated in areas outside the metro-urban class than rural population as a whole.

The excess of rural poverty population in nonmetro-urban areas is wholly accounted for by the difference in degree of concentration in declining areas. Moreover, while similar proportions of the total population are located in nonmetro-urban stable areas and declining areas, and the rural populations are evenly balanced between such areas, a larger proportion of the poverty population and an even greater proportion of the <u>rural</u> poverty population is located in declining areas than in the stable areas.

Table 3 indicates that while percent rural tends to be least in the metro-urban areas and most in the nonmetro-urban areas, at a minimum, one-fifth of the total in even the least urban areas, is comprised of urban population. Moreover, the proportion of metro-urban population that is rural is not insignificant. Also notable is that for every urbanism-stability area index class, the urban proportion in the poverty population falls below the urban proportion in the total population.

TABLE 2. PERCENT DISTRIBUTION OF TOTAL AND RURAL POPULATION BY URBANISM-STABILITY AREA INDEX BY POVERTY STATUS, UNITED STATES, 1960

Type of Area	Percent			
	Total Population		Population in Poverty	
	Total	Rural	Total	Rural
Total	100	100	100	100
Metro-Urban	76	40	60	28
Metro	64	26	47	15
Nonmetro	12	15	13	13
Stable, Nonmetro-Urban	13	30	18	30
No Migration Loss	5	11	6	9
With Migration Loss	8	19	12	21
Declining, Nonmetro-Urban	11	29	22	43
Level-of-Living Index $\geq$ 100	4	10	5	11
Level-of-Living Index $<$ 100	7	19	17	32

TABLE 3. PERCENTAGE URBAN AND RURAL OF POPULATION CLASSIFIED BY URBANISM-STABILITY AREA INDEX BY POVERTY STATUS, UNITED STATES 1960

Type of Area	Percent					
	Total Population			Poverty Population		
	Total	Urban	Rural	Total	Urban	Rural
Total	100	70	30	100	55	45
Metro-Urban	100	84	16	100	79	21
Metro	100	88	12	100	85	15
Nonmetro	100	63	37	100	57	43
Stable, Nonmetro-Urban	100	31	69	100	25	75
No Migration Loss	100	33	67	100	28	72
With Migration Loss	100	30	70	100	24	76
Declining, Nonmetro-Urban	100	20	80	100	15	85
Level-of-Living Index $\geq$ 100	100	20	80	100	15	85
Level-of-Living Index $<$ 100	100	20	80	100	15	85

Prevalence of Poverty

Table 4 shows the percent of population in poverty by urban-rural residence by urbanism-stability status. On the whole, the prevalence of poverty is greater in predominantly rural areas than in the metro-urban areas. However, predominantly rural areas without loss due to migration had urban-like rates of poverty prevalence. Moreover, the urban population residing in predominantly rural areas where there was either (1) loss due to migration, but not absolute loss, or (2) in declining areas where the level-of-living index was above 100, the prevalence of poverty is in the metro-urban range.

The prevalence of poverty shown in table 4, is especially high in predominantly rural declining areas with low level-of-living index values. While only 7 percent of the U.S. population lived in such areas in 1960, some 19 percent of the rural population was so located and about 32 percent of the rural poverty population lived in such areas. On the other hand, only 2 percent of the urban population and 5 percent of the urban poverty population resided in such areas (table 5).

Conclusions

Although 45 percent of the U.S. population in 1960 defined as in poverty was comprised of rural

TABLE 4. PERCENT OF POPULATION IN POVERTY, BY URBAN-RURAL RESIDENCE, POVERTY STATUS AND URBANISM-STABILITY AREA INDEX, UNITED STATES, 1960

Type of Area	Percent in Poverty				
	Total	Urban	Rural		
			Total	Nonfarm	Farm
Total	22	17	33	32	36
Metro-Urban	17	16	23	22	26
Metro	16	16	20	19	22
Nonmetro	25	22	28	28	28
Stable, Nonmetro-Urban	29	23	32	32	33
No Migration Loss	23	20	25	26	24
With Migration Loss	33	26	36	36	36
Declining, Nonmetro-Urban	45	34	48	49	46
Level-of-Living Index $\geq$ 100	32	24	34	35	32
Level-of-Living Index $<$ 100	52	39	55	56	54

TABLE 5. PERCENT DISTRIBUTION OF TOTAL AND URBAN POPULATION BY URBANISM-STABILITY AREA INDEX, BY POVERTY STATUS, UNITED STATES, 1960

Type of Area	Percent			
	Total Population		Poverty Population	
	Total	Urban	Total	Urban
Total	100	100	100	100
Metro-Urban	76	91	60	86
Metro	64	80	47	72
Nonmetro	12	11	13	14
Stable, Nonmetro-Urban	13	6	18	8
No Migration Loss	5	2	6	3
With Migration Loss	8	4	12	5
Declining, Nonmetro-Urban	11	3	22	6
Level-of-Living Index ≥ 100	4	1	5	2
Level-of-Living Index < 100	7	2	17	5

population, examination of the population distribution according to a simple typology of areas permits a closer approximation of the extent of poverty outside areas of direct urban influence. Of the rural population in poverty, some 28 percent reside in metro-urban areas and an additional 30 percent are found in nonmetro-urban stable areas.

Thus, anti-poverty programs oriented toward alternatives of encouraging out-migration or economic development are probably more reasonably assessed against some 22 percent of the population in poverty (which includes an urban component) rather than the 45 percent figure with which we began (table 1).

Further analysis of the data not covered in the paper is planned in order to evaluate color differentials and variations by region as to the way in which the conventional view of urban-rural residence intersects with the view suggested by the urbanism-stability area classification described.

REFERENCES

[1] Orshansky, Mollie. "Who's Who Among the Poor: A Demographic View of Poverty." Social Security Bulletin, July, 1965.

[2] Robinson, Warren C. "Changes in the Rural Population of the United States by Metropolitan and Nonmetropolitan Status, 1900 to 1960." Rural Sociology, 30 (June, 1965), pp. 166-183.

[3] U.S. Bureau of the Census. County and City Data Book, 1962. U.S. Government Printing Office, Washington 25, D. C. 1962.

[4] Whitney, Vincent Heath. "Changes in the Rural-Nonfarm Population, 1930-1950." American Sociological Review, 25 (June, 1960), pp. 363-368.

STOCHASTIC ASPECTS OF THE LABOR FORCE PROCESS

Haskel Benishay, Northwestern University*

I. Introduction

This paper develops a model of the labor force-employment process within the framework of which the variability which is expected to occur normally in the number of labor force participants, in the number of employed, in accessions and in retirements is highlighted and emphasized. The basic approach consists of grafting a stochastic labor force-employment process onto a stochastic population process. The major improvements over previous work are the following three:[1] (1) the development of results under more realistic assumptions in regard to accessions and retirements (2) the explicit incorporation of employment into the model albeit under simplified assumptions (3) the provision of an outline of a more realistic model which includes multiple entries and exits from the labor force and employment.

II. Notation

The notation and assumptions of the model as well as the description of the population-labor force-employment process are presented below.

Let the time interval from point in time t-1 to point in time t be t, so that t represents either a time interval or its end point. Generally, the symbol t will represent a time interval when flows are treated and the end point of this interval when stocks are considered.

Let E, V, before parentheses represent the expectation and variance of the variable in parentheses.

*Portions of this presentation and related materials were presented in a faculty seminar in Northwestern University Economics Department in spring 1967, and at the U.S. Department of Labor, Bureau of Labor Statistics in the summer of 1967. Comments made in response to these presentations were useful in providing a perspective in regard to the relative importance of various phases of this work. Support for some of the research embodied in this presentation was provided by the U.S. Department of Labor, Manpower Administration.

[1]This presentation is made against the background of a previous paper with similar but more limited aims [1].

N_t is an independent exogenous random variable representing the number of live births during interval t.

S is an endogenous random variable representing the number of people alive at a point in time t.

F is an endogenous random variable representing the number of individuals in the labor force at point in time t.

D_d is the possible span of life at birth, for any individual. D_d is a discrete variable representing intervals (units) of time which can consist only of integers which range between a minimum of 1 (one) and a maximum of m. The probability that $D_d = i$ where i takes on consecutively, values from 1 to m, is P_{di}, i.e., $P(D_d=i)=P_{di}$, i=1...m.

d deaths is an endogenous random variable representing the number of individuals who die during a time interval.

a_{lv} represents live accessions to the labor force during a time interval. Lower case v represents life here and henceforth. Within the context of this model a_{lv} is an endogenous random variable.

P_{ali} is the probability that an individual will join the labor force (become an accession) i periods from birth. e is the first possible number of time intervals after birth at which and after which accessions occur (say 14 years after birth). The i in this case can assume values from e to m-1. D_{al} is the random variable defined by $e \le i \le m-1$, and by P_{ali}.

r_{lv} represents live retirements from the labor force during time intervals. It is an endogenous variable determined within the model.

P_{crli} is the <u>conditional</u> probability that an individual will retire from the labor force i periods after birth given that he has already acceded to the labor force.

F_v is an endogenous random variable representing the number of live individuals who are in the labor force at the end of time intervals.

w_{lvF} is an endogenous random variable representing the number of live individuals in the labor force who are becoming employed for the first time during a time interval.

W_{1vF} is an endogenous random variable representing the number of live people within the labor force who are employed at the end of a time interval.

U is an endogenous random variable representing the number of live people within the labor force who are unemployed at the end of a time interval.

P_{w1i} is the probability that an individual will become employed, i periods from time of birth. It is assumed that the first possible number of periods (after birth) at which and after which an individual may become employed is e, the same number of time intervals after birth at which first accession to the labor force may occur (14 years after birth). Thus i in this case can also assume only discrete integer values between e and m-1. Additionally, the probability P'_{w1i}, that an individual is employed at t, is the cumulative probability of P_{w1i}. D_{w1i} is the random variable defined by possible values $e \leq i \leq m-1$ and probabilities P_{w1i}.

The exogenous variables under consideration are assumed serially and mutually independent, with one exception. The conditional probabilities of retirement at i are P_{cr1i} for members of of the labor force at i-1 and zero for non-members at i-1. It is logically impossible to retire from the labor force without having entered first. Retirement from the labor force is thus dependent on previous entry into it.

The exogenous variables in this model are:

1. Births, N.
2. The life span distribution, D_d.
3. The distribution of accession times, birth to accession D_{a1}.
4. The distribution of (first) employment times, birth to first employment, D_{w1}.
5. The conditional probabilities of retirements, P_{cr1i}.

A random birth process and a random death process take place together to form the population process. Each time interval a random number drawn from N, determines births. Next, N random draws from the life span distribution, D_d, determine the length of life i for each of the N births. This process is viewed as continuing repeatedly until process equilibrium (steady state) is reached and beyond into a phase in which various stages of the process occur continually through time. Every period births, and deaths from births in previous periods, occur. Every period individuals born and not yet dead comprise the living population. The labor force-employment process is "grafted" onto the population process. Every period accessions come into being from birth cohorts which came into the world 14 to 30 years or so before this period. Every period individuals who have joined the labor force in the past, retire. Period after period individuals are becoming employed for the first time. Period after period individuals originating from preceding birth cohorts come to the end of their life.

III. The Endogenous Variables: Concrete Definitions

The life span random variable D_d may be viewed as a succession of dichotomous variables X_{di}, applicable to a succession of points in time at i periods from birth, for each of which the two possibilities for an individual at birth are: 1. dying i periods after birth, defined identically as unity (one). 2. dying not after i periods, defined identically as zero. Clearly the probability of unity is P_{di} the probability of zero $(1-P_{di})$. Also, $E(X_{di})$, $V(X_{di})$, are simply P_{di} and $P_{di}(1-P_{di})$ respectively, as is true for the binomial distribution. This view becomes useful for the technical definition of, d, deaths.

Deaths, d, is conceived as a sum of the m terms below, each one of which being the number of individuals born i periods before t and dying during t.

(1.1)

$$d = \sum_{j=1}^{N_{t-1}} X_{d1j} + \sum_{j=1}^{N_{t-2}} X_{d2j} + \cdots + \sum_{j=1}^{N_{t-m}} X_{dmj}.$$

Each term above is the sum of a random number N of random draws from X_{di}.

The technical definition of S, the number of individuals alive at t, requires a definition of a second binomial variable, Y_{di} whose parameter is Q_{di}, where Q_{di} is the sum of the previously defined P_{di} from P_{di} to P_{dm}, i.e.:

(1.2)

$$Q_{d1} = P_{d1}+P_{d2}+P_{d3}+ \cdots +P_{dm}$$

$$Q_{d2} = P_{d2}+P_{d3}+ \cdots +P_{dm}$$

$$Q_{d3} = P_{d3} + \dots + P_{dm}$$

$$\vdots$$

$$Q_{dm} = P_{dm}$$

Y_{di} is a binomial variable whose value of unity represents the state of being alive after i-0' periods from time of birth and whose value of 0 represents the state of being dead after i-0' or less where 0' is a number as close to zero as is conceivable.

Q_{di} is the probability that an individual will remain a live member of the population, S, after i-0' periods from the time it is born, while $(1-Q_{di})$ is the probability that an individual will no longer be in S, i-0' periods or less from birth. Births are thereby dichotomously divided into those who die and leave the population after i-0' periods or less and those who still remain part of the population after i-0'.

The number of people in the population can also be expressed as a sum of terms each of which being a random number of random draws in the explicit way below.

(1.3)

$$S = \sum_{j=1}^{N_{t-1}} Y_{d1j} + \sum_{j=1}^{N_{t-2}} Y_{d2j} + \dots + \sum_{j=1}^{N_{t-m}} Y_{dmj}.$$

Each term is a sum of a randum number of births at t, t-1...t-m+1 which are randomly "still alives" at point in time t.

Accessions, a_{1v}, is similarly conceived to be a sum of terms originating from birth cohorts before period t which produce accessions in period t. For the purpose of expressing a_{1v} under this view additional notation and discussion are necessary.

X_{a1vi} is a binomial variable whose value of unity represents the state of being alive and entering the labor force i periods from time of birth and whose value of zero represents not being in this state. The probability of X_{a1vi} given previous assumptions, is P_{a1i} $Q_{d(i+1)}$ and the range of i is between e and m-1. Accessions, a_{1v}, can thus be viewed as the sequence of terms below:

(1.4)

$$a_{1v} = \sum_{j=1}^{N_{t-e}} X_{a1vej} + \sum_{j=1}^{N_{t-(e+1)}} X_{a1v(e+1)j} + \dots + \sum_{j=1}^{N_{t-(m-1)}} X_{a1v(m-1)j}$$

Each term is a sum of randum number N_{t-i} of random draws from X_{a1vi}.

Live retirements r_{1v}, is also conceived as a sum of terms which result from past birth cohorts. X_{r1vi} is defined as a binomial variable whose value of unity represents being alive, having already acceded to the labor force, and retiring. The probability P_{r1vi} of X_{r1vi} can be derived from some previous and a few additional definitions. Define P_{cr1i} as the conditional probability that an individual retires at i periods from birth, given that he has already acceded and not yet retired. Thus, the probability that an individual acceded and is retiring alive at i is for e+1:

(1.51)

$$P_{r1v(e+1)} = P_{cr1(e+1)} P_{a1e} Q_{d(e+2)}$$

for e+2:

(1.52)

$$P_{r1v(e+2)} = P_{cr1(e+2)} [P_{a1e}(1-P_{cr1(e+1)}) + P_{a1(e+1)}] Q_{d(e+3)}$$

for e+3:

(1.53)

$$P_{r1v(e+3)} = P_{cr1(e+3)} [P_{a1e}(1-P_{cr1(e+1)}) (1-P_{cr1(e+2)})$$

$$+ P_{al(e+1)}(1-P_{crl(e+2)})$$

$$+ P_{al(e+2)}]Q_{d(e+4)}$$

$$\cdots \qquad \cdots \qquad \cdots$$

for (m-1):

(1.5(m-1))

$$P_{rlv(m-1)} = P_{crl(m-1)} [P_{ale}(1-P_{crl(e+1)}) (1-P_{crl(e+2)}) \cdots (1-P_{crl(m-2)}) \cdot + P_{al(e+1)}(1-P_{crl(e+2)}) (1-P_{crl(e+3)}) \cdots (1-P_{crl(m-2)}) + \cdots + P_{al(m-2)}]Q_{dm}{}^{\circ}$$

The logic of the above expression is straight-forward. The conditional probability of retiring (first term) is multiplied by the probability of having acceded but not yet retired (terms in square parenthesis) which in turn is multiplied by the probability of being alive (last term). The probability of having acceded but not yet retired by i after birth is essentially the cumulative probability of having acceded by i less the probability of having already retired by i.

With the help of the preceding results, r_{lv}, live retirements, can be written as:

(1.6)

$$r_{lv} = \sum_{j=1}^{N_{t-e-1}} X_{rlv(e+1)j} + \sum_{j=1}^{N_{t-e-2}} X_{rlv(e+2)j} + \cdots + \sum_{j=1}^{N_{t-m+1}} X_{rlv(m-1)j}.$$

And each term, again, is a random number N_{t-i} of random draws from X_{rlvi}.

Given the previous results, the labor force, F_v, can be conceived also as the sum of random sums of random variables where the number of terms in each random sum is the random number of births in t-1. X_{Fvi}, is defined as a binomial variable representing the state of being in the labor force and alive i periods after birth. The probability P_{Fvi} of X_{Fvi} can be gleaned from previous results. Being in the labor force means having entered but not left the labor force and being alive. The probability of such an event has already been provided partially in expressions (1.5i) above. Delete from each of these expressions the first term, P_{crli}, and what remains is the probability P_{Fvi} of X_{Fvi}, i.e., of having entered the labor force, not having left it, and being alive.

The labor force, F_v can be viewed as a sum of random draws from a random variable as were other endogenous variables.

(1.7)

$$F_v = \sum_{j=1}^{N_{t-e}} X_{Fvej} + \sum_{j=1}^{N_{t-(e+1)}} X_{Fv(e+1)j} + \cdots + \sum_{j=1}^{N_{t-(m-1)}} X_{Fv(m-1)j}.$$

And now we address ourselves to employment. We treat two aspects of employment, employment inflow and employment stock, i.e., the number of people becoming employed each period and the number of employed people outstanding at a point in time. The former is represented by w_1, the latter by W_1. Two assumptions are made in respect to employment: (1) employment is a state into which an individual enters once (hence the 1 in w_1). (2) First entries into the state of employment occur e periods after birth at which entry into the labor force is assumed to commence.

As has been indicated earlier, we represent the probability of becoming employed i periods from time of birth by P_{wli}. Consequently, P'_{wli} the probability of being employed at i, as a result of entering the labor force at i or before i is the cumulative probability of joining the employed anew (becoming employed for the first time) at i and at periods smaller than i.

Thus:

(1.81)

$$P'_{w1e} = P_{w1e}$$

(1.82)

$$P'_{w1(e+1)} = P_{w1e} + P_{w1(e+1)}$$

$$\cdots \qquad \cdots$$

(1.8(m-1))

$$P'_{w1(m-1)} = P_{w1e} + P_{w1(e+1)} + \cdots + P_{w(m-1)} .$$

The probability of the binomial variable whose value of unity stands for newly joining the employed, being in the labor force, and being alive, i periods from birth, to which we shall refer as, X_{w1vFi}, is simply the product $(P_{w1i})(P_{Fvi})$. Thus the total number of live people 'inflowing' into state of being employed within the labor force, w_1, can be expressed as follows:

(1.9)

$$w_{1vF} = \sum_{j=1}^{N_{t-e}} X_{w1vFej} + \sum_{j=1}^{N_{t-e-1}} X_{w1vF(e+1)j} + \cdots + \sum_{j=1}^{N_{t-m+1}} X_{w1vF(m-1)j} .$$

Similarly, the probability of the binomial variable whose value of unity stands for having joined the employed, being in the labor force, and being alive i periods from birth, to which we shall refer as X_{W1vFi}, is the product $(P'_{w1i})(P_{Fvi})$. Consequently, the stock of live people who are in the labor force and employed is expressed as:

(2.0)

$$W_{1vF} = \sum_{j=1}^{N_{t-e}} X_{W1vFej} + \sum_{j=1}^{N_{t-e-1}} X_{W1vF(e+1)j} + \cdots + \sum_{j=1}^{N_{t-m+1}} X_{W1vF(m-1)j} .$$

Unemployment, U, is a by product of preceding expressions. The variable, X_{Ui} represents the state of being, at i periods after birth, unemployed, in the labor force and alive. Its probability is the product of $(1-P'_{w1i})(P_{Fvi})$ the term in the first parenthesis of the product being the probability of being unemployed by i, the term in the second representing the probability of being alive and in the labor force.

U can then be expressed as:

(2.1)

$$U = \sum_{j=1}^{N_{t-e}} X_{Uej} + \sum_{j=1}^{N_{t-e-1}} X_{U(e+1)j} + \cdots + \sum_{j=1}^{N_{t-m+1}} X_{U(m-1)j} .$$

IV. Some Results

The expectation and variance of the endogenous variables treated in the preceding section have the same general form. Hence, we shall provide the expectation and variance of one variable, death, d, and indicate that the other results may be expressed similarly.

The expectation and variance of deaths was derived in a previous paper as follows [1,2]:

(2.2)

$$E(d) = E(N) \sum_{i=1}^{m} P_{di}$$

(2.3)

$$V(d) = V(N) \sum_{i=1}^{m} (P_{di})^2$$

$$+ E(N) - E(N) \sum_{i=1}^{m} (P_{di})^2.$$

The expectation and variance of the other endogenous variables are similarly derived. The difference being the substitution of the probabilities of the respective variables for P_{di}.

V. Some Further Developments

The treatment of r_{1v} and a_{1v}, presented earlier, points the way to a possible solution of some of the problems connected with the construction of a logical and probabilistic framework for the phenomena of multiple entries and retirements into and out of the labor force as well as multiple entries into and exits from the state of employment. It is useful to note that accessions are assumed here to occur concurrently with retirements, (at least from e+1 to (m-1) for each cohort), an assumption which is not usually made in traditional work in this area. It may be also interesting to note that retirements are expressed, for the case of concurrent entry and retirement, in such a way as to take account of the fact that retirements in t-1 cannot retire again in t. In the same spirit we assume in further extensions of this work that first entry, first retirement, second entry, second retirement, etc., both into and out of the labor force and into and out of employment, can occur simultaneously. We then treat second entry in relation to first retirement, second retirmenet in relation to second entry, etc., essentially, in the same fashion as r_{1v}, first retirement, was treated in this work in relation to, a_{1v}, first accession. This provides a sensible framework for the evaluation of the phenomena of multiple entries and exits into and out of the labor force and employment. Work in this direction has been conducted and will be continued.

Additional work has been and is being carried out to investigate the serial correlations and the cross and cross-serial correlations of the endogenous random variables, with the hope that it may become useful in forecasting labor force, employment, and their components.

References

[1] Benishay, H., "On the Construction of Stochastic Working Life Tables," Proceedings of the American Statistical Association, Business and Economics Statistics, 1965. Pp. 332-337.

[2] Benishay, H., "Parameters and Relations of Stochastically Lagged and Disaggregative Time Series," Econometrica, 1967.

XII

CONTRIBUTED PAPERS II

Chairman, GEORGE J. STOLNITZ, University of Indiana

THE BIAS AND STABILITY OF JACK-KNIFE VARIANCE ESTIMATOR IN RATIO ESTIMATION

R.P. Chakrabarty and J.N.K. Rao
University of Georgia and Texas A & M University

Summary

The Jack-Knife variance estimator $v_2(r)$ (say) based on splitting the sample at random into g groups is applied to estimation of the variance of the ratio estimator $r=\bar{y}/\bar{x}$ of the population ratio $R=\bar{Y}/\bar{X}$. Assuming a linear regression of y on x where x has a gamma distribution it is shown that the exact bias of $v_2(r)$ is a decreasing function of g. The exact bias of $v_2(r)$ with g = n is less than that of $v_1(r)$, the customary variance estimator of r, for moderate sample sizes. The exact stability of $v_2(r)$ for the special case of g = 2 is shown to be less than that of $v_1(r)$. The asymptotic stability of $v_1(r)$ is also discussed for a bivariate normal distribution.

1. Introduction

Ratio estimators are often employed in sample surveys for estimating the population mean $\bar{Y}$ of a characteristic of interest 'y' or the population ratio $R=\bar{Y}/\bar{X}$ utilizing an auxiliary variate 'x' that is positively correlated with 'y'. The estimate of the variance associated with an estimator is often used in drawing statistical inferences (e.g. confidence limits on the estimand). It is, therefore, desirable that a variance estimator should be as stable as possible. In this paper we investigate the bias and stability of the Jack-Knife variance estimator and the customary variance estimator in ratio estimation.

We shall confine ourselves to simple random sampling and assume that the population size N is infinite to simplify the discussion. From a simple random sample of n pairs (y_i, x_i) we have the customary ratio estimator of R as

$$r = \bar{y}/\bar{x} \qquad (1)$$

where $\bar{y}$ and $\bar{x}$ are the sample means of y and x respectively. As an estimator of V(r), the variance of r, it is customary to take

$$v_1(r) = (s_y^2 - 2rs_{yx} + r^2 s_x^2)/(n\bar{X}^2) \qquad (2)$$

where s_y^2 and s_x^2 are sample mean squares and s_{yx} is the sample covariance. It is known that the bias of $v_1(r)$ is of order 1/n.

Let the sample of size n be divided at random into g groups, each of size p so that n=pg. Let

$$r_Q = g^{-1} \sum_1^g r_{Qj} \qquad (3)$$

where

$$r_{Qj} = gr - (g-1)r_j' \qquad (4)$$

and r_j' is the customary ratio estimator calculated from the sample after omitting the jth group.

Tukey (1958) has pointed out that the estimators like r_{Qj} (called pseduo - values) may, to a good approximation, be treated as though they are g independent estimators. Therefore, we can use the simple estimator

$$v_2(r) = g^{-1}(g-1)^{-1} \sum_1^g (r_{Qj} - r_Q)^2 \qquad (5)$$

as an estimator of V(r) since $r_Q = \sum_1^g r_{Qj}/g$. Tukey has called this general procedure, described here in the context of ratio estimation, the 'Jack-Knife'.

The stability of a variance estimator may be judged by its coefficient of variation. Kokan (1963) has investigated the large-sample stabilities of $v_1(r)$ and the unbiased variance estimator:

$$v(\bar{y}/\bar{X}) = s_y^2/(n\bar{X}^2) \qquad (6)$$

where $\bar{y}/\bar{X}$ is the ratio estimator of R not using the sample x-information. He has shown that the coefficient of variation of $v_1(r)$ is <u>always</u> larger than that of $v(\bar{y}/\bar{X})$ for a bivariate normal distribution and this property also holds for a bivariate log normal distribution for certain ranges of the parameters. Lauh and Williams (1963) have made a Monte Carlo study of the stabilities of $v_1(r)$ and $v_2(r)$ with g=n for small samples. Assuming that the regression of y on x is through the origin and C_x is small, they have shown that the Monte Carlo variances of $v_1(r)$ and $v_2(r)$ are about the same when x has a normal distribution whereas the variance of $v_2(r)$ is considerably smaller than that of $v_1(r)$ when x has an exponential distribution.

Recently Rao and Beegle (1966) investigated the small sample properties of $v(\bar{y}/\bar{X})$, $v_1(r)$ and $v_2(r)$ by a Monte Carlo study. They have shown that under the Lauh and Williams' model with x normal, the coefficient of variation of $v_2(r)$ decreases as g increases. The coefficients of variation of $v_2(r)$ with g=n and $v_1(r)$ are essentially equal. Further $v_1(r)$ and $v_2(r)$ with g=n are quite stable compared to $v(\bar{y}/\bar{X})$. They have also considered the general regression model where the regression of y on x does not pass through the origin, and C_x is large. Under this non-ideal condition also the coefficient of variation of $v_2(r)$ decreases with g so that g=n is the optimum choice. The coefficients of variation of $v_2(r)$ (with g=n) and $v_1(r)$ are again essentially equal, but both are considerably larger than the coefficient of variation of $v(\bar{y}/\bar{X})$. Rao and Beegle conclude that caution is needed in the indiscriminate use of ratio extimators.

2. Stabilities of Variance Estimators $v(\bar{y}/\bar{X})$ and $v_1(r)$: Asymptotic Theory Assuming a Bivariate Normal Distribution.

Kokan (1963) used the formula for the relative variance of $v_1(r)$ given by Hansen et al (1953 page 177) to compare the stability of $v_1(r)$ with that of $v(\bar{y}/\bar{X})$. This formula was derived by substituting R for r in $v_1(r)$ as a large sample approximation. We will show that this approach is not correct. The correct approach is to expand r in terms of $\delta x = (\bar{x}-\bar{X})/\bar{X}$ and $\delta y = (\bar{y}-\bar{Y})/\bar{Y}$ and find the variance of $v_1(r)$ for large samples. Using this approach and utilizing the theory of cumulants and k-statistics (Kendall and Stuart, 1958) to find the variance and covariances of sample moments it can be shown that the relative variance (CV^2) of $v_1(r)$ is

$$CV^2[v_1(r)] = \frac{2}{n} + \frac{4}{n}\,\frac{[C_x^2 - \rho C_y C_x]^2}{[C_y^2 - 2\rho C_y C_x + C_x^2]}$$

$$= \frac{2}{n} + \frac{4[\text{Bias}(r)]^2}{V(r)} \tag{7}$$

to terms of order n^{-1}, where C_y, C_x are coefficients of variation (CV) of y and x respectively and ρ is the coefficient of correlation between y and x. The relative variance of $v(\bar{y}/\bar{X})$ is

$$CV^2[v(\bar{y}/\bar{X})] = \frac{2}{n} \tag{8}$$

to terms of order n^{-1}.

From (7) and (8) we have

$$CV[v(\bar{y}/\bar{X})] \leq CV[v_1(r)]. \tag{9}$$

The equality sign in (9) holds only when the regression of y on x is a straight line through the origin. Thus in large samples, with simple random sampling from a bivariate normal population, the coefficients of variation of the variance estimators $v(\bar{y}/\bar{X})$ and $v_1(r)$ are equal only if the regression of y on x is a straight line through the origin; otherwise CV of $V(\bar{y}/\bar{X})$ is always less than that of $v_1(r)$.

It is interesting to note that the Monte Carlo results for small samples obtained by Rao and Beegle (1966) agree with the asymptotic results obtained here, namely, $v(\bar{y}/\bar{X})$ and $v_1(r)$ are equally stable if the regression of y on x is through the origin; otherwise $v_1(r)$ is always less stable. Further, since $|\text{Bias}\ (r)|/\sigma r \leq C_x/\sqrt{n}$ we have to terms of order 1/n.

$$CV^2[v_1(r)] \leq \frac{2}{n} + \frac{4C_x^2}{n} \tag{10}$$

Kokan had obtained the expression on r.h.s. of (10) for $CV^2[v_1(r)]$ using the general formula given by Hansen at el (1953), which clearly over estimates the coefficient of variation of $v_1(r)$ so far as the large sample approximation to terms of order n^{-1} is concerned. As a result, he had found CV of $v_1(r)$ to be considerably higher than that of $v(\bar{y}/\bar{X})$ even when the regression is through the origin.

The relative stability of two variance estimators may be judged by the ratio of their relative variances. The relative stability of $v_1(r)$ compared to $v(\bar{y}/\bar{X})$ is given by

$$E_1 = CV^2[v(\bar{y}/\bar{X})]/CV^2[v_1(r)]$$

$$= [1+2C_y^2K^2(K-\rho)^2/(1-2\rho K+K^2)]^{-1} \tag{11}$$

$$= 1 \text{ if } K = \rho \quad \text{(regression through the origin)}$$

$$< 1 \text{ otherwise} \tag{12}$$

where $K = C_x/C_y$. The relative stability clearly depends on ρ, C_y and $K = C_x/C_y$. The stability of $v_1(r)$ relative to that of $v(\bar{y}/\bar{X})$ is of interest only when the estimator r is more efficient than $\bar{y}/\bar{X}$ (i.e. when $\rho > K/2$). Consequently, the numerical values of E_1 for selected values of $\rho(>K/2)$ K and C_y are presented in Table 1. It will be seen from Table 1 that for fixed C_y, E_1 decreases as $|\rho-K|$ (i.e. departure from regression through the origin) increases. The stability of $v_1(r)$ is low when $C_x = KC_y$ is large.

3. Stability of Variance Estimators $v(\bar{y}/\bar{X})$ and $v_1(r)$: Exact Theory assuming x has a gamma distribution.

In this section we assume that $y_i=\alpha+\beta x_i+u_i$, where u_i's are independent normals with mean zero and variance $n\delta$ (δ is of order n^{-1}) and the variates x_i/n have the gamma distribution with parameter h so that $\bar{x} = \Sigma x_i/n$ has the gamma distribution with parameter m = nh. Under this model we derive the formulae for the variance estimators and investigate their stabilities. All our results are exact for any sample size, n.

The variance of $\bar{y}/\bar{X}$ under the model is given by

$$V(\bar{y}/\bar{X}) = \frac{\delta}{m^2} + \frac{\beta^2}{m} \tag{13}$$

and $v(\bar{y}/\bar{X})$ is an unbiased estimator of $V(\bar{y}/\bar{X})$. The variance of r is

$$V(r) = \frac{\alpha^2}{(m-1)^2(m-2)} + \frac{\delta}{(m-1)(m-2)}. \tag{14}$$

The bias of $v_1(r)$ as an estimator of V(r) can be shown to be

$$\text{Bias}\ [v_1(r)] = -\frac{(5m^2-5m+2)\alpha^2}{m^2(m^2-1)(m-1)(m-2)}$$

$$-\frac{2(m^2+2m-2)\delta}{m^2(m^2-1)(m-2)}$$

$$= c_3\alpha^2 + c_4\delta \quad \text{(say)}. \tag{15.}$$

We note that for finding the variances of $v(\bar{y}/\bar{X})$ and $v_1(r)$, expected values of some functions of sample moments are needed. The method of evaluating these expectations is same as that of Rao and Webster (1966). The details of

Table 1. Asymptotic relative stability of $v_1(r)$ for selected values of K, C_y and ρ (>k/2).

K	C_y	ρ = .3	ρ = .5	ρ = .7	ρ = .9
.5	1.0	98	100	96	81
	2.0	92	100	87	52
	3.0	84	100	75	33
1.0	0.5	-	89	93	98
	1.0	-	67	77	91
	2.0	-	33	45	71
1.5	0.5	-	-	-	58
	1.0	-	-	-	25
	2.0	-	-	-	8

evaluating these expectations, which involve some tedious algebra, are omitted and only the final results are given. The variance of $v(\bar{y}/\bar{X})$ is given by

$$V[v(\bar{y}/\bar{X})] = \frac{2\delta^2}{m^4(n-1)} + \frac{4\beta^2\delta}{m^3(n-1)} + \frac{\beta^4}{m^3}[\theta(m+1)(m+2)(m+3)-m]. \qquad (16)$$

The variance of $v_1(r)$ can be shown to be

$$V[v_1(r)] = \frac{\delta^2}{m^4}[3\ \theta + \frac{(n+1)(m+3)}{(n-1)(m+1)} - \frac{(m+2)^2}{(m+1)^2}] + \frac{2\alpha^2\delta}{m^4}[3\ \theta + \frac{(2m-n+3)}{(n-1)(m+1)^2}] + \frac{\alpha^4}{m^4}[\theta - \frac{1}{(m+1)^2}] \qquad (17)$$

where

$$\theta = \frac{[(n+1)(m+6)-12]}{(n-1)(m+3)(m+2)(m+1)}. \qquad (18)$$

From (15) and (17) the MSE of $v_1(r)$ can be obtained as

$$MSE[v_1(r)] = c_5\ \delta^2 + c_6\alpha^4 + c_7\alpha^2\delta \quad \text{(say)} \qquad (19)$$

where the coefficients c_5, c_6 and c_7 are functions of m and n.

Further, we note that in terms of the model

$$\alpha = \bar{Y}[(K-\rho)/K]$$
$$\beta = \bar{Y}[\rho/(Km)]$$
$$\delta = \bar{Y}^2[(1-\rho^2)/(K^2m)] \qquad (20)$$

where $K = C_x/C_y$. Now, using (13) through (18) and (20) the relative variances (i.e. CV^2) of the variance estimators $v(\bar{y}/\bar{X})$ and $v_1(r)$ can be obtained as functions of K, ρ, m and n. At present, we are evaluating CV^2 of the variance estimators to compare their exact stabilities for different values of K, ρ, m and n.

4. Bias of the Jack-Knife Variance Estimator.

In this section we investigate the bias of $v_2(r)$ and compare with that of $v_1(r)$ under the model of section 3. The Jack-Knife variance estimator $v_2(r)$ can be written as

$$v_2(r) = g^{-1}(g-1)^{-1}\sum_1^g (r_{Qj} - r_Q)^2$$

$$= \frac{(g-1)}{g}\,[\alpha^2\sum_1^g\{\frac{1}{\bar{x}'_j} - \frac{1}{g}\sum_1^g\frac{1}{\bar{x}'_j}\}^2 + \sum_1^g\{\frac{\bar{u}'_j}{\bar{x}'_j} - \frac{1}{g}\sum_1^g\frac{\bar{u}'_j}{\bar{x}'_j}\}^2 + 2\alpha\sum_1^g\{\frac{1}{\bar{x}'_j} - \frac{1}{g}\sum_1^g\frac{1}{\bar{x}'_j}\}\{\frac{\bar{u}'_j}{\bar{x}'_j} - \frac{1}{g}\sum_1^g\frac{\bar{u}'_j}{\bar{x}'_j}\}] \qquad (21)$$

where $\bar{u}'_j$ and $\bar{x}'_j$ are the sample means obtained after omitting the j th group. Now, since $[(g-1)\bar{x}'_j]/g$ has the gamma distribution with parameter $(n-p)h=[m(g-1)]/g$, we have

$$E[\frac{1}{\bar{x}'^2_j}] = \frac{(g-1)^2}{[m(g-1)-g][m(g-1)-2g]}.$$

For $\bar{u}'_j$'s we have the following expected values:

$$E(\bar{u}'^2_j) = \frac{g}{g-1}\delta$$

and

$$E[\bar{u}'_j\ \bar{u}'_j] = \frac{g(g-2)}{(g-1)^2}\delta \quad ;\ i \neq j, g \geq 3.$$

Using these expected values it can be shown that the expected value of $v_2(r)$ for $g \geq 3$ (the special case of $g = 2$ is discussed in the next section) is

$$E[v_2(r)] = \alpha^2 \{\frac{(g-1)^4}{g[m(g-1)-g][m(g-1)-2g]} - \frac{(g-1)^2}{g} E(\frac{1}{\bar{x}'_i \ \bar{x}'_j})\} +\delta\{\frac{(g-1)^3}{[m(g-1)-g][m(g-1)-2g]} - (g-2) \ E(\frac{1}{\bar{x}'_i \ \bar{x}'_j})\} . \quad (22)$$

From Rao and Webster (1966) we have, for integer m,

$$\frac{g^2}{(g-1)^2} E[\frac{1}{\bar{x}'_i \ \bar{x}'_j}]=\Gamma(2a+b-2)\Gamma^{-2}(a)\Gamma^{-1}(b)C(a,b) \quad (23)$$

where

$$C(a,b)= \sum_{k=0}^{a-2} (-1)^k \frac{\Gamma^2(a-k-1)\ \Gamma(b+k)}{\Gamma(2a+b-k-2)} +(-1)^{a-1}[2 \sum_{k=1}^{a+b-2} (-1)^{k+1}\frac{1}{(a+b-k-1)^2} + (-1)^{a+b} \frac{\Pi^2}{6}] \quad \text{if } a \geq 2 \quad (24)$$

$$= 2 \sum_{k=1}^{b-1} (-1)^{k+1} \frac{1}{(b-k)^2} + (-1)^{b+1} \frac{\Pi^2}{6} \quad \text{if } a=1,\ b \geq 2 \quad (25)$$

$$= \frac{\Pi^2}{6} \quad \text{if } a=1,\ b=1 \quad (26)$$

and $a=m/g$ and $b=[m(g-2)]/g$.

Now, the bias of $v_2(r)$ as an estimator of $V(r)$ is

$$\text{Bias}[v_2(r)]=E[v_2(r)]-V(r)=c_1 \ \alpha^2+c_2\delta \ \text{(say)}. \quad (27)$$

Using (22) through (26) the coefficients c_1 and c_2 can be expressed explicitly as functions of g and m. However, since the resulting expression would not be in a closed form, it is difficult to investigate analytically the behavior of the bias of $v_2(r)$ as a function of g for fixed m. Therefore, we have made a numerical investigation and the results are presented in Table 2. We find from Table 2 that the bias of $v_2(r)$ decreases monotonically as g increases for fixed m so that the bias is minimum when g=n.

We now compare the bias of $v_2(r)$ with that of the customary variance estimator $v_1(r)$ given by (15). The absolute values of the coefficients c_3 and c_4 in the formula for Bias $[v_1(r)]$ decrease as m (≥ 3) increases. These coefficients have been calculated to compare with those in the formula for the bias of $v_2(r)$ and are presented in Table 2. The bias of $v_1(r)$ should be compared with that of $v_2(r)$ (with g=n) since the bias of $v_2(r)$ is minimum when g=n. From Table 2 we find that the absolute bias of $v_2(r)$ with g=n is less than that of $v_1(r)$ for $n \geq 6$ whenever m>8.

5. Stability of the Jack-Knife Variance Estimator.

In this section we investigate the stability of the Jack-Knife variance estimator $v_2(r)$ under the model of section 3 and compare it with that of $v_1(r)$. The variance of $v_2(r)$ is defined by

$$V[v_2(r)] = E[v_2(r)]^2 - E^2[v_2(r)]. \quad (28)$$

For the case of g=2, the means obtained from half-samples are independent and therefore the variance formulas are relatively simple. We have

$$v_2(r) = \frac{1}{4}[\alpha^2(\frac{1}{\bar{x}'_1} - \frac{1}{\bar{x}'_2})^2 + (\frac{\bar{u}'_1}{\bar{x}'_1} - \frac{\bar{u}'_2}{\bar{x}'_2})^2 + 2\alpha(\frac{1}{\bar{x}'_1} - \frac{1}{\bar{x}'_2}) (\frac{\bar{u}'_1}{\bar{x}'_1} - \frac{\bar{u}'_2}{\bar{x}'_2})] \quad (29)$$

where $\bar{x}'_1$ and $\bar{x}'_2$ are means of first and second half-samples respectively and they are independent gamma variates each with parameter m/2. Therefore, we have

$$E(\frac{1}{\bar{x}'^t_1}) = E(\frac{1}{\bar{x}'^t_2}) = \frac{1}{\prod_{i=1}^{t}(m-2i)}; \ t \geq 1.$$

Consequently the expected value of $v_2(r)$ is

$$E[v_2(r)] = \frac{\alpha^2}{(m-2)^2(m-4)} + \frac{\delta}{(m-2)(m-4)} . \quad (30)$$

The bias of $v_2(r)$ as an estimator of $V(r)$ is

$$\text{Bias}[v_2(r)]= \frac{(4m-7)\alpha^2}{(m-1)^2(m-2)^2(m-4)} + \frac{3\delta}{(m-1)(m-2)(m-4)} . \quad (31)$$

Thus the bias of $v_2(r)$ with g=2 decreases as m increases.

Now from (29) we have

$$16E[v_2(r)]^2= E[\alpha^4(\frac{1}{\bar{x}'_1} - \frac{1}{\bar{x}'_2})^4 + (\frac{\bar{u}'_1}{\bar{x}'_1} - \frac{\bar{u}'_2}{\bar{x}'_2})^4 + 6\alpha^2(\frac{1}{\bar{x}'_1} - \frac{1}{\bar{x}'_2})^2(\frac{\bar{u}'_1}{\bar{x}'_1} - \frac{\bar{u}'_2}{\bar{x}'_2})^2].$$

Table 2. The coefficients c_1, c_2, c_3 and c_4 in Bias $[v_2(r)] = c_1\alpha^2 + c_2\delta$ and Bias $[v_1(r)] = c_3\alpha^2 + c_4\delta$ for selected values of m and g.

m	g	$\lvert c_1\rvert$x10^6	$\lvert c_2\rvert$x10^6	m	$\lvert c_3\rvert$x10^6	$\lvert c_4\rvert$x10^6
8	2*	3543	17857			
8	4	1961	9082	8	1665	6448
8	8	1500	6513			
10	2	1061	6944	10	634	2980
10	10	479	2573			
12	2	423	3409	12	292	1612
12	4	273	1931			
12	6	233	1571			
12	12	197	1261			
16	2	108	1190	16	88	626
16	4	74	699			
16	8	58	514			
16	16	52	435			
20	2	39	548	20	35	305
20	10	21	227			
20	20	19	198			
24	2	17	296	24	16	171
24	6	11	147			
24	12	9	119			
24	24	9	106			
32	2	5	115	32	5	69
32	16	3	44			
32	32	2	41			

*Note: formula for Bias $[v_2(r)]$ for g = 2 is given in section 4.

On Simplification, this reduces to

$$E[v_2(r)]^2 = \frac{3\alpha^4}{(m-2)^2(m-4)^2(m-6)(m-8)} + \frac{3(m^2-10m+28)\delta^2}{(m-2)^2(m-4)^2(m-6)(m-8)} + \frac{6\alpha^2\delta}{(m-2)(m-4)^2(m-6)(m-8)} . \qquad (32)$$

We can obtain $E^2[v_2(r)]$ from (30). Finally the variance of $v_2(r)$ is obtained as

$$V[v_2(r)] = \alpha^4\left[\frac{3}{(m-2)^2(m-4)^2(m-6)(m-8)} - \frac{1}{(m-2)^4(m-4)^2}\right] + \frac{2(m^2-8m+18)\delta^2}{(m-2)^2(m-4)^2(m-6)(m-8)} + \frac{4(m^2+m-18)\alpha^2\delta}{(m-2)^3(m-4)^2(m-6)(m-8)} . \qquad (33)$$

From (31) and (33) we can obtain MSE of $v_2(r)$ as

$$MSE[v_2(r)] = c_8\delta^2 + c_9\alpha^4 + c_{10}\alpha^2\delta \quad \text{(say)} \qquad (34)$$

where the coefficients c_8, c_9 and c_{10} are functions of m(>8) only.

We have evaluated the coefficients in (34) for selected values of m and those in MSE of

Table 3. The coefficients $c_5 \ldots c_{10}$ in MSE Formulas, $MSE[v_1(r)] = c_5\delta^2 + c_6\alpha^4 + c_7\alpha^2\delta$, $MSE[v_2(r)] = c_8\delta^2 + C_9\alpha^4 + c_{10}\alpha^2\delta$ for selected values of m and n.

m	n	$c_5 x10^6$	$c_6 x10^{10}$	$c_7 x10^{10}$	m	$c_8 x10^6$	$c_9 x10^{10}$	$c_{10} x10^{10}$
10	2	251	16736	510772	10	4184	1571043	25103949
10	10	37	6376	125380				
12	2	116	5420	196618	12	871	181480	3622623
12	4	41	2593	80736				
12	6	26	2028	57560				
12	12	13	1566	38597				
16	2	35	966	44448	16	130	11594	323972
16	4	12	406	17153				
16	8	5	246	9354				
16	16	3	182	6235				
20	2	14	259	14163	20	37	1795	64537
20	10	2	49	2326				
20	20	0.8	36	1547				
24	2	7	88	5586	24	14	444	19336
24	6	13	22	1332				
24	12	0.6	13	752				
24	24	0.3	10	500				
32	2	2	16	1294	32	4	52	3154
32	16	0.1	2	128				
32	32	0.07	1	85				

$v_1(r)$, given by (19), for selected values of m and n. They are presented in Table 3. It will be seen from Table 3 that the MSE of $v_2(r)$ with g=2 is considerably larger than that of $v_1(r)$. We conclude that the Jack-Knife variance estimator $v_2(r)$ with g=2 is not very stable. At the present time the investigation of the stability of $v_2(r)$ for general g is in progress and the results will be reported in a subsequent paper.

References

Hansen, M. H., Hurwitz, W. N. and Madow, W. G. (1953). "Sample Survey Methods and Theory", Vol. I, John Wiley and Sons, New York.

Kendall, M. G. and Stuart, A. (1958). "The Advanced Theory of Statistics", Vol. I, Hafner Publishing Company, New York.

Kokan, A. R. (1963). "A Note On the Stability of the Estimates of Standard Errors of the Ordinary Mean Estimate and the Ratio Estimate", Calcutta Statist. Assoc. Bull. 12, 149-58.

Lauh, E. and Williams, W. H. (1963). "Some Small Sample Results for the Variance of a Ratio", Proc. Amer. Statist. Assoc., (Social Statistics Section), 273-83.

Rao, J. N. K. and Beegle, L. D. (1966). "A Monte Carlo Study of Some Ratio Estimators", Proc. Amer. Statist. Assoc., (Social Statistics Section), 443-450.

Rao, J. N. K. and Webster, J. T. (1966). "On Two Methods of Bias Reduction in the Estimation of Ratios", Biometrika, 53, 571-577.

Tukey, J. W. (1958). "Bias and Confidence In Not-Quite Large Samples", Ann. Math. Statist. (Abstract), 29, 614.

This research is supported by the National Science Foundation, under Grant GP-6400.

THE ESTIMATION OF DOMAIN SIZES WHEN SAMPLING FRAMES ARE INTERLOCKING

Robert S. Cochran, Statistics Department, University of Wyoming

I. INTRODUCTION

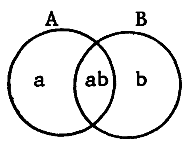
The concept of interlocking sampling frames has been discussed by Hartley (1962) and by Cochran (1964). In these papers estimates of general y characteristics have been presented for situations where two frames overlap and the population frequency in each of the three domains is known. Hartley (1962) also discussed the estimation of the population total for y when the domain frequencies are not known. The purpose of this paper is to consider the estimation of the number in the domains created by the overlapping of two or three frames.

Bryant and King (1960) treated the problem when three frames overlap by using the modified minimum chi-square estimation technique. In the research for this paper some corrections in their approach were made and some comparisons were also made between their estimators and those obtained using multiple-frame procedures.

II. SIMPLE RANDOM SAMPLING FROM EACH OF TWO FRAMES

A. Sample Size Given and Weights to be Determined

The two frames being sampled are frames A and B with frequencies N_A and N_B respectively. From these frames two independent random samples of size n_A and n_B are selected without replacement. Because there are some members of the population in both frames the three domains a, b and ab are created by the use of the two frames. Those members of the population that are just in frame A are in domain a, those that are just in frame B are in domain b, and those that belong to two frames are in domain ab.



After the samples have been drawn the elements selected are classified into their proper domain. In this way we have

$$n_a + n'_{ab} = n_A \qquad \text{and} \qquad n''_{ab} + n_b = n_B \ .$$

From this information the estimates of N_a, N_b and N_{ab} are to be computed. It is obvious that since N_A and N_B are known it is only necessary to concentrate on the estimation of one of these with the other two being obtained by subtraction.

Without loss of generality we will estimate N_{ab} directly and obtain the estimates of N_a and N_b by subtraction as

$$\hat{N}_a = N_A - \hat{N}_{ab} \qquad \text{and} \qquad \hat{N}_b = N_B - \hat{N}_{ab} \ .$$

The number of distinct elements in the population can also be estimated as

$$\hat{N} = N_A + N_B - \hat{N}_{ab} \ .$$

When confronted with the information from two independent samples about the relative frequency of the overlap area it seems only natural that the best way to estimate the frequency of this area is by combining this information. From frame A there is

$$\hat{N}'_{ab} = \frac{N_A}{n_A} n'_{ab}$$

and from frame B there is

$$\hat{N}''_{ab} = \frac{N_B}{n_B} n''_{ab} \ .$$

If the sampling fractions, n_i/N_i, are small enough and the populations are large enough these estimates have variances

$$V(\hat{N}'_{ab}) = \frac{N_A^2}{n_A}\left(\frac{N_{ab}}{N_A}\right)\left(1 - \frac{N_{ab}}{N_B}\right) = \frac{N_A^2}{n_A}\alpha(1-\alpha)$$

and

$$V(\hat{N}''_{ab}) = \frac{N_B^2}{n_B}\left(\frac{N_{ab}}{N_B}\right)\left(1 - \frac{N_{ab}}{N_B}\right) = \frac{N_B^2}{n_B}\beta(1-\beta) \ .$$

Combining these two independent estimators yields the multiple frame estimator

$$\hat{N}_{ab} = p\hat{N}'_{ab} + q\hat{N}''_{ab}, \qquad \text{where} \qquad p+q = 1 \ .$$

While it has been developed here as a weighted average of two independent estimators it can also be developed by using the multiple frame approach of Hartley (1962).

Using the well-known principle of linear functions of independent random variables the variance of $\hat{N}_{ab}$ is easily seen to be

$$V(\hat{N}_{ab}) = p^2 V(\hat{N}'_{ab}) + q^2 V(\hat{N}''_{ab})$$
$$= p^2 \frac{N_A^2}{n_A}\alpha(1-\alpha) + q^2 \frac{N_B^2}{n_B}\beta(1-\beta) \ .$$

The value of p that will minimize this variance is

$$p_0 = \frac{V(\hat{N}''_{ab})}{V(\hat{N}'_{ab}) + V(\hat{N}''_{ab})} \ .$$

Cochran (1965) shows some results of using a nonoptimum p in terms of the loss in precision for the estimator.

Bryant and King (1960) did not deal with two-frame situations. However, Cochran (1965) did use their procedure and came up with the result that for two frames the estimators are algebraically equivalent and have the same estimates of their variances when p_0 was estimated from the sample information.

B. Sample Size and Weights to be Determined

In the application of these techniques to most surveys the estimation of the number in the separate domains is only one of several pieces of information desired from the survey. With this in mind the sample sizes drawn are usually selected to give maximum information on some other variable or to satisfy some other restraint. However, when the estimation of the frequency in each of the several domains is either the only quantity of interest or is the most important quantity some interesting results can be given.

In the case of two frames where the quantity N_{ab} is to be estimated it was previously given that

$$\hat{N}_{ab} = p\hat{N}'_{ab} + q\hat{N}''_{ab}$$

and

$$V(\hat{N}_{ab}) = p^2 \frac{N_A^2}{n_A}\alpha(1-\alpha) + q^2 \frac{N_B^2}{n_B}\beta(1-\beta) .$$

Assuming n_A and n_B to have been previously determined the optimum value for p becomes

$$p_0 = \frac{\frac{N_B^2}{n_B}\beta(1-\beta)}{\frac{N_A^2}{n_A}\alpha(1-\alpha) + \frac{N_B^2}{n_B}\beta(1-\beta)} .$$

Now setting the partial derivative of $V(N_{ab})$ with respect to n_A and n_B subject to a cost condition

$$C = n_A C_A + n_B C_B$$

equal to zero yields

$$n_A = p\left(\frac{N_A^2\,\alpha(1-\alpha)}{C_A}\right)^{1/2} \quad \text{and} \quad n_B = q\left(\frac{N_B^2\,\beta(1-\beta)}{C_B}\right)^{1/2} .$$

The constant λ can be shown to be

$$\lambda = C^{-1}[p(N_{ab}N_a C_A)^{1/2} + q(N_{ab}N_b)^{1/2}]$$

Substituting the optimum n_A and n_B into the expression for the optimum p yields

$$p = \frac{p[\frac{1}{N_A^2\,\alpha(1-\alpha)C_A}]^{1/2}}{p[\frac{1}{N_A^2\,\alpha(1-\alpha)C_A}]^{1/2} + q[\frac{1}{N_B^2\,\beta(1-\beta)C_B}]^{1/2}} .$$

This expression when solved for p has two solutions, $p = 0$ and $p = 1$, unless

$$N_B^2\,\beta(1-\beta)C_B = N_A^2\,\alpha(1-\alpha)C_A .$$

In this unlikely case any value for p, $0 \leq p \leq 1$ will be a solution.

In order to determine conditions that will indicate which of the values, $p = 0$ or $p = 1$, actually gives the minimum variance consider the variance equation under each of these choices.

When $p = 1$ the sample size should be

$$n_A = \frac{C}{C_A}$$

and the variance becomes

$$V_1 = \frac{N_A^2\,\alpha(1-\alpha)C_A}{C}$$

When $p = 0$ the sample size should be

$$n_B = \frac{C}{C_B}$$

and the variance becomes

$$V_0 = \frac{N_B^2\,\beta(1-\beta)C_B}{C} .$$

Thus the question can be settled by considering the relationship between $\frac{1-\beta}{\beta}C_B$ and $\frac{1-\alpha}{\alpha}C_A$ or $N_b C_B$ and $N_a C_A$. When $\frac{1-\beta}{\beta}C_B > \frac{1-\alpha}{\alpha}C_A$, then V_1 is the smaller. When all sampling costs are the same ($C_A = C_B$) the above can be shown to imply that $N_B > N_A$. Therefore with equal costs of sampling, sample entirely from the smallest frame.

When the sampling costs are not the same the relationship between $C_B N_b$ and $C_A N_a$ can sometimes be derived from the relationships between C_B and C_A, N_B and N_A, and $C_B N_B$ and $C_A N_A$. Whenever

$$N_A > N_B \quad \text{and} \quad C_A \geq C_B$$

then

$$N_A C_A > N_B C_B \quad \text{and} \quad N_a C_A > N_b C_B$$

and the obvious decision is to sample from the small cheap frame. Whenever

$$C_B > C_A \quad \text{and} \quad N_A C_A \geq N_B C_B \quad \text{then} \quad C_A N_a > C_B N_b .$$

The third possibility is

$$C_B > C_A \quad \text{and} \quad N_B C_B > N_A C_A .$$

In this case the result depends upon the unknown N_{ab}. $C_A N_a$ will be larger whenever

$$\frac{N_B C_B - N_A C_A}{C_B - C_A} < N_{ab} .$$

III. SIMPLE RANDOM SAMPLING FROM THREE FRAMES

When sampling from three frames, A, B and C, there are seven domains. Of these, it is only necessary to directly estimate the number of units in four. The number of units in the remaining three can be estimated by subtraction from the known domain sizes. Without loss of generality let these four be N_{ab}, N_{ac}, N_{bc} and N_{abc}, the number of units in the areas of overlap.

Using an obvious extension of the notation and procedures of the two-frame case above the following estimates are obtained:

$$\hat{N}_{ab} = p_{ab}\frac{N_A}{n_A}n'_{ab} + q_{ab}\frac{N_B}{n_B}n''_{ab}$$

$$\hat{N}_{ac} = p_{ac}\frac{N_A}{n_A}n'_{ac} + q_{ac}\frac{N_C}{n_C}n''_{ac}$$

$$\hat{N}_{bc} = p_{bc} \frac{N_B}{n_B} n'_{bc} + q_{bc} \frac{N_C}{n_C} n''_{bc}$$

and

$$\hat{N}_{abc} = p_A \frac{N_A}{n_A} n'_{abc} + p_B \frac{N_B}{n_B} n''_{abc} + p_C \frac{N_C}{n_C} n'''_{abc} .$$

The variances of the quantities are

$$V(\hat{N}_{ab}) = p_{ab}^2 \frac{N_A^2}{n_A} \alpha_1(1-\alpha_1) + q_{ab}^2 \frac{N_B^2}{n_B} \alpha_2(1-\alpha_2),$$

$$\alpha_1 = \frac{N_{ab}}{N_A}; \quad \alpha_2 = \frac{N_{ab}}{N_B}$$

$$V(\hat{N}_{ac}) = p_{ac}^2 \frac{N_A^2}{n_A} \gamma_1(1-\gamma_1) + q_{ac}^2 \frac{N_C^2}{n_C} \gamma_2(1-\gamma_2),$$

$$\gamma_1 = \frac{N_{ac}}{N_A}; \quad \gamma_2 = \frac{N_{ac}}{N_C}$$

$$V(\hat{N}_{bc}) = p_{bc}^2 \frac{N_B^2}{n_B} \beta_1(1-\beta_1) + q_{bc}^2 \frac{N_C^2}{n_C} \beta_2(1-\beta_2),$$

$$\beta_1 = \frac{N_{bc}}{N_B}; \quad \beta_2 = \frac{N_{bc}}{N_C}$$

and

$$V(\hat{N}_{abc}) = p_A^2 \frac{N_A^2}{n_A} \delta_1(1-\delta_1) + p_B^2 \frac{N_B^2}{n_B} \delta_2(1-\delta_2) + p_C^2 \frac{N_C^2}{n_C} \delta_3(1-\delta_3), \qquad \delta_1 = \frac{N_{abc}}{N_A};$$

$$\delta_2 = \frac{N_{abc}}{N_B}; \quad \delta_3 = \frac{N_{abc}}{N_C} .$$

The values of the p's that minimize these variances are:

$$p_{ab} = \frac{V(\hat{N}''_{ab})}{V(\hat{N}'_{ab}) + V(\hat{N}''_{ab})}, \qquad q_{ab} = 1-p_{ab}$$

$$p_{ac} = \frac{V(\hat{N}''_{ac})}{V(\hat{N}_{ac}) + V(\hat{N}''_{ac})}, \qquad q_{ac} = 1-p_{ac}$$

$$p_{bc} = \frac{V(\hat{N}''_{bc})}{V(\hat{N}'_{bc}) + V(\hat{N}''_{bc})}, \qquad q_{bc} = 1-p_{bc}$$

$$p_A = \frac{\frac{1}{V(\hat{N}'_{abc})}}{\frac{1}{V(\hat{N}'_{abc})} + \frac{1}{V(\hat{N}''_{abc})} + \frac{1}{V(\hat{N}'''_{abc})}}$$

$$= \frac{V(\hat{N}''_{abc})\,V(\hat{N}'''_{abc})}{V(\hat{N}'_{abc})V(\hat{N}''_{abc})+V(\hat{N}'_{abc})V(\hat{N}'''_{abc})+V(\hat{N}''_{abc})V(\hat{N}'''_{abc})}$$

$$p_B = \frac{\frac{1}{V(\hat{N}''_{ab})}}{\frac{1}{V(\hat{N}'_{abc})} + \frac{1}{V(\hat{N}''_{abc})} + \frac{1}{V(\hat{N}'''_{abc})}}$$

$$p_C = \frac{\frac{1}{V(\hat{N}'''_{abc})}}{\frac{1}{V(\hat{N}'_{abc})} + \frac{1}{V(\hat{N}''_{abc})} + \frac{1}{V(\hat{N}'''_{abc})}} .$$

The estimation of quantities such as N_a can now be carried out by subtraction of the estimates of N_{ab}, N_{ac}, and N_{abc} from the known frame size N_A,

$$\hat{N}_a = N_A - (\hat{N}_{ab} + \hat{N}_{ac} + \hat{N}_{abc}) .$$

Cochran (1965) shows the variance of $\hat{N}_a$ to be

$$\begin{aligned} V(\hat{N}_a) = \frac{N_A^2}{n_A} \Big\{ & p_{ab}^2\, \alpha_1(1-\alpha_1) + p_{ac}^2\, \gamma_1(1-\gamma_1) \\ & + p_A^2\, \delta_1(1-\delta_1) - 2\ 2p_{ab}\, p_{ac}\, \alpha_1\gamma_1 \\ & - 2\ p_{ab}p_A\alpha_1\delta_1 - 2\ p_{ac}p_A\gamma_1\delta_1 \Big\} \\ & + \frac{N_B^2}{n_B} \Big\{ q_{ab}^2\, \alpha_2(1-\alpha_2) + p_B^2\, \delta_2(1-\delta_2) \\ & \quad - 2\ p_B\, q_{ab}\, \alpha_2\, \delta_2 \Big\} \\ & + \frac{N_C^2}{n_C} \Big\{ q_{ac}^2\, \gamma_2(1-\gamma_2) + p_C^2\, \delta_3(1-\delta_3) \\ & \quad - 2\ p_C\, q_{ac}\, \gamma_2\, \delta_3) \Big\} . \end{aligned}$$

As an example of the multiple-frame approach consider the 1964 data of the following table.

License Frame	1964 Population			Est.	Var.
	Deer	Elk	A'lope		
Deer only	1,637			22,425	
Elk only		765		7,277	
Antelope only			278	1,915	
Deer-Elk	1,023	1,402		13,222	44,830
Deer-Antelope	353		549	4,454	16,704
Elk-Antelope		107	115	954	4,171
Deer-Elk-Ant.	48	720	768	6,372	14,004
Total Sample	3,497	2,994	1,710		
Pop. Size	46,473	27,825	13,695	56,619	

For the deer-elk overlap

$$\hat{N}'_{de} = \frac{46,473}{3,497} \cdot 1023 = 13,596$$

$$\hat{N}''_{de} = \frac{27,825}{2,994} \cdot 1402 = 13,030$$

$$v(\hat{N}'_{de}) = 145,074; \quad v(\hat{N}''_{de}) = 64,416$$

$$p_{de} = .34 .$$

For the deer-elk-antelope overlap

$$\hat{N}'_{dea} = \frac{46,473}{3,497} \cdot 484 = 6,433$$

$$\hat{N}''_{dea} = \frac{27,825}{2,994} \cdot 720 = 6,692$$

$$\hat{N}'''_{dea} = \frac{13,695}{1,710} \cdot 768 = 6,151$$

$$v(\hat{N}'_{dea}) = 74,359; \quad v(\hat{N}''_{dea}) = 47,168;$$

$$v(\hat{N}'''_{dea}) = 27,146; \; p_D = .9; \; p_E = .31; \; p_A = .50 .$$

The approach of Bryant and King (1960) leads to modified minimum chi-square estimates for the three-frame case, the solution of four simultaneous equations in four unknowns. In matrix notation this is

$$A\hat{N} = Y$$
$$\hat{N} = A^{-1} Y$$

The A matrix of coefficients is made up of the partial derivations of

$$\begin{aligned}
\chi^2 = {} & \frac{\left(n'_{ab} - \frac{n_A}{N_A} N_{ab}\right)^2}{n'_{ab}} + \frac{\left(n''_{ab} - \frac{n_B}{N_B} N_{ab}\right)^2}{n''_{ab}} \\
& + \frac{\left(n'_{ac} - \frac{n_A}{N_A} N_{ac}\right)^2}{n'_{ac}} + \frac{\left(n''_{ac} - \frac{n_C}{N_C} N_{ac}\right)^2}{n''_{ac}} \\
& + \frac{\left(n'_{bc} - \frac{n_B}{N_B} N_{bc}\right)^2}{n'_{bc}} + \frac{\left(n''_{bc} - \frac{n_C}{N_C} N_{bc}\right)^2}{n''_{bc}} \\
& + \frac{\left(n'_{abc} - \frac{n_A}{N_A} N_{abc}\right)^2}{n'_{abc}} + \frac{\left(n''_{abc} - \frac{n_B}{N_B} N_{abc}\right)^2}{n''_{abc}} \\
& + \frac{\left(n'''_{abc} - \frac{n_C}{N_C} N_{abc}\right)^2}{n'''_{abc}}
\end{aligned}$$

with respect to N_{ab}, N_{ac}, N_{bc} and N_{abc}. The Y vector contains the constants arising from the differentiation process. They derive the variance of these estimates to be of the form

$$[F]\ [\sigma_{rs}]\ [F]'$$

where

$$[F] = F'_{n'_{ab}},\ F'_{n''_{ab}},\ \ldots,\ F'_{n'''_{abc}} \quad {}_{4x9}$$

$$F'_{n_i} = -A^{-1} \frac{\partial A}{\partial n_i} A^{-1} Y + A^{-1} \frac{\partial Y}{\partial n_i}$$

and $[\sigma_{rs}]$ is the 9 x 9 variance-covariance matrix for the number of observations appearing in the overlap areas.

No specific analytical comparisons were made between these estimates and their variances with the multiple-frame type of estimates. However, some numerical comparisons were made using information from the 1960 through 1964 big game studies conducted for the Wyoming Game and Fish Commission by the University of Wyoming.

The figure for these five years indicates close agreement between the two estimators. However, in all but one instance (1962, E-2D) the estimate of the variance of the multiple-frame estimate was less than the estimate of the variance of the minimum chi-square estimate. In some cases there was an appreciable gain using the multiple-frame estimator.

Year	Class	Estimate: Multiple Frame	Estimate: Minimum χ^2	Variances: Multiple Frame	Variances: Minimum χ^2
1960	D-E	14,333	14,397	58,452	66,461
	D-2D	1,446	1,408	8,631	12,500
	E-2D	454	439	2,420	2,720
	D-E-2D	1,466	1,442	5,545	8,555
1961	D-E	16,006	15,865	68,937	73,259
	D-2D	1,332	1,417	7,792	10,660
	E-2D	243	237	1,882	2,194
	D-E-2D	1,179	1,172	5,040	7,883
1962	D-E	11,633	11,272	108,013	115,100
	D-2D	1,048	1,035	11,861	12,583
	E-2D	2,124	2,015	20,893	15,544
	D-E-2D	1,475	1,332	11,833	14,017
1963	D-E	12,179	11,852	81,635	136,079
	D-A	5,629	5,794	27,981	125,882
	E-A	875	880	6,316	8,567
	D-E-A	6,104	6,058	22,626	91,786
1964	D-E	13,222	13,450	44,830	86,697
	D-A	4,454	4,379	16,704	61,979
	E-A	954	925	4,171	5,350
	D-E-A	5,372	6,299	14,004	57,822

IV. LITERATURE CITED

Bryant, Edward C. and King, Donald W. (1960). "Estimation from Populations Identified by Overlapping Sample Frame." Unpublished paper presented at the American Statistical Association meeting, Palo Alto, California.

Cochran, Robert S. (1964). "Multiple Frame Sample Surveys." Proceedings of Social Science Section of American Statistical Association Meetings, Chicago, Illinois.

________. (1965). "Theory and Application of Multiple Frame Surveys." Unpublished Ph.D. thesis, Iowa State University, Ames, Iowa.

Hartley, H. O. (1962). "Multiple Frame Surveys." Proceedings of the Social Science Section of the American Statistical Association Meetings, Minneapolis, Minnesota.

MULTIPLE CONTRAST METHODS FOR ANALYTICAL SURVEYS[I]

Leonard A. Marascuilo, University of California at Berkeley

Maryellen McSweeney, Michigan State University

Introduction

Multistage sampling procedures and nonresponse of sampled units frequently make the analysis of data generated by analytical surveys extremely difficult. If there are only two independent subpopulations of particular interest, the analysis of survey data is not especially complex since the subpopulation parameters may be estimated from the data. If in addition, the sample sizes are large, it is always possible to test the null hypothesis,

$$H_o: \theta_1 = \theta_2$$

against the alternative,

$$H_1: \theta_1 \neq \theta_2$$

by computing

$$Z = \frac{\hat{\theta}_1 - \hat{\theta}_2}{\sqrt{SE^2_{\hat{\theta}_1} + SE^2_{\hat{\theta}_2}}}$$

and referring the observed Z value to standard normal curve tables. If H_o is rejected, a point estimate of the parameter difference is given by $(\hat{\theta}_1 - \hat{\theta}_2)$ and a $(1 - \alpha)\%$ interval estimate is given by

$$(\hat{\theta}_1 - \hat{\theta}_2) \pm Z_{\frac{\alpha}{2}} \sqrt{SE^2_{\hat{\theta}_1} + SE^2_{\hat{\theta}_2}}$$

For studies involving more than two subpopulations comparable analytical methods have not been reported. Investigation of the technical literature shows that Gold (1963) and Goodman (1964) have extended the simultaneous confidence interval method of Scheffé (1959) to certain special cases associated with the parameters of contingency tables and the parameters of Markov Chains. Marascuilo (1966) has extended their model to include multiple confidence intervals for correlation coefficients and for sample averages from analysis of variance designs in which the variances are unequal. Since analogous situations occur in survey research studies, this extension should be of considerable value in the analysis of survey data.

In this paper a proof of the chi-square analog of Scheffé's Theorem is given. From the results of this proof a simple-to-compute test statistic is proposed for the test of the hypothesis

$$H_o: \theta_1 = \theta_2 = \dots = \theta_K = \theta_o$$

against the alternative

$$H_1: H_o \text{ is false.}$$

The post hoc multiple comparison procedures associated with the rejection of the hypothesis H_o are indicated. These methods are used to test and identify the sources of differences in attitudes expressed by adult citizens toward the integration of de facto segregated schools in three different socio-economic subpopulations of an urban American community.

Chi-square Analog of Scheffé's Theorem

Consider a univariate model in which there are K treatments, conditions, or populations. Let the parameters of the model be represented by $\underset{\sim}{\theta}' = (\theta_1, \theta_2, \dots, \theta_K)$. Let $\hat{\underset{\sim}{\theta}}' = (\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_K)$ be a set of large sample efficient estimators of these unknown parameters. Furthermore, let the covariance matrix for these estimators be of rank $q \leq K$. It is known from large sample theory that

$$U = (\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta})' (\underbrace{\mathrm{Cov}\ (\hat{\theta})})^{-1} (\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta})$$

has an asymptotic chi-square distribution with q degrees of freedom since it is the exponent in the asymptotic K-variate normal distribution of $\hat{\underset{\sim}{\theta}}' = (\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_K)$. Since the rank of U is q, it is possible to find a set of q linearly independent estimable functions

$$\psi_p = \sum_{k=1}^{K} a_k \theta_k = \underset{\sim}{a}' \underset{\sim}{\theta}$$

that will span the space of all contrasts of the form

$$\psi = \sum_{p=1}^{q} c_p \psi_p = \underset{\sim}{c}' \underset{\sim}{\psi} \qquad p = 1, 2, \dots, q$$

Let the set of all possible contrasts of the form ψ be denoted by L. A set of estimates for the ψ_p is

$$\hat{\psi}_p = a_1 \hat{\theta}_1 + a_2 \hat{\theta}_2 + \dots + a_K \hat{\theta}_K = \underset{\sim}{a}' \hat{\underset{\sim}{\theta}}$$

Since the $\hat{\psi}_p$ are linearly independent functions of asymptotically multi-variate normal random variables with a covariance matrix of rank q, they are also asymptotically multivariate normal with exponent given by

$$Q = (\hat{\underset{\sim}{\psi}} - \underset{\sim}{\psi})' (\underbrace{\mathrm{Cov}\ (\hat{\psi})})^{-1} (\hat{\underset{\sim}{\psi}} - \underset{\sim}{\psi})$$

As a result Q must have an asymptotic chi-square distribution with q degrees of freedom. Thus, a $(1 - \alpha)\%$ confidence ellipsoid for the point $\underset{\sim}{\psi}' = (\psi_1, \psi_2, \ldots, \psi_q)$ is given by $Q \leq \chi_q^2 (1 - \alpha)$.

This confidence ellipsoid serves as the basis for the analog of Scheffé's Theorem. The proof of this theorem parallels, as one would expect, the proof of Scheffé's Theorem. The notation used is that of Scheffé so that the two proofs may be easily compared. The analog of Scheffé's Theorem reads as follows:

<u>Theorem</u>. The probability is $(1 - \alpha)$ in the limit that simultaneously for all $\psi \in L$

$$\hat{\psi} - \sqrt{\chi_q^2 (1-\alpha)} \sqrt{\mathrm{Var}(\hat{\psi})} \leq \psi \leq \hat{\psi} + \sqrt{\chi_q^2 (1-\alpha)} \sqrt{\mathrm{Var}(\hat{\psi})}$$

<u>Proof</u>. The inequality that defines the asymptotic confidence ellipsoid for the point $\underset{\sim}{\psi}' = (\psi_1, \psi_2, \ldots, \psi_q)$ is

$$(\underset{\sim}{\hat{\psi}} - \underset{\sim}{\psi})' \, (\underbrace{\mathrm{Cov}\ (\hat{\psi})})^{-1} (\underset{\sim}{\hat{\psi}} - \underset{\sim}{\psi}) \leq \chi_q^2 \ (1 - \alpha)$$

The point $\underset{\sim}{\psi}'$ is in the ellipsoid if, and only if, it lies between all pairs of parallel planes of support of the ellipsoid. If $\underset{\sim}{C}' = (C_1, C_2, \ldots, C_q)$ is an arbitrary nonzero vector, Scheffé has shown that the point $\underset{\sim}{\psi}'$ lies between the two planes of support of the ellipsoid orthogonal to $\underset{\sim}{C}$ if, and only if,

$$\left| \underset{\sim}{C}' \, (\underset{\sim}{\psi} - \underset{\sim}{\hat{\psi}}) \right| \leq \sqrt{\underset{\sim}{C}' \, \underset{\sim}{M}^{-1} \, \underset{\sim}{C}}$$

In this case,

$$\underset{\sim}{M} = \frac{1}{\chi_q^2 \ (1-\alpha)} (\underbrace{\mathrm{Cov}\ (\hat{\psi})})^{-1}$$

Thus

$$\underset{\sim}{C}' \underset{\sim}{M}^{-1} \underset{\sim}{C} = \underset{\sim}{C}' \left(\frac{1}{\chi_q^2 \ (1-\alpha)} (\underbrace{\mathrm{Cov}\ (\hat{\psi})})^{-1} \right)^{-1} \underset{\sim}{C}$$

$$= \underset{\sim}{C}' \chi_q^2 \ (1-\alpha) \ \underbrace{\mathrm{Cov}\ (\hat{\psi})} \ \underset{\sim}{C}$$

$$= \chi_q^2 \ (1-\alpha) \ \left[\underset{\sim}{C}' \underbrace{\mathrm{Cov}\ (\hat{\psi})} \ \underset{\sim}{C} \right]$$

Since any contrast ψ in L can be estimated by

$$\hat{\psi} = \sum_{p=1}^{q} c_p \hat{\psi}_p = \underset{\sim}{c}' \underset{\sim}{\hat{\psi}},$$

the variance of the estimate is given by

$$\mathrm{Var}\ (\hat{\psi}) = \underset{\sim}{c}' \ \underbrace{\mathrm{Cov}\ (\hat{\psi})} \ \underset{\sim}{c}.$$

Therefore

$$\underset{\sim}{C}' \underset{\sim}{M}^{-1} \underset{\sim}{C} = \chi_q^2 \ (1 - \alpha) \ \mathrm{Var}\ (\hat{\psi})$$

and

$$\sqrt{\underset{\sim}{C}' \underset{\sim}{M}^{-1} \underset{\sim}{C}} = \sqrt{\chi_q^2 \ (1 - \alpha)} \ \sqrt{\mathrm{Var}\ (\hat{\psi})}$$

Since

$$\left| \underset{\sim}{C}' \underset{\sim}{\psi} - \underset{\sim}{C}' \underset{\sim}{\hat{\psi}} \right| = \left| \psi - \hat{\psi} \right|$$

the last inequality actually states that simultaneously for all $\psi \in L$ the probability in the limit is $(1 - \alpha)$ that

$$\left| \psi - \hat{\psi} \right| \leq \sqrt{\chi_q^2 \ (1 - \alpha)} \ \sqrt{\mathrm{Var}\ (\hat{\psi})}$$

This completes the proof.

As with the F-test, the Chi-square test will reject H_0 if, and only if, the estimate $\hat{\psi}$ of at least one ψ is significantly different from zero. Equivalently, if H_0 is rejected there is at least one contrast in the $\hat{\psi}_p$ that is significantly different from zero.

<u>Derivation of the test statistic to test</u>
<u>H_0: $\theta_1 = \theta_2 = \ldots = \theta_K = \theta_0$</u>

In most surveys the subpopulations consist of population strata or domains of investigation that usually comprise mutually exclusive subsets of the total universe of interest. For this reason, the estimates of the parameters within the individual subpopulations are statistically independent, so that $\mathrm{Cov}\ (\hat{\theta}_i, \hat{\theta}_j) = 0$ for $i \neq j$ and

$$U = (\underset{\sim}{\hat{\theta}} - \underset{\sim}{\theta})' \, (\underbrace{\mathrm{Cov}\ (\hat{\theta})})^{-1} \, (\underset{\sim}{\hat{\theta}} - \underset{\sim}{\theta})$$

reduces to

$$\sum_{k=1}^{K} \frac{(\hat{\theta}_k - \theta_k)^2}{\mathrm{Var}\ (\hat{\theta}_k)}$$

which is asymptotic χ^2 with $q = K$.

To test the hypothesis

$$H_0: \theta_1 = \theta_2 = \ldots = \theta_K = \theta_0$$

it is only necessary to evaluate U under H_0 and determine whether or not $U > \chi_K^2 \ (1 - \alpha)$. If U is too large, H_0 is rejected.

For most applications, the exact value of θ_0 is unknown and must be estimated. An easy-to-obtain estimate is the one that minimizes U. This estimate is given by

$$\hat{\theta}_o = \sum_{k=1}^{K} \frac{1}{\text{Var}(\hat{\theta}_k)} \hat{\theta}_k \Bigg/ \sum_{k=1}^{K} \frac{1}{\text{Var}(\hat{\theta}_k)} =$$

$$\sum_{k=1}^{K} W_k \hat{\theta}_k \Bigg/ \sum_{k=1}^{K} W_k$$

If this estimate is substituted into U and if the resulting expression is denoted as U_o, it follows that

$$U_o = \sum_{k=1}^{K} \frac{(\hat{\theta}_k - \hat{\theta}_o)^2}{\text{Var}(\hat{\theta}_k)} = \sum_{k=1}^{K} W_k (\hat{\theta}_k - \hat{\theta}_o)^2$$

Familiar analysis of variance methods can be employed to show that U_o is asymptotically chi-square with (K - 1) degrees of freedom. Therefore, a simple decision rule that may be used for testing H_o is: reject H_o if $U_o > \chi^2_{K-1}(1-\alpha)$ and do not reject H_o if $U_o < \chi^2_{K-1}(1-\alpha)$.

If the variances are unknown and the sample sizes are large, the large sample estimates of the variances can be substituted into the final result with little loss. This also applies to the estimate of $\hat{\theta}_o$ which would then be equal to

$$\hat{\hat{\theta}}_o = \sum_{k=1}^{K} \frac{1}{\text{var}(\hat{\theta}_k)} \hat{\theta}_k \Bigg/ \sum_{k=1}^{K} \frac{1}{\text{var}(\hat{\theta}_k)} =$$

$$\sum_{k=1}^{K} \hat{W}_k \hat{\theta}_k \Bigg/ \sum_{k=1}^{K} \hat{W}_k$$

In addition, the test statistic would be

$$U_o' = \sum_{k=1}^{K} \hat{W}_k (\hat{\theta}_k - \hat{\hat{\theta}}_o)^2$$

Examples

By means of a principal component analysis based on 1960 census data, the 28 census tracts of Berkeley, California, were partitioned into three mutually exclusive subpopulations representing high, medium, and low socio-economic status areas. Within each census tract a two per cent sample of adults was selected. The following sampling procedure was used to obtain independent samples for each census tract. A city block was chosen at random with probability proportional to the block size reported in the 1960 census data, and a simple random sample of six households was taken for the selected block. This process was repeated for additional blocks within a census tract until the number of adults in the sample was estimated to equal two per cent of the 1960 census tract adult population. The sampling procedure was repeated in each of the 28 census tracts. Since the population magnitude varied across census tracts, the sample sizes over census tracts ranged from 25 to 86 adults. The removal of wrong addresses and vacant houses from the sampling frame reduced the actual sampling fraction to 1.88 per cent and the initial sample size to 1,392 adults.

The survey was begun on April 15, 1964, with the mailing of letters and questionnaires to the 742 randomly selected households in the community. On April 29, follow-up letters were mailed to all nonresponding households. Between May 11th, and 19th, a random sample of 1/3 of the remaining nonrespondents was interviewed by trained female personnel from the Survey Research Center of the University of California. Usable information was ultimately obtained from 971 adults of the originally selected sample.

One of the items appearing on the questionnaire read as follows:

> For some (elementary) schools the committee suggested that lines be changed so that the percentage of nonwhite and white children in these schools would be more like the percentage for the entire school system.
>
> (1) ___ I agree
>
> (2) ___ I disagree
>
> (3) ___ I am not sure

The "I am not sure" category of response was excluded in the analysis of the data. The analysis of this item considered the effect of socio-economic status on attitudes toward increasing the racial integration of the schools by means of boundary changes. In particular, it was hypothesized that members of the low SES Negro areas of the community would show the strongest support for the boundary changes designed to effect school integration while the greatest opposition would be expressed by the high SES white areas.

The sampling unit for this survey was the household, but the unit for analysis was the individual respondent. Consequently the number of adults per household was a random variable, the value of which was undetermined until data were obtained for each household. Since the number of adults per household was unknown prior to sampling, the per cent agreeing to the change in school boundaries was estimated by a separate ratio estimate, $\hat{P}_h$, for each census tract. Furthermore separate ratio estimates, $\hat{P}_{hi}$, were required for each wave of response within a single census tract because the responses to the original letter, the follow-up letter, or the personal interview produced an artificial stratification of the respondents for each census tract. Despite the small sample sizes within strata, no appreciable differences between the separate and combined ratio estimates were found. Separate ratio estimates were chosen in preference to combined estimates on the basis of greater simplicity of computation and explication. As a result, the final parameter estimate for each of the three subpopulations defined by principal component analysis involved primary stratification of the census tracts together with the artificial within-tract stratification based on the wave of response.

For a subpopulation defined by principal component analysis:

1. $$\hat{p} = \sum_{h=1}^{L} \left(\frac{n_h}{n}\right) \hat{p}_h$$ $h = 1, 2, \ldots, L$ census tracts

where

2. $$\hat{p}_h = \left(\frac{n_{h1}}{n_h}\right)\hat{p}_{h1} + \left(\frac{n_{h2}}{n_h}\right)\hat{p}_{h2} + \left(\frac{n_h - n_{h1} - n_{h2}}{n_h}\right)\hat{p}_{h3}$$

$i = 1, 2, 3$ waves of response

and

3. $$\hat{p}_{hi} = \frac{\sum_{j=1}^{n_{hi}} a_{hij}}{\sum_{j=1}^{n_{hi}} m_{hij}}$$

$j = 1, 2, \ldots, n_{hi}$ households in census tract h that answer in the ith. wave of response

a_{hij} = number of adults in household j who answered "I agree."

m_{hij} = number of adults in household j

The approximate variance is estimated by

4. $$SE^2_{\hat{p}} = \sum_{h=1}^{L} \left(\frac{n_h}{n}\right)^2 SE^2_{\hat{p}_h}$$

where

5. $$SE^2_{\hat{p}_h} = \left(\frac{n_{h1}}{n_h}\right)^2 SE^2_{\hat{p}_{h1}} + \left(\frac{n_{h2}}{n_h}\right)^2 SE^2_{\hat{p}_{h2}} + \left(\frac{n_h - n_{h1} - n_{h2}}{n_h}\right)^2 SE^2_{\hat{p}_{h3}}$$

and

6. $$SE^2_{\hat{p}_{hi}} = \frac{1}{n_{hi}\ \bar{m}^2_{hi}} \sum_{j=1}^{n_{hi}} \frac{(a_{hij} - \hat{p}_{hi}\, m_{hij})^2}{n_{hi} - 1}$$

In Table 1 the distribution of response to the question by subpopulation is shown. If binomial estimates of the variances are used, the hypothesis of equal proportions agreeing in the three subpopulations will be rejected since $\chi^2 = 94.80$ exceeds $\chi^2_2(.95) = 5.99$. However, binomial estimates and the chi-square test of homogeneity are inappropriate because the responses within a cluster (household) are not independent, but positively correlated.

The appropriate ratio estimates of the parameters for the three subpopulations are given in Table 2. For each subpopulation the estimated variance of the proportion agreeing is considerably larger for the ratio estimate than for the corresponding binomial estimate.

For these data,

$$\hat{\hat{p}}_o = \frac{\sum_{k=1}^{3} \hat{W}_k \hat{p}_k}{\sum_{k=1}^{3} \hat{W}_k}$$ $k = 1, 2, 3$ subpopulations defined by principal component analysis

$$\hat{\hat{p}}_o = \frac{691.1039(.862) + 394.4510(.586) + 652.4490(.348)}{691.1039 + 394.4510 + 652.4490} = .606$$

and

$$U'_o = 691.1039(.862-.606)^2 + 394.4510(.586-.606)^2 + 652.4490(.348-.606)^2$$

$$U'_o = 88.88$$

Table 1: Distribution of Responses by Subpopulation and Binomial Estimates of the Parameters Based on Proportional Allocation

Response	Subpopulation: Low SES	Medium SES	High SES	Total
Agree	170	179	105	454
Disagree	30	126	157	313
Total	200	305	262	767
Per Cent Agreement (binomial estimate)	.850	.586	.401	.592
Variance (binomial estimate)	.00063	.00079	.00091	

Table 2: Ratio Estimates of the Parameters by Subpopulation

Parameter Estimated	Subpopulation: Low SES	Medium SES	High SES	Total
Per Cent Agreement	.862	.586	.348	.606
Variance	.001447	.002535	.001533	
Weight, $\hat{W}_k$	691.1039	394.4510	652.4490	

Where the estimate of the total is given by $\hat{\hat{p}}_o = \dfrac{\sum_{k=1}^{K} \hat{W}_k \hat{p}_k}{\sum_{k=1}^{K} \hat{W}_k} = .606$

Table 3: 95% Confidence Intervals for the Set of Simple Contrasts

Contrast	Value of Contrast $\hat{P}_k - \hat{P}_{k'}$	Estimated Variance of Contrast	Lower Limit of Confidence Interval	Upper Limit of Confidence Interval	Significance
Low vs. Medium	.862-.586	.001447 + .002535	.122	.430	Sig.
Low vs. High	.862-.348	.001447 + .001533	.377	.651	Sig.
Medium vs. High	.586-.348	.002535 + .001533	.082	.394	Sig.

Since $U_0' > \chi^2_2(.95) = 5.99$, H_0 is rejected. Thus there is reason to believe that at least one linear contrast of the parameters is significantly different from zero.

For this study, the general form of the $\binom{3}{2}$ or 3 simple contrasts is given by

$$\hat{\psi}_{kk'} = \hat{p}_k - \hat{p}_{k'}$$

$$k \neq k'$$

with the estimated variance given by

$$\text{var}(\hat{\psi}_{kk'}) = \text{var}(\hat{p}_k) + \text{var}(\hat{p}_{k'})$$

These contrasts and their estimated variances are summarized in Table 3. All three contrasts are statistically significant from zero at the overall .05 level.

Although these hypothesis testing and multiple contrast techniques have been illustrated for the case of three independent subpopulations, their range of possible application in analytical surveys is far broader. For example, the hypothesis of equality of a set of domain means could be tested by these techniques. If the domains are defined by the strata of a stratified sampling procedure, the estimates of the domain means and of their variances given by Cochran (1963,pp.148-149) could be substituted into the test statistic U_0'. If the hypothesis of equal domain means is rejected because $U_0' > \chi^2_{K-1}(1-\alpha)$, then statistically significant sources of differences could be determined by use of the post hoc procedure suggested in this paper.

Furthermore it should be noted that the general theorem permits one to test hypotheses and determine simultaneous confidence intervals for analytical surveys in which the parameter estimates are not independent. An example of correlated ratio estimates in a survey in which the sampling unit consists of clusters of households is suggested by Cochran (1963, p. 182). A test of the hypothesis that the proportion of men who smoke is equal to the proportion of women who smoke could be based on U_0'. The test statistic would be given by

$$U_0' = (\hat{\theta}_1 - \hat{\hat{\theta}}_0,\ \hat{\theta}_2 - \hat{\hat{\theta}}_0) \begin{pmatrix} \text{var}(\hat{\theta}_1) & \text{cov}(\hat{\theta}_1, \hat{\theta}_2) \\ \text{cov}(\hat{\theta}_1, \hat{\theta}_2) & \text{var}(\hat{\theta}_2) \end{pmatrix}^{-1} \begin{pmatrix} \hat{\theta}_1 - \hat{\hat{\theta}}_0 \\ \hat{\theta}_2 - \hat{\hat{\theta}}_0 \end{pmatrix}$$

where the estimate of θ_0 which minimizes U_0' would be

$$\hat{\hat{\theta}}_0 = \frac{\hat{\theta}_1 \text{var}(\hat{\theta}_2) + \hat{\theta}_2 \text{var}(\hat{\theta}_1) - (\hat{\theta}_1 + \hat{\theta}_2)\text{cov}(\hat{\theta}_1, \hat{\theta}_2)}{\text{var}(\hat{\theta}_1) + \text{var}(\hat{\theta}_2) - 2\,\text{cov}(\hat{\theta}_1, \hat{\theta}_2)}$$

The extension of this test to three or more domains could also be based on U_0' where the elements of the covariance matrix could be obtained by the formulas given by Keyfitz (1957) or by Kish and Hess (1959). If the hypothesis of equality of the ratios were to be rejected, sources of the differences in the parameters could be determined by the post hoc procedure outlined above. The estimated variances of linear contrasts in the ratios could be obtained by substitution of the elements of the covariance matrix into the formula

$$\text{var}(\hat{\psi}) = \sum_{k=1}^{K} \text{var}(\hat{\theta}_k) + \sum_{k \neq k'}\sum \text{cov}(\hat{\theta}_k, \hat{\theta}_{k'})$$

Summary

The analysis of data generated by analytical surveys is compounded by complex sampling procedures and the nonresponse of sampled units. The problem is significantly greater when the number of subpopulations of interest exceeds two. On the basis of a chi-square analog of Scheffé's Theorem a simple multiple contrast or confidence interval procedure can be generated that can be used to identify possible parameter differences provided that the null hypothesis of no parameter differences has been rejected. This method should prove to be of considerable use to scientists whose major research methodology involves survey sampling.

References

COCHRAN, WILLIAM G. (1963). Sampling Techniques. New York: John Wiley and Sons, Inc.

GOOD, R. Z. (1963). Tests auxiliary to χ^2 tests in a Markov Chain. Ann. Math. Statist. 31, 56-74.

GOODMAN, LEO A. (1964). Interactions in multi-dimensional contigency tables. Ann. Math. Statist. 35, 716-725.

KEYFITZ, NATHAN (1957). Estimates of sampling variance where two units are selected from each stratum. J. Amer. Statist. Ass. 52, 503-510.

KISH, LESLIE & HESS, IRENE (1959). On variances of ratios and their differences in multistage samples. J.Amer. Statist. Ass. 54, 416-446.

MARASCUILO, LEONARD A. (1966). Large-sample multiple comparisons. Psychol. Bull. 65, 280-290.

SCHEFFÉ, HENRY (1959). The Analysis of Variance. New York: John Wiley and Sons, Inc.

Footnotes

1. This report was prepared at the Institute of Human Learning which is supported by grants from the National Science Foundation and the National Institutes of Health.

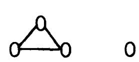
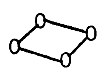
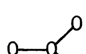
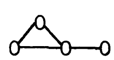
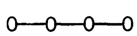
THE VARIANCE OF AN ESTIMATE OF LINKAGE DENSITY FROM A SIMPLE RANDOM SAMPLE OF GRAPH NODES

C. H. Proctor, North Carolina State University

In order to estimate the density of linkages in a finite graph one may select a simple random sample of nodes and determine for each pair of nodes selected whether or not there is a link between them. The sample proportion of linked pairs is an unbiassed estimate of the population proportion or linkage density. The variance of this estimate can be expressed as a function of certain graph moments and an unbiassed estimate of the variance can be found. The purpose of this paper is to describe the variance formulae.

The results were worked out to aid in interpreting some data on an acquaintanceship network in Wake Forest, North Carolina. The research was supported by the Institute of Statistics, Raleigh Section at North Carolina State University and is described rather informally in a mimeographed paper [4]. For such data the nodes are called actors and the graph is referred to as a social network [1]. This somewhat sociologically specialized terminology will be retained in favor of the more general graph theory one. It should be noted that only one link may join a pair of actors and it is supposed that both actors must be interviewed in order to determine whether a link is present.

In the following discussion ℓ's will denote sample proportions and λ's population proportions; n is written for the sampled number of actors and N is the population number. The quantity ℓ_{21} is the proportion of linked actor-pairs in the sample. The quantity ℓ_{32} is the sample proportion of actor-triples with two links. The quantity λ_{41} is the population proportion of actor-quadruples with one link. That the first subscript refers to the number of actors and the second to links can be inferred from these examples. When dealing with actor-quadruples containing two, three and four links a third, alphabetic, subscript is added to distinguish the cases as follows:

Proportion	Structure	Proportion	Structure
λ_{42a}	0-0-0 0	λ_{43c}	
λ_{42b}	0-0 0-0	λ_{44a}	
λ_{43a}		λ_{44b}	
λ_{43b}			

The expressions for the variance of ℓ_{21} and an unbiassed estimate of this variance turn out as follows:

$$V(\ell_{21}) = E(\ell_{21}^2) - [E(\ell_{21})]^2$$

$$= \{\lambda_{21}/\binom{n}{2} + 2(n-2)(\lambda_{32} + 3\lambda_{33})/3\binom{n}{2} + (n-2)(n-3)[\lambda_{42b} + \lambda_{43b} + \lambda_{44b} + 2(\lambda_{44a} + \lambda_{45}) + 3\lambda_{46}]/6\binom{n}{2}\} - \lambda_{21}^2 \qquad (1)$$

$$= \lambda_{21}\left[\frac{1}{\binom{n}{2}} - \frac{1}{\binom{N}{2}}\right] + \frac{2}{3}\left\{(\lambda_{32} + 3\lambda_{33})\left[\frac{n-2}{\binom{n}{2}} - \frac{N-2}{\binom{N}{2}}\right]\right\} + \frac{1}{6}[\lambda_{42b} + \lambda_{43b} + \lambda_{44b} + 2(\lambda_{44a} + \lambda_{45}) + 3\lambda_{46}]\left[\frac{(n-2)(n-3)}{\binom{n}{2}} - \frac{(N-2)(N-3)}{\binom{N}{2}}\right]. \qquad (2)$$

The expression within curly brackets in (1) was found by first writing ℓ_{21}^2 as $[2\sum_{t>u}^{n} a_{tu}/n(n-1)]^2$ in which a_{tu} equals one or zero according as to whether or not the t^{th} drawn and u^{th} drawn actors are linked. Then the square of the summation was expanded and terms of three kinds were collected. The types were: a_{tu}^2, $a_{tu}a_{tu'}$, and $a_{tu}a_{t'u'}$, where a prime denotes a subscript unequal to the unprimed one. The expected value was then taken using the facts that $E(a_{tu}^2) = \lambda_{21}$, $E(a_{tu}a_{tu'}) = \lambda_{32}/3 + \lambda_{33}$, and

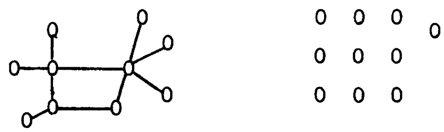
$$E(a_{tu}a_{t'u'}) = (\lambda_{42b} + \lambda_{43b} + \lambda_{44b})/3 + 2(\lambda_{44a} + \lambda_{45})/3 + \lambda_{46},$$

while the numbers of the three kinds of terms are $n(n-1)/2$, $n(n-1)(n-2)$, and $n(n-1)(n-2)(n-3)/4$ respectively.

In order to get the final form of $V(\ell_{21})$ in (2) the quantity λ_{21}^2 was written $[2\sum_{i>j}^{N} A_{ij}/N(N-1)]^2$ where A_{ij} equals 1 (0) if actor i is linked (not linked) to j and expanded as was done for ℓ_{21}^2. In this case the i and j subscripts refer to population identification numbers and the A_{ij}'s are not random variables.

An estimate of $V(\ell_{21})$ may be calculated using the corresponding ℓ quantities in place of the λ quantities in (2). This quantity will be denoted $v(\ell_{21})$. The fact that $E(\ell_{ks}) = \lambda_{ks}$ for any number of actors k and any structural subscript s insures that $E(v(\ell_{21})) = V(\ell_{21})$. To prove that $E(\ell_{ks})$ equals λ_{ks} one first writes ℓ_{ks} as the sum of $\binom{n}{k}$ indicator-of-structure-s variables divided by $\binom{n}{k}$ and then notes that the expected value of each and every indicator variable is λ_{ks}. This is an "argument of symmetry" [3]. Other properties of the estimate $V(\ell_{21})$ are not yet known but the study of higher moments of the distributions of both ℓ_{21} and $v(\ell_{21})$ will undoubtedly be greatly facilitated by the work of D. E. Barton and F. N. David on graph moments [2].

A small scale numerical example may help to illustrate the computation of $v(\ell_{21})$. A questionnaire was sent to a simple random sample of 20 names from the about 2,000 names in the North Carolina State University Staff Directory in 1964, and pairs of persons were said to be linked if each reported they had "spoken" to the other. The sociogram of linkages (note the 10 isolates) was as follows:



From these data one can calculate the following:

$\ell_{21} = 10/190 = .052632$

$\ell_{32} = 20/1140 = .017544$

$\ell_{33} = 0$

$\ell_{42b} = 3/4845 = .00061920$

$\ell_{43b} = 20/4845 = .00412797$

$\ell_{44b} = 0$

$\ell_{44a} = 1/4845 = .00020640$

$\ell_{45} = 0$, $\lambda_{46} = 0$

Thus $v(\ell_{21}) = (.052632)(.0052627)$
$+ .0116960(.0937373)$
$+ (.000859995)(-.385475)$
$= .00104.$

The estimate of density thus suffers a estimated coefficient of variation of 61%. Since the effect on the variance of an increase in sample size is roughly inversely proportional to $\binom{n}{2}$, it follows that in order to reduce the coefficient of variation to 6% would require an increase in sample size from 20 to over 200. For n = 200 the estimated coefficient of variation is still about 9%.

If a sample of 190 pairs of names, involving 380 or somewhat fewer persons, had been drawn as a simple random sample of the 1,999,000 pairs in the population the variance of the estimated density could be estimated as

$$pq/n = (.052632)(.947368)/190 = .00026$$

In so far as the simple random sample of actors (actor-SRS) also contains data on 190 pairs there appears to be a loss in precision for this population when using the actor-SRS rather than what may be called a pair-SRS of the same number of pairs. Of course, the cost of making the 190 observations would normally be greatly increased in the pair-SRS method over the actor-SRS.

[1] Chapters by Eric R. Wolf, J. Clyde Mitchell and Adrian C. Mayer in Banton, M. (ed.) The Social Anthropology of Complex Societies, A.S.A. Monographs 4, London: Travistock Publications; Frederick A. Praeger, Publishers, New York, 1965.

[2] Barton, D. E. and F. N. David, "The Random Intersection of Two Graphs," F. N. David (ed.), Research Papers in Statistics, John Wiley and Sons, Inc., New York, 1966, pp. 445-459.

[3] Cochran, W. G., Sampling Techniques, second edition, John Wiley and Sons, Inc., New York, 1963, page 22.

[4] Proctor, C. H., "Two Measurement and Sampling Methods for Studying Social Networks," Institute of Statistics Project Report - 1966, N. C. State University, Raleigh, North Carolina.

DESIGNING SEVERAL FACTOR ANALYTICAL SURVEYS

Gordon Booth and J. Sedransk

Iowa State University

1. Introduction

In many sample surveys the principal objective is to compare several sectors of a finite population. Specifically, there may be several factors (criteria of classification) of interest and each of these factors may have been divided into several categories. Then, for each factor, one may wish to compare these categories. For instance, a proposed study of hospitals in Iowa includes the factors "size of hospital" and "type of ownership." To take a simple example, one may wish to contrast large and small hospitals as well as public and private ones. (Here, for the factor "size of hospital," there are two categories - large and small.) Assuming this general specification of the problem, one must proceed to suggest comparisons of interest, and appropriate estimators. Then it is possible to select the sample so that specified precision for the estimates of the comparisons is attained at minimum cost, or maximum precision is achieved subject to a given budget.

For simplicity of presentation, attention is restricted to two-factor studies with two categories for each criterion of classification. For many of the topics discussed, extensions to more complex surveys are immediately clear.

In Section 2, it is assumed that one may select an independent random sample in each of the population sectors to be studied. (Thus, using the illustrative example, one may choose a random sample of Iowa's small, public hospitals independently of the large, private ones.) Then, utilizing a general type of comparison, a procedure to obtain the optimal sample size allocation is given. Since electronic computer algorithms may be needed to calculate the optimal allocation, approximate solutions are suggested. In a numerical study these approximate solutions are contrasted with the corresponding exact solutions.

If one cannot select independent samples in the subpopulations under study, a double sampling procedure may be feasible. Utilizing the results obtained in Section 2, a double sampling scheme is given in Section 3. This entails the formulation of a complex sampling rule for selecting the main sample from the large preliminary sample. A procedure for choosing the optimal sizes for the two samples (assuming a given budget) is also presented.

2. Single-phase Sampling

2.1. Comparisons

The two factors under investigation, each having two categories, may be represented by a 2 x 2 table with the (i,j)-th cell denoting the i-th category of the first factor, and the j-th category of the second factor. There are N_{ij} elements in the (i,j)-th cell (i,j = 1, 2), from which a random sample of size n_{ij} is selected. The sample mean is denoted by $\bar{y}_{ij}$, and the (population) within-cell variance by σ^2_{ij}. Let $N_{i.} = \sum_j N_{ij}$, $N_{.j} = \sum_i N_{ij}$, $N = \sum_i \sum_j N_{ij}$, with analogous definitions for $n_{i.}$, $n_{.j}$, and n.

The type of comparison to be employed depends on the assessment of the presence of interaction between the factors. If the interaction is deemed to be sufficiently small so that it may be neglected, and if a linear model is assumed to provide an adequate representation of the population, the contrasts of interest are readily apparent from the assumed model. Making such assumptions, Sedransk (1967) has obtained optimal sample size allocations which are applicable to studies where equal precision is required for all comparisons. He considers the two, three and four-factor cases where each factor has two categories of interest to the investigator.

However, if the interaction effect is large, and if one wishes to make an overall comparison, the choice of type of comparison is not so evident. Denoting the first factor by α and the second by τ, the two categories for each factor may be compared by considering

$$D_\alpha = \{N_{.1}(\mu_{11} - \mu_{21})/N\} + \{N_{.2}(\mu_{12} - \mu_{22})/N\}$$

and

$$D_\tau = \{N_{1.}(\mu_{11} - \mu_{12})/N\} + \{N_{2.}(\mu_{21} - \mu_{22})/N\} \tag{2.1.1}$$

where μ_{ij} is the population mean for the (i,j)-th cell.

Unbiased estimators for D_α and D_τ are given by

$$\hat{D}_\alpha = \{N_{.1}(\bar{y}_{11} - \bar{y}_{21})/N\} + \{N_{.2}(\bar{y}_{12} - \bar{y}_{22})/N\}$$

and

$$\hat{D}_\tau = \{N_{1.}(\bar{y}_{11} - \bar{y}_{12})/N\} + \{N_{2.}(\bar{y}_{21} - \bar{y}_{22})/N\}. \tag{2.1.2}$$

Rewriting $\hat{D}_\alpha$ as

$$\left(\frac{N_{.1}}{N} \bar{y}_{11} + \frac{N_{.2}}{N} \bar{y}_{12} \right) - \left(\frac{N_{.1}}{N} \bar{y}_{21} + \frac{N_{.2}}{N} \bar{y}_{22} \right) , \tag{2.1.3}$$

it may be noted that each of the two terms in (2.1.3) is an estimator similar to one suggested by Yates [1960, p. 134, (2)], except that, in Yates' example, some of the weights $\{N_{i.}, N_{.j}\}$ are to be estimated from the sample.

The other estimator cited by Yates [1960, p. 134, (1)] as appropriate if interaction is present, can be obtained from (2.1.2) by setting $N_{.j} = N_{i.} = N/2$ for (i,j = 1, 2). Finally, all results obtained by using (2.1.2) are applicable to <u>any</u> pre-specified choice of weights $\{W_1, 1-W_1, W_2, 1-W_2\}$ with W_1 replacing $N_{.1}/N$, $(1-W_1)$ replacing $N_{.2}/N$, etc.

The "proportionately weighted estimators" given by (2.1.2) are considered extensively in the sequel. They often provide a reasonable mode of comparison when overall contrasts are required. Also, the difficulties in obtaining optimal sample size allocations are well illustrated by assuming comparisons of this form. As noted above, the choice of particular (pre-specified) weights in (2.1.2) does not affect the ensuing analysis.

In some situations, it may be preferable to consider "simple effects" rather than composite comparisons such as (2.1.1). (Here, the "simple effects" are $\mu_{11} - \mu_{12}$, $\mu_{21} - \mu_{22}$, $\mu_{11} - \mu_{21}$ and $\mu_{12} - \mu_{22}$.) Such an approach is certainly more reasonable if the main objective is to select (separately) for each category of the α-factor the "better" category of the τ-factor. However, it may be unsatisfactory to look only at the simple effects if one wishes to obtain overall appraisals.

2.2. Procedures to obtain optimal allocations

For the proportionately weighted estimators, it is desired to find those values of the $n_{ij} \geq 0$ that

I. minimize $\sum_i \sum_j c_{ij} n_{ij}$

subject to

$$\mathrm{Var}(\hat{D}_\alpha) = \sum_i \sum_j N_{.j}^2 \, \sigma_{ij}^2/(N^2 n_{ij}) = V_1$$

and

$$\mathrm{Var}(\hat{D}_\tau) = \sum_i \sum_j N_{i.}^2 \, \sigma_{ij}^2/(N^2 n_{ij}) = V_2 \tag{2.2.1}$$

where c_{ij} is the cost of sampling an element in the (i,j)-th cell and V_1, V_2 are constants specified in advance.

This problem is equivalent to the one considered by Cochran (1963, pp. 123-124). Using standard calculus methods, there is no simple explicit solution for the optimal values of the n_{ij}, and Cochran presents a complex iterative method that may be used to obtain the solution.

Alternatively, one may reformulate problem I:

II. minimize $\sum_i \sum_j c_{ij} n_{ij}$

subject to

$$\mathrm{Var}(\hat{D}_\alpha) \leq V_1, \quad \mathrm{Var}(\hat{D}_\tau) \leq V_2 ,$$

$$0 \leq n_{ij} \leq N_{ij} \quad \text{for} \quad i,j = 1, 2. \tag{2.2.2}$$

This is a convex programming problem, and numerical solutions may be obtained by using an appropriate electronic computer algorithm. For example, Hartley and Hocking (1963) describe a method of convex programming by tangential approximation.

Since the convex programming approach is the more general one, it is preferable to the specification denoted by I, but neither is completely adequate. The convex programming method depends on the availability of appropriate computer facilities, and the procedure described by Cochran is somewhat cumbersome to carry out. Moreover, neither approach yields explicit algebraic expressions for the optimal values of the n_{ij}. (Such expressions are very useful in planning a double sampling procedure of the type described in Section 3.) Hence, it appears profitable to explore an alternative formulation (III) of the problem, and compare the resultant optimal allocation with the corresponding one obtained by using the convex programming method. Thus, the $n_{ij} \geq 0$ are to be selected to

III. minimize $\sum_i \sum_j c_{ij} n_{ij}$

subject to

$$W_1 \mathrm{Var}(\hat{D}_\alpha) + (1-W_1) \mathrm{Var}(\hat{D}_\tau) = V^* \tag{2.2.3}$$

where, to approximate problem formulation I or II, one might set $W_1 = V_2/(V_1 + V_2)$ and $V^* = 2V_1V_2/(V_1 + V_2)$.

Then, it is easily shown that the optimal value of n_{ij}, n_{ij}^*, is given by

$$n_{ij}^* = \left[\frac{\sigma_{ij}^2}{c_{ij}} \right]^{1/2} \left[W_1 N_{.j}^2 + (1-W_1) N_{i.}^2 \right]^{1/2} \cdot \frac{\{\sum\sum \sigma_{ij} [c_{ij} \{W_1 N_{.j}^2 + (1-W_1) N_{i.}^2\}]^{1/2}\}}{N^2 V^*} \tag{2.2.4}$$

For each of thirty numerical examples, a comparison of the allocation given by (2.2.4) with the corresponding optimal allocation from the convex programming approach (II) is given in Section 2.3.

For the "simple effects" estimators, one might choose those values of the n_{ij} that minimize $\Sigma\Sigma c_{ij} n_{ij}$ subject to $\mathrm{Var}(\bar{y}_{11} - \bar{y}_{12}) = \mathrm{Var}(\bar{y}_{21} - \bar{y}_{22}) = \mathrm{Var}(\bar{y}_{11} - \bar{y}_{21}) = \mathrm{Var}(\bar{y}_{12} - \bar{y}_{22}) = \bar{V}$. The optimal allocation can be obtained by standard methods.

2.3. Evaluation of an approximation to the optimal sample size allocation

In this section, it is desired to ascertain the efficacy of using the allocation given by (2.2.4) as an approximation to the optimal allocation of the n_{ij} obtained by using the "convex programming" approach (II). There is no loss of generality by taking $c_{ij} = 1$ for $(i,j = 1, 2)$, and for simplicity, it is assumed that $V_1 = V_2 = \bar{V}'$. Then, the optimal allocation is obtained by using a convex programming algorithm to

$$\text{minimize} \quad \Sigma\Sigma n_{ij}$$

subject to

$$\mathrm{Var}(\hat{D}_\alpha) = \Sigma\Sigma N_{.j}^2 \sigma_{ij}^2/(N^2 n_{ij}) \leq \bar{V}'$$

$$\mathrm{Var}(\hat{D}_\tau) = \Sigma\Sigma N_{i.}^2 \sigma_{ij}^2/(N^2 n_{ij}) \leq \bar{V}'$$

$$0 \leq n_{ij} \leq N_{ij} \quad \text{for } i,j = 1, 2. \qquad (2.3.1)$$

The approximation to this optimal allocation is obtained by

$$\text{minimizing} \quad \Sigma\Sigma n_{ij}$$

subject to

$$\{\mathrm{Var}(\hat{D}_\alpha) + \mathrm{Var}(\hat{D}_\tau)\}/2 = \bar{V}'. \qquad (2.3.2)$$

From (2.2.4), the approximate optimal allocation is given by

$$n_{ij}^* = \sigma_{ij}(N_{.j}^2 + N_{i.}^2)^{\frac{1}{2}} \{\Sigma\Sigma\sigma_{ij}(N_{i.}^2 + N_{.j}^2)^{\frac{1}{2}}\}/2N^2\bar{V}'. \qquad (2.3.3)$$

For each of thirty numerical examples, the following general procedure was used: (1) Applying the Hartley-Hocking technique to (2.3.1), the optimal allocation $\{n_{ij}\}$ was obtained; (2) The approximate optimal allocation $\{n_{ij}^*\}$ was calculated from (2.3.3); (3) $\mathrm{Var}(\hat{D}_\alpha)$ and $\mathrm{Var}(\hat{D}_\tau)$ were computed using each of the allocations given by (1) and (2) above; (4) The total sample size ($n = \Sigma\Sigma n_{ij}$) was computed for each of the allocations (1) and (2). Now define as follows:

$C_{(i)}$ = total cost (sample size) using the allocation given under (i) for i = 1, 2,

$V_{(2)} = \mathrm{Max}\,[\mathrm{Var}(\hat{D}_\alpha), \mathrm{Var}(\hat{D}_\tau)]$, computed using the allocation given under (2),

$$V_{(1)} = \begin{cases} \mathrm{Var}(\hat{D}_\alpha), \text{ computed using the allocation under (1), if } V_{(2)} = \mathrm{Var}(\hat{D}_\alpha) \\ \mathrm{Var}(\hat{D}_\tau), \text{ computed using the allocation under (1), if } V_{(2)} = \mathrm{Var}(\hat{D}_\tau), \end{cases}$$

$$S_{.j} = \sum_i \sigma_{ij}^2 \quad \text{and} \quad S_{i.} = \sum_j \sigma_{ij}^2 ,$$

$$R_1 = \mathrm{Max}\,[N_{.1}/N_{.2}, N_{.2}/N_{.1}] \quad \text{and}$$

$$R_2 = \mathrm{Max}\,[N_{1.}/N_{2.}, N_{2.}/N_{1.}] ,$$

$$R_3 = \mathrm{Max}\,[S_{.1}/S_{.2}, S_{.2}/S_{.1}] \quad \text{and}$$

$$R_4 = \mathrm{Max}\,[S_{1.}/S_{2.}, S_{2.}/S_{1.}] ;$$

(5) The following values were computed for each of the numerical examples:

(a) $P_c = (C_{(1)} - C_{(2)})(100)/C_{(1)}$ = percent decrease in cost by using (2.3.3) rather than the solution to (2.3.1).

(b) $P_v = (V_{(2)} - V_{(1)})(100)/V_{(1)}$ = percent increase in variance by using (2.3.3) rather than the solution to (2.3.1).

(c) $R_N = \mathrm{Max}\,[R_1/R_2, R_2/R_1]$

(d) $R_S = \mathrm{Max}\,[R_3/R_4, R_4/R_3]$.

It is easily shown that P_v and P_c are invariant under changes in scale. That is, if each of the N_{ij} is multiplied by a constant k_1, and each of the σ_{ij}^2 by a constant k_2, the values of P_v and P_c are not altered. Thus, the numerical examples presented in Table 1 cover a very wide range of values of the N_{ij} and σ_{ij}^2.

The results of the thirty numerical examples, for the most part, encourage the use of the approximate solution (2.3.3). For many of the examples, the allocations from (2.3.3) and (2.3.1) were essentially the same, and most of the values of P_v and P_c in Table 1 are small. (In choosing the examples, the objective was to identify those cases where P_v is large.) Note that large positive values of P_c are accompanied by large positive values of P_v. In such situations, $\mathrm{Max}\,[\mathrm{Var}(\hat{D}_\alpha), \mathrm{Var}(\hat{D}_\tau)]$ is much larger

Table 1

Values of P_v and P_c for thirty specifications* of the N_{ij} and σ_{ij}^2.

Example	N_{11}	N_{12}	N_{21}	N_{22}	σ_{11}^2	σ_{12}^2	σ_{21}^2	σ_{22}^2	P_c	P_v
1	1	1	1	1	1	1	1	1	0.00	0.00
2	1	1	1	1	1	1	1	50	0.00	0.00
3	1	1	1	1	1	1	1	1000	0.00	0.00
4	1	1	1	1	1	10	1	10	-0.06	-0.06
5	1	1	1	1	1	100	1	100	0.02	0.02
6	1	1	1	1	1	1000	1	1000	0.00	0.00
7	1	1	1	1	100	1	100	1	0.02	0.02
8	1	1	1	1	8	4	2	1	-0.06	-0.06
9	1	1	1	90	1	1	1	90	0.01	0.01
10	1	1	1	90	1	1	1	900	0.34	0.23
11	1	1	1	100	100	100	100	1	0.00	0.00
12	1	1	1	1000	1	1	1	1	0.17	0.02
13	1	10	1	10	1	1	1	1	0.62	5.24
14	1	100	1	100	1	1	1	1	1.89	10.05
15	1	100	1	100	1	100	1	100	31.72	52.49
16	1	100	1	100	100	1	100	1	38.29	71.02
17	1	1000	1	1000	1	1	1	1	1.87	10.50
18	1	1000	1	1000	1000	1	1000	1	46.13	89.45
19	1	100	100	1	100	1	1	100	0.17	0.16
20	1	100	100	1	1	100	100	1	0.08	0.08
21	1	10	100	1000	1	10	100	1000	10.89	18.83
22	100	100	100	1	1	1	1	50	0.00	0.00
23	100	100	100	1	1	1	1	100	0.00	0.00
24	100	1	1	100	100	1	1	100	0.17	0.16
25	1000	100	10	1	1000	100	10	1	10.95	18.84
26	1000	100	10	1	1	10	100	1000	23.68	52.85
27	2	2	1	1	4	3	2	1	4.08	6.86
28	21	7	9	7	1	1	1	1	0.00	0.00
29	245	213	119	117	99036	99036	91023	91023	0.00	0.00
30	26	18	46	51	5208	833	3333	1875	1.75	2.68

*Most of the sets of N_{ij} and σ_{ij}^2 have been scaled for convenience of presentation.

than Min $[\mathrm{Var}(\hat{D}_\alpha), \mathrm{Var}(\hat{D}_\tau)]$ under allocation (2). Thus, the reduction in total cost obtained by using the approximate solution is achieved at the cost of having one estimate with variance exceeding $\bar{V}'$ by a substantial amount. (Note that rounding errors in two examples cause P_c and P_v to be slightly negative.)

To identify the situations where P_v is large, the indices R_N and R_S are used. In Table 2, note that when R_N is significantly larger than one, P_v is generally quite large. Also, when <u>both</u> R_N and R_S are simultaneously much greater than one, both P_c and P_v are larger than they would be if only R_N were greater than one. (Compare example 14 with 15, and 17 with 18.) Note that when R_N exceeds one, even if R_S is equal to one, P_v may be moderately large (examples 13, 14 and 17). Finally, R_S alone is not a good indicator of the conditions under which P_v is large (examples 4, 5, 6 and 7). From the above, the following procedure is suggested: (1) Compute R_N. If R_N is near one, use (2.3.3). If R_N is much larger than one (perhaps, conservatively, larger than two), compute R_S. (2) If R_S is near one, then the allocation given by (2.3.3) may still be satisfactory. If R_S is also much larger than one, (2.3.3) is likely to be unsatisfactory.

Note that (except for rounding errors) the reduction in cost (P_c) is never as large as the increase in variance (P_v). Nevertheless, if P_v is only moderately large (examples 14, 17 and 27), then the resultant reduction in cost might make the use of (2.3.3) attractive. (This may be true for those examples with R_N large and $R_S = 1$.) Finally, it should be observed that it will always be possible to calculate R_N <u>prior</u> to sampling.

Table 2

Indices of the effectiveness of (2.3.3).

Example	R_N	R_S	P_c	P_v
1	1.00	1.00	0.00	0.00
2	1.00	1.00	0.00	0.00
3	1.00	1.00	0.00	0.00
4	1.00	10.00	-0.06	-0.06
5	1.00	100.00	0.02	0.02
6	1.00	1000.00	0.00	0.00
7	1.00	100.00	0.02	0.02
8	1.00	2.00	-0.06	-0.06
9	1.00	1.00	0.01	0.01
10	1.00	1.00	0.34	0.23
11	1.00	1.00	0.00	0.00
12	1.00	1.00	0.17	0.02
13	10.00	1.00	0.62	5.24
14	100.00	1.00	1.89	10.05
15	100.00	100.00	31.72	52.49
16	100.00	100.00	38.29	71.02
17	1000.00	1.00	1.87	10.50
18	1000.00	1000.00	46.13	89.45
19	1.00	1.00	0.17	0.16
20	1.00	1.00	0.08	0.08
21	10.00	10.00	10.89	18.83
22	1.00	1.00	0.00	0.00
23	1.00	1.00	0.00	0.00
24	1.00	1.00	0.17	0.16
25	10.00	10.00	10.95	18.84
26	10.00	10.00	23.68	52.85
27	2.00	1.55	4.08	6.86
28	1.22	1.00	0.00	0.00
29	1.76	1.09	0.00	0.00
30	2.12	2.72	1.75	2.68

3. Two-phase Sampling

If the elements in the four sub-populations (represented by the cells in a 2 x 2 table) are not identifiable in advance, one cannot sample independently in each of them. Thus, for the example cited earlier, it is assumed that there is no comprehensive list of the small, public hospitals in Iowa. However, one may select a large preliminary sample, and identify the sub-population to which each sampled element belongs. Then, for each sub-population, a sub-sample is selected for further study. Such a "two-phase" or "double" sampling procedure will be useful if the cost of identifying an element is small relative to the cost of securing the necessary information in the main survey.

For simplicity of presentation, only the proportionately weighted estimators ($\hat{D}_\alpha$, $\hat{D}_T$) are considered, and the single "composite" precision statement given by (3.1) is utilized[1]. The numerical analysis in Section 2.3 suggests that if equal precision is desired for $\hat{D}_\alpha$ and $\hat{D}_T$,

$$V = \{Var(\hat{D}_\alpha) + Var(\hat{D}_T)\}/2 \qquad (3.1)$$

is a reasonable composite precision statement.

It is assumed, at first, that the weights in D_α and D_T are selected prior to sampling. It may be noted that the ensuing analysis is not affected by selecting weights other than those given in (2.1.1). If the weights given in (2.1.1) are to be estimated from the sample results, modifications in the theory are necessary. These are given in the last part of this section.

Further, it is assumed that a given budget (C^*) is available to the investigator, and that the sizes of the preliminary sample (n') and main sample (n) are related by the cost equation

$$C^* = c'n' + cn \qquad (3.2)$$

where c' is the cost of identifying an element, and c is the cost of securing the necessary information from an element in the main sample.

The overall objective is to find that value of n' which satisfies (3.2), and minimizes V. This is most easily accomplished by evaluating V for a series of values of n' satisfying (3.2), and selecting that value of n' giving the minimum value of V.

To evaluate V one must first derive the sampling rule (S.R.) to be used. For fixed, but arbitrary, values of n' and n, the S.R. specifies, for any preliminary sample, how the sub-sampling is to be carried out. Let the random variable n'_{ij} denote the number of elements in the preliminary sample that are members of the (i,j)-th cell ($\Sigma\Sigma n'_{ij} = n'$). Then, conditional on the observed n'_{ij}, the S.R. gives the sample sizes (denoted by n_{ij}) for the main sample.

For a given S.R.,

$$V = \frac{1}{2}[E\{Var(\hat{D}_\alpha|\{n'_{ij}\})\} + Var\{E(\hat{D}_\alpha|\{n'_{ij}\})\} + E\{Var(\hat{D}_T|\{n'_{ij}\})\} + Var\{E(\hat{D}_T|\{n'_{ij}\})\}]$$

$$= \frac{1}{2}E[\sum_{ij}\sum\left(\frac{N^2_{.j} + N^2_{i.}}{N^2}\right)\frac{\sigma^2_{ij}}{n_{ij}}] = E[\Sigma\Sigma\, g^2_{ij}/n_{ij}] \qquad (3.3)$$

$$= \{E[V(n_{ij})]\}/2 \qquad (3.4)$$

where $\{n'_{ij}\}$ represents a fixed set of the n'_{ij} and the expectation in (3.3) and (3.4) is taken over all possible sets of n'_{ij} with the values of the n_{ij} determined by the S.R.

From (3.4) it is clear that, for a given

[1] The general procedure outlined in Section 3 is similar to the one given by Sedransk (1965) who investigated comparisons among two and three sub-populations.

set of n'_{ij}, one should select the n_{ij} to minimize $V(n_{ij})$ - that is, to minimize the sum of the two conditional variances. The "complete" sampling rule (described below) achieves this objective by utilizing the standard optimal allocation, $n_{ij} \propto g_{ij}$, when it is possible to do so, and by making appropriate modifications when the observed n'_{ij} are too small. (The convexity of $V(n_{ij})$ suggests the adjustments to be made.) Define

$$g^2_{ij} = (N^2_{.j} + N^2_{i.})\sigma^2_{ij}/2N^2 \quad \text{for} \quad i,j = 1, 2,$$

$$a_{ij} = ng_{ij}/\Sigma\,\Sigma\, g_{ij} \quad \text{for} \quad i,j = 1, 2,$$

$$b_{ij} = (n - n'_{11})g_{ij}/(g_{12} + g_{21} + g_{22}) \quad \text{for} \quad (i,j) \neq (1,1),$$

$$d_{ij} = (n - n'_{11} - n'_{12})g_{ij}/(g_{21} + g_{22}) \quad \text{for} \quad (i,j) = (2,1),(2,2),$$

$$\bar{n} = n - n'_{11} - n'_{12} \quad \text{and} \quad \bar{n}' = n'_{21} + n'_{22}\ .$$

The complete S.R. is defined as follows. (Note that when using the S.R., a re-labeling of the n'_{ij} may be necessary.)

(1) If $n'_{ij} \geq a_{ij}$ for all (i,j), let

$$n_{ij} = a_{ij} \text{ for all } (i,j).$$

(2) If only one $n'_{ij} < a_{ij}$, say $n'_{11} < a_{11}$, but if $n'_{ij} \geq a_{ij}$ for all $(i,j) \neq (1,1)$, let

$$n_{11} = n'_{11}.$$

For the determination of the remaining n_{ij}, there are three possible cases to be considered.

Case one:

If $n'_{ij} \geq b_{ij}$ for all $(i,j) \neq (1,1)$, let

$$n_{ij} = b_{ij} \quad \text{for all } (i,j) \neq (1,1).$$

Case two:

If only one of the $n'_{ij} < b_{ij}$, for $(i,j) \neq (1,1)$, say $n'_{12} < b_{12}$, but if $n'_{21} \geq b_{21}$ and $n'_{22} \geq b_{22}$, let

$$n_{12} = n'_{12}$$

$$\begin{aligned} n_{21} &= d_{21}, \quad \text{if } d_{21} < n'_{21} \leq \bar{n}' - (\bar{n} - d_{21}) \\ &= n'_{21}, \quad \text{if } n'_{21} \leq d_{21} \\ &= \bar{n} - n'_{22}, \quad \text{if } \bar{n}' - (\bar{n} - d_{21}) < n'_{21} \end{aligned}$$

$$n_{22} = \bar{n} - n_{21}.$$

Case three:

If only one of the $n'_{ij} > b_{ij}$, for $(i,j) \neq (1,1)$, say $n'_{22} > b_{22}$, but if $n'_{12} < b_{12}$ and $n'_{21} < b_{21}$, let

$$n_{12} = n'_{12}, \qquad n_{21} = n'_{21}$$
$$n_{22} = n - n'_{11} - n'_{12} - n'_{21}.$$

(3) If two of the $n'_{ij} < a_{ij}$, say $n'_{11} < a_{11}$ and $n'_{12} < a_{12}$, then utilize the allocation given by Case two of (2).

(4) If only one of the $n'_{ij} > a_{ij}$, say $n'_{22} > a_{22}$, then let

$$n_{11} = n'_{11}, \quad n_{12} = n'_{12}, \quad n_{21} = n'_{21},$$
$$n_{22} = n - n_{11} - n_{12} - n_{21}.$$

Thus, the complete S.R. specifies, for any preliminary sample, the exact subsampling procedure. Having selected a random sample of n' elements, the investigator would identify each of the elements and ascertain the values of the n'_{ij}. From the S.R., the number of elements (n_{ij}) to be selected from the (i,j)-th cell for the main sample would be determined. The S.R. is somewhat cumbersome to present, but simple to use for a single problem.

To find the optimal value of n', one must be able to evaluate (assuming the S.R. given above) $2V = E\,[V(n_{ij})]$ for specified values of n' and n. Noting that the expectation in (3.4) refers to all possible sets of the n'_{ij} (with the values of the n_{ij} determined by the S.R.), it is clear that an analytical evaluation is not feasible. Moreover, even with specified values of the necessary parameters and the aid of a computer, a complete numerical evaluation of $E\,[V(n_{ij})]$ is impractical because of the extremely large number of possible sets of the n'_{ij}.

In lieu of a complete evaluation of $E[V(n_{ij})]$, one may use Monte Carlo sampling to estimate $E[V(n_{ij})]$. Thus, to determine the n'_{ij}, a sample of size n' is drawn from the multinomial distribution with inclusion probabilities $\{\pi_{ij} = N_{ij}/N\}$. Using the S.R., the n_{ij} are found, and $V(n_{ij})$ calculated. This procedure is replicated K-1 times, and the sample mean ($\overline{V}$) and variance of $V(n_{ij})$ calculated. The entire procedure is repeated for different values of n', and the "optimal value" of n' is chosen on the basis of minimum $\overline{V}$. If one selects a large value for K (perhaps, K = 500) it appears, from some numerical examples, that the Monte Carlo procedure provides an unequivocal choice for the optimal value of n'. The sample standard error of $\overline{V}$ is so small, and the pattern of the $\overline{V}$ (as n increases) is so regular that an error in selecting the optimal n' value is very unlikely. Even if K = 100, one is unlikely to make a costly error in selecting the optimal value of n'. To illustrate this, consider the following specification: $C^* = 220$, $c' = 1$, $c = 10$, $\pi_{11} = .15$, $\pi_{12} = .20$, $\pi_{21} = .25$, $\pi_{22} = .40$, $g_{11} = 1$, $g_{12} = 1$, $g_{21} = 1$, $g_{22} = 4$, K = 100. Then, for each trial value of n', the following quantities are given in Table 3: the corresponding value of n, $\overline{V}$ and $s(\overline{V})$ = (Monte Carlo) sample estimate of the standard error of $\overline{V}$. The best choice for n' is seen to be 30.

Table 3

Monte Carlo estimates of V for various values of n'.

n'	n	$\overline{V}$	$s(\overline{V})$
20	20	3.27	0.082
30	19	2.68	0.015
40	18	2.79	0.006
50	17	2.95	0.001

To estimate $E[V(n_{ij})]$, one must specify values for the π_{ij}. Although some of the marginal totals of the N_{ij} might be known, the N_{ij} will usually be unknown. However, in many situations, reasonable estimates of the π_{ij} can be made. It must be emphasized that once n' is determined, the sampling rule depends only on the weights in D_α and D_τ. (As noted earlier, such weights do not have to be functions of the N_{ij}.)

If a computer is not available, one may roughly approximate $\{E[V(n_{ij})]\}/2$ using a procedure given by Sedransk[2] (1965). Thus, assuming that $\pi_{11} \le \pi_{12} \le \pi_{21} \le \pi_{22}$, consider the following approximate determination of the values of the $E(n_{ij})$:

$$E(n_{11}) = ng_{11}/g^*, \quad \text{if} \quad E(n'_{11}) > ng_{11}/g^*$$
$$= E(n'_{11}), \quad \text{otherwise;}$$

$$E(n_{12}) = (n-En_{11})g_{12}/(g^*-g_{11}),$$
$$\text{if} \quad E(n'_{12}) > (n-En_{11})g_{12}/(g^*-g_{11})$$
$$= E(n'_{12}), \quad \text{otherwise;}$$

$$E(n_{21}) = (n-En_{11}-En_{12})g_{21}/(g_{21}+g_{22}),$$
$$\text{if} \quad E(n'_{21}) > (n-En_{11}-En_{12})g_{21}/(g_{21}+g_{22})$$
$$= E(n'_{21}), \quad \text{otherwise;}$$

$$E(n_{22}) = n - En_{11} - En_{12} - En_{21} \tag{3.5}$$

where $g^* = \Sigma\Sigma g_{ij}$, $E(n'_{ij}) = n'\pi_{ij}$, and for the special case of n' = n, take $E(n_{ij}) = E(n'_{ij})$.

Finally, approximate $\{E[V(n_{ij})]\}/2$ by

$$V' = \frac{1}{2}\Sigma\Sigma\left(\frac{N^2_{\cdot j} + N^2_{i\cdot}}{N^2}\right)\frac{\sigma^2_{ij}}{E(n_{ij})}. \tag{3.6}$$

Then, one may evaluate V' [using (3.5)] for different choices of n', and choose the "optimal" value of n' as before. For the cases considered by Sedransk (1965), such a procedure proved to be very satisfactory. To further investigate the utility of this approximate method, 22 examples have been worked using both the approximate method, and the Monte Carlo sampling procedure. At least 100 Monte Carlo replications were used for each trial value of n' considered. The examples were chosen to represent the wide variety of possible relationships among the $\pi_{ij} = N_{ij}/N$ and g_{ij}.

In Table 4, the following quantities are presented for each example: (1) The values of the g_{ij} and π_{ij}. Using the definition of g_{ij}, it is easily verified that the examples include a number of patterns of the σ^2_{ij}; (2) The "optimal" values of n' as determined by the approximate and Monte Carlo methods (n'_a and n'_0, respectively); (3) The estimate, A, from Monte Carlo sampling, of the per cent increase in the variance, $V = E[V(n_{ij})]/2$, by using the "optimal" n' value from the approximate method rather than the optimal n' obtained from the Monte Carlo calculations; (4) The ratio $B = \overline{V}/V'$ evaluated at the "optimal" value of n' as determined by using the approximate method. (Recall that $\overline{V}$ is the Monte Carlo estimate of V.) B indicates the utility of using V' as an estimator for V, the value of n' chosen being the most appropriate

[2] For further details, see Section 4.2.2 of that paper.

Table 4

Efficiency of a procedure to estimate the optimal sample sizes: $C^* = 220 = n' + 10n$.

Ex.	g_{11}	g_{12}	g_{21}	g_{22}	π_{11}	π_{12}	π_{21}	π_{22}	n'_a	n'_0	A*	B**	C***
1	1	1	1	4	.15	.20	.25	.40	30	30	0.0	1.04	1.22
2	4	1	1	1	.15	.20	.25	.40	60	60	0.0	1.05	2.24
3	4	4	2	2	.15	.20	.25	.40	40	40	0.0	1.09	1.49
4	4	3	2	1	.15	.20	.25	.40	40	50	5.4	1.14	1.70
5	1	2	3	4	.15	.20	.25	.40	20	30	15.7	1.22	1.16
6	2	2	4	4	.15	.20	.25	.40	20	30	16.5	1.22	1.17
7	4	1	4	1	.15	.20	.25	.40	40	50	2.1	1.12	1.85
8	1	1	1	4	.10	.40	.10	.40	30	30	0.0	1.07	
9	4	1	1	1	.10	.40	.10	.40	70	80	3.7	1.13	
10	4	3	2	1	.10	.40	.10	.40	50	60	0.2	1.16	
11	2	2	4	4	.10	.40	.10	.40	50	50	0.0	1.09	
12	4	1	4	1	.10	.40	.10	.40	60	70	4.8	1.22	
13	1	1	1	4	.20	.20	.30	.30	30	40	4.4	1.14	
14	4	1	1	1	.20	.20	.30	.30	50	50	0.0	1.06	
15	4	3	2	1	.20	.20	.30	.30	30	40	3.2	1.12	
16	2	2	4	4	.20	.20	.30	.30	20	30	12.8	1.22	
17	4	1	4	1	.20	.20	.30	.30	30	40	5.7	1.11	
18	1	1	1	4	.10	.10	.20	.60	30	30	0.0	1.08	
19	4	1	1	1	.10	.10	.20	.60	70	70	0.0	1.12	
20	4	3	2	1	.10	.10	.20	.60	60	70	1.0	1.18	
21	2	2	4	4	.10	.10	.20	.60	30	40	4.1	1.17	
22	4	1	4	1	.10	.10	.20	.60	50	60	9.1	1.18	

*A is the (Monte Carlo) estimate of the per cent increase in variance, V, by using the "optimal" n' value from the approximate method rather than the optimal n' obtained from the Monte Carlo calculations.

**B is the ratio $\overline{V}/V'$ calculated at the "optimal" value of n' as determined by using the approximate method.

***C is the ratio $\overline{V}_{20}/\overline{V}_{opt}$ where $\overline{V}_{20}$ is the (Monte Carlo) estimate of V for n' = n = 20, and $\overline{V}_{opt}$ is the corresponding estimate of V at the optimal value of n'.

one if the approximate method is employed; (5) The ratio $C = \overline{V}_{20}/\overline{V}_{opt}$ where $\overline{V}_{20}$ is the (Monte Carlo) estimate of V for n' = n = 20, and $\overline{V}_{opt}$ is the corresponding estimate of V for $n' = n'_0$.

The results (presented in Table 4) are very similar to those found by Sedransk (1965, pp. 996-999), and indicate that the loss of precision by using the "optimal" n' value from the approximate method is generally small. The values of A range from 0.0% to 16.5% with only three examples having values exceeding 10.0%. The cases where A is large coincide with values of $n'_a = 20$. This may be explained by the observation that increasing n' a little from the single-phase sampling position (n' = n) usually reduces the variance considerably, whereas further increases in n' produce smaller reductions in the variance. Thus, when the true optimum value of n' is only slightly larger than n' = n = 20, taking $n'_a = 20$ may result in a large increase in variance. To be conservative, if the value of n'_a is near n (i.e., single-phase sampling is indicated to be optimal), a further investigation is indicated. However, in most of the examples the optimum is flat; that is, $\overline{V}$ varies insubstantially with n' as n' moves in either direction from its optimal value. From this discussion, it appears that if n'_a is not (nearly) equal to n, one may be confident about using the approximate method.

For some specifications of the π_{ij} and choices of n', the probability of obtaining a zero value for at least one of the n'_{ij} may not be negligible. In such cases, one may wish to take a larger value of n' than that given by n'_0 or n'_a. For the examples presented in Table 4, and including all trial values of n' considered, all estimates are based on samples having $n'_{ij} > 0$ for i, j = 1, 2.

From B, it is clear that V' underestimates $\overline{V}$ (and, therefore, should underestimate V) for all examples. This agrees with the findings of Sedransk (1965, p. 998). The ratio C, presented for the first seven examples, indicates the efficacy of double sampling. If one used single-phase sampling (i.e., selected $n' = n = 20$), rather than double sampling with $n' = n_0'$, the per cent increase in variance ranged from 16 to 124. Greater increases could be expected for examples where n_0' is very large.

Finally, it should be noted that Sedransk (1965, pp. 998-999) has presented some numerical evidence that the double sampling procedure is robust with regard to specifying the values of the g_{ij}. That is, the loss in unconditional precision because of choosing a non-optimal value of n' will generally be moderate, even if there are fairly large errors in specifying the g_{ij}. Also, he indicates that all of these conclusions seem to persist if the cost relationship c/c' is altered.

If the approximate method is considered to be unsatisfactory for some situations, one might continue to use V' as given in (3.6), but approximate the $E(n_{ij})$ more closely. (The first order Taylor series approximation is likely to be satisfactory since, in most applications, n' and n will be large.) Using an "approximate" S.R. akin to that given by Sedransk (1965, p. 994), one may obtain $E(n_{11})$, $E(n_{12}) = E\{E(n_{12}|n_{11})\}$, and $E(n_{21}) = E\{E(n_{21}|n_{11},n_{12})\}$. This is most easily accomplished by using normal approximations to the distributions of the n_{ij}', but one will have to use some numerical integration to evaluate a few of the terms in $E(n_{21})$.

If the proportionately weighted estimators are employed with the weights estimated from the preliminary sample, the estimators of D_α and D_τ are given by

$$\hat{D}_\alpha' = \{n'_{.1}(\overline{y}_{11} - \overline{y}_{21})/n'\} + \{n'_{.2}(\overline{y}_{12} - \overline{y}_{22})/n'\}$$

$$\hat{D}_\tau' = \{n'_{1.}(\overline{y}_{11} - \overline{y}_{12})/n'\} + \{n'_{2.}(\overline{y}_{21} - \overline{y}_{22})/n'\} \, . \tag{3.7}$$

Then, following the procedure used to derive (3.3),

$$[\mathrm{Var}(\hat{D}_\alpha') + \mathrm{Var}(\hat{D}_\tau')]/2 = \frac{1}{2}\,[E\{\Sigma\Sigma(\frac{(n'_{.j})^2 + (n'_{i.})^2}{(n')^2})\frac{\sigma_{ij}^2}{n_{ij}}\} + V[\frac{n'_{.1}}{n'}(\mu_{11} - \mu_{21}) + \frac{n'_{.2}}{n'}(\mu_{12} - \mu_{22})] + V[\frac{n'_{1.}}{n'}(\mu_{11} - \mu_{12}) + \frac{n'_{2.}}{n'}(\mu_{21} - \mu_{22})]] \tag{3.8}$$

where E and V in (3.8) refer to all possible sets of the n_{ij}' with the corresponding n_{ij} determined by the S.R.

It is clear from (3.8) that the choice of a S.R. depends only on the term in curly brackets. However, this term is equivalent to (3.3) with $(n'_{.j}/n')$ and $(n'_{i.}/n')$ replacing $(N_{.j}/N)$ and $(N_{i.}/N)$. Hence, the complete S.R. given below (3.4) may be utilized for the current problem with the appropriate substitutions for $(N_{i.}/N)$ and $(N_{.j}/N)$.

If the estimators given by (3.7) are used, it is difficult to determine the optimal value of n'. For a fixed value of n', if one uses rough approximations for the π_{ij}, Monte Carlo sampling may be employed to estimate the first term in (3.8). As suggested above, this procedure may be repeated for several different values of n', and the "optimal" value thus located. However, the variance terms in (3.8) depend on n', the π_{ij} and the μ_{ij}. Since the values of the μ_{ij} (and the π_{ij}) may determine the true optimal value of n', it appears that rough estimates of the two variance terms in (3.8) should be combined with the Monte Carlo estimates of the first term to approximate $\frac{1}{2}\,[\mathrm{Var}(\hat{D}_\alpha') + \mathrm{Var}(\hat{D}_\tau')]$ more closely.

4. An Additional Application of the Two-phase Sampling Procedure

In classical double sampling for stratification, one selects a simple random sample of size n' with n_h' elements subsequently identified as being members of stratum h ($h = 1, 2, \ldots, L$). Then, one may select (by simple random sampling) a subsample of n_h elements from the n_h' elements found to belong to stratum h. Assuming that n' is sufficiently large so that the probability of obtaining any $n_h' > 0$ is negligible, $\overline{y}_{st}^* = \sum_h n_h' \overline{y}_h/n'$ is an unbiased estimator of the population mean with

$$\mathrm{Var}(\overline{y}_{st}^*) = V[\sum_h n_h' \overline{Y}_h/n'] + E\{\sum_h (n_h')^2 S_h^2/(n')^2 n_h\} - E[\sum_h (n_h')^2 S_h^2/(n')^2 N_h] \tag{4.1}$$

where E and V in (4.1) refer to all sets of the n_h' ($n_h' \neq 0$) with the corresponding n_h determined by some sampling rule.

From (4.1), it is clear that for given values of n' and n, one should choose the sampling rule to minimize the term in curly brackets.

(None of the other terms are functions of the n_h.) This sampling rule (for L = 4 strata) can be obtained immediately from the one presented in Section 3. Finally, note the similarity between (4.1) and (3.8). (The two expressions are essentially identical if the f.p.c. terms are omitted from (4.1) or included in (3.8).) Thus, the problem of finding the optimal value of n' (assuming a linear cost function with prespecified budget) for double sampling with stratification is identical with that presented earlier.

Acknowledgment

This research has been supported, in part, by the U. S. Office of Education under Contract OEC-3-6-002041-2041.

References

Cochran, W. G. (1963). Sampling Techniques, Second Edition, New York: John Wiley.

Hartley, H. O. and Hocking, R. (1963). "Convex programming by tangential approximation," Management Science, 9, 600-612.

Sedransk, J. (1965). "A double sampling scheme for analytical surveys," Jour. Amer. Stat. Assoc., 60, 985-1004.

Sedransk, J. (1967). "Designing some multifactor analytical studies," Jour. Amer. Stat. Assoc., 62, (to appear).

Yates, F. (1960). Sampling Methods for Censuses and Surveys, Third Edition, London: Griffin.

FROM IDENTITIES TO NORMAL EQUATIONS: AN EASY APPROACH TO LEAST SQUARES

Irving H. Siegel, The W. E. Upjohn Institute for Employment Research*

A paper I presented six years ago at a meeting sponsored by the New York Academy of Sciences sought "to exhibit the classical method of least squares without recourse to the conventional summary normal equations" [1]. It emphasized two procedures that explicitly introduce the n unknown residuals into the n observation equations. According to one of these procedures, the rectangular n x m observation matrix is expanded by a simple rule into a much larger invertible square matrix. A supermatrix system equivalent to conventional normal equations is immediately obtained, and it becomes possible to delegate all arithmetic processing to computer specialists. The second procedure is to set up "normal identities" in an obvious manner and then to eliminate certain summary terms that contain residuals and that an adjustment process might reasonably be expected to reduce to zero. The result is a conventional system of normal equations or something very similar to it -- plus some footnote information on the minimum value of the sum of squared residuals. Having already explored my supermatrix approach in some degree [2], I return at this time to my normal-identity approach.

Let us consider the familiar least-squares case of fitting the line $y = a + bx$ with only the y_i subject to error. Here, the number of unknown constants is $m = 2$ and the number of observations is n. For simplicity, we assume that the observations have equal weight.

Actually, each observation equation is an identity containing an additional term, the variable residual r_i. Explicitly introducing this term, we write $y \equiv a + bx_i + r_i$. Since there are n observations, we have n residuals.

We now proceed to develop normal identities from the observation identities. We multiply each observation identity by the coefficient of a (i.e., by 1) and sum to obtain the first normal identity. Next, we multiply each observation identity by the coefficient of b (i.e., by x_i) and sum to obtain the second normal identity. Finally, we multiply each observation identity by r_i and sum to obtain the third normal identity.

The resulting system looks like this:

$$\begin{aligned} \Sigma y &\equiv na + b\Sigma x + \Sigma r \\ \Sigma xy &\equiv a\Sigma x + b\Sigma x^2 + \Sigma xr \\ \Sigma ry &\equiv a\Sigma r + b\Sigma xr + \Sigma r^2 \quad . \end{aligned}$$

The determinant of the right-hand side is axisymmetric. The unknowns include a, b, and the terms in r.

How may this system be solved? One obvious scheme is to assume $\Sigma r = 0$ and $\Sigma xr = 0$ and restrict attention to the first two lines, since only two unknowns, a and b, really need to be found. The third line is redundant for solution, but it states a consistency condition, a necessary implication of the adjustment process.

Our two assumptions, it might be noted, tell us exactly the same thing that the two normal equations do. The assumptions tell us with reference to residuals what the normal equations tell us with reference to the unknowns of primary interest.

The third line, which simply states that $\Sigma ry = \Sigma r^2$, is a mathematical footnote. It tells us what our assumptions mean with respect to the minimum value of the sum of squared residuals. Indeed, it tells us what we mean by "least squares" in this case. If this information is deemed exceptionable, if this implication offends common sense or some prior principle, the process should be reconsidered or different sums should be eliminated.

The normal identity approach makes it clear that the simple condition $\Sigma r = 0$ arises in least-squares adjustment only when a free constant (such as a) exists. For any other linear model (e.g., the one used by Gauss in his Theoria Motus to illustrate the adjustment process), the residuals always appear with unequal weights in the normal identities. Hence, in such instances, only sums of weighted residuals may be set equal to zero -- even though the observations themselves are unweighted. Furthermore, in the absence of a free constant, a fitted line cannot pass through the point of unweighted means of the observed values.

*The author's views do not necessarily represent positions of The W. E. Upjohn Institute for Employment Research.

Much more complicated cases of curve-fitting may also be investigated with the aid of normal identities. Let us consider briefly the case in which both variables, the y_i and the x_i, are subject to error. This problem has attracted the attention of many statisticians over a long span of time. A paper published in 1959 provides an impressive bibliography -- 53 items appropriately extending from A (Adcock, 1878) to Z (Zucker, 1947). This bibliography, however, must still be far from exhaustive for the period covered [3].

We again start with $y_i = a + b_i$ and do not weight the observations. Inserting s_i for the residual corresponding to y_i and inserting t_i for the residual corresponding to x_i, we obtain $y_i + s_i \equiv a + b(x_i + t_i) \equiv a + bx_i + bt_i$ as the prototype observation identity.

This time, we obtain five normal identities as we subject the observation identities to multiplication, in turn, by 1, x_i, y_i, s_i, and t_i and sum the results. The whole system looks like this:

$$\begin{aligned}
\Sigma y + \Sigma s &\equiv na + b\Sigma x + b\Sigma t\\
\Sigma xy + \Sigma xs &\equiv a\Sigma x + b\Sigma x^2 + b\Sigma xt\\
\Sigma y^2 + \Sigma ys &\equiv a\Sigma y + b\Sigma yx + b\Sigma yt\\
\Sigma sy + \Sigma s^2 &\equiv a\Sigma s + b\Sigma sx + b\Sigma st\\
\Sigma ty + \Sigma ts &\equiv a\Sigma t + b\Sigma tx + b\Sigma t^2 .
\end{aligned}$$

What assumptions could we reasonably make in order to solve this system? The sums of unweighted residuals presumably should be made equal to zero: $\Sigma s = \Sigma t = 0$. We may also suppose the independence of: (1) the two sets of residuals and (2) the observed values of one variable and the residuals associated with the other. Thus, we also assume $\Sigma xs = \Sigma yt = \Sigma st = 0$.

After making these simplifications, we are left with this pattern of equations:

$$\begin{aligned}
\Sigma y &= na + b\Sigma x\\
\Sigma xy &= a\Sigma x + b\Sigma x^2 + b\Sigma xt\\
\Sigma y^2 + \Sigma ys &= a\Sigma y + b\Sigma yx\\
\Sigma sy + \Sigma s^2 &= 0\\
0 &= b\Sigma xt + b\Sigma t^2
\end{aligned}$$

The fourth and fifth lines tell us what the adjustment process, if it is accepted, means for the values of Σs^2 and Σt^2. Unless some relationship between these sums of squared residuals is posited, we cannot solve the system, since we still have too many unknowns.

If we assume that $\Sigma s^2 = k\Sigma t^2$, we arrive at the quadratic equation that is often shown in the literature as the key to complete solution. This additional assumption entails $\Sigma sy = \Sigma xt$. Substituting in the first three summary equations and simplifying, we obtain this expression:

$$b^2M_{xy} + b(kM_x - M_y) - kM_{xy} = 0,$$

where M_{xy}, M_x, and M_y refer to the moments appearing in the familiar formulas for the correlation coefficient and for the variances of x and y. The quadratic expression may be solved readily for b; and, giving different values to the parameter k, we obtain various special cases of interest [4]. If we set k = 1, we have the well-known case of orthogonal regression, which is usually discussed in terms of polar coordinates and solved with respect to the tangent of an angle [5].

Obviously, the normal-identity approach is versatile, and it should have both pedagogic and theoretical interest. It seems to treat the adjustment process as a deterministic, rather than as a probabilistic, one; but a transition from "mathematics" to "statistics" is made via the assumptions. Since the assumptions refer to summary terms involving residuals, a range of choices may be explored advantageously when more than one variable is subject to error. The "errors-in-variables" model, moreover, is nowadays contrasted in econometrics with the "errors-in-equations" model, and the normal-identity tool ought to be useful in the investigations pursued [6].

REFERENCES

[1] I. H. Siegel, "Least Squares 'Without Normal Equations'," Transactions of the New York Academy of Sciences, February 1962, pp. 362-371.

[2] See these papers by I. H. Siegel: "Simplified Least-Squares Approach for an Age of Computers," 1966 Social Statistics Section Proceedings of the American Statistical Association, pp. 398-400; "Deferment of Computation in the Method of Least Squares," Mathematics of Computation, April 1965, pp. 329-331; and "Least Squares with Less Effort," 1964 Business and Economic Statistics Section Proceedings of the American Statistical Association, pp. 284-285.

[3] The bibliography appears at the end of a paper by Albert Madansky, "The Fitting of Straight Lines When Both Variables Are Subject to Error," Journal of the American Statistical Association, March 1959, pp. 172-205. It is called "a complete survey of the literature" by E. Malinvaud (Statistical Methods of Econometrics, Rand McNally, Chicago, 1966, p. 362), but many relevant items are not included, such as: Mansfield Merriman, A Text-Book in the Method of Least Squares, 8th ed., Wiley, New York, 1911, pp. 127-128, 216-217; a paper by Merriman in Report of the Superintendent of the U.S. Coast and Geodetic Survey, Fiscal Year 1890, pp. 687-690; Henry Schultz's "Comments" on R. G. D. Allen's paper in Economica, May 1939, pp. 202-204; L. J. Reed, "Fitting Straight Lines," Metron, No. 3, 1921, pp. 54ff.; and P. J. Dwyer and M. S. Macphail, "Symbolic Matrix Derivatives," Annals of Mathematical Statistics, December 1948, especially pp. 531-532.

[4] See, for example, J. Johnston, Econometric Methods, McGraw-Hill, New York, 1960, pp. 150-162, as well as Allen's Economica paper, pp. 191-201.

[5] See, for example, Madansky, loc. cit., p. 202; Y. V. Linnik, Method of Least Squares and Principles of the Theory of Observation, Pergamon, New York, 1961, pp. 11-13, 320-322; and Reed, loc. cit.

[6] Malinvaud, op. cit., Chapter 10; and Johnston, op. cit., Chapter 6.

AN EMPIRICAL STUDY OF THE STABILITIES OF ESTIMATORS AND VARIANCE ESTIMATORS IN p.p.s. SAMPLING

J. N. K. Rao and D. L. Bayless
Texas A & M University

1. Introduction

Several methods of probability proportional to size (p.p.s.) sampling without replacement have been proposed in recent years. However, not much is known with regard to the stabilities of estimators of the population total, and practically nothing is known with regard to the stabilities of their variance estimators. Therefore, in the present paper, we make an empirical study of the stabilities of estimators and variance estimators for the important case of sample size n=2, using several natural as well as artificial populations. The artificial populations are deliberately chosen to represent situations more extreme than those normally encountered in practice and they provide better discrimination among the methods.

We have chosen only those methods (excepting one) which satisfy the following essential requirements: (a) Variance of the estimator should be smaller than that of the customary estimator in p.p.s. sampling with replacement. (b) A non-negative, unbiased variance estimator should be available. (c) Computations involved should not be cumbersome. Requirement (b) eliminates the systematic method of Madow (1949) and Hartley (1966). We have not included the asymptotic methods (valid for large or moderate population size N) of Hartley and Rao (1962), Rao (1963) and Hajek (1964), although they satisfy the above requirements. Similarly, Stevens' (1958) method is excluded as it is applicable only when the units in the population are or can be grouped with respect to the sizes x_j such that units in a group have the same size.

Based on the above considerations, we have selected the following methods for the present study: (1) The methods of Brewer (1963), Carroll and Hartley (1964), Fellegi (1963), Rao (1963, 65), Durbin (1967), and Hanurav (1967), all using the Horvitz-Thompson estimator and satisfying $\pi_j=2x_j/X$ where π_j is the probability of selecting the j^{th} population unit ($j=1,2,\ldots,N$) in the sample and $X=\Sigma x_j$, (2) the methods of Des Raj (1956) and Murthy (1957), (3) the method of Rao, Hartley and Cochran (1962) and (4) Lahiri's (1951) method using a ratio estimator.

The methods of Brewer, Durbin and Rao (1965) are equivalent in that they have the same joint probabilities of selection π_{ij} and, hence, the same variance. However, Durbin's method has the advantage of allowing for rotation of the sample. For convenience, we denote this group of methods as Brewer's method. The methods of Rao (1963) (investigated in detail by Carroll and Hartley) and Fellegi are equivalent and both possess rotational properties. Again for convenience we denote this group as Fellegi's method. Hanurav has proposed the criterion $\varphi = \min \pi_{ij}/(\pi_i\pi_j) > \beta$, for β sufficiently away from zero to improve the stability of the variance estimator and developed a method which satisfies this criterion except when the largest size $\tilde{x}_N$ is markedly different from the next largest size $\tilde{x}_{N-1}$. He has, however, not shown whether his method satisfies requirement (a), although it appears highly probable. The methods of Brewer and Fellegi also seem to satisfy Hanurav's criterion, except when $\tilde{x}_N/X$ is very close to 1/2. For instance, if $\tilde{x}_N/X \leq 1/3$ and the other $x_i/X \leq 1/4$, $\varphi > 0.3$ for Brewer's method; the above bound, however, is conservative and the actual value of φ is normally much larger.

The method of Des Raj depends on the order in which the units are drawn and it is known that Murthy's estimator is uniformly more efficient than that of Des Raj. The requirements (a) and (b) are not satisfied by Lahiri's estimator. Nevertheless we have included it in veiw of the recent work by Godambe (1966) based on concepts other than efficiency.

The methods in (1) have an advantage over the others in that the estimates become self-weighting with equal work loads within the selected primaries whereas the others require random work loads. We shall, however, not consider this as a factor in the choice among the methods.

The computations involved in applying the methods of Brewer, Murthy, Des Raj, Lahiri or Rao, Hartley and Cochran (R.H.C.) are very simple and about the same amount. Hanurav's method is slightly more involved whereas Fellegi's method involves simple, iterative calculations. In any case, the choice among these methods based on computational simplicity is not very realistic, especially when a high-speed computer is available.

We supplement our empirical study with a semi-theoretical study based on a super-population approach in which the finite population is regarded as being drawn from an infinite super-population. The results obtained apply only to the average of all finite populations that can be drawn from the super-population. We assume the following, often used, super-population model for the comparison of estimators:

$$y_i = \beta x_i + e_i, \; i=1,\dots,N$$
$$\epsilon(e_i|x_i) = 0, \; \epsilon(e_i^2|x_i) = ax_i^g \tag{1}$$
$$\epsilon(e_ie_j|x_i,x_j) = 0, \; a > 0, \; g \geq 0$$

where ϵ denotes the average over all the finite populations that can be drawn from the super-population. For the comparison of variance estimators we further assume that e_i's are normally distributed. In most practical situations, $1 \leq g \leq 2$. Some theoretical results are available on the relative efficiencies of the estimators (Rao, 1966) under the above model, but no guidelines are available with regard to the relative magnitudes. Nothing is known on the stabilities of the variance estimators under the super-population model.

2. Formulae

Let Y denote the population total of the characteristics of interest $y_i (i=1,2,\dots,N)$. For the methods in group (1), the Horvitz-Thompson estimator of Y is

$$\hat{Y}_1 = \frac{1}{2}\left(\frac{y_1}{p_1} + \frac{y_2}{p_2}\right) \tag{2}$$

with variance

$$V_1 = \sum_{i<j}^{N}\sum \left(p_ip_j - \frac{\pi_{ij}}{4}\right)\left(\frac{y_i}{p_i} - \frac{y_j}{p_j}\right)^2 \tag{3}$$

and variance estimator (Yates-Grundy)

$$v_1 = \frac{4p_1p_2-\pi_{12}}{4\pi_{12}}\left(\frac{y_1}{p_1} - \frac{y_2}{p_2}\right)^2 \tag{4}$$

where 1 and 2 denote the two units in the sample and $p_i = x_i/X$. The variance of the variance estimator is $Ev_1^2 - V_1^2$ where

$$16[Ev_1^2] = \sum_{i<j}^{N}\sum \frac{(4p_ip_j-\pi_{ij})^2}{\pi_{ij}}\left(\frac{y_i}{p_i} - \frac{y_j}{p_j}\right)^4. \tag{5}$$

The formulae for π_{ij} for the various methods can be obtained from Durbin (1967), Fellegi (1963) and Hanurav (1967).

For Des Raj's method, the estimator of Y is

$$\hat{Y}_2' = \frac{1}{2}\left[y_1' \frac{(1+p_1')}{p_1'} + y_2' \frac{(1-p_1')}{p_2'}\right] \tag{6}$$

with variance

$$V_2' = \frac{1}{4}\sum_{i<j}^{N}\sum p_ip(2-p_i-p_j)\left(\frac{y_i}{p_i} - \frac{y_j}{p_j}\right)^2 \tag{7}$$

and variance estimator

$$v_2' = \frac{(1-p_1')^2}{4}\left(\frac{y_1}{p_1} - \frac{y_2}{p_2}\right)^2 \tag{8}$$

where (y_1',p_1') and (y_2',p_2') denote the y- and p-values of the units selected at the first and second draws respectively. The variance of the variance estimator is $Ev_2'^2 - V_2'^2$ where

$$16[Ev_2'^2] = \sum_{i=0}^{3}\sum_{j=0}^{4} (-1)^{1-i-j}\binom{3}{i}\binom{4}{j} \cdot A_{j,i+j-4}\, A_{4-j,3-j} \tag{9}$$

where

$$A_{ij} = \sum_{t=1}^{N} \frac{y_t^i}{p_t^j}. \tag{10}$$

Murthy's estimator of Y is

$$\hat{Y}_3 = \frac{1}{2-p_1-p_2}\left[(1-p_2)\frac{y_1}{p_1} + (1-p_1)\frac{y_2}{p_2}\right] \tag{11}$$

with variance

$$V_3 = \sum_{i<j}^{N}\sum p_ip_j \frac{(1-p_i-p_j)}{(2-p_i-p_j)}\left(\frac{y_i}{p_i} - \frac{y_j}{p_j}\right)^2 \tag{12}$$

and variance estimator

$$v_3 = \frac{(1-p_1)(1-p_2)(1-p_1-p_2)}{(2-p_1-p_2)^2}\left(\frac{y_1}{p_2} - \frac{y_2}{p_2}\right)^2. \tag{13}$$

Also

$$Ev_3^2 = \sum_{i<j}^{N}\sum p_ip_j \frac{(1-p_i)(1-p_j)(1-p_i-p_j)^2}{(2-p_i-p_j)^3} \cdot \left(\frac{y_i}{p_i} - \frac{y_j}{p_j}\right)^4. \tag{14}$$

In the Rao-Hartley-Cochran method, the population is split at random into 2 groups of sizes N_1 and N_2 $(N_1+N_2=N)$ and a sample of size one is drawn with probabilities proportional to p_t from each of the two groups independently. Their estimator of Y is

$$\hat{Y}_4 = \frac{y_1}{p_1}P_1 + \frac{y_2}{p_2}P_2 \tag{15}$$

where $P_i = \sum_{\text{group } i} p_t$, $(i=1,2)$. The variance and variance estimator of $\hat{Y}_4$ are given by

$$V_4 = 2c_0c_1(A_{21} - A_{10}^2) \tag{16}$$

and

$$v_4 = c_0 \sum_1^2 P_i\left(\frac{y_i}{p_i} - \hat{Y}_4\right)^2$$

$$= c_0P_1P_2\left(\frac{y_1}{p_1} - \frac{y_2}{p_2}\right)^2 \tag{17}$$

respectively, where the A_{ij} are given by (10) and

$$c_0 = \frac{N_1^2+N_2^2-N}{N^2-N_1^2-N_2^2},\ c_1 = \frac{N^2-N_1^2-N_2^2}{N(N-1)}\ . \tag{18}$$

The derivation of Ev_4^2 is very tedious but straight forward. We have

$$\begin{aligned} c_0^{-2}[Ev_4^2] = {} & c_1[4A_{30}A_{10}-4A_{10}A_{31}+3A_{20}^2-6A_{40}] \\ & + c_2[A_{41}+2A_{42}A_{0,-2}-A_{42}+12A_{40} \\ & - 6A_{21}A_{20}+6A_{21}A_{2,-1}+3A_{21}^2-3A_{21}^2A_{0,-2} \\ & + A_{43}A_{0,-3}-12A_{30}A_{10}+12A_{31}A_{10} \\ & + 4A_{10}A_{32}A_{0,-2}-4A_{32}A_{1,-2}-8A_{31}A_{1,-1} \\ & + A_{43}-2A_{43}A_{0,-2}-4A_{32}A_{10}+4A_{32}A_{1,-1}] \\ & +c_3[2A_{42}-2A_{41}-A_{42}A_{0,-2}-6A_{20}^2 \\ & + 12A_{21}A_{20}-12A_{21}A_{2,-1}-3A_{21}^2 \\ & + 3A_{21}^2A_{0,-2}-2A_{43}A_{0,-3}+3A_{43}A_{0,-2} \\ & + 8A_{30}A_{10}-8A_{31}A_{10}+4A_{32}A_{10} \\ & - 4A_{32}A_{10}A_{0,-2}+8A_{31}A_{1,-1} \\ & - 8A_{32}A_{1,-1}-A_{43}+8A_{32}A_{1,-2}] \end{aligned} \tag{19}$$

where

$$c_2 = \frac{N_1N_2(N_1+N_2-2)}{N(N-1)(N-2)},\ c_3 = \frac{N_1N_2(N_1^2+N_2^2-3N+4)}{N(N-1)(N-2)(N-3)}. \tag{20}$$

Lahiri's estimator of Y is

$$\hat{Y}_5 = (y_1+y_2)/(p_1+p_2) \tag{21}$$

with variance

$$V_5 = \frac{1}{N-1}\sum_{i<j}^{N}\sum \frac{(y_i+y_j)^2}{p_i+p_j} - Y^2 \tag{22}$$

and variance estimator

$$v_5 = \hat{Y}_5^2 - \frac{1}{p_1+p_2}\left[(y_1-y_2)^2+2Ny_1y_2\right] \tag{23}$$

which takes negative values. Further

$$Ev_5^2 = \frac{1}{N-1}\sum_{i<j}^{N}\sum(p_i+p_j)\left[\hat{Y}_5^2 - \frac{1}{p_i+p_j}\left\{(y_i-y_j)^2 + 2Ny_iy_j\right\}\right]^2. \tag{24}$$

The variance of the customary estimator in p.p.s. sampling with replacement is given by

$$V_6 = (A_{21}-A_{10}^2)/2. \tag{25}$$

The variance estimator is

$$v_6 = \frac{1}{4}\left(\frac{y_1}{p_1} - \frac{y_2}{p_2}\right)^2 \tag{26}$$

and

$$8[Ev_6^2] = A_{43}+3A_{21}^2-4A_{32}A_{10}. \tag{27}$$

In (18) and (20) we have taken $N_i=N/2$ when N is even and $N_1=(N-1)/2$ and $N_2=(N+1)/2$ when N is odd.

3. Empirical Results

We have chosen 7 artificial and 20 natural populations for the empirical study. Table 1 gives the source, nature of y and x, coefficients of variation (C.V.) of y and x, correlation ρ and the ratio φ for the methods of Brewer, Fellegi and Hanurav. It is clear from Table 1 that we have a wide variety of populations with N ranging from 4 to 35, and C.V.(x) from 0.14 to 1.26 and ρ from 0.59 to 0.999. For the natural populations, Hanurav's criterion is satisfied by the methods of Brewer, Fellegi and Hanurav, except possibly for population 5 where $\tilde{x}_N$ is markedly different from $\tilde{x}_{N-1}$ and close to X/2. In general, Hanurav's ratio appears slightly larger than Fellegi's which in turn is slightly larger than Brewer's ratio. Turning to artificial populations, we see that all the three ratios are very close to zero for population 7 in which $\tilde{p}_N=0.49$ and $\tilde{p}_{N-1}=0.30$. For population 6, we took $\tilde{p}_N=0.42$ and $\tilde{p}_{N-1}=0.40$ so that Hanurav's ratio is considerably larger than Brewer's ratio. In any case, it is clear from these examples that none of the three methods guarantee that φ will be sufficiently away from zero for all populations. It also appears that, under Hanurav's criterion, the stabilities of the variance estimators should be about equal for these three methods. We shall, however, provide direct evidence on this point by computing the coefficients of variation of the variance estimators.

Table 1. Description of the populations

No	Source	y	x	N	C.V.(y)	C.V.(x)	ρ	φ Fellegi	φ Brewer	φ Hanurav
Artificial Populations										
1	Cochran (1963)	Artificial	Artificial	5	0.57	0.50	0.87	0.44	0.41	0.52
2	Cochran (1963)	Artificial	Artificial	5	0.68	0.50	0.997	0.44	0.41	0.52
3	Yates & Grundy (1953)	Artificial	Artificial	4	0.67	0.52	0.995	0.39	0.35	0.40
4	Yates & Grundy (1953)	Artificial	Artificial	4	0.50	0.52	0.88	0.39	0.35	0.40
5	Fellegi (1963)	Artificial	Artificial	6	0.64	0.25	0.93	0.52	0.52	0.53
6	Present Authors	Artificial	Artificial	4	0.72	0.74	0.999	0.33	0.21	0.46
7	Present Authors	Artificial	Artificial	4	0.78	0.80	0.997	0.07	0.05	0.06
Natural Populations										
1	Horvitz & Thompson (1952)	No. of households	Eye-estimated no.of households	20	0.44	0.40	0.87	0.49	0.49	0.50
2	DesRaj (1965) (Modification of 1)	No. of households	Eye-estimated no.of households	20	0.44	0.41	0.66	0.49	0.49	0.50
3	Rao (1963)	Corn acreage in 1960	Corn acreage in 1958	14	0.39	0.43	0.93	0.49	0.49	0.50
4	Kish (1965) blocks 1-10	No.of rented dwelling units	Total no.of dwelling units	10	1.45	1.15	0.99	0.36	0.32	0.42
5	Kish (1965) blocks 11-20	No.of rented dwelling units	Total no.of dwelling units	10	1.25	1.26	0.98	0.26	0.21	0.23
6	Cochran (1963)	Wt. of peaches	eye-estimated wt. of peaches	10	0.19	0.17	0.97	0.53	0.53	0.53
7	Hanurav (1967)	Population in 1967	Population in 1957	20	0.30	0.30	0.97	0.50	0.50	0.51
8	Hanurav (1967)	Population in 1967	Population in 1957	19	0.45	0.44	0.97	0.47	0.47	0.50
9	Hanurav (1967)	Population in 1967	Population in 1957	16	0.66	0.65	0.99	0.47	0.47	0.49
10	Hanurav (1967)	Population in 1967	Population in 1957	17	0.51	0.52	0.96	0.48	0.48	0.50
11	Cochran (1963)	No.of persons per block	No. of rooms per block	10	0.15	0.14	0.65	0.53	0.53	0.54
12	Cochran (1963) Cities 1-16	No. of people in 1930	No. of people in 1920	16	0.98	0.98	0.99	0.44	0.43	0.50

Table 1 continued

13	Cochran (1963) Cities 17-32	No. of people in 1930	No. of people in 1920	16	1.14	1.19	0.98	0.39	0.38	0.40
14	Cochran (1963) Cities 33-49	No. of people in 1930	No. of people in 1920	17	0.79	0.91	0.97	0.46	0.45	0.49
15	Sukhatme (1954) Villages 1-10	No.of wheat acres in 1937	No. of wheat acres in 1936	10	0.65	0.59	0.98	0.46	0.45	0.49
16	Sukhatme (1954) Villages 11-20	No. of wheat acres in 1937	No. of wheat acres in 1936	10	0.94	0.93	0.99	0.41	0.39	0.48
17	Sampford (1962)	Oats acreage in 1957	Total acreage in 1947	35	0.71	0.71	0.83	0.49	0.49	0.50
18	Sukhatme (1954) Circles 1-20	Wheat acreage	No. of villages	20	0.63	0.50	0.59	0.48	0.48	0.50
19	Sukhatme (1954) Circles 21-40	Wheat acreage	No. of villages	20	0.61	0.46	0.76	0.48	0.48	0.50
20	Sukhatme (1954) Circles 81-89	Wheat acreage	No. of villages	9	0.47	0.65	0.69	0.45	0.44	0.47

3.1. Stabilities of the estimator.

We first consider the stabilities of the estimators. Table 2 gives the percent gains in efficiency of the estimators over Brewer's estimator (i.e., [V(Brewer's est.)/V(est.)-1]x100), for the populations of Table 1. The following tentative conclusions can be drawn from Table 2: (1) For the natural populations, the efficiencies of Hanurav's, Brewer's and Fellegi's estimators are essentially identical; for the artificial populations, however, Hanurav's estimator appears slightly less efficient than the latter. (2) Murthy's estimator is consistently more efficient than the R.H.C. estimator and the gains are considerable for the artificial populations and the natural populations with small N and moderately large C.V.(x) (Natural pops. 4 and 5). For the natural populations, the R.H.C. estimator compares favorably with the estimators of Brewer, Fellegi and Hanurav. (3) The loss in efficiency of Des Raj's estimator over Murthy's estimator is very small for the natural populations, excepting populations 4 and 5. It is, however, considerable for the artificial populations. (4) Lahiri's estimator is considerably more efficient than the others when one or two units in the population have large x_i relative the x_i of the remaining units and samples containing these units give good estimates of Y (e.g., natural pops. 12-14). However, it is considerably less efficient for other populations and, in fact, less efficient than the customary estimator in sampling with replacement for six of the natural populations. (5) For the natural populations, Murthy's estimator appears more efficient than those of Brewer, Fellegi and Hanurav (gains range from -2 to 18%). However, for the artificial populations it is not clear cut. In any case, it appears on the whole that Murthy's estimator compares favorably with those of Brewer, Fellegi and Hanurav.

3.2. Stabilities of the variance estimators.

We now compare the stabilities of the variance estimators. Table 3 gives the percent gains in efficiency of the variance estimators over Brewer's variance estimator (i.e., 100 x [C.V.^{2}(Brewer's variance estimator)/C.V.^{2}(variance estimator)-1]) for the populations of Table 1. The following tentative conclusions can be drawn from Table 3: (1) Lahiri's variance estimator is considerably less efficient than the others for natural as well as artificial populations. Henceforth, we shall exclude Lahiri's method from further discussion. (2) Stabilities of Murthy's and Des Raj's variance estimators are essentially equal. It is, however, not true that Murthy's variance estimator will always be more stable than the latter. (3) The R.H.C. variance estimator is more efficient than Murthy's and other variance estimators. (4) Murthy's variance estimator is consistently more efficient than those of Brewer, Fellegi and Hanurav. The gains are considerable for several of the artificial as well as natrual populations. (5) For the

Table 2. Percent gains in efficiency* of the estimators over Brewer's estimator.

Pop. No.	Hanurav	Fellegi	Murthy	Des Raj	R.H.C.	Lahiri	With Rep.
			Artificial Populations				
1	4	1	- 8	-19	-23	-31	-39
2	-8	-2	5	- 2	- 3	1	-22
3	-3	-2	-10	-23	-15	-22	-44
4	1	3	-15	-32	-29	-41	-52
5	-1	-0	- 1	- 6	- 3	- 9	-22
6	-4	-1	5	-21	-19	-37	-46
7	-0	-1	61	32	40	19	- 7
			Natural Populations				
1	+0	+0	- 1	- 1	- 2	-16	- 7
2	-0	-0	- 0	- 0	- 1	-11	- 6
3	-0	-0	1	+ 0	1	1	- 7
4	-3	-2	4	- 1	- 7	-31	-17
5	-0	+0	18	12	7	5	- 5
6	-0	-0	+ 0	- 1	+ 0	1	-11
7	-0	-0	+ 0	+ 0	+ 0	1	- 5
8	+0	+0	- 0	- 0	- 2	-13	- 7
9	+0	+0	- 0	- 1	- 3	-17	-10
10	+0	-0	- 0	- 1	- 2	-12	- 8
11	+0	+0	1	- 1	1	4	-10
12	+0	+0	6	5	4	34	- 3
13	+0	-0	9	8	7	511	- 1
14	-0	-0	4	4	3	33	- 3
15	1	+0	- 2	- 4	- 6	-17	-17
16	1	+0	7	4	3	15	- 9
17	-0	-0	+ 0	- 0	- 1	-17	- 4
18	-0	-0	1	+ 0	- 0	- 4	- 5
19	+0	+0	+ 0	+ 0	- 0	- 2	- 6
20	-1	-0	6	4	5	28	- 7

*+0 and -0 indicate that the actual values are positive and negative respectively.

natural populations, stabilities of Brewer's, Fellegi's and Hanurav's variance estimators are about equal, except that, for populations 12 and 16, the gains in efficiency of Hanurav's variance estimator over that of Brewer are 15% and 17% respectively. However, for these two populations Murthy's estimator performs considerably better. For the artificial populations, Hanurav's variance estimator appears less stable than that of Brewer. In this connection, it is interesting to note that Hanurav's variance estimator is less stable than Brewer's for the artificial population 6, although his ratio φ is considerably larger than Brewer's φ (see Table 1). This clearly shows that Hanurav's criterion does not always guarantee a more stable variance estimator — the stability also depends on the differences $y_i/x_i - y_j/x_j$. (6) For all populations, the R.H.C. variance estimator is more efficient than the customary variance estimator in sampling with replacement. This is, however, not true with regard to the other variance estimators. (7) Gains in efficiency of the R.H.C. variance estimator over Murthy's are not large, excepting few extreme cases.

4. Formulae under the super-population model.

Using the model (1) we get the following average variances of the estimators:

$$\epsilon V_1 = \frac{aX^g}{2} \sum_{1}^{N}(1-2p_i)p_i^{g-1} \tag{28}$$

$$\epsilon V_2 = \frac{aX^g}{2} \sum_{i<j}^{N}\sum p_i^{g-1}p_j(2-p_i-p_j) \tag{29}$$

$$\epsilon V_3 = 2aX^g \sum_{i<j}^{N}\sum p_i^{g-1}p_j \frac{1-p_i-p_j}{2-p_i-p_j} \tag{30}$$

$$\epsilon V_4 = ac_0c_1X^g \sum_{1}^{N}(1-p_i)p_i^{g-1}. \tag{31}$$

In view of our results in Section 3 we have not included Lahiri's method here, but it is known that $\epsilon V_5 \gtrless \epsilon V_1$ according as $g \gtrless 1$.

It is clear from (28) that all methods with $\pi_i=2p_i$ and using the Horvitz-Thompson estimator have the same average variance. It is also known

Table 3. Percent gains in efficiency of the variance estimators over Brewer's variance estimator.

Pop. No.	Hanurav	Fellegi	Murthy	Des Raj	R.H.C.	Lahiri	With Rep.
			Artificial Populations				
1	- 5	2	61	57	53	- 99	-10
2	-16	-4	10	- 0	28	-100	-26
3	- 7	-5	31	14	74	-100	-16
4	8	8	418	440	512	- 99	164
5	- 2	-1	- 0	- 2	2	- 94	-30
6	-10	-5	185	252	789	-100	200
7	- 3	-8	2548	2755	6083	- 99	2363
			Natural Populations				
1	1	+0	- 3	- 3	- 5	- 99	-13
2	- 0	-0	+ 0	- 0	- 1	- 89	- 7
3	- 0	-0	6	6	10	-100	- 6
4	- 2	2	38	24	59	- 99	8
5	- 5	7	301	303	508	- 97	322
6	- 0	-0	3	3	7	-100	-12
7	0	+0	1	1	2	-100	- 5
8	- 0	+0	1	1	1	-100	- 9
9	- 2	-0	4	4	5	-100	- 9
10	- 1	-0	4	4	9	-100	- 3
11	+ 0	+0	3	3	5	- 99	-12
12	15	2	22	21	39	-100	19
13	- 1	+2	39	34	54	26	30
14	2	+0	8	7	15	- 96	2
15	2	-0	13	11	20	-100	- 6
16	17	5	38	38	75	-100	36
17	1	+0	2	2	4	- 97	- 3
18	- 0	+0	4	4	8	- 90	- 3
19	- 0	+0	5	5	9	- 90	+ 0
20	- 1	1	16	13	27	- 87	3

that $\epsilon V_4 \gtrless \epsilon V_1$ according as $g \gtrless 1$ and $\epsilon V_4 = \epsilon V_1$ if g=1; $\epsilon V_3 < \epsilon V_4$ if g=2; $\epsilon V_1 < \epsilon V_3$ if g=2 and $\epsilon V_1 > \epsilon V_3$ if g=1.

For the comparison of variance estimators, we further assume that the e_i are normally distributed so that $\epsilon(e_i^4) = 3a^2x_i^{2g}$. The most appropriate measure of the stability of v_i appears to be $\epsilon[C.V.^2(v_i)]$, i.e., average $(C.V.)^2$ of the variance estimator. However, since $\epsilon[C.V.^2(v_i)]$ is the expectation of the ratio of two random variables, the evaluation is difficult. We have, therefore, used the alternative measure

$$\frac{\epsilon E[v_i-\epsilon V_i]^2}{(\epsilon V_i)^2} = \frac{\epsilon[EV_i^2]-(\epsilon V_i)^2}{(\epsilon V_i)^2} \qquad (32)$$

which is readily evaluable. Notice that (32) actually measures the variability of v_i around the average variance ϵV_i. We, however, expect that (32) and $\epsilon[C.V.^2(v_i)]$ would lead to same conclusions.

To evaluate (32) we need the following formulae:

$$a^{-2}\epsilon[Ev_1^2] = \frac{3}{16} X^{2g} \sum_{i<j}^{N}\sum \frac{(4p_ip_j-\pi_{ij})^2}{\pi_{ij}} \cdot (p_i^{g-2}+p_j^{g-2})^2 \qquad (33)$$

$$a^{-2}\epsilon[Ev_2^2] = \frac{3}{8} X^{2g} \sum_{i<j}^{N}\sum p_ip_j((1-p_i)^3 \cdot (p_i^{g-2}+p_j^{g-2})^2 \qquad (34)$$

$$a^{-2}\epsilon[Ev_3^2] = 3X^{2g} \sum_{i<j}^{N}\sum p_ip_j \cdot \frac{(1-p_i)(1-p_j)(1-p_i-p_j)^2}{(2-p_i-p_j)^3} \cdot (p_i^{g-2}+p_j^{g-2})^2 \qquad (35)$$

and

$$a^{-2}\epsilon[Ev_4^2] = c_0^2X^{2g}[3c_1(B_{20}^2-4B_{41}) + 3c_2(4B_{02}B_{42}-B_{02}B_{21}^2+2B_{21}B_{2,-1}$$

Table 4. Percent gains in average efficiency of the estimators over Brewer's estimator (under the super-population model for g=1.0, 1.5, 1.75, 2.0)

	g=1		g=1.5			g=1.75			g=2.0		
Pop. No.	Murthy	Des Raj	Murthy	Des Raj	R.H.C.	Murthy	Des Raj	R.H.C.	Murthy	Des Raj	R.H.C.
				Artificial Populations							
1,2	4	-4	1	- 8	-11	+0	- 9	-13	-1	-11	-14
3,4	6	-9	2	-15	- 7	-1	-17	-11	-3	-20	-14
5	1	-4	+0	- 4	- 1	+0	- 4	- 1	-0	- 5	- 2
				Natural Populations							
1	+ 0	+0	+0	- 0	- 0	+0	- 0	- 1	-0	- 0	- 1
2	+ 0	+0	+0	- 0	- 0	+0	- 0	- 1	-0	- 0	- 1
3	1	+0	+0	- 0	- 1	+0	- 1	- 1	-0	- 1	- 2
4	10	5	4	- 2	-10	1	- 6	-15	-2	-10	-19
5	12	6	5	- 2	-11	-0	- 8	-17	-6	-14	-23
6	+ 0	-1	+0	- 1	- 0	+0	- 1	- 0	-0	- 1	- 0
7	+ 0	-0	+0	- 0	- 0	+0	- 0	- 0	-0	- 0	- 1
8	1	+0	+0	- 0	- 1	+0	- 0	- 1	-0	- 0	- 1
9	2	1	1	- 0	- 1	+0	- 1	- 2	-0	- 1	- 3
10	1	+0	+0	- 0	- 1	+0	- 0	- 2	-0	- 1	- 2
11	+ 0	-1	+0	- 1	- 0	+0	- 1	- 0	-0	- 1	- 0
12	4	2	2	+ 0	- 3	1	- 1	- 5	-1	- 3	- 7
13	6	4	3	1	- 4	1	- 2	- 7	-2	- 5	-11
14	3	2	1	+ 0	- 3	1	- 1	- 5	-0	- 2	- 6
15	2	+0	1	- 1	- 2	+0	- 2	- 3	-0	- 3	- 5
16	6	3	2	- 2	- 7	1	- 4	-10	-1	- 6	-12
17	1	1	+0	+ 0	- 1	+0	- 0	- 1	-0	- 0	- 2
18	1	+0	+0	- 0	- 1	+0	- 0	- 1	-0	- 1	- 1
19	1	+0	+0	- 0	- 1	+0	- 0	- 1	-0	- 0	- 1
20	3	+0	1	- 2	- 5	+0	- 3	- 6	-1	- 4	- 8

$$
\begin{aligned}
&- 2B_{21}B_{20}+B_{21}^{2}+B_{43}-3B_{42}+13B_{41}-8B_{40}+B_{0,-3}B_{43} \\
&- 2B_{0,-2}B_{43})+3c_3(B_{0,-2}B_{21}^{2}-4B_{0,-2}B_{42} \\
&- 2B_{20}^{2}-4B_{21}B_{2,-1}+4B_{21}B_{20}-B_{21}^{2}-B_{43}+4B_{42} \\
&- 10B_{41}+12B_{40}-2B_{0,-3}B_{43}+3B_{0,-2}B_{43})] \qquad (35)
\end{aligned}
$$

where $c_0,\ldots,c_3$ are as before and

$$B_{0j} = \sum_{1}^{N} p_t^{-j},\quad B_{2j} = \sum_{1}^{N} p_t^{g-j}$$

and

$$B_{4j} = \sum_{1}^{N} p_t^{2g-j}. \qquad (36)$$

5. Empirical results under the super-population model.

5.1. Stabilities of the estimators.

Table 4 gives the percent gains in average efficiency of the estimators over Brewer's estimator (i.e., 100 x [ϵV(Brewer's est.)/ϵV(est.)-1]) for the populations of Table 1 (excluding artificial populations 6 and 7) and g=1.0, 1.5, 1.75 and 2.0. The following tentative conclusions can be drawn from Table 4: (1) Murthy's estimator is more efficient than the Horvitz-Thompson estimator (i.e., Brewer's, Fellegi's and Hanurav's) for $g \leq 1.75$; however, the gains are small for $g \geq 1.5$. Moreover, the losses in efficiency over the latter for g=2 are small. (2) Murthy's estimator is consistently more efficient than the R.H.C. estimator and the gains are considerable for several populations. (3) Gains in efficiency of Murthy's estimator over Des Raj's are considerable for populations with small N or moderately large C.V.(x). (4) Des Raj's estimator is less efficient than the Horvitz-Thompson estimator for $g \geq 1.5$.

5.2. Stabilities of the variance estimators.

Using the measure (32) we have computed the percent gains in average efficiency of the variance estimators over Brewer's for g=1.0, 1.5, 1.75 and 2.0, and the results are given in Table 5. The following tentative conclusions can be drawn from Table 5: (1) As before, the stabilities of Murthy's and Des Raj's variance estimators are essentially equal. (2) The R.H.C. variance estimator is consistently more efficient than Murthy's and other variance estimators. However, as before, the gains over Murthy's variance estimator are not large, excepting for few extreme cases. (3) Murthy's variance estimator is consistently

Table 5. Percent gains in average efficiency of the variance estimators over Brewer's variance estimator (under the assumption of a super-population model for $g = 1.0, 1.5, 1.75, 2.0$).

Pop. No.	g=1.0 Murthy	Des Raj	R.H.C.	Hanurav	Fellegi	g=1.5 Murthy	Des Raj	R.H.C.	Hanurav	Fellegi
				Artificial Populations						
1,2	50	51	82	15	5	43	43	62	8	4
3,4	155	165	247	8	9	157	162	219	6	6
5	8	9	13	+ 0	1	6	6	9	+ 0	+0
				Natural Populations						
1	3	3	5	+ 0	+0	2	2	3	+ 0	+0
2	3	3	6	+ 0	+0	2	2	3	+ 0	+0
3	6	6	10	1	+0	4	4	6	+ 0	+0
4	70	67	130	15	9	87	86	143	12	8
5	277	268	433	- 4	8	370	362	543	- 3	4
6	2	2	3	+ 0	+0	1	1	2	+ 0	+0
7	2	2	3	+ 0	+0	1	1	2	+ 0	+0
8	8	8	16	+ 0	+0	2	2	4	+ 0	+0
9	10	10	18	2	+0	8	8	13	1	+0
10	8	8	16	1	+0	4	4	7	+ 0	+0
11	1	1	2	+ 0	+0	1	1	1	- 0	+0
12	24	23	42	13	2	22	22	36	8	1
13	54	52	85	- 0	2	68	65	97	+ 0	1
14	16	16	30	7	1	14	14	24	4	1
15	17	17	31	3	1	13	13	21	1	+0
16	34	33	70	9	3	35	34	60	7	3
17	4	4	8	1	+0	3	3	5	+ 0	+0
18	5	5	9	0	+0	3	3	5	+ 0	+0
19	4	4	8	0	+0	2	2	4	+ 0	+0
20	26	26	45	3	1	22	22	34	1	1

more efficient than those of Brewer, Fellegi and Hanurav and the gains are considerable for several of the artificial as well as the natural populations. (4) Fellegi's variance estimator is consistently more efficient than Brewer's; however, the gains are small. The efficiencies of Hanurav's and Fellegi's variance estimators are essentially equal for $g \geq 1.75$ although the latter is consistently more efficient for g=2. Hanurav's variance estimator is slightly more efficient for $g \leq 1.5$.

6. Concluding Remarks.

It appears that our results under the super-population model are in agreement with those from the empirical study using the actual y-data. The following major conclusions may be drawn from our studies: (1) Murthy's method is preferable over the other methods when a stable estimator as well as a stable variance estimator are required. (2) The R.H.C. variance estimator is the most stable, but the R.H.C. estimator might lead to significant losses in efficiency. (3) Hanurav's method does not lead to significant improvements over Fellegi's or Brewer's methods with regard to stability of the variance estimator.

It should be noted that, for the case of $n > 2$, some of these methods are either not applicable (e.g., Brewer's method) or become computationally cumbersome (e.g., Murthy's method when n is moderate). Therefore, the case of $n > 2$ could lead to completely different conclusions. A detailed investigation of the stabilities for $n > 2$ is underway and the results will be reported in a subsequent paper.

Acknowledgment.

This research is supported by the National Science Foundation, under Grant GP-6400.

References.

Brewer, K. R. W. (1963), "A model of systematic sampling with unequal probabilities", Australian Journal of Statistics, 5, 5-13.

Carroll, J. L. and Hartley, H. O. (1964), "The symmetric method of unequal probability sampling without replacement", unpublished to date; abstract in Biometrics, (1964), 20, 908-9.

Cochran, W. G. (1963), Sampling Techniques, 2nd Ed., New York: John Wiley and Sons.

Des Raj (1956), "Some estimators in sampling with varying probabilities without replacement", Journal of the American Statistical Association, 51, 269-84.

/

Table 5 continued

	g=1.75					g=2.00				
Pop. No.	Murthy	Des Raj	R.H.C.	Hanurav	Fellegi	Murthy	Des Raj	R.H.C.	Hanurav	Fellegi
				Artificial Populations						
1,2	38	37	49	5	3	32	30	36	1	2
3,4	152	153	195	4	5	144	141	170	3	3
5	5	5	6	-0	+0	4	3	4	-0	+0
				Natural Populations						
1	1	1	2	+0	+0	1	+ 0	1	-0	+0
2	1	1	2	+0	+0	1	+ 0	1	-0	+0
3	3	3	4	+0	+0	2	1	2	-0	+0
4	82	79	112	7	5	70	63	76	3	3
5	400	377	518	-0	2	406	361	457	-1	+0
6	1	1	1	-0	+0	+ 0	+ 0	1	-0	+0
7	1	1	1	-0	+0	+ 0	+ 0	+ 0	-0	+0
8	1	1	1	-0	+0	+ 0	+ 0	+ 0	-0	+0
9	6	6	9	+0	+0	4	3	5	-0	+0
10	2	2	4	+0	+0	1	1	1	-0	+0
11	+ 0	+ 0	1	-0	+0	+ 0	+ 0	+ 0	-0	+0
12	19	18	27	4	1	13	11	15	-2	+0
13	66	61	86	+0	1	58	51	66	-0	+0
14	11	10	17	1	+0	7	5	8	-1	+0
15	10	9	14	1	+0	7	6	8	-0	+0
16	28	26	40	4	2	19	17	22	+0	1
17	2	2	3	+0	+0	+ 0	+ 0	1	-0	+0
18	2	2	3	+0	+0	1	1	1	-0	+0
19	2	1	3	+0	+0	1	1	1	-0	+0
20	18	17	25	+0	1	13	12	15	-0	+0

Des Raj (1965), "Variance estimation in randomized systematic sampling with probability proportional to size", Journal of the American Statistical Association, 60, 278-84.

Durbin, J. (1967), "Estimation of sampling errors in multistage surveys", Applied Statistics, 16, 152-64.

Fellegi, I. P. (1963), "Sampling with varying probabilities without replacement: rotating and non-rotating samples", Journal of the American Statistical Association, 58, 183-201.

Godambe, V. P. (1966), "A new approach to sampling from a finite population-I", Journal of the Royal Statistical Society, Series B, 28, 310-19.

Hájek, J. (1964), "Asymptotic theory of rejective sampling with varying probabilities from a finite population", Annals of Mathematical Statistics, 35, 1491-523.

Hanurav, T. V. (1967), "Optimum utilization of auxiliary information: πPS sampling of two units from a stratum", Journal of the Royal Statistical Society, Series B, 29, 374-91.

Hartley, H. O. and Rao, J. N. K. (1962), "Sampling with unequal probabilities and without replacement", Annals of Mathematical Statistics, 33, 350-74.

Hartley, H. O. (1966), "Systematic sampling with unequal probability and without replacement", Journal of the American Statistical Association, 61, 739-48.

Horvitz, D. G. and Thompson, D. J. (1952), "A generalization of sampling without replacement from a finite universe", Journal of the American Statistical Association, 47, 663-85.

Kish, L. (1965), Survey Sampling. New York: John Wiley and Sons.

Lahiri, D. B. (1951), "A method for sample selection providing unbiased ratio estimates", International Statistical Institute Bulletin, 33, No. 2, 133-40.

Madow, W. G. (1949), "On the theory of systematic sampling II", Annals of Mathematical Statistics, 20, 333-54.

Murthy, M. N. (1957), "Ordered and unordered estimators in sampling without replacement", Sankhyā, 18, 379-90.

Rao, J. N. K., Hartley, H. O., and Cochran, W. G. (1962), "On a simple procedure of unequal probability sampling without replacement", Journal of the Royal Statistical Society, Series B, 24, 482-91.

Rao, J. N. K. (1963), "On three procedures of unequal probability sampling without replacement", Journal of the American Statistical Association, 58, 202-15.

Rao, J. N. K. (1965), "On two simple schemes of unequal probability sampling without replacement", Journal of the Indian Statistical Association, 3, 173-80.

Rao, J. N. K. (1966), "On the relative efficiency of some estimators in PPS sampling for multiple characteristics", Sankhya, 28, 61-70.

Sampford, M. R. (1962). An introduction to sampling theory. Edinburgh and London: Oliver and Boyd Ltd.

Stevens, W. L. (1958), "Sampling without replacement with probability proportional to size", Journal of the Royal Statistical Society, Series B, 20, 393-7.

Sukhatme, P. V. (1954). Sampling theory of surveys with applications. Ames, Iowa: Iowa State College Press.

Yates, F. and Grundy, P. M. (1953), "Selection without replacement from within strata with probability proportional to size", Journal of the Royal Statistical Society, Series B, 15, 253-61.

XIII

CONTRIBUTED PAPERS III

Chairman, ROBERT D. GROVE, National Center for Health Statistics

INDIANS AND SMUDGES ON THE CENSUS SCHEDULE

by Donald S. Akers and Elizabeth A. Larmon
U. S. Bureau of the Census

This is an explanation of how certain anomalies in the reported characteristics of American Indians in the 1960 Census are to be explained by systematic smudging of the Census schedules.

The Bureau of the Census has long recognized that the census counts are subject to error. Every effort is made to identify the errors, to correct them if possible, and to seek means of preventing their recurrence. One of the most useful means of identifying errors is to review the census results for reasonableness. To trace the source of the error, once its existence is discovered, may at times require an intensive exercise in deductive logic as well as an intimate knowledge of census procedures.

One such exercise in identification of a census error by deduction, this time by experts outside the Bureau, received some publicity a few years ago. Coale and Stephan in their article "The Case of the Indians and the Teen-Age Widows,"1/ relating to the 1950 Census, noted that the tabulated number of widowed males under age 20 was excessive and that the excess was greater the younger their age. They were able to demonstrate that the error occurred when the census information was transferred to punch cards. When the card was punched one column off, an excessive number of widows and Indians would be generated, all of whom would be in their teens or twenties. Coale and Stephan concluded by saying:

> "The Bureau of the Census changed over in 1960 to data sensing machinery to transcribe information onto magnetic tape, and the specific problem of a shift in columns is no longer relevant to census operations. The new set of processing operations poses new problems of error control for the Bureau, and may possibly cause misleading figures to show up in new and unsuspected ways in small cells. Users must continue to regard such data with special care."

As expected, a variety of irregularities and inconsistencies have appeared in the 1960 Census counts. Some relate to the counts of Indians once more. They are small compared to the total population but are large enough to distort the statistics for the Indian population.

Only the total number of Indians and their sex distribution were tabulated from the complete count. Their age distribution and all other characteristics were tabulated from the 25-percent sample and published in the report Nonwhite Population by Race, U.S. Census of Population: 1960, Final Report PC(2)-1C.

In these tabulations there is a marked excess of Indians at ages 55 to 59 (Table A); the number in this age group exceeds those 50 to 54 years old by one half. The only reasonable explanation for this excess is census error.

TABLE A.--AMERICAN INDIANS 35 YEARS OLD AND OVER, BY AGE: 1960
(Based on 25-percent sample)

Age	Number	Year of birth
Total, 35 and over	162,783	1920 and before
35-39	28,389	1920-24
40-44	22,929	1915-19
45-49	21,711	1910-14
50-54	20,767	1905-09
55-59	31,560	1900-04
60-64	11,830	1895-99
65-69	9,975	1890-94
70-74	6,857	1885-89
75 and over	8,765	1884 and before

At first, it was thought that the error in age might be related to an excess that had been observed in the count of the total population at age 59 (Table B). However, there are incongruities in the characteristics of Indians at these

TABLE B.--POPULATION 50 TO 64 YEARS OLD, BY AGE, ACCORDING TO SAMPLE AND COMPLETE COUNT: 1960

Age	Sample Count	Complete Count
50 to 54	9,696,502	9,605,954
50	2,035,449	-
51	1,979,215	-
52	1,951,935	-
53	1,856,757	-
54	1,873,146	-
55 to 59	8,595,947	8,429,865
55	1,801,394	-
56	1,702,574	-
57	1,697,037	-
58	1,536,568	-
59	1,858,374	-
60 to 64	7,111,897	7,142,452
60	1,504,160	-
61	1,441,041	-
62	1,386,263	-
63	1,369,684	-
64	1,410,749	-

ages which are not to be found in the total population. There are quite unreasonable excesses in the percent single and the percent counted as Other Relative at these ages (Table C). Half of the employed at these ages have neither occupation nor industry of worker reported.

TABLE C.--SELECTED CHARACTERISTICS OF AMERICAN INDIANS, 35 YEARS OLD AND OVER, BY AGE: 1960
(Based on 25-percent sample)

Characteristic	35 to 44	45 to 64	65 and over
Percent single:			
Male	13.7	25.9	7.3
Female	7.5	22.0	3.0
Percent other relative	12.9	27.2	18.4
Percent of employed:			
Occupation not reported	10.1	51.6	11.8
Industry not reported	9.3	33.8	11.2

These peculiarities in the age distribution and characteristics of the Indian population seem to be limited to a few areas, places where Indians represent a relatively small proportion of the total population. The excess at age 55 to 59 over other ages, and the anomalies in other characteristics are largely limited to the urban population and are most striking in a few States, such as New York, Illinois, Florida, Kansas and Texas.

Early efforts to explain the error failed, but in 1967, the investigation was reopened, when the Census Bureau was asked to make projections of the Indian population by age. The first step in the investigation was to obtain a printout for the State of Kansas showing the detailed information for all Indians 45 to 69 years old on the computer tape for the 25-percent sample. A portion of this printout, presented in Table D, shows the number of Indians in Kansas aged 45 to 69 by single years of age. Instead of a concentration at age 59, which would be expected if year of birth was rounded to 1900, there were excesses at ages 57 and 58. There were 90 Indians aged 57 and 58 where about 12 might have been reasonably expected. Of the 90 cases, 85 were recorded as born in 1902. Clearly, some bias was operating in favor of one particular year of birth.

It also appeared that in most of these cases only race (Indian) and year of birth (1902) were reported. For most of the 85 cases, the other 3 complete-count items were blank. Sex had to be allocated for 76 of the cases, marital status for 75, and relationship for 84.

The cases were scattered throughout the State. Generally, each Indian born in 1902 was the only Indian recorded in this enumeration district. In only one instance were there as many as four in the district.

Each piece of evidence made the case more puzzling until it was suggested that the schedules might have been smudged in some systematic way during transcription to the sample questionnaires. This explanation proved to be the correct one.

The 1960 Census schedules were designed for use by FOSDIC (Film Optical Sensing Device for Input to Computer). Entries were made by filling in the appropriate circles on the schedule with pencil. It was recognized from the first that smudging could be an important source of error. The FOSDIC scanner has no way of telling an accidental smudge on the circle from a purposeful marking by the enumerator. There would be little chance of a smudge causing an error if it occurred for an item already filled by the enumerator, because the enumerator's entry would be darker than the smudge. The FOSDIC scanner selects only one circle for each item, the one with the darkest marking.

The FOSDIC schedules in the 1960 Census were printed on both sides and were bound in spiral notebooks. When the enumerator made an entry, his pencil pressed the reverse side of the sheet against the sheet below. The point where pressure was applied on the sheet below could be a FOSDIC circle if the sheets were aligned exactly, for the sheets were identical. If there was already a

TABLE D.--TALLY FROM 25-PERCENT SAMPLE FOR INDIANS IN KANSAS 45 TO 69 YEARS OLD, BY SINGLE YEARS OF AGE: 1960

Age	Cases
45	8
46	10
47	11
48	7
49	13
50	9
51	9
52	12
53	8
54	8
55	7
56	5
57	60
58	30
59	7
60	3
61	3
62	4
63	5
64	3
65	0
66	3
67	5
68	5
69	4

pencil mark in this circle, it could leave a smudge on the reverse side of the sheet being marked. If the smudge happened to fall on a blank circle on the reverse side of the sheet, the FOSDIC scanner might read it as an entry.

It so happened that the circles for certain characteristics in the panel of questions for the first and third individuals on the FOSDIC questionnaire (Form PH-3 and PH-4) were back to back. The left-hand page of the FOSDIC book had the panel of housing questions at the top and the first population panel at the bottom. The right-hand page contained two population panels. Thus, the first population panel, where information for the head of the household was normally entered, was backed by a population panel for the preceding household. The FOSDIC circles for most of the 100-percent items in these two panels coincided exactly. The alignment of the circles is shown in Table E. An inspection of the table will show that if the first individual on successive sheets is coded as Head of household, white, born in the decade 1890-1899, the reverse side of the top sheet may show smudges on the circles for an Indian, born in the 1900's, and in the specific year, 2.

The three smudges were sufficient to make the computer mistake a blank panel for a person. The FOSDIC scanner cannot tell whether a name has been entered in a panel, so the computer must rely on the presence of a coded entry to tell whether the enumerator had meant to enter someone or had intended to leave the panel blank. It was recognized that smudging on a blank panel might lead to the inclusion of the panel in the count of population. To minimize such spurious counting, the computer was instructed to count a panel as a person only if there were at least two entries among the five characteristics reported in the complete count (relationship to Head of household, sex, race, age, and marital status) of which at least one was required to be relationship, sex, or race. By this rule, the smudges on Indian race and 1902 year of birth were just sufficient to generate a fictitious person.

This explanation for the error in the Indian data was confirmed by an examination of the census schedules. The microfilms for selected enumeration districts where the printout recorded one or more Indians born in 1902 were examined for an entry of Indian. No actual entry of an Indian was found, but there were eleven cases observed where smudges had been left on Indian, and on the decade of 1900-1909 and the specific year 2, for date of birth. In every case, there were entries of head, white, and the 1890 decade on both the reverse side of the same sheet and on the face of the preceding sheet. In one case, the smudges fell on a panel where there were actual entries, so that the FOSDIC scanner would have ignored the smudges, but in the other cases they fell on blank panels. On three of these, the smudges must have been sufficiently faint for FOSDIC to pass over them, because only seven cases were recorded in the printout. Smudges were also found which marked FOSDIC circles off-center or fell in spaces where there were no circles. These, of course,

TABLE E.--THE ALIGNMENT OF CIRCLES ON THE FOSDIC SAMPLE SCHEDULE

First Population Panel	Third Population Panel
Relationship:	Specific year of birth:
Head	2
Wife	1
Son/Daughter	0
Relative	-
Nonrelative	-
	Decade of birth:
Inmate	1960
Sex:	
Male	1930
Female	1920
Color or race:	
White	1900
Negro	--
Indian	1890
Japanese	1870
Chinese	1860
Filipino	1850
Other	--
Decade of birth:	Color or race:
1850	Filipino
1860	Chinese
1870	Japanese
1880	---
1890	Indian
1900	White
1910	---
	Sex:
1920	Female
1930	Male
1940	---
1950	---
	Relationship:
1960	Inmate
Specific year of birth:	
0	Son/Daughter
1	Wife
2	Head
3	---
4	---
5	---
6	---
7	---
8	---
9	---

had no effect on the count.

We estimate that there should have been successive entries of Head of household, white, born in 1890-1899, with an intervening blank panel, about 140,000 times in the sample schedules. We also estimate that spurious entries of Indian born in 1902 were recorded by the FOSDIC scanner about 3,500 times. So, the spurious entries were counted about 2.5 percent of the times that they could have occurred. (Multiplied by 4, the sampling ratio, the 3,500 recorded smudges would add 14,000 to the count of Indians).

Once the computer decided that it had found an Indian born in 1902, it ascribed many of the other characteristics, according to detailed specifications for assigning missing information. According to these specifications, persons with relationship missing were most often assigned to the category "Other Relative", and to single marital status. Occupation and industry characteristics were not assigned, however. This procedure explains the anomalous distribution of characteristics of Indians already noted and shown in Table C.

Having explained the excess of Indians at age 55 to 59 and their peculiar distribution of characteristics, we were still faced with the question of why the error was concentrated in certain areas. When one considers the nature of the error, it is reasonable to expect that it would occur with greatest frequency in areas with the largest population. The larger the total population, the larger the number of white heads of households born in 1890-99 and the greater the probability that such individuals would appear on successive schedules, creating the situation which could produce spurious Indians born in 1902. Indeed, most of the States with the largest excesses, in terms of absolute numbers, were the States with the largest population (Table F). The relative excess depended on the size of the Indian population in relation to the white population. Generally, the smaller the ratio of Indians to whites, the greater the distortion caused by the spurious additions. Thus, the estimated excess, in absolute numbers, is only about half as large in Illinois as in California, but the relative error in Illinois is three times as great as in California, because Illinois has a much smaller proportion of Indians.

TABLE F.--SELECTED STATISTICS RELATING TO THE EXCESS OF INDIANS 55 TO 59 YEARS OLD IN THE SAMPLE, BY STATE
(States with 2,500 or more Indians)

State	Total Population	Ratio of Indian to White Population	Estimated Excess of Indians 55 to 59 Years Old 1/	
			Number	Percent
States				
Alaska	226,167	.083	-70	-17.5
Arizona	1,302,161	.071	435	24.8
California	15,717,204	.003	1,340	100.8
Colorado	1,753,947	.003	157	155.5
Florida	4,951,560	.001	653	315.5
Idaho	667,191	.008	39	25.2
Illinois	10,081,158	.001	744	322.1
Kansas	2,178,611	.002	314	251.2
Louisiana	3,257,022	.002	105	69.5
Michigan	7,823,194	.001	212	49.2
Minnesota	3,413,864	.005	441	95.9
Mississippi	2,178,141	.003	51	53.7
Montana	674,767	.033	152	29.3
Nebraska	1,411,330	.004	124	59.6
Nevada	285,278	.025	28	13.7
New Mexico	951,023	.064	390	32.8
New York	16,782,304	.001	2,225	232.5
North Carolina	4,556,155	.011	-8	-0.9
North Dakota	632,446	.019	-12	-3.7
Oklahoma	2,328,284	.031	474	19.4
Oregon	1,768,687	.005	239	101.7
South Dakota	680,514	.040	100	14.5
Texas	9,579,677	.001	428	203.8
Utah	890,627	.008	111	116.8
Washington	2,853,214	.008	509	87.2
Wisconsin	3,951,777	.004	307	68.4
Wyoming	330,066	.012	64	63.4

1/ The expected number was estimated by linear interpolation.

Tabulations of Indian statistics from the 25-percent sample were made only for States with 2,500 or more Indians. If sample statistics were available for States such as Pennsylvania and Ohio, with very large white populations but fewer than 2,500 Indians, the distortions in their data for Indians would, undoubtedly, be found to be very extreme.

The concentration of the error in urban areas has a similar explanation. Urban areas contained about 70 percent of the population in 1960 and, presumably, should have received about that proportion of the error in the Indian statistics. Moreover, since the proportion of the population which is Indian is only 0.1 percent in the urban areas but is 0.7 percent in rural areas the rate of error should be much greater in urban areas.

Although the distortions in the Indian data had been explained, the investigation was not complete until the extent of other errors due to smudging could be determined. The alignment of circles shown in Table E shows that a white head of household, if born in the decade of the 1900's would produce a smudge adding to the white population born in 1902 and, if born in the 1870's, could add to the Japanese population born in 1902. However, in the first case, the actual number of white persons born in 1902 is so large (about 1.5 million) that the estimated additions to this group as a result of smudging (about 23,000) would have a minor effect. In the second case, the number of whites born in the 1870's who were household heads in 1960 was relatively small and could have generated only a small number of smudges, if the rate of occurrence was the same as for the smudges that produced the Indians. Other potential sources of error due to smudging were considered, but none were found to meet the conditions necessary to produce noticeable distortions in the data.

Although this kind of systematic smudging has distorted the characteristics of the population, it has not added to the total count. The sample was adjusted to the complete count by a ratio estimation procedure. The Indian picked up by smudging had an equal chance with all other nonwhites in the same age-sex group of being retained in the sample count for this area when it was adjusted to the complete count.

The data on Indians from the complete count in 1960 must certainly be more accurate than from the sample count, for they were not affected by the smudging error. This is not to say that the complete count is entirely accurate. There are particular problems in enumerating Indians that would lead one to expect that even without processing errors the Indian data might be less accurate than the data for most other ethnic groups in the population.

Errors are inevitable in a project as involved as taking the census. The reduction of the number of errors depends in part on the ingenuity of the staff of the Bureau of the Census but also in part on the cost. Some types of errors could be virtually eliminated but the cost would be prohibitive. Each expenditure to reduce the likelihood of error must be balanced against the gain in accuracy to be expected.

Footnotes

1/ Coale, Ansley J., and Stephan, Frederick F., "The Case of the Indians and the Teen-Age Widows", Journal of the American Statistical Association, 57 (1962), 338-47.

Source Notes for Tables

Table A.--U.S. Bureau of the Census., U.S. Census of Population: 1960, Subject Reports. Nonwhite Population by Race. Final Report, PC(2)-1C.

Table B.--U.S. Bureau of the Census, U.S. Census of Population: 1960. Detailed Characteristics. United States Summary. Final Report PC(1)-1D.

Table C.--U.S. Bureau of the Census, U.S Census of Population: 1960, Subject Reports. Nonwhite Population by Race. Final Report, PC(2)-1C.

Table D.--U.S. Bureau of the Census. Unpublished Data.

Table F.-U.S. Bureau of the Census, U.S. Census of Population: 1960, General Population Characteristics, United States Summary, Final Report PC(1)-1B, and Subject Reports, Nonwhite Population by Race, Final Report PC(2)-1C.

MEASURING THE PROBABILITY OF EVENTUALLY BEARING n LIVE BIRTHS: AN EXTENSION OF FERTILITY TABLES*

Chai Bin Park, University of Hawaii

As everyone must die someday, it is meaningless to ask the probability of eventually dying after a given age. However, in the case of birth, the situation is different. Since not all women will have even one birth, it is of interest to contemplate, for instance, how large is the probability that a woman will eventually bear a birth after a given age, or what are the probabilities that a woman who has already had two births will eventually bear two, three or any additional number of births in her remaining lifetime.

In this paper, a methodology of computing the probability that a woman of parity m at age x will eventually bear a total of n births during her lifetime will be discussed. The term birth here refers to live birth only.

Fertility Tables

Using the age-parity specific birth probability, it is easy to construct a fertility table for each order of birth. The *n*th order birth will be born only once, if ever, to the women of parity n-1; this is analogous to the situation of death in the construction of a life table. The moment she bears her *n*th child she "dies" from the cohort of n-1 parity.

In the construction of a set of fertility tables by order of birth, we consider a hypothetical cohort of women for each parity. For simplicity, we assume they live through their childbearing age. In a fertility table (see Table 1) we introduce the following symbols:

Column (1). x - Exact year of age of women as in a life table. In a fertility table, only childbearing age may be of concern.

Column (2). $w_{n-1,x}$ - This is the number of women of the cohort of parity n-1 at exact age x, corresponding to the l_x column in a life table. We consider a radix of 100,000 women at the beginning of the childbearing age for each parity. As the women who deliver their *n*th birth will become the women of parity n, these new mothers will leave the cohort. Thus, in each of the successive ages the cohort will be depleted by the number of births born in the previous age. (In actuality, the mothers of these births will become the new access to the parity of the next rank.)

Column (3). $f_{n,x}$ - This is the age-parity specific birth probability which is the backbone in preparing the fertility table as the probability of death in the life table. It is the probability that a woman of parity n-1 at exact age x will bear a child before reaching exact age x+1.[1] This specific measure of fertility is now available for the U. S.[2]

Column (4). $b_{n,x}$ - The number of the *n*th order births born to the women of the cohort of parity n-1 between ages x and x+1 when they are subjected to the corresponding age-parity specific birth probability given in Column (3). that is,

$$b_{n,x} = f_{n,x} \, w_{n-1,x} \, .$$

$$w_{n-1,x+1} = w_{n-1,x} - b_{n,x} \, .$$

Column (5). $B_{n,x}$ - This indicates the total number of the *n*th order births born to the women of the cohort of n-1 parity after age x. It is obtained by summing up b_n from the bottom of Column (4) successively till x, inclusive, i.e.,

$$B_{n,x} = \sum_{i=x} b_{n,i} \, .$$

Column (6). $F_{n,x}$ - This value shows the probability that a woman of parity n-1 at age x will eventually bear her *n*th child in her remaining lifetime. In short, it may be called the probability of the *n*th order birth, or the next birth, after age x to a woman of n-1 parity. The value is obtained by dividing $B_{n,x}$ by $w_{n-1,x}$.

An illustration is given in Table 1 for the first order birth based on the experience of the U. S. women in 1957. Obviously it is a period table.

There may be a few more functions in the fertility table if we wish to include the average year of sterility before the next birth, the counterpart in a life table being the life expectancy. These functions are not considered in the present paper since our interest centers in the probability of additional births after the mother reaches a certain age.

Measurement of the Probability of n Births

Thus, the probability that a woman of a given parity and a given age will eventually bear *an* additional child in her lifetime may be easily obtained. However, in actuality this is not the probability of bearing *exactly* one more

---

* This study is, in part, supported by a grant of the Agency for International Development to the University of Hawaii, Grant No. AID/csd-1439.

birth but that of bearing *at least* one more birth, because some of the women who proceed to the parity of the next rank by having an additional child will bear further births.

Our next question is this: What is the probability that a woman of a given parity and age will eventually bear several additional births in her life? We introduce a notation ${}_nF_{m,x}$ to represent the probability that a woman of parity m at age x will eventually bear a total of at least n births in her lifetime; in short, we may call it the probability of n births after age x to a woman of m parity. Therefore, our notation $F_{n,x}$ in the above fertility table is, in fact, ${}_nF_{n-1,x}$. We assume that no two births can occur in a year to a woman.

Let us first consider the simplest case of ${}_2F_{0,x}$, i.e., the probability of having at least 2 births from a woman with no previous experience of childbirth at age x. Notice that:

$$1-f_{1,x} = \Pr\{\text{a woman of 0 parity and age x will not bear a child before she reaches age x+1}\}$$

$$= w_{0,x+1} / w_{0,x}$$

Now,

$${}_2F_{0,x} = \Pr\{\text{the first birth in (x, x+1) and the second birth any time after x+1}\} \text{ or } \Pr\{\text{no birth till x+1 and the first birth in (x+1, x+2) and the second birth any time after x+2}\} \text{ or } \Pr\{\text{no birth till x+2 and the first birth in (x+2, x+3) and the second birth any time after x+3}\} \text{ or } \cdots$$

$$= f_{1,x} F_{2,x+1} + (1-f_{1,x}) f_{1,x+1} F_{2,x+2} + (1-f_{1,x})(1-f_{1,x+1}) f_{1,x+2} F_{2,x+3} + \cdots$$

$$= \frac{b_{1,x}}{w_{0,x}} F_{2,x+1} + \frac{w_{0,x+1}}{w_{0,x}} \frac{b_{1,x+1}}{w_{0,x+1}} F_{2,x+2} + \frac{w_{0,x+1}}{w_{0,x}} \frac{w_{0,x+2}}{w_{0,x+1}} \frac{b_{1,x+2}}{w_{0,x+2}} F_{2,x+3} + \cdots$$

Therefore,

$${}_2F_{0,x} = \frac{1}{w_{0,x}} \sum_{i=x} b_{1,i} F_{2,i+1}. \qquad (1)$$

For the computation of ${}_3F_{0,x}$ we need to consider all possible paths that 3 births can take place through the childbearing age after x. Noticing that we already have $F_{3,x}$, our attention centers on the different paths taken by the first two births only. For a woman with no previous childbirth till age x, the earliest possible two births will occur in the manner of the first birth occurring during ages x and x+1 and the second birth during ages x+1 and x+2; there is only one possible path for such a sequence. Therefore, the probability is:

$$f_{1,x} f_{2,x+1} F_{3,x+2} = \frac{b_{1,x}}{w_{0,x}} \frac{b_{2,x+1}}{w_{1,x+1}} F_{3,x+2}.$$

There are two possible paths if the second birth is to occur during ages x+2 and x+3: (a) the first birth during x and x+1 and no birth during x+1 and x+2 and the second birth during x+2 and x+3 or (b) no birth till x+1 and two successive births in the following two years. The corresponding probabilities are:

$$f_{1,x} (1-f_{2,x+1}) f_{2,x+2} F_{3,x+3} = \frac{b_{1,x}}{w_{0,x}} \frac{b_{2,x+2}}{w_{1,x+1}} F_{3,x+3}.$$

and

$$(1-f_{1,x}) f_{1,x+1} f_{2,x+2} F_{3,x+3} = \frac{b_{1,x+1}}{w_{0,x}} \frac{b_{2,x+2}}{w_{1,x+2}} F_{3,x+3}.$$

In a similar manner, we see there are three different paths for the second birth taking place during x+3 and x+4 and the probabilities may be written as follows:

$$f_{1,x} (1-f_{2,x+1}) (1-f_{2,x+2}) f_{2,x+3} F_{3,x+4} = \frac{b_{1,x}}{w_{0,x}} \frac{b_{2,x+3}}{w_{1,x+1}} F_{3,x+4}.$$

$$(1-f_{1,x}) f_{1,x+1} (1-f_{2,x+2}) f_{2,x+3} F_{3,x+4} = \frac{b_{1,x+1}}{w_{0,x}} \frac{b_{2,x+3}}{w_{1,x+2}} F_{3,x+4}.$$

$$(1-f_{1,x}) (1-f_{1,x+1}) f_{1,x+2} f_{2,x+3} F_{3,x+4} = \frac{b_{1,x+2}}{w_{0,x}} \frac{b_{2,x+3}}{w_{1,x+3}} F_{3,x+4}.$$

There are 4 probabilities to be multiplied by $F_{3,x+5}$, in other words, 4 different paths the second birth may occur during x+4 and x+5. They are:

$$\frac{b_{1,x}}{w_{0,x}} \frac{b_{2,x+4}}{w_{1,x+1}}, \quad \frac{b_{1,x+1}}{w_{0,x}} \frac{b_{2,x+4}}{w_{1,x+2}}, \quad \frac{b_{1,x+2}}{w_{0,x}} \frac{b_{2,x+4}}{w_{1,x+3}},$$

$$\frac{b_{1,x+3}}{w_{0,x}} \frac{b_{2,x+4}}{w_{1,x+4}}.$$

Therefore,

$$
\begin{aligned}
{}_3F_{0,x} &= \frac{1}{w_{0,x}} \Big[\frac{b_{1,x}}{w_{1,x+1}} (b_{2,x+1}\, F_{3,x+2} \\
&\quad + b_{2,x+2}\, F_{3,x+3} + b_{2,x+3}\, F_{3,x+4} + \cdots) \\
&\quad + \frac{b_{1,x+1}}{w_{1,x+2}} (b_{2,x+2}\, F_{3,x+3} + b_{2,x+3}\, F_{3,x+4} \\
&\quad + \cdots) + \cdots \Big] \\
&= \frac{1}{w_{0,x}} \Big[\frac{b_{1,x}}{w_{1,x+1}} \Big(\sum_{j=x+1} b_{2,j}\, F_{3,j+1}\Big) \\
&\quad + \frac{b_{1,x+1}}{w_{1,x+2}} \Big(\sum_{j=x+2} b_{2,j}\, F_{3,j+1}\Big) + \cdots \Big].
\end{aligned}
$$

That is,

$$
{}_3F_{0,x} = \frac{1}{w_{0,x}} \Big[\sum_{i=x} \frac{b_{1,i}}{w_{1,i+1}} \Big(\sum_{j=i+1} b_{2,j}F_{3,j+1}\Big)\Big]. \quad (2)
$$

In a similar manner, it can be shown that:

$$
{}_4F_{0,x} = \frac{1}{w_{0,x}} \Big[\sum_{i=x} \frac{b_{1,i}}{w_{1,i+1}} \Big\{\sum_{j=i+1} \frac{b_{2,j}}{w_{2,j+1}} \Big(\sum_{k=j+1} b_{3,k}\, F_{4,k+1}\Big)\Big\}\Big]. \quad (3)
$$

In general,

$$
{}_nF_{0,x} = \frac{1}{w_{0,x}} \Big[\sum_{i=x} \frac{b_{1,i}}{w_{1,i+1}} \Big\{\sum_{j=i+1} \frac{b_{2,j}}{w_{2,j+1}} \Big(\cdots \frac{b_{n-2,p}}{w_{n-2,p+1}} \Big[\sum_{q=p+1} b_{n-1,q}\, F_{n,q+1}\Big]\Big)\Big\}\Big]. \quad (4)
$$

and

$$
{}_nF_{m,x} = \frac{1}{w_{m,x}} \Big[\sum_{i=x} \frac{b_{m+1,i}}{w_{m+1,i+1}} \Big\{\sum_{j=i+1} \frac{b_{m+2,j}}{w_{m+2,j+1}} \cdots \Big(\sum_{q=p+1} b_{n-1,q}\, F_{n,q+1}\Big)\Big\}\Big]. \quad (5)
$$

Obviously, the probability that a woman of parity m at age x will bear *exactly* a total of n children in her lifetime will be given by ${}_nF_{m,x} - {}_{n+1}F_{m,x}$. If we denote the probability of at least n births at the beginning of the childbearing age by F_n, the theoretical distribution of the women of completed fertility may be easily computed. The proportion of the childless women will be given by $1-F_1$ and the proportion of n births by $F_n - F_{n+1}$.

Application

Material to illustrate the procedure above are provided by a publication of P. K. Whelpton and A. A. Campbell.[3] The age-parity specific birth probability was calculated for the calendar year of 1957 when the U. S. women showed the highest fertility rate in recent years. In their publication, P. K. Whelpton and A. A. Campbell presented the central and the cumulative birth rates by order of birth and by age for each birth cohort of women. The cumulative rate enables one to derive age-parity specific birth probabilities.

Let $Q_{i,x}(y)$ be the cumulative birth rate of the *i*th order birth by a woman reaching age x on January 1 of the calendar year y. Then the age-parity specific birth probability for age x, parity i-1, and calendar year y, $f_{i,x}(y)$, may be obtained as follows:

$$
f_{i,x}(y) = \frac{Q_{i,x+1}(y+1) - Q_{i,x}(v)}{Q_{i-1,x}(y) - Q_{i,x}(y)}
$$

In particular,

$$
f_{1,x}(y) = \frac{Q_{1,x+1}(y+1) - Q_{1,x}(y)}{1-Q_{1,x}(y)}
$$

This appears to be consistent with the method used by the National Center for Health Statistics in computing the age-parity specific birth probability.[4]

With these age-parity specific birth probabilities thus obtained and certain fertility functions shown in Table 1, ${}_nF_{m,x}$ may be computed. As n becomes larger in comparison with m, the procedure becomes progressively tedious. As an example, the computational procedure of ${}_3F_{0,x}$ is illustrated in Table 2 and the explanation for each column of the Table is given below.

Column (1). x - Exact year of age of women as in Table 1.

Column (2). $b_{2,x+1}$ - The number of the second order births born to women of the parity-one-cohort between ages x+1 and x+2. The entry of this column is transcribed from Column (4) of an appropriate fertility table for the second order birth such as Table 1 shifting a row upward.

Column (3). $F_{3,x+2}$ - The probability of the third order birth after age x+2 for the woman of parity 2. It is transcribed from Column (6) of an appropriate fertility table for the third order birth shifting two rows upward.

Column (4). $b_{2,x+1}\ F_{3,x+2}$ - The product of the preceding two columns for the row of x. (This shows the number of the third order births eventually born after age x+2 if there are $b_{2,x+1}$ women of parity 2 at age x+2 and they are exposed to the given series of $f_{3,x}$ afterwards.)

Column (5). $\Sigma\ b_2\ F_3$ - The value obtained by adding the values in Column (4) from the bottom till x, inclusive. (It gives the total number of the third order births eventually born to all the women joining parity 2 from the parity-1-cohort in successive ages x+2 on. As evident, the number of women joining to parity 2 at age x is given by $b_{2,x-1}$.)

Column (6). $b_{1,x}$ - The number of the first order births born to the women of the parity-0-cohort between ages x and x+1. The entry of this column is transcribed from Column (4) of an appropriate fertility table for the first order birth such as Table 1.

Column (7). $w_{1,x+1}$ - The number of women of the parity-1-cohort at age x+1. It is transcribed from an appropriate fertility table shifting a row upward.

Column (8). $b_{1,x}\ /\ w_{1,x+1}$ - The quotient of the two preceding columns for the row of x. (It is the ratio of the number of women who would have joined to parity 1 from the parity-0-cohort at age x+1 to the number of women remaining in the parity-1-cohort at that age.)

Column (9). $\frac{b_1}{w_1}\ \Sigma\ b_2\ F_3$ - The product of Columns (5) and (8). (It represents the number of the third order births which would have eventually been born to $b_{1,x}$ women joining to parity 1 from the parity-0-cohort at age x+1.)

Column (10). $\Sigma\ \frac{b_1}{w_1}\ \Sigma\ b_2\ F_3$ - The sum of the values of Column (9) from the bottom till x. (The value shows the total number of the third order births eventually born to $w_{0,x}$ women of the parity-0-cohort at age x.)

Column (11). $w_{0,x}$ - The number of the women of the parity-0-cohort at age x, transcribed from Column (2) of an appropriate fertility table for the first order birth.

Column (12). ${}_3F_{0,x}$ - The quotient of the preceding two columns for the row of x. It is the probability that at least three births will eventually be born to a woman of 0 parity at age x when she is exposed to the given age-parity specific birth probabilities.

The attached figures show some of the curves of ${}_nF_{m,x}$ computed from the age-parity specific birth probabilities of the U. S. women experienced in 1957.

Summary

This paper is intended to present a method of measuring the probability that a woman of a given parity at age x will eventually have at least n additional births in her remaining lifetime.

A hypothetical cohort of women of equal radix for each parity is considered and the number of women of n parity at age x is denoted by $w_{n,x}$. We assume no mortality.

Define:

$f_{n,x}$ = Pr{a woman of n-1 parity at age x will bear her *n*th live birth between ages x and x+1}.

$b_{n,x}$ = the number of the *n*th order children born to the women between ages x and x+1

$$= f_{n,x}\ w_{n-1,x}.$$

Clearly,

$$w_{n-1,x+1} = w_{n-1,x} - b_{n,x}.$$

and

$(1-f_{n,x})$ = Pr{a woman of n-1 parity at age x will be sterile until age x+1}

$$= w_{n-1,x+1}\ /\ w_{n-1,x}.$$

It can easily be seen that the probability $F_{n,x}$ that a woman of n-1 parity at age x will eventually bear the next birth in her life is as follows:

$$F_{n,x} = \sum_{i=x} b_{n,i}\ /\ w_{n-1,x}.$$

The above functions, i.e., $w_{n,x}$, $f_{n,x}$, $b_{n,x}$ and $F_{n,x}$, may be arrayed in the form of a life table.

To find the probability ${}_nF_{m,x}$, that a woman of m parity at age x will eventually bear a total of at least n live births, let us first consider:

$$\begin{aligned} {}_2F_{0,x} &= f_{1,x}\ F_{2,x+1} + (1-f_{1,x})\ f_{1,x+1}\ F_{2,x+2} \\ &\quad + (1-f_{1,x})\ (1-f_{1,x+1})\ f_{1,x+2}\ F_{2,x+3} + \cdots \\ &= \frac{1}{w_{0,x}} \sum_{i=x} b_{1,i}\ F_{2,i+1}. \end{aligned}$$

Similarly,

$${}_3F_{0,x} = \frac{1}{w_{0,x}} \sum_{i=x} \frac{b_{1,i}}{w_{1,i+1}} \left(\sum_{j=i+1} b_{2,j}\ F_{3,j+1}\right).$$

and

$$ {}_nF_{0,x} = \frac{1}{w_{0,x}} \sum_{i=x} \frac{b_{1,i}}{w_{1,i+1}} \left[\sum_{j=i+1} \frac{b_{2,j}}{w_{2,j+1}} \left\{ \Sigma \cdots \left(\sum_{q=p+1} b_{n-1,q}\, F_{n,q+1} \right) \right\} \right]. $$

In general,

$$ {}_nF_{m,x} = \frac{1}{w_{m,x}} \sum_{i=x} \frac{b_{m+1,i}}{w_{m+1,i+1}} \left[\sum_{j=i+1} \frac{b_{m+2,j}}{w_{m+2,j+1}} \left\{ \sum_{k=j+1} \cdots \left(\sum_{q=p+1} b_{n-1,q}\, F_{n,q+1} \right) \right\} \right]. $$

To illustrate, some ${}_nF_{m,x}$ curves are given based on the U. S. women experience in 1957, when the fertility rate was the highest in the post-war period.

References

1 National Center for Health Statistics, *Fertility Measurement: A Report of the United States National Committee on Vital and Health Statistics.* U. S. Public Health Service, Publication No. 1000, Series 4, No. 1, Washington, D. C., Nov. 1965.

2 See *Vital Statistics of the United States*, Volume 1, U. S. Public Health Service.

3 Whelpton, P. K. and Campbell, A. A. "Fertility tables for birth cohorts of American Women, part 1" *Vital Statistics--Special Reports*, Vol. 51, No. 1, Public Health Service, Washington, D. C., Jan. 1960.

4 Campbell, A. A. Personal Communication, March 21, 1967.

TABLE 1

FERTILITY TABLE FOR THE FIRST ORDER BIRTH, U. S. WOMEN, 1957

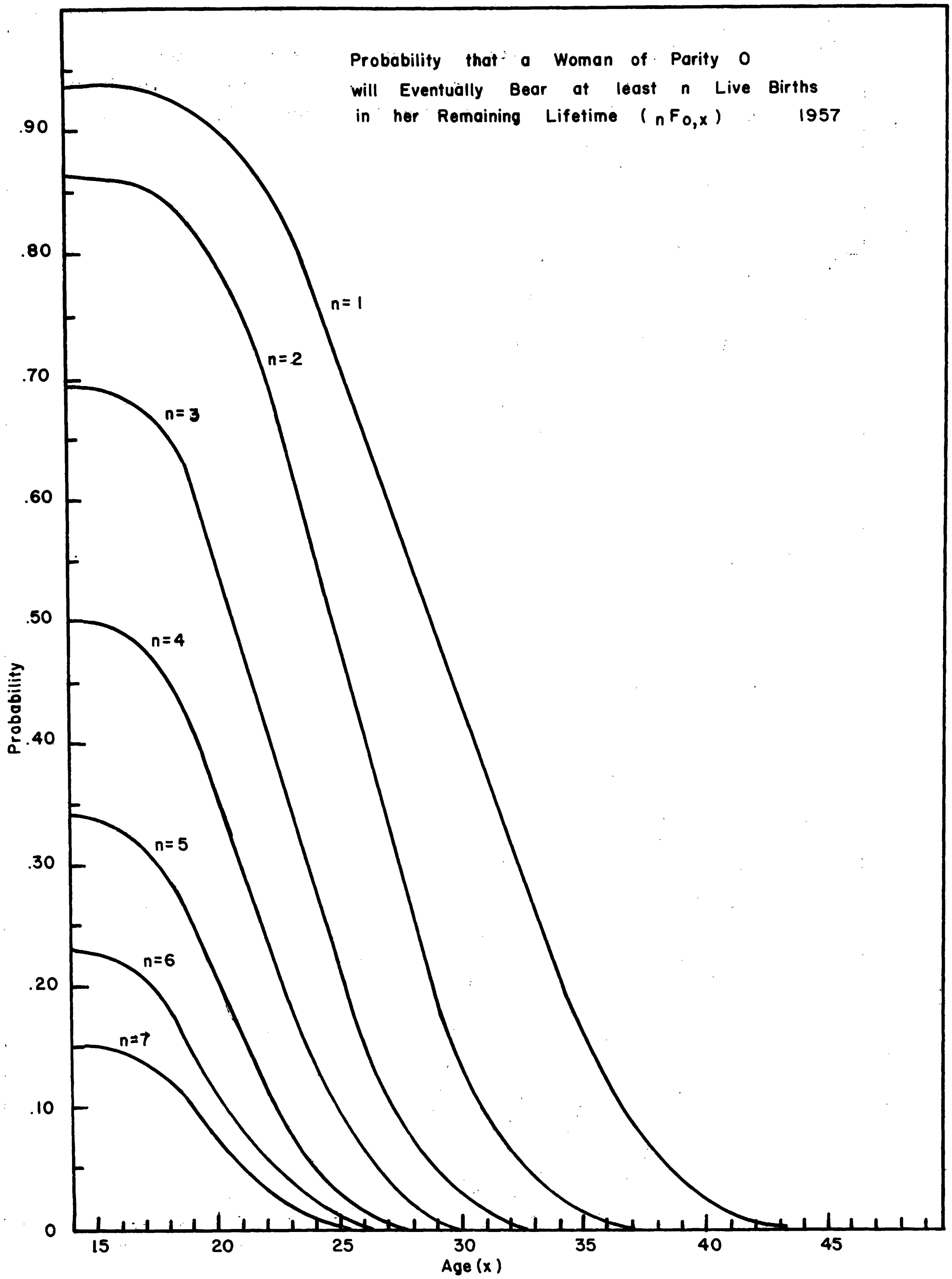
Age	No. of women of 0 parity at age x	Birth probability of the 1st order birth in age x	No. of the 1st order births born	Total no. of 1st order births born after age x	Probability of the 1st order birth after age x
x	$w_{0,x}$	$f_{1,x}$	$b_{1,x}$	$B_{1,x}$	$F_{1,x}$
(1)	(2)	(3)	(4)	(5)	(6)
14	100,000	.0048	480	93,787	.9379
15	99,520	.0141	1,403	93,307	.9376
16	98,117	.0390	3,826	91,904	.9367
17	94,291	.0728	6,864	88,078	.9341
18	87,427	.1116	9,757	81,214	.9289
19	77,670	.1575	12,233	71,457	.9200
20	65,437	.1782	11,661	59,224	.9050
21	53,776	.1951	10,492	47,563	.8845
22	43,284	.2059	8,912	37,071	.8565
23	34,372	.2077	7,139	28,159	.8192
24	27,233	.1971	5,363	21,020	.7719
25	21,865	.1730	3,783	15,652	.7158
26	18,082	.1479	2,674	11,869	.6564
27	15,408	.1361	2,097	9,195	.5968
28	13,311	.1172	1,560	7,098	.5332
29	11,751	.1030	1,210	5,538	.4713
30	10,541	.0853	899	4,328	.4106
31	9,642	.0738	712	3,429	.3556
32	8,930	.0682	609	2,717	.3043
33	8,321	.0577	480	2,108	.2533
34	7,841	.0494	387	1,628	.2076
35	7,454	.0463	345	1,241	.1665
36	7,109	.0376	267	896	.1260
37	6,842	.0342	234	629	.0919
38	6,608	.0224	148	395	.0598
39	6,460	.0132	85	247	.0382
40	6,375	.0097	62	162	.0254
41	6,313	.0066	42	100	.0158
42	6,271	.0039	24	58	.0092
43	6,247	.0025	16	34	.0054
44	6,231	.0015	9	18	.0029
45	6,222	.0008	5	9	.0014
46	6,217	.0003	2	4	.0006
47	6,215	.0002	1	2	.0003
48	6,214	.0001	1	1	.0002
49	6,213	.0001	0	0	.0001
50	6,213	.0000	0	0	.0000

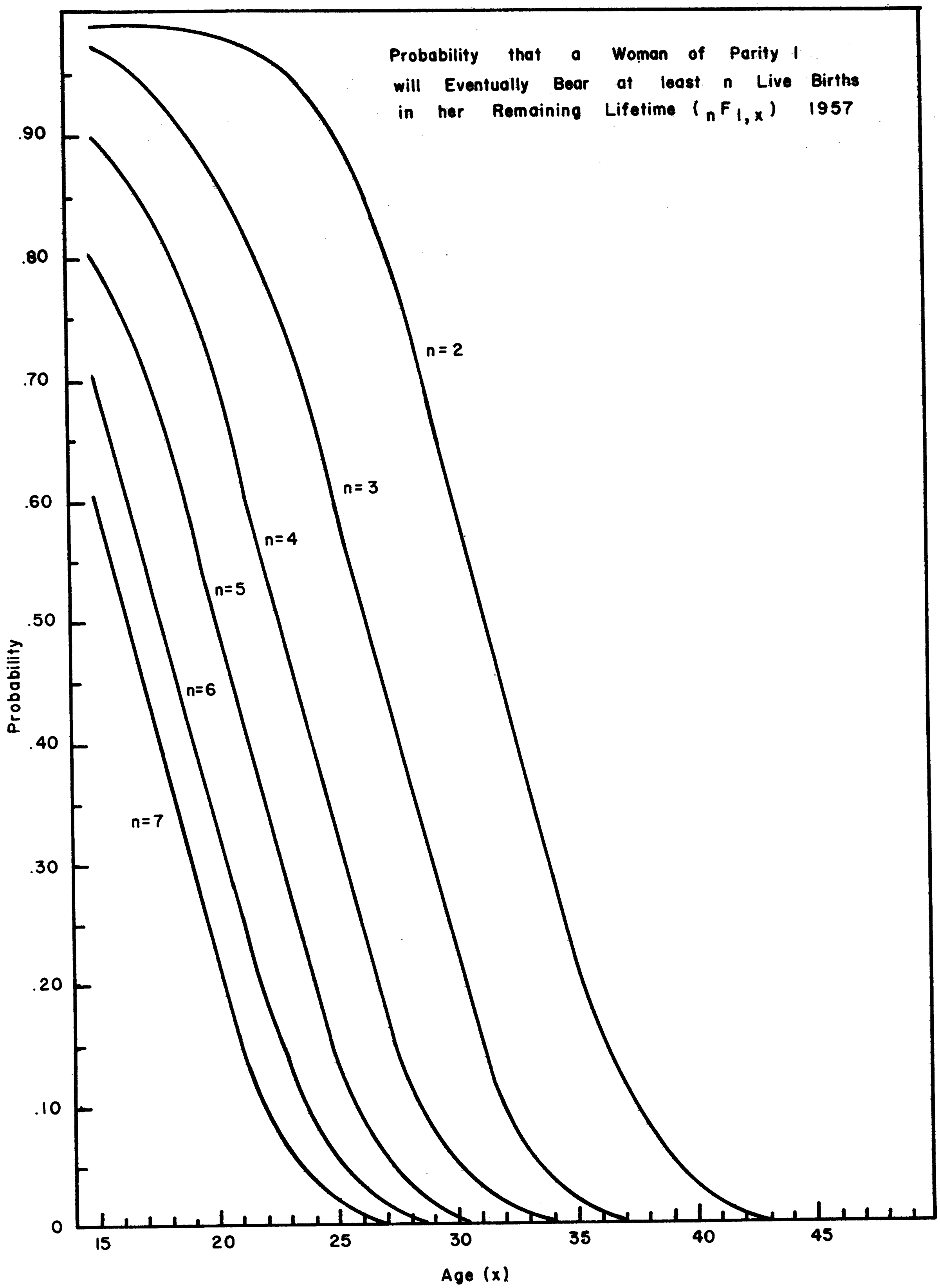
TABLE 2

CALCULATION OF $_3F_{0,x}$ — U. S. WOMEN, 1957

$$_3F_{0,x} = \frac{1}{w_{0,x}} \left[\sum_{i=x} \frac{b_{1,i}}{w_{1,i+1}} \left(\sum_{j=i+1} b_{2,j} F_{3,j+1} \right) \right]$$

X	$b_{2,x+1}$	$F_{3,x+2}$	$b_{2,x+1}F_{3,x+2}$	$\Sigma\ b_2F_3$	$b_{1,x}$	$w_{1,x+1}$	$\frac{b_{1,x}}{w_{1,x+1}}$	$\frac{b_1}{w_1}\ \Sigma\ b_2F_3$	$\Sigma\ \frac{b_1}{w_1}\ \Sigma\ b_2F_3$	$w_{0,x}$	$_3F_{0,x}$
(1)	(2)	(3)	(4)=(2)(3)	(5)=Σ(4)	(6)	(7)	(8)=(6)/(7)	(9)=(5)(8)	(10)=Σ(9)	(11)	(12)=(10)/(11)
14	26,000	.9950	25871.0400	97313.2190	480	100,000	.0048	467.1034	69298.8634	100,000	.6930
15	21,445	.9919	21272.1533	71442.1990	1403	74,000	.0190	1354.5437	68831.7600	99,520	.6916
16	17,380	.9875	17162.7500	50170.0257	3826	52,555	.0728	3652.3779	67477.2163	98,117	.6877
.											
.											
.											
.											
.											
.											
.											
.											
38	5	.0474	.2371	.3715	148	180	.8222	.3054	.3917	6,608	.0001
39	3	.0293	.0880	.1344	85	175	.4857	.0653	.0863	6,460	.0000
40	2	.0176	.0317	.0464	62	172	.3605	.0167	.0210	6,375	.0000
41	1	.0098	.0108	.0146	42	170	.2471	.0036	.0043	6,313	.0000
42	1	.0048	.0029	.0039	24	169	.1420	.0006	.0007	6,271	.0000
43	0	.0024	.0007	.0010	16	168	.0952	.0001	.0001	6,247	.0000
.											
.											
.											
.											
.											
.											


Probability that a Woman of Parity 0
will Eventually Bear at least n Live Births
in her Remaining Lifetime ($_nF_{0,x}$) 1957
n= 1
n=2
n=3
n=4
n=5
n=6
n=7
Probability
.90
.80
.70
.60
.50
.40
.30
.20
.10
0
15
20
25
30
35
40
45
Age (x)


Probability that a Woman of Parity I will Eventually Bear at least n Live Births in her Remaining Lifetime ($_nF_{1,x}$) 1957
n = 2
n = 3
n = 4
n = 5
n = 6
n = 7
Probability
.90
.80
.70
.60
.50
.40
.30
.20
.10
0
15
20
25
30
35
40
45
Age (x)

UTILIZING AVAILABLE SOCIAL STATISTICS TO DELINEATE COMMUNITY ACTION PROGRAM SITES

Mary G. Powers, Fordham University

Introduction

Persons involved in social action programs in our large cities frequently turn to the urban sociologist for assistance in the preparation of studies or to review the results of such studies. This paper grew out of one such cooperative effort which raised the perennially thorny issues of just how one delineates a neighborhood or sub-community within a large city.

The Juvenile Court Community Development Project in New York City is a 2-year demonstration project designed to test the significance of an area-focused, community development strategy for programs in juvenile corrections.[1] The program's aim is to demonstrate how a community-oriented diagnostic process and program might be utilized by the juvenile court. This means focusing on geographic areas where delinquents are concentrated instead of on offenders as individuals.

Once the decision has been made to launch area-focused programs, the problem of selecting sites and of describing such areas in terms relevant to the programs becomes a critical one. The present project is now located in the East Tremont section of the Bronx as shown on Map I. Part of the process of selecting and delimiting this area and of deciding that it is, in fact, a recognizable "sub-community" and/or "neighborhood," will be described in the rest of this paper.

Because of the nature of the specific community action program involved, the delineation and description of the project site or target area involved four separate steps.[2] The first task was to determine the geographic distribution of delinquency and PINS[3] cases from the Bronx coming to the attention of the Family Court.

The second task was to describe, in terms of relevant social and demographic characteristics, the different sub-areas of the Bronx in which the

---

[1]This project (Grant # 66015, OJDYD-HEW) is directed by John M. Martin, Institute for Social Research, Fordham University. Results of the project will appear in two forthcoming books from Random House by John M. Martin, Joseph P. Fitzpatrick, Robert E. Gould, M.D. and Associates, The Analysis of Delinquent Behavior: A Structural Approach and Case Studies in The Analysis of Delinquent Behavior.

[2]A complete description of all four steps is found in John M. Martin, Mary G. Powers, Selma H. Stevens and Others, Area Selection for a Correctional Community Development Program and Area Selection for a Correctional Community Development Program, Supplement # 1. and Madeline H. Engel, Robert E. Gould, M.D., John M. Martin and Others, Illustrative Case History: Henry Robinson. Juvenile Court Community Development Project, Fordham University, 1966 and 1967.

[3]In 1962, the Family Court of New York State underwent marked change. One of the changes was the creation of a new designation, PINS (Persons in Need of Supervision), for certain types of youth who were previously handled as delinquents. Paraphrased, the new definitions are as follows:

(a) Juvenile delinquent means a person over seven and less than sixteen years of age who commits any act which, if done by an adult, would constitute a crime;
(b) PINS means a male less than sixteen years of age and a female less than eighteen years of age who is an habitual truant or who is incorrigible, ungovernable or habitually disobedient and beyond the lawful control of parent or other lawful authority.

juvenile cases identified were found to be concentrated. One the basis of the data yielded by these two procedures, likely communities or neighborhoods were identified and quantitatively described. Census tract data were used initially and brought up to date to some extent with other available data, such as local health and school statistics, which show population changes between 1960 and 1966.

In addition to these data, the community was described from two other less quantitative perspectives. The third task was to obtain a community profile of the area in which it seemed that the project would be located. This was done by direct observation of the area and through interviews with key persons in a wide variety of institutional structures in the area. A final task was to provide a description of what Professor Sweetser once called the "personal neighborhood"[4] from the point of view of delinquents themselves. This was done through intensive sociogenic case studies of individual delinquents from the area who are known to the Family Court.

This paper will focus mainly on the first two steps which involve problems of comparing, combining, and integrating statistical data from numerous and disparate sources.

Location of Juvenile Cases Referred for Probation Investigation

The geographic distribution of the delinquent population served by the Bronx Office of Probation was determined by plotting on tract maps all delinquent and PINS cases referred to that office for investigation from January 1, 1965 to December 31, 1965.

[4]Frank L. Sweetser, Jr., Neighborhood Acquaintance and Association, A Study of Personal Neighborhoods. New York: Columbia University Press, 1951.

The data for this plot were secured from the record books of the Bronx Office of Probation serving the Family Court of New York City, in that borough. To provide some measure of change in these patterns, it was decided that the same plot also would be constructed for all cases investigated from January 1, 1963 to December 31, 1963. The year 1963 was selected because: 1) it was the first full year of operation following the new Family Court Act of New York State, passed in 1962; and 2) it was close to 1960, the year of the last U.S. Census which was used in the demographic analysis.

The areas in which delinquency and PINS cases were highly concentrated in 1965 and 1963 are shown on Map II. Six areas containing contiguous high referral census tracts were identified. These were delineated as possible alternative target areas for basing the present project.

The six areas marked off on the map contain 22 census tracts. This represents 5.9% of the total (374) number of tracts in the Bronx. These 22 tracts contained a total of 605 cases in 1965, or 44.5% of the total number of delinquent and PINS cases (1361) referred to the Bronx Office of Probation for investigation in 1965. In other words, approximately 6% of the Bronx census tracts contained approximately 45% of the juveniles referred to the Office of Probation on delinquent and PINS petitions in 1965. Only one of the six areas experienced less than a 50% increase in cases between 1963 and 1965 -- Area 6. Area 5 experienced the highest rate of increase -- over 200%. This high increase in delinquency suggested that Area 5 was probably experiencing rapid change and that it should be examined more closely as a possible location for the project.

Selected Social and Economic Characteristics of The Bronx and of Six Potential Neighborhoods

As the Bronx census tracts including the largest number of delinquency and PINS cases in 1963 and 1965 were being delineated, the 374 borough tracts were also being analyzed in terms of their social and economic characteristics.

The variables used to characterize the areas were taken from several sources, but the first analysis was based mainly on the census tract reports and some special tabulation of tract data made by the Bureau of Labor Statistics in connection with the President's Committee on Youth Unemployment.[5] The variables were selected on two basis: 1) from empirical evidence (largely the work of Calvin Schmid and associates)[6] that they delineated distinct social areas; 2) variables of special relevance to delinquency research as suggested by reported empirical studies.[7]

Two of the best known typologies devised to provide analytic frameworks to study the social structure of the American city are those constructed by Tryon, and by Shevky and his collaborators.[8] Both have been used in the analysis of the ecological distribution of crime. Both have also been criticized for lack of a theoretical basis. In exploring the utility of the indices in research on crime, Schmid developed a similar set of indices based on the logic of modern statistical techniques.[9] He also found that all his indices, as well as those of Tryon and of Shevky and associates, were highly correlated with a few individual census tract variables. The six variables so described by Schmid were utilized in the present analysis of Bronx tracts as well as eight others of special interest to the project. For example, the per cent foreign stock and the per cent Puerto Rican were included as measures of ethnic status along with per cent Negro, because they are significant components of the Bronx population. Thus, the basic social and economic variables may be described as follows:

A. Indices of Family Status:

1. % of population under 16 years.
2. % married of the population 14 years and over.
3. Mean population in household.

B. Indices of Socioeconomic Status:

1. % of professional and technical workers in the male labor force.
2. Median grade school completed by persons 25+.

---

[5]Sources include: U.S. Censuses of Population and Housing: 1960 Census Tracts, Final Report PHS (1)-104, Part 1. U.S. Bureau of the Census. Washington, D.C.: U.S. Government Printing Office, 1962; Income, Education and Unemployment in Neighborhoods: N.Y.C.: The Bronx, U.S. Department of Labor, Bureau of Labor Statistics, Washington, D.C.: U.S. Government Printing Office, 1963.

[6]Calvin Schmid, "Urban Crime Areas: Part I," American Sociological Review 25: 527-542 (August, 1960); "Urban Crime Areas: Part II," American Sociological Review 25:655-678 (October, 1960).

[7]Kenneth Polk, "Juvenile Delinquency and Social Areas," Social Problems 5:214-217 (1957-1958); Karl Schuessler, "Components of Variation in City Crime Rates," Social Problems 9: 314-323 (1962); Sarah L. Boggs, "Urban Crime Patterns, " American Sociological Review 30:899-908 (December, 1965).

[8]Robert C. Tryon, Identification of Social Areas by Cluster Analysis. Berkeley (California): Univ. of California Press, 1955; Eshref Shevky and Wendell Bell, Social Areas Analysis, Stanford (California): Stanford University Press, 1945, among others.

[9]Calvin Schmid and Kiyoshi Tagushira, "Ecological and Demographic Indices, A Methodological Analysis," Demography 1:194-211 (1964).

C. Indices of Ethnic Status:

1. % Negro of total population.
2. % foreign stock of total population.
3. % Puerto Rican of total population.

D. Measures of Population Structure:

1. Sex ratio.
2. Dependency ratio.

Because of the nature of the study, the tracts were also described in terms of:

1. The proportion of the 14-17 year old age group enrolled in school.
2. The proportion of unemployed males.
3. The proportion of movers in the population.
4. The proportion of the population who moved in from outside the Standard Metropolitan Statistical Area (SMSA).

Values for the 14 variables were computed for each of the tracts in the Bronx with some population in 1960, and these values were converted into quartiles and mapped on tract maps of the Bronx. The clusters of areas which fell into the relevant extreme quartile with respect to any of the variables could be identified, and those with broadly similar social profiles were delimited.[10]

[10]Though a more refined analysis might be made using factor analysis to describe more precisely the relationship between delinquency rates and the various social and economic indices, this was not done since the primary interest and need of the project was to compare, in terms of broad social profiles, those areas with high delinquency rates with all others. The maps used were prepared by Joseph F. Scheuer and Terrence R. McGovern.

Table I describes the six potential target areas and the Bronx as a whole in 1960. Of the six areas selected on the basis of heavy concentration of delinquency and PINS cases, Area 5 seemed to offer the best potential in terms of a large and increasing number of delinquents. In addition, according to census data for 1960, it was not atypical of the Bronx in terms of the demographic and socioeconomic characteristics of the inhabitants, especially with respect to race and ethnicity. Each of the other areas included proportionately heavier Puerto Rican and Negro populations than the Bronx as a whole. Geographically, in 1960, Area 5 was a kind of "border" or transition area in terms of population composition.

Changes in Population Composition in The Bronx and The Target Areas: 1960 To 1966

The first step in observing the target community in 1966 was a tour of the area by automobile. In this way the primary and secondary business areas were identified, the centers of larger commercial establishments and the location of smaller neighborhood stores. Initial observation of the public schools indicated intensive use. Quonset hut facilities suggested overcrowding; and late afternoon dismissals were evidence of double sessions. To the external observer, the housing appeared to be in reasonably good condition. The area is sprinkled by clusters of one and two family houses. Puerto Rican and Negro residents were very much in evidence; white residents appeared to be in the older age groups.

Close observation of various blocks in the area and subsequent analysis of data obtained from case studies of delinquents from the area and interviews with key personnel from institutions within the area suggested the nature of the ethnic change between 1960 and 1966. A large part of the East Tremont section had a predominantly middle-class Jewish population until after World War II. Since then, however, it has been in rapid transition with Puerto

Ricans, and increasing numbers of Negroes replacing the Jewish population, especially during the past 10 years. For example, as one informant pointed out, the East Tremont YM-YWHA (an affiliate of the Federation of Jewish Philanthropies) had moved to its present location in the target area in 1962. By 1966 it had decided to give up its location and to move elsewhere. Four or five synagogues also closed in recent years.

An attempt was made to determine the extent of this change in the Target Area by examining school and health statistics and other data collected since 1960 by social agencies, city government, business and so forth. Because the available statistics did not describe the target area precisely, it was necessary to look at changes in the Bronx as a whole and in whatever smaller sub-areas for which data were available, and from this make some inferences as to changes in the Target Area.

Two sources provided a more recent picture of population change for New York City as a whole and for the boroughs separately. These are the Population Health Survey of 1964 conducted by the New York City Department of Health, and the annual estimates of the population of New York City and Westchester County prepared by Consolidated Edison of New York, Inc. The city-wide Population Health Survey provided an estimate of the non-institutional population in New York City and the individual boroughs in 1964. The estimate of 7,558,500 from the 1964 survey is about 2% less than the 7,706,300 shown in the 1960 census. According to the Survey estimates, Manhattan and the Bronx lost population, Queens and Richmond gained, and Brooklyn remained relatively stable.[11] The survey results for the Bronx as a whole indicate a drop of 7.4% in the non-institutional population resulting from a loss of 18% of the white population other than Puerto Rican, and a gain of 21 and 32 per cents respectively in the non-white and Puerto Rican populations.

The Health Survey findings are not directly comparable with the annual estimates of the Population of New York City prepared by Consolidated Edison Co. because the latter focuses on total population changes, compared to the non-institutional population in the Health Survey. Con-Edison estimates show a population loss for New York City between 1950 and 1960 -- particularly of the white component of the population. Between 1960 and 1962, however, according to their estimates, the population remained practically constant, "...the composite effect of a continuing but slower decline in Manhattan and Brooklyn, a slight gain in the Bronx and continuing increase in Queens and Richmond."[12] By 1963, according to these estimates the downward population trend had ended and an upward trend, to which all the boroughs contributed, resumed.

"By the end of 1965 continuation of the upward trend that had been established in 1963 had resulted in raising the population of every borough in the city above the 1960 census figure."[13] The increase was felt to result from a combination of several short-run factors such as the World's Fair and the spurt of building resulting from attempts to be covered under the old code which was less restrictive than the present one. A slow downtrend is expected to resume by the end of the decade. For the Bronx, the estimated gain has been from a total population of 1,425,000 in 1960 to 1,430,000 in 1963 and 1964 to 1,460,000 in 1966. The 1966

---

[11] N.Y.C. Dept. of Health, Population Health Survey, Report No. P-1 "Population Characteristics, 1964," N.Y.C., April, 1966.

[12] Consolidated Edison Company of New York, Inc., Population of New York City and Westchester County, January 1, 1966. (N.Y. System Engineering Dept., September, 1966), p. 1.

[13] Ibid.

figure represents a 2.5% increase over the 1960 census population. The Bronx includes three "meter districts," or smaller areas for which data are also supplied. Target Area 5 and the surrounding contiguous tracts which make up most of the East Tremont area, are located in the northern-most part of District 10 and the southern part of District 11, both of which show a very low rate of growth compared to District 12, the area east of the Bronx River, -- 0.2%, 1.0% and 7.0% respectively. Even with the overall population growth, the estimated population per occupied dwelling unit was lower in 1966 than in 1960 in all three Bronx Meter Districts as seen in Table II. This suggests either less crowding of families or the inmigration of single persons and couples with few children compared to those who left. An examination of school statistics does not suggest fewer children in the area, but more -- a fact to be discussed shortly. Most likely, there was less crowding because of the restrictions of public housing projects which opened since 1960.

In summary, it appears that the Bronx as a whole experienced either a slight upturn in population trends during the sixties or at least a diminution in the rate of population decline. We turn now to look at the area in which the project is located, and the surrounding tracts. The larger area corresponds to what has been historically identified as the East Tremont Section of the Bronx; the project is located in a part of this area.

Target Area 5 and East Tremont

Data on population characteristics for the intercensal period and for areas smaller than the borough are available from two sources -- and then only indirectly: the Board of Education annual ethnic survey describing the school population, and the Department of Health Annual Vital Statistics report.[14] Both sources provide information only on ethnic characteristics of the population. Both sources have been used here to get some idea of the extent of population change in the target area.

The school statistics are directly relevant since the potential project clientele are school age children. The Board of Education estimate of the color and ethnic composition of the school population provide some measure of change in numbers and ethnic composition over time, since most pupils attend schools in the district in which they reside -- especially elementary schools. Although school district and census tract boundary lines do not generally coincide, an examination of the composition of the school population in schools in and around the target area should also provide an indication of overall population change in the area. Eight such schools were identified and the proportion of pupils listed as "Puerto Rican," "Negro," or "other" was examined each year from 1957 (the first year for which data are available) through 1965. The eight schools are:

P.S. 6	P.S. 67
P.S. 57	P.S. 92
P.S. 58	J.H. 44
P.S. 59	J.H. 118

They are all located in that part of the East Tremont Section of the Bronx immediately surrounding the project target area.

Between 1957 and 1965 the number of pupils enrolled in all of these schools increased -- substantially in

[14] A long list of persons representing public and private agencies and groups concerned with current population data for New York City was contacted and each indicated that no new data was obtained in the inter-censal period except for school and health statistics.

several cases. The total enrollment figures are shown in Table III. The increase in enrollment could result from any of several factors: the movement into the area of families with large numbers of children, or of families including a larger proportion who send their children to public rather than parochial schools.

Several of the schools did experience a decline between 1957 and 1962 or 1963, but have increased since then, which suggests that the Con-Edison estimates of the Bronx population showing a reversal of the downward population trend beginning around 1963 may be a more accurate indicator of what has happened than the Public Health Survey figures for the non-institutional population.

In all of the schools in the area the percentage of pupils who were Negro or Puerto Rican increased whereas the proportion of "others" declined. With the exception of P.S. 58, which had 51.6% Puerto Rican pupils in 1957, all of the schools showed a decline in the "other" population from between 2/3 and 4/5 of the total in 1957 to less than 1/5 in 1965, with the exception of P.S. 57 which still included 27% "other" in 1965. The exact figures are shown on Table IV. All of this suggests a tendency for younger white families with school age children to 1) move out of the area and be replaced by Puerto Rican and Negro families or 2) not to move into the spaces vacated by older persons dying or moving to smaller quarters.

The school population, which is the age group of primary interest to programs concerned with delinquency prevention, has shifted from a predominance of "other" pupils to a predominance of Puerto Rican and Negro pupils. The increase in numbers in the schools also suggest considerable overcrowding in the public schools.[15]

[15] Project interviews with key institutional personnel show that the reverse situation exists in the parochial schools. Their numbers have declined since 1960.

The Department of Health of the City of New York issues an annual report of vital statistics for New York City as a whole, for the five boroughs, and for health center districts and health areas within each borough. The health areas are the smallest units for which data are available and these usually include several census tracts. All of the statistics are reported for the white and non-white components of the population. The birth statistics have been further broken down to show the number of births to Puerto Rican mothers. Because the population change in the East Tremont area seems to include a dramatic increase in the size of the Puerto Rican population, this further breakdown was extremely useful.

The number of births to non-white and Puerto Rican mothers was examined and the proportion of the total they represented each year for 1950-1964 was computed. The results are shown on Tables V and VI. These figures show a continuous increase in the proportion of total births classified as either "non-white" or "to Puerto Rican mothers," except for Area 18 which shows a continuous increase in the proportion of non-white but a decline from a high point of 62% "to Puerto Rican Mother" in 1960 to 49% in 1964. This may be an area where a Negro population is currently displacing the Puerto Rican population.

A certain amount of this change is probably due simply to a difference in the age structure of the population. That is, the Puerto Rican and Negro population may be and probably is considerably younger than the white non-Puerto Rican population. However, the change in percentages of total births which were non-white or to Puerto Rican mothers between 1950 and 1964 is too great to be explained solely by age differences. Again the data suggest a shift in population composition with respect to ethnicity -- from an area of predominantly white non-Puerto Rican families to area of predominantly Negro[16]

[16] Negroes make up most of the non-white total.

and Puerto Rican families.

Thus the combination of statistics suggests that the population of Target Area 5, which was selected initially as the project site because 1) it had a heavy concentration of delinquency and PINS cases and 2) yet was somewhat typical of the Bronx as a whole with respect to ethnic and socioeconomic characteristics, had undergone considerable change between 1960 and 1966. The population size appears to have remained relatively stable, but the ethnic composition shifted to a predominance of Puerto Rican and non-white families. On the basis of these data the project site was expanded somewhat, with the two tracts of Area 5 remaining the "core tracts" of the project's program.

Summary and Conclusion

In summary, the selection of a program site was based on the utilization of statistics from the courts which showed the geographic location and concentration of delinquents, and on an analysis of census and other data which provided a description of the social and demographic characteristics of the areas of heavy concentration. Because the entire borough was undergoing rapid change in the years between 1960 and 1966, current data to describe these changes were urgently needed. The change in size and ethnic composition of the population was documented by combining available health, school, and public utility data. We were unable to document any change in socioeconomic level which may also have occurred.

All of this suggests ways of utilizing statistics from a number of sources to provide current inter-censal descriptions of urban sub-areas. Any kind of community action program, and, indeed, many other types of programs, need up-to-date descriptions of their program sites and clientele or target populations. In the inter-censal years it is especially difficult to provide accurate current descriptions of small sub-areas, particularly in rapidly changing urban areas. The present study describes one method of making some intelligent "guestimates" of the mid-censal social and demographic characteristics of such areas.

TABLE I

SELECTED SOCIAL AND ECONOMIC CHARACTERISTICS OF THE BRONX AND OF SIX POTENTIAL AREAS

Variable	The Bronx	Area #1	Area #2	Area #3	Area #4	Area #5	Area #6
Total Population	1,424,815	82,312	35,014	30,148	31,617	18,795	9,051
Per Cent of Population	100.00	5.78	2.46	2.12	2.22	1.32	0.64
Per Cent Foreign Stock	50.60	30.50	29.30	18.40	31.20	60.50	24.90
Per Cent Negro	11.50	21.70	26.70	30.50	20.30	7.20	42.20
Per Cent Puerto Rican	13.10	34.40	51.40	47.70	41.90	10.10	12.20
Sex Ratio	88.80	89.90	88.00	84.80	89.90	90.50	86.00
Per Cent Males Married Aged 14 & Older	68.20	62.10	64.90	62.10	65.80	67.70	77.10
Per Cent Under Age 16	25.50	29.80	29.30	36.10	29.40	24.30	52.20
Dependency Ratio	503.60	526.90	489.30	654.10	536.30	536.70	1139.70
Mean Population in Household	3.02	3.24	3.35	3.53	3.30	2.92	4.41
Per Cent Enrolled in School 14-17 Year Group	93.00	83.00	92.00	71.00	87.00	99.00	92.00
Median Grade Completed for Population 25 Yrs. and Older	9.50	8.60	8.60	7.90	8.40	8.70	10.30
Per Cent Unemployed in Male Civilian Labor Force	4.90	7.20	8.70	8.90	7.40	5.90	6.00
Per Cent Professional & Technical Workers in Male Civilian Labor Force	8.70	3.20	2.60	2.20	3.40	7.10	1.50
Per Cent Moved Between 1955-1960	39.90	37.80	37.30	42.80	39.60	33.60	37.60
Per Cent of Population Moved into SMSA Between 1955-1960	1.87	2.05	2.15	2.33	2.19	1.78	0.78

TABLE II

ESTIMATED POPULATION PER OCCUPED DWELLING UNIT
IN THE BRONX BY METER DISTRICT, 1960 AND 1966

Area	1966	1960
Bronx Total	2.93	3.07
Meter District 10	3.05	3.20
Meter District 11	2.71	2.84
Meter District 12	3.08	3.23

TABLE III

TOTAL ENROLLMENT IN EIGHT PUBLIC SCHOOLS IN THE EAST
TREMONT SECTION OF THE BRONX, 1957-1965

Schools	Number of Pupils									
	1957	1958	1959	1960	1961	1962	1963	1964	1965	% Change 1957-65
P.S. 6	962	959	1001	1065	1114	1148	1282	1344	1604	66.7
P.S. 57	782	810	874	1005	1078	1114	1168	1238	1397	78.6
P.S. 58	545	789	863	921	980	1002	1050	1183	1080	98.2
P.S. 59	903	908	895	884	923	942	1112	1265	1053	16.6
P.S. 67	1189	1124	1135	1273	1423	1485	1623	1864	1987	67.1
P.S. 92	1207	1231	1257	1213	1390	1498	1586	1863	2021	67.4
J.H. 44	1060	845	856	838	875	525	1119	1120	1148	12.6
J.H. 118	1155	1189	1102	1048	970	979	975	1285	1300	8.3

TABLE IV

ETHNIC COMPOSITION OF THE SCHOOL POPULATION IN 8 PUBLIC SCHOOLS IN THE BRONX

Schools	1957	1958	1959	1960	1961	1962	1963	1964	1965
P.S. 6									
Total	962	959	1001	1065	1114	1148	1282	1344	1604
Per Cent	100	100	100	100	100	100	100	100	100
Puerto Rican	14.8	16.7	22.3	29.1	34.6	42.3	50.9	57.1	62.6
Non-white	6.4	7.4	7.7	9.2	11.1	13.9	17.6	23.4	25.0
Other	78.8	75.9	70.0	61.7	54.2	43.7	31.5	19.6	12.2
P.S. 57									
Total	782	810	874	1005	1078	1114	1168	1238	1397
Per Cent	100	100	100	100	100	100	100	100	100
Puerto Rican	9.3	11.5	14.9	19.1	26.0	31.7	36.9	44.1	44.8
Non-white	11.9	13.1	16.7	16.8	20.3	19.6	20.9	27.3	27.9
Other	78.8	75.4	68.4	64.1	53.7	48.7	42.2	28.6	27.2
P.S. 58									
Total	545	789	863	921	980	1002	1050	1183	1080
Per Cent	100	100	100	100	100	100	100	100	100
Puerto Rican	51.6	62.1	67.2	70.7	68.0	65.3	62.3	61.8	55.9
Non-white	9.7	11.7	11.9	12.5	16.0	22.1	27.0	33.6	38.6
Other	38.7	26.2	20.9	16.8	16.0	12.7	10.7	4.6	5.4
P.S. 59									
Total	903	908	895	884	923	942	1112	1265	1053
Per Cent	100	100	100	100	100	100	100	100	100
Puerto Rican	25.0	30.2	33.0	36.8	39.2	43.9	42.8	44.7	49.9
Non-white	12.2	12.4	14.4	16.9	19.4	23.0	28.1	32.2	30.4
Other	62.8	57.4	52.6	46.4	41.4	33.0	29.0	23.1	19.5

TABLE IV (Continued)

Schools	1957	1958	1959	1960	1961	1962	1963	1964	1965
P.S. 67									
Total	1189	1124	1135	1273	1423	1485	1623	1864	1987
Per Cent	100	100	100	100	100	100	100	100	100
Puerto Rican	12.4	15.2	20.1	25.7	32.9	36.8	42.1	48.0	52.5
Non-white	7.2	9.1	10.0	11.2	13.8	19.6	22.4	29.6	34.0
Other	80.4	75.7	70.0	63.2	53.3	43.6	35.5	22.4	13.3
P.S. 92									
Total	1207	1231	1257	1213	1390	1498	1586	1863	2021
Per Cent	100	100	100	100	100	100	100	100	100
Puerto Rican	18.9	21.0	26.3	28.7	32.4	35.8	40.6	47.7	47.8
Non-white	9.5	9.4	9.9	11.5	16.4	20.6	27.2	33.3	35.7
Other	73.7	69.6	63.8	59.8	51.2	43.7	32.2	19.0	16.3
J.H. 44									
Total	1060	845	856	838	875	525	1119	1120	1148
Per Cent	100	100	100	100	100	100	100	100	100
Puerto Rican	14.3	16.8	18.0	26.3	28.1	34.2	40.4	45.3	65.1
Non-white	10.0	12.7	12.3	14.0	14.5	17.3	22.0	26.3	21.4
Other	75.7	70.5	69.7	59.8	57.4	48.5	37.6	28.4	13.4
J.H. 118									
Total	1155	1189	1102	1048	970	979	975	1285	1300
Per Cent	100	100	100	100	100	100	100	100	100
Puerto Rican	27.6	34.1	39.0	49.0	51.2	60.3	63.5	62.7	61.0
Non-white	5.6	7.1	8.4	7.7	11.6	20.8	19.4	25.2	26.5
Other	66.8	58.8	52.5	43.2	37.1	18.9	17.1	12.2	12.4

TABLE V

Per Cent Non-white of Total Live Births in Selected Health Areas of the Bronx, 1950-1964

Year	Health Areas				
	17	18	19	20	21.1
1950	1.93	4.67	3.49	3.69	2.95
1951	1.69	5.19	3.54	3.57	3.37
1952	3.32	6.88	5.27	4.58	2.58
1953	4.97	7.60	7.63	4.23	3.54
1954	3.79	10.00	7.90	5.39	3.31
1955	6.43	13.95	9.93	8.06	4.59
1956	10.00	10.81	8.72	10.65	5.82
1957	10.39	17.02	11.70	12.44	3.20
1958	11.89	11.38	12.82	11.72	7.71
1959	14.08	16.42	12.83	12.66	6.54
1960	14.96	13.51	18.21	15.40	11.91
1961	13.39	18.16	21.63	16.34	15.45
1962	21.58	23.54	21.93	18.36	19.37
1963	26.48	30.24	23.82	26.17	22.27
1964	29.65	36.24	28.18	34.14	29.68

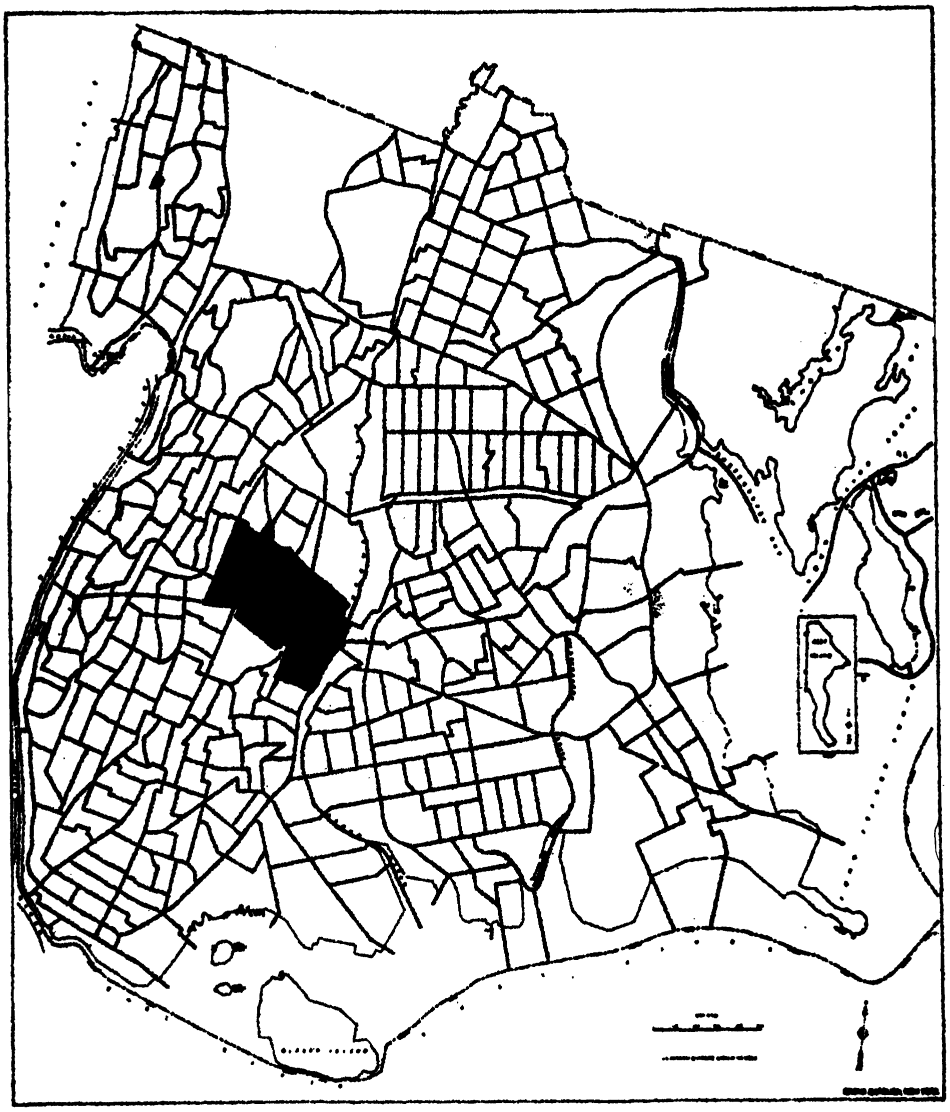
TABLE VI

Per Cent of Total Live Births to Puerto Rican Mothers
In Selected Health Areas of the Bronx, 1950-1964

Year	Health Areas				
	17	18	19	20	21.1
1950	1.93	7.16	1.16	2.37	.98
1951	2.91	13.49	2.43	4.08	2.36
1952	4.60	17.66	3.89	6.28	3.55
1953	7.18	28.07	5.96	9.52	4.96
1954	6.50	34.24	8.54	11.26	6.62
1955	10.89	38.75	9.09	10.75	8.52
1956	10.50	48.91	12.08	13.31	14.56
1957	14.35	52.97	10.63	13.38	14.74
1958	14.07	56.43	15.38	17.45	15.75
1959	14.55	59.90	18.11	22.96	18.75
1960	20.90	62.23	19.73	26.60	27.97
1961	25.86	61.92	25.42	35.19	39.13
1962	30.21	61.04	30.39	33.15	42.53
1963	24.78	57.08	35.30	40.48	48.46
1964	30.57	48.76	42.04	39.17	48.59

MAP I

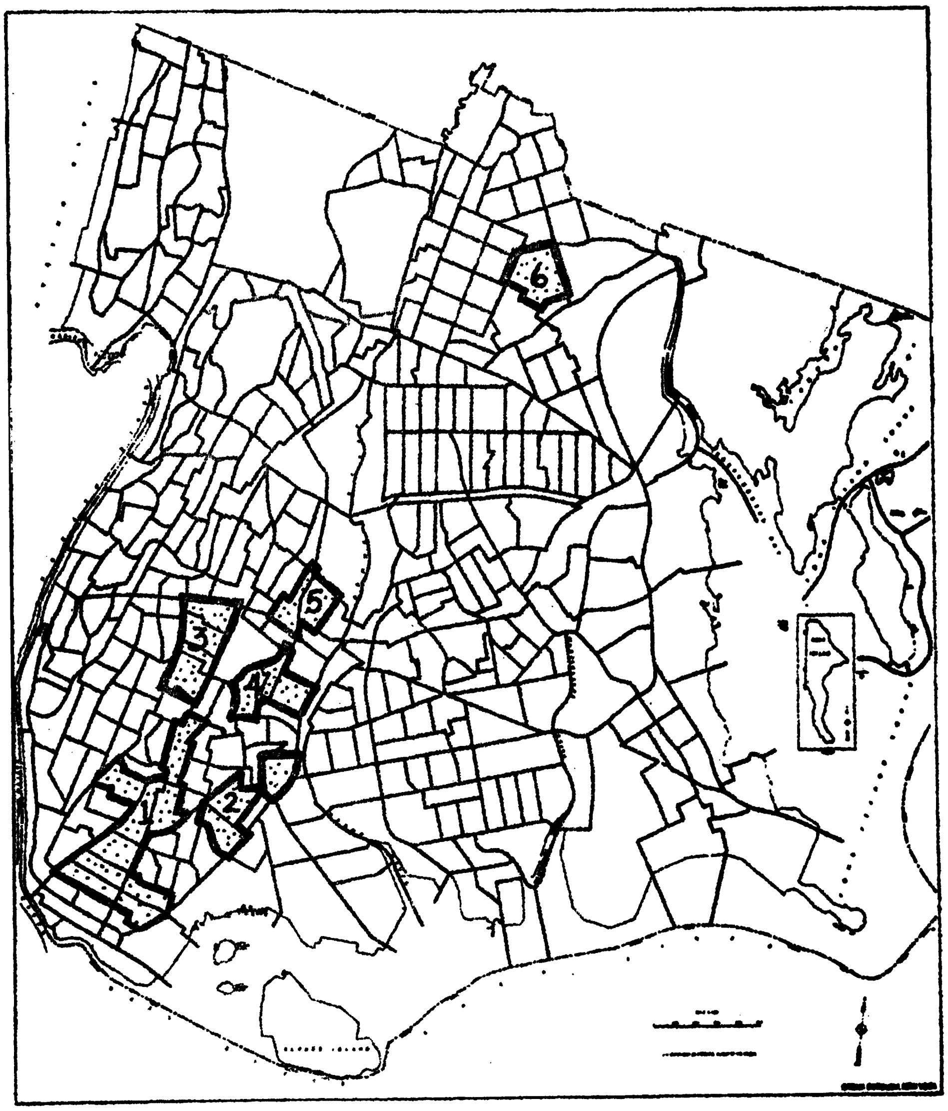
East Tremont Project Area Within Bronx Borough



Juvenile Court Community Development Project, 1966

MAP II

Six Alternative Target Areas



Juvenile Court Community Development Project, 1966

MIGRATION AND THE RISK OF DYING

Herbert I. Sauer, National Center for Chronic Disease Control, Public Health Service

Geographic differences in the risk of dying have been well demonstrated for middle-aged whites in the United States (1-7). Death rates are low in the Great Plains area generally and high near the East Coast of the U.S., with the middle-aged men in the highest-rate areas having twice the risk of those in the lowest-rate areas. While various hypotheses have been suggested to account for factors responsible for these differences, the evidence to support these hypotheses is generally tenuous.

In a prior study (6), we have shown that: If it were possible to identify the factors responsible for the low rates for all causes of death in the four lowest-rate areas (having a population exceeding one million) and if these factors could be applied to the U.S. over-all, then for those under age 65 alone there would be 100,000 fewer deaths per year.

While our concern is deaths from all causes, we at the Ecology Field Station of the Heart Disease Control Program are even more concerned with the cardiovascular-renal diseases, including coronary heart disease.

HYPOTHESES

The question has been raised as to whether the differences in death rates could be due to migration patterns (5). Stated more specifically, one hypothesis suggests that low death rates in North Dakota, Nebraska and other Great Plains areas are observed because men with heart disease and other chronic ailments move away, presumably to "retirement" States.

Our objective is to focus on this and other questions related to migration and ecological patterns. For example, does migration of itself affect the risk of dying? Is the risk of dying related to the area in which one is born? What differences in mortality are observed for different migration streams?

MATERIALS AND METHODS

Deaths of the native-born were tabulated by State of birth and by State of usual residence, by age, sex and race for middle-aged individuals. By arrangements made with the National Center for Health Statistics (NCHS) tabulations were obtained for coronary heart disease deaths in 1950 and for all deaths and 17 cause categories, 1959-1961.

The age groups selected were those used by the Bureau of the Census in its State of birth and lifetime migration reports (8-10). Some rates were calculated by NCHS and some by our Program.

All rates presented are for middle-aged native-born whites, by sex. Age-specific death rates for age groups 45-64, 35-74, and 40-69 are all age-adjusted by the direct method by ten-year age groups to the total U.S. population in these age groups in 1950.

In any study of morbidity or mortality rates, consideration must be given as to whether the population accepted as the "population at risk" is identified with reasonable accuracy. In this study, several factors are pertinent.

a. The population is estimated from the 25-percent sample enumeration. The resulting sample error is generally quite small for all those living in a specific State of birth or for all lifetime migrants to or from a State. For specific streams of migration sample error is somewhat greater. Probably more important is the systematic bias which may easily arise from such sampling; for example, the U.S. population of white males, age 65-74, as calculated from the 25-percent sample is approximately two percent less than the complete enumeration (11). In the age group 45-64 this difference is generally less than one percent.

b. In the 1960 Census, those with no indication of nativity were classified as native-born (10) but they obviously could not be classified as to State of birth. This group together with a miscellaneous group were proportionately much more numerous in the population than on death certificates. This introduced a bias in the conventional calculation of death rates, --for age groups 45 to 74 probably one to three percent higher. * Comparisons of rates are presented in such a way that errors should be negligible, generally not more than two percent. **

---

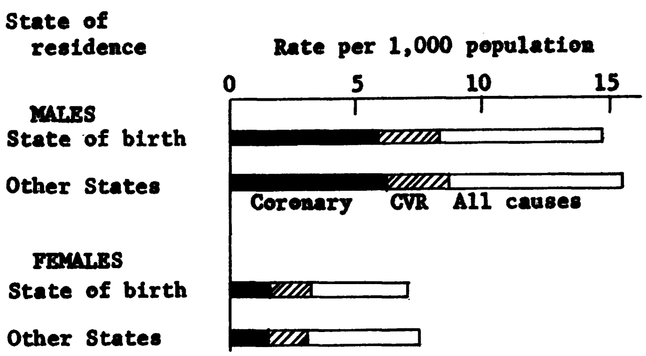
*For Alaska, Nevada and Rhode Island the rates may be about four percent higher than "true" rates, and for the District of Columbia, about seven percent.

**To study the magnitude of the bias, the rates of whites residing in their State of birth and of residents not born in the State were adjusted, by age and sex. Both population and deaths with

AGE, SEX AND CAUSE OF DEATH

Death rates for middle-aged whites residing in their State of birth are almost identical to those residing in States other than State of birth (Figure 1). The native-born whites who

Figure 1--DEATH RATES FOR RESIDENTS OF STATE OF BIRTH AND OF OTHER STATES BY SEX, NATIVE WHITES, AGE 45-64, UNITED STATES, 1959-1961



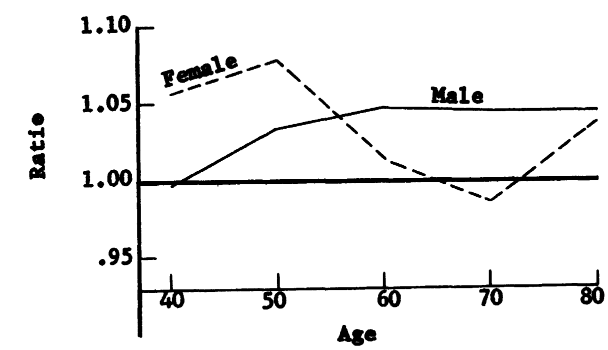
are residing in States other than State of birth at the time of the census enumeration are defined by the Bureau of the Census as "lifetime migrants" (10) and for convenience are hereafter frequently referred to as merely "migrants". A native-born decedent whose residence at time of death is not his State of birth is similarly classified as a lifetime migrant.

The category, "residents of State of birth" at time of census enumeration or at death, includes those who have never moved; but it also includes intra-state migrants and those who have moved across State lines and then returned to re-establish residence in their State of birth prior to census or death; this group may be called "non-migrants".

Our operational definitions thus are obviously more limited than those which need to be used in a general theory of the effects of migration on health or of migration (12).

In order to observe more clearly the slight differences that are recorded, we may arbitrarily accept the age-sex-specific rates for those residing in their State of birth as unity, calculate the ratios of the lifetime migrants to the "non-migrants", and plot them on a scale to magnify the differences (Figure 2).

Figure 2--RATIO OF DEATH RATES OF WHITES: "LIFETIME MIGRANTS" divided by "RESIDING STATE OF BIRTH", ALL CAUSES, BY AGE AND SEX, 1959-1961



For all causes of deaths, the rates for male migrants are about four percent higher than for those residing in State of birth, whereas females show a mixed pattern of rates. For the cardiovascular-renal causes, male migrants tend to have a rate nominally higher, and female migrants nominally lower, than those residing in their State of birth (Figure 3).

The ratios vary somewhat by cause showing migrants' rates for hypertension and hypertensive heart disease to be about ten percent lower than rates for "non-migrants" (Figure 4). The ratios show higher migrant rates for lung cancer, accidental and violent causes and chronic respiratory diseases than for those living in State of birth. The high migrant rate for chronic respiratory diseases (chronic bronchitis, emphysema, asthma, and other non-specific chronic respiratory diseases) is due in part to the high rate of migrants in Arizona. But even if this group were excluded, the ratio would still remain

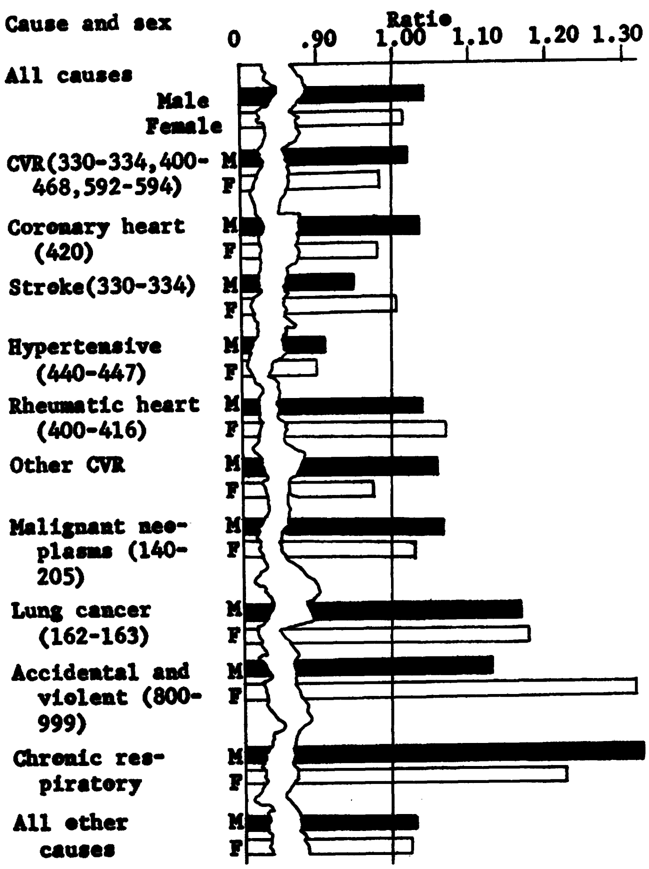
---

State of birth or nativity not specified (or classified as "other") were proportionately distributed to the specific categories. One of the larger differences as a result of adjustment is a comparison of Nebraska with California: The Nebraskans born in the State have unadjusted rates 14.5 percent lower than the California-born, as compared with 12.6 percent lower, derived from adjusted rates. For many comparisons, such as Missouri with New York, the result is essentially the same (within one-half percent), whether adjusted or unadjusted rates are used.

For rates of those who have moved out of their State of birth, no practical adjustment procedure has yet been developed. Therefore, no formal adjustment factors are used in the presentation of rates in this paper for either group.

Figure 3--RATIO OF DEATH RATES OF WHITES: "LIFE-TIME MIGRANTS" divided by "RESIDING STATE OF BIRTH", CVR DISEASES, BY AGE AND SEX, 1959-1961

Figure 4--RATIO OF DEATH RATES OF WHITES AGE 35-74: "LIFETIME MIGRANTS" divided by "RESIDING STATE OF BIRTH", SELECTED CAUSES OF DEATH, BY SEX, 1959-1961



high. The high ratio for "accidental and violent" causes for white females is due largely to the low rates for those residing in their State of birth; the death rates of migrants are only about 10 points higher, producing a very high ratio. Even so, accidents and violence (and to a lesser extent, malignant neoplasms) are largely responsible for the elevated all-causes ratios for white females, age 35-54 (Figure 2).

BY STATE

We may consider the specific hypothesis that the Dakotas and Nebraska have low death rates because the men with chronic ailments move away; or the more general hypothesis that all differences in death rates are due to selective migration. Comparison of death rates by State of birth without regard to residence would be a test of these hypotheses. If the more general hypothesis is true, we would expect no difference in rates between states. These calculations are presented for males 45-64 for cardiovascular-renal causes and show a pattern of death rates very similar to those tabulated by State of residence (Table 1), the product moment correlation being +0.91.

For all causes of death, a similar parallel is obtained, with a correlation of +0.88 in death rates by State of residence with rates by State of birth. (For these and all subsequent correlations, n = 46 States, with Alaska, Hawaii, Nevada and Wyoming excluded because of small numbers in one of the following categories: "residing in State of birth" or "residing in States other than State of birth"--that is with one standard deviation of random error exceeding seven percent, for white males, age 45-64, CVR diseases.) These are obviously not independent correlations, but they do test the hypothesis proposed.

In order to avoid such dependence, rates have also been calculated by State for those residing in State of birth and for those whose State of residence is "other States" (Table 2). Those living in States other than State of birth obviously moved away at some time after birth and in this setting are viewed as out-migrants. The correlation between these two variables is +0.80.

The six States with the highest rates for those residing in State of birth have rates approximately as high for those born in the State who have moved on to other States (Figure 5). The six States with the lowest rates for those residing in State of birth also had among the lowest rates for those who had moved to other States.

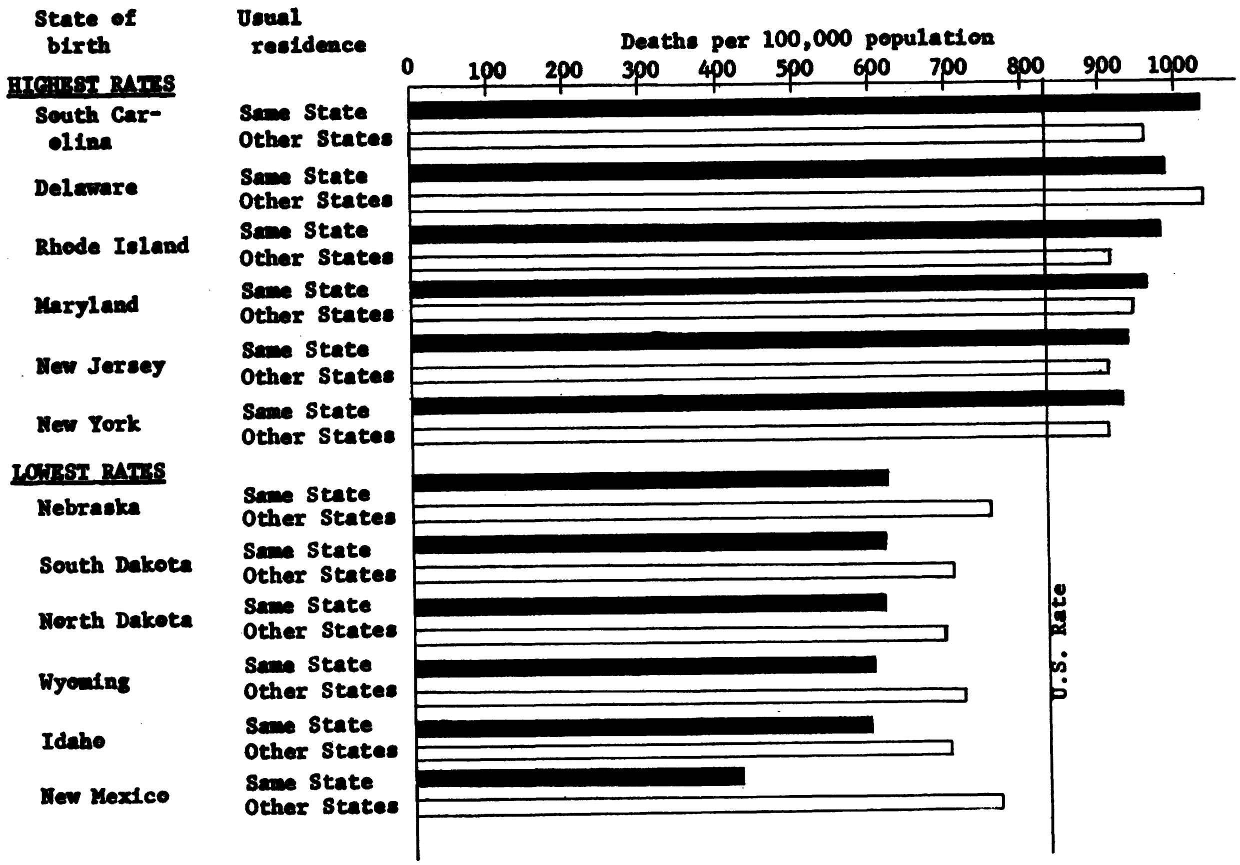
Table 1--CARDIOVASCULAR-RENAL (CVR) DISEASES DEATH RATES BY STATE OF RESIDENCE AND BY STATE OF BIRTH, NATIVE WHITE MALES, AGE 45-64, 1959-1961

State	State of residence	State of birth
UNITED STATES	836.0	836.0
NEW ENGLAND		
Maine	876.1	855.0
New Hampshire	866.0	882.1
Vermont	783.3	817.3
Massachusetts	914.2	916.1
Rhode Island	992.7	958.5
Connecticut	818.0	845.7
MIDDLE ATLANTIC		
New York	931.4	927.8
New Jersey	942.5	934.5
Pennsylvania	925.0	925.4
SOUTH ATLANTIC		
Delaware	879.4	1004.7
Maryland	932.1	958.3
Dist. of Columbia	1068.3	939.8
Virginia	863.7	899.0
West Virginia	843.8	840.4
North Carolina	921.3	904.0
South Carolina	1005.6	1015.2
Georgia	917.1	921.7
Florida	865.6	824.7
EAST NORTH CENTRAL		
Ohio	836.4	850.8
Indiana	832.7	854.5
Illinois	900.4	868.9
Michigan	820.8	827.5
Wisconsin	752.1	776.5
EAST SOUTH CENTRAL		
Kentucky	784.5	802.0
Tennessee	761.1	804.7
Alabama	796.7	800.9
Mississippi	757.4	779.4
WEST NORTH CENTRAL		
Minnesota	708.2	714.8
Iowa	720.8	758.2
Missouri	767.8	793.6
North Dakota	672.4	667.3
South Dakota	692.0	675.4
Nebraska	658.4	697.9
Kansas	690.1	713.4
WEST SOUTH CENTRAL		
Arkansas	732.0	736.7
Louisiana	903.8	871.0
Oklahoma	711.2	715.1
Texas	756.4	740.4
MOUNTAIN		
Montana	770.4	827.7
Idaho	681.8	662.1
Wyoming	707.3	694.1
Colorado	684.5	727.5
New Mexico	571.8	608.8
Arizona	797.1	688.4
Utah	642.5	705.6
Nevada	1007.4	809.7
PACIFIC		
Washington	791.8	691.8
Oregon	766.6	705.4
California	823.3	735.9
Alaska	705.6	611.7
Hawaii	977.8	1034.3

Table 2--CVR DISEASES DEATH RATES BY STATE OF BIRTH BY PLACE OF USUAL RESIDENCE, WHITE MALES, AGE 45-64, 1959-1961

State of birth	Place of usual residence	
	State of birth	Other states
UNITED STATES	829.4	847.9
NEW ENGLAND		
Maine	857.1	851.3
New Hampshire	904.4	856.2
Vermont	748.0	881.5
Massachusetts	921.1	904.4
Rhode Island	983.2	918.0
Connecticut	810.8	920.6
MIDDLE ATLANTIC		
New York	931.8	916.3
New Jersey	941.4	916.3
Pennsylvania	924.2	928.2
SOUTH ATLANTIC		
Delaware	987.0	1039.0
Maryland	961.3	946.6
Dist. of Columbia	1148.2	864.7
Virginia	866.9	963.2
West Virginia	826.4	859.2
North Carolina	912.7	878.3
South Carolina	1037.7	963.6
Georgia	924.6	916.2
Florida	838.1	773.0
EAST NORTH CENTRAL		
Ohio	839.7	879.9
Indiana	835.7	888.5
Illinois	892.0	825.6
Michigan	811.8	875.6
Wisconsin	746.6	845.5
EAST SOUTH CENTRAL		
Kentucky	761.2	853.1
Tennessee	755.2	882.9
Alabama	767.0	858.8
Mississippi	750.9	819.2
WEST NORTH CENTRAL		
Minnesota	680.2	769.0
Iowa	701.9	817.4
Missouri	745.5	847.4
North Dakota	620.4	698.3
South Dakota	621.4	710.8
Nebraska	623.7	757.1
Kansas	642.3	770.3
WEST SOUTH CENTRAL		
Arkansas	700.3	763.0
Louisiana	885.6	827.2
Oklahoma	631.5	771.2
Texas	727.6	778.0
MOUNTAIN		
Montana	888.5	791.3
Idaho	600.0	702.9
Wyoming	605.4	736.3
Colorado	641.2	787.2
New Mexico	431.7	771.1
Arizona	582.2	764.1
Utah	629.9	807.1
Nevada	762.9	845.8
PACIFIC		
Washington	685.9	701.3
Oregon	662.7	770.7
California	729.1	788.1
Alaska	259.7	724.9
Hawaii	1183.9	870.8

Figure 5--CVR DISEASES DEATH RATES BY STATE OF BIRTH BY PLACE OF USUAL RESIDENCE, STATES WITH HIGHEST AND LOWEST RATES, WHITE MALES, 45-64, 1959-1961



The death rates for those living in State of birth, or "non-migrants," show a consistently high correlation with out-migrants, for each sex separately, for all causes and coronary heart disease, as well as for the CVR diseases (Table 3).

Those residents not born in a State who moved into and became residents of a State prior to death may be thought of as in-migrants. Rates for this group also show a rather high correlation with "non-migrants."

Out-migrants also show substantial correlations with in-migrants. While somewhat lower, they are all statistically significant at the .01 level. Loosely speaking, those who spend the early portion of their lives in a State show some tendency toward having death rates similar to those who spend the latter part of their lives in the State.

WEST NORTH CENTRAL

The West North Central States generally have low death rates, and this is particularly true for the "non-migrants" in this area, those residing in State of birth. But what about the death rates of those born in these States who migrate to various other parts of the United States when compared as in-migrants to various areas? Those age 35-74 who moved to East Coast States, had a death rate about 20 percent lower than those born in their respective East Coast States (Figure 6).

The migrants from each of the West North Central States to each division of the East Coast States (New England, Middle Atlantic and South Atlantic) had lower rates than did those born there, for all causes and for coronary heart disease as well as for CVR diseases--in spite of

Table 3--CORRELATION COEFFICIENTS OF DEATH RATES OF LIFETIME MIGRANTS AND RESIDENTS OF STATE OF BIRTH FOR SELECTED CAUSES AND BY SEX, MIDDLE-AGED WHITES, 1959-1961 and 1950

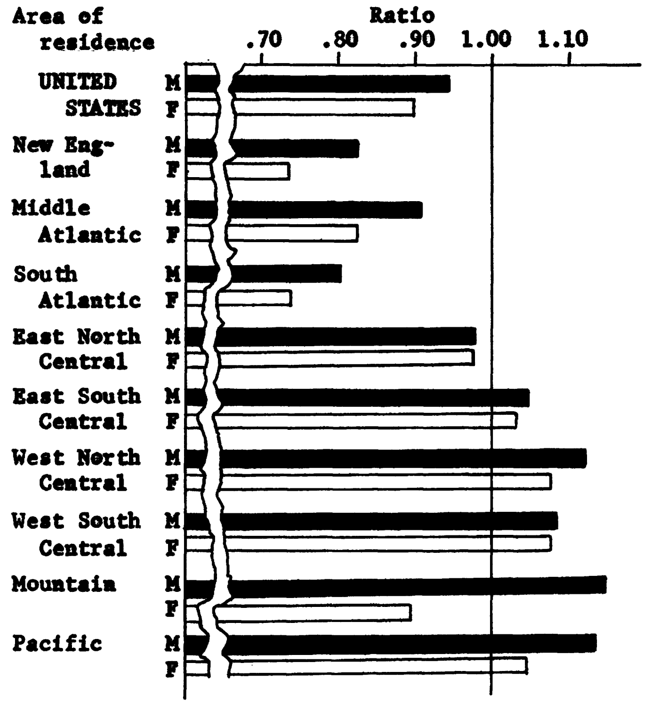
x = "Non-migrants": Living in State of birth (or born in State of residence).
y = Out-migrants: Born in State but living elsewhere (or lifetime migrants to other States).
z = In-migrants: Residents not born in State (or lifetime migrants moving into the State).
n = 46 States, excluded are Alaska, Hawaii, Nevada, Wyoming and District of Columbia.

Year, age, cause and sex	r_{xy} "Non-migrants" with out-migrants	r_{xz} "Non-migrants" with in-migrants	r_{yz} Out-migrants with in-migrants
1959-1961, age 45-64			
All causes			
Male	.75	.60	.46
Female	.76	.65	.40
CVR			
Male	.80	.70	.48
Female	.73	.68	.45
Coronary heart disease			
Male	.75	.80	.50
Female	.72	.64	.44
1950, age 40-69			
Coronary heart disease			
Male	.82	.79	.67
Female	.60	.77	.66
1959-61, age 35-74			
All causes			
Male	.81	.72	.58
1959-61, age 65-74			
All causes			
Male	.75	.74	.63

small numbers and resulting random error. (A partial exception is the Missouri-born.) Thus, the people moving from the northern part of the Great Plains to the East Coast had death rates intermediate between their State of birth and their State of residence.

Those born in the West North Central States who moved to States in other parts of the country, who thus became in-migrants, show a more mixed pattern of rates, similar to or slightly higher than the "non-migrant" rates in the various geographic divisions of States to which they have moved.

Figure 6--RATIO OF CVR DISEASES DEATH RATES: "IN-MIGRANTS FROM WEST NORTH CENTRAL STATES" divided by "NON-MIGRANTS", BY GEOGRAPHIC DIVISION OF RESIDENCE BY SEX, AGE 35-74, 1959-1961



TO RETIREMENT AREAS

Out-migrants, age 45-64, from the Northeast to Florida have slightly lower death rates, for CVR as well as all causes, than those remaining in the Northeast, but have slightly higher rates than those residents born in Florida. Rates for migrants from other areas present a mixed pattern.

Out-migrants to Florida, age 65-74, have particularly low death rates for this age group (Figure 7), either in comparison with those born in Florida or with "non-migrants", --those who remained in their State of birth. Most pronounced are the rates for those from the Northeast and East North Central States--for males approximately 23 percent and for females 37 percent below the rate of the "non-migrants".

Female migrants to Florida, age 65-74, thus show a pattern of low rates even more clearly than do males (Figure 8), and females, age 55-64, also present evidence of low rates. However, these patterns of low rates for whites do not apply to nonwhite migrants to Florida.

Figure 7--CVR DEATH RATES FOR OUT-MIGRANTS TO FLORIDA AND "NON-MIGRANTS" BY GEOGRAPHIC DIVISION OF BIRTH, MALES, AGE 65-74, 1959-1961

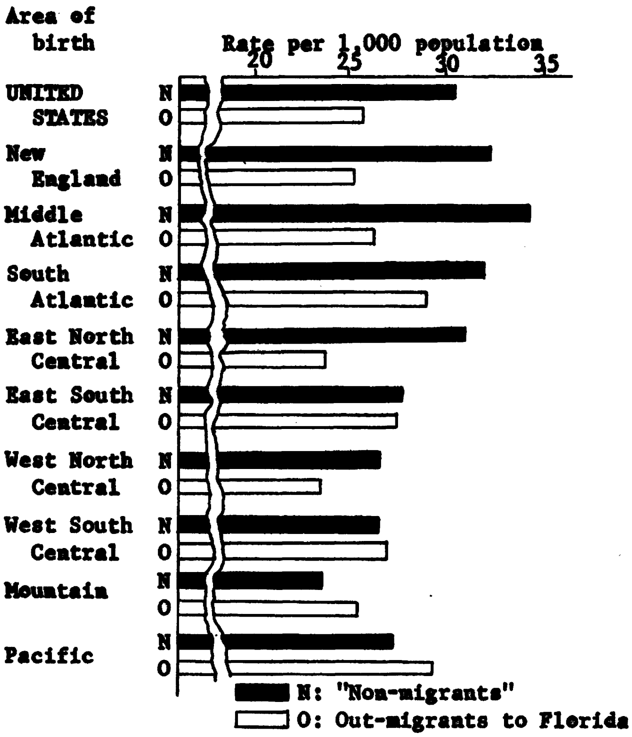
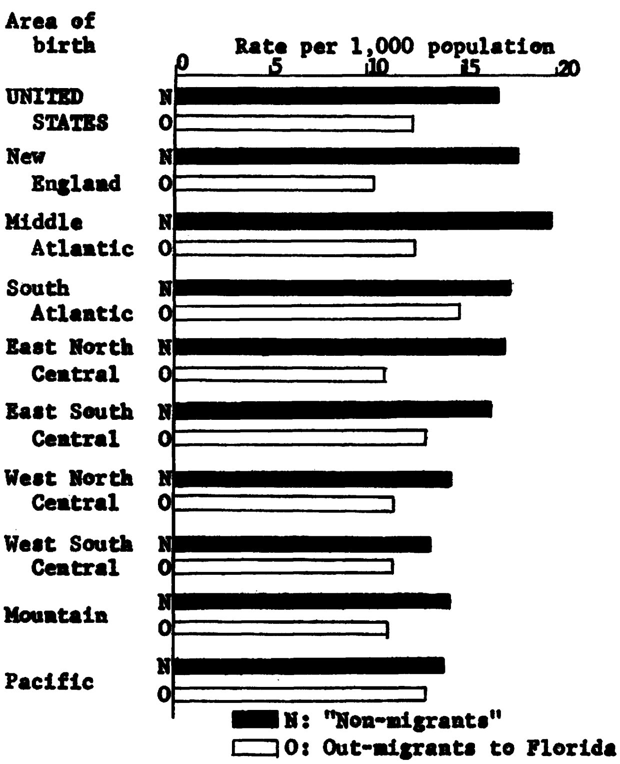


Figure 8--CVR DEATH RATES FOR OUT-MIGRANTS TO FLORIDA AND "NON-MIGRANTS" BY GEOGRAPHIC DIVISION OF BIRTH, FEMALES, AGE 65-74, 1959-1961



White female migrants to Arizona had slightly lower rates than did those remaining in their State of birth, for all causes as well as CVR causes, but middle-aged male migrants had rates very slightly higher than did those remaining in their State of birth.

Migrants to California in 1959-1961 definitely had higher rates than those born in California--10 to 15 percent higher for the CVR diseases. The rates for those migrating to California from most areas were higher than for those remaining in their State of residence, except for the East Coast. Migrants from the latter had death rates nominally lower than the rates for those remaining along the East Coast, but they still had the highest rates among those migrating to California.

Migrants to California, age 65-74, had patterns very similar to those for age 45-64, which is in marked contrast to the data for Florida.

Since the health status of individuals in different streams of migration and in different age-sex groups may be different, no conclusions should be drawn about the relative merits of various retirement areas.

TO INDUSTRIAL AREAS

Another major wave of migration has been to the Middle Atlantic States, and the migrant death rates are similar to the high rates for those born in these States. They thus tend to be slightly higher than rates of those remaining in their State of birth (Figure 9).

Those from the West North Central States have CVR death rates 15 to 20 percent higher than those remaining in their State of birth, but these rates are still 5 to 20 percent lower than the rate for those born in the Middle Atlantic States (for ages 45-64 and 65-74, by sex).

Migrants to the East North Central States present a similar but less pronounced pattern of higher rates than for those remaining in their State of birth--except that migrants from the Northeast had rates slightly lower than those remaining in the Northeast.

DISCUSSION

Standard vital statistics procedures have been used to tabulate and calculate the rates and ratios presented. The presence of both random and systematic error has been recognized. In addition, one may speculate: Possibly the informant for the death certificate may sometimes not know the State of birth and assume it to be the State in which the decedent was last residing--while the informant to the census enumerator will report otherwise. An alternate possibility is that the individual while living may wish to be classified as born in the State even though he was actually born elsewhere. (This latter line of thought is suggested by data for nonwhites in New York State.)

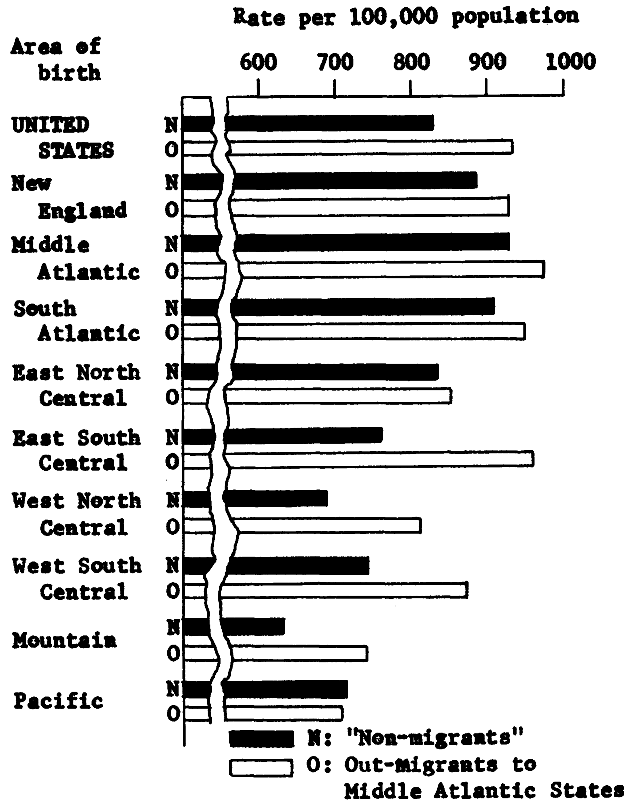
To achieve an adequate comprehension of the meaning inherent in the comparisons is particularly difficult. The comparisons of U.S. totals for lifetime migrants with those born in State of residence are in part geographic comparisons, because in some sections of the country the population consists largely of the population born there. Within geographic divisions of States there also is a degree of heterogeneity New Jersey, for example, has many in-migrants, while Pennsylvania's population was largely born there.

State data also may lack homogeneity--as is shown by the Omaha metropolitan State economic area: Of the population, all ages, born in Nebraska, approximately 20 percent lives in the Omaha area, but of those who have moved into Nebraska almost 40 percent lives in the Omaha area (13). In-migrants to Nebraska have a death rate 21 percent higher than those born in the State (native white males age 45-64). If migrants to a metropolitan area experience the high mortality observed for the metropolitan area as a whole, then the application of indirect methods of adjustment suggest that the difference in rate would be not more than 14 percent, and possibly considerably less.

Similarly, a substantial portion of the out-migrants from rural Nebraska has moved to metropolitan areas (14). Thus, it would be reasonable to expect Nebraska out-migrants to have a higher death rate than those remaining in the State, as is shown in Table 2.

The low Florida rates for migrants of retirement age suggest several possibilities: Those moving to Florida for retirement (a) are, as a group, in better health than those who remain in their home communities, to be near relatives, friends, a family physician, hospital, and other resources known to them; (b) have a higher socio-economic status, with its accompanying lower risk of dying, as inferred from studies in metropolitan areas in the U.S. (15-17); (c) are in some instances likely to return north either when a chronic illness sets in or when a spouse dies, or (d) may be younger, --that is, may be individuals age 65-69 to a substantially greater extent than individuals age 65-74 generally. An alternative hypothesis is that there is

Figure 9--CVR DISEASES DEATH RATES OF OUT-MIGRANTS TO MIDDLE ATLANTIC STATES AND "NON-MIGRANTS", BY GEOGRAPHIC DIVISION OF BIRTH, MALES, AGE 45-64, 1959-1961



something inherent in Florida living for those of retirement age which reduces their risk of dying, but which does not appreciably affect those prior to retirement.

A hypothesis within the realm of possibility is that a higher proportion of migrants generally will be aggressive "Type A" individuals (who have higher coronary rates) as defined by Friedman and associates (18), as compared with those who remain in their State of birth--more generally the "Type B" person.

Syme and associates have well demonstrated higher rates for those with high cultural mobility (19). If both of these characteristics are present to a substantially greater extent in migrants than in those remaining in their State of birth, then the migrant mortality rates would be expected to be higher than present tabulations indicate. On the other hand, some individuals (such as ministers, teachers, construction workers, etc.) migrate in order to remain in the same type of work.

Further, the educational level of lifetime migrants to large metropolitan areas generally is much higher than for those born in State of residence, but for those in the Pacific States, migrants consistently have lower educational achievement levels (20). Migrants to California have higher rates in relation to the rates for the California-born than would be anticipated from the experience for other parts of the country. One may speculate that this is related to their lower educational levels, particularly in view of the relationship observed for selected urban areas between socio-economic status and death rates (15-17).

While lifetime migrants generally have death rates slightly higher than those remaining in their State of birth, these differences are negligible compared to the high rates for migrants from abroad, or the foreign-born, in 1900 (as compared with native whites then) (21). In the U.S., Norwegian and Swedish-born middle-aged males have much higher death rates than those who remained in Norway and Sweden (22).

SUMMARY

1. Lifetime migrants--those living at time of death in a State different than State of birth--appear to have slightly higher death rates than do those living in State of birth.

2. Migration generally has tended toward metropolitan areas, which on the average have higher death rates than do non-metropolitan areas.

3. Available evidence does not support the hypothesis that geographic differences in death rates are primarily due to migration.

4. Lifetime migrants to Florida from the Northeast and North Central regions who are of retirement age have particularly low death rates. The pattern of rates suggests the possibility that selective migration in some way plays an important role.

REFERENCES

(1) Enterline, P.E., and Stewart, W.H. (1956). Geographic Patterns in Deaths from Coronary Heart Disease. Public Health Rep 71, (9), 849-55.

(2) American Heart Association, National Heart Institute, and Heart Disease Control Program, PHS, DHEW (1965). Cardiovascular Diseases in the U.S. Facts and Figures.

(3) National Office of Vital Statistics. Death Rates for Selected Causes by Age, Color and Sex: United States and Each State, 1949-1951. Vital Stat Special Report Vol. 49.

(4) Sauer, H.I., and Enterline, P.E. (1959). Are Geographic Variations in Death Rates for the Cardiovascular Diseases Real? J Chronic Dis 10, (6), 513-524.

(5) Sauer, H.I. (1962). Epidemiology of Cardiovascular Mortality--Geographic and Ethnic. Amer J Public Health 52, (1), 94-105.

(6) Sauer, H.I., and Moore, D.M. (1966). Area and Migration Differences in the Risk of Dying. Presented at the American Public Health Assn. annual meeting, November 2, 1966.

(7) Duffy, E.A., and Carroll, R.E. (In Press). United States Metropolitan Mortality, 1959-1961. Environmental Health Services, Public Health Service, DHEW.

(8) U.S. Bureau of the Census. United States Census of Population: 1950. Vol. IV, Special Reports, Pt.4, Chapter 4, State of birth.

(9) U.S. Bureau of the Census. United States Census of Population: 1960. Subject Reports, State of Birth. Final Report PC(2)-2A.

(10) U.S. Bureau of the Census. United States Census of Population: 1960. Subject Reports, Lifetime and Recent Migration. Final Report PC(2)-2D.

(11) U.S. Bureau of the Census. United States Census of Population: 1960. General Population Characteristics, United States Summary. Final Report PC(1)-1B to 52B.

(12) Lee, E.S. (1966). A Theory of Migration. Demography 3, (1), 47-57.

(13) U.S. Bureau of the Census. United States Census of Population: 1960. Subject Reports, Mobility for Metropolitan Areas. Final Report PC(2)-2C.

(14) U.S. Bureau of the Census. United States Census of Population: 1960. Subject Reports, Migration Between State Economic Areas. Final Report PC(2)-2E.

(15) Yeracaris, Constantine A. (1955). Differential Mortality, General and Cause-Specific In Buffalo, 1939-41. J Am Stat Assoc 50, (4), 1235-1247.

(16) Lilienfeld, A.M. (1956). Variations in Mortality from Heart Disease by Race, Sex and Socio-economic Status. Public Health Rep 71, (6), 545-552.

(17) Winkelstein, et al. (1967). The Relationship of Air Pollution and Economic Status to Total Mortality and Selected Respiratory System Mortality in Men. Archives of Environmental Health 14, (1), 162-171.

(18) Friedman, M. (1964). Behavior Pattern and Its Pathogenetic Role in Clinical Coronary Artery Disease. Geriatrics 19, (8), 562-567.

(19) Syme, S.L., Hyman, M.N. and Enterline, P.E. (1965). Cultural Mobility and the Occurrence of Coronary Heart Disease. J Health and Human Behavior 6, 178-189.

(20) Lee, E.S. Unpublished analyses of 1960 census data.

(21) Linder, F.E. and Grove, R.D. (1943). Vital Statistics Rates in the United States 1900-1940. United States Government Printing Office, Washington, D.C.

(22) Krueger, D.E., and Moriyama, I.M. (1967). Mortality of the Foreign Born. Am J Public Health 57, (3), 496-503.

A STUDY OF THE RELATIONSHIP BETWEEN SUICIDE RATE AND AGE IN THE U.S. (1914 to 1964)

David W. Walker, TRW Systems

Abstract

A study is presented comparing the functional relationships of suicide rate to age in the United States for four race-sex subgroups, white male, white female, nonwhite male, and nonwhite female. Regression coefficients or *b* values are used as a measure of the functional relationship for each year in each race-sex subgroup. Changes in these functional relationships during the period 1914-1964 are examined along with the hypothesis that these functional relationships are sensitive to the level of general business activity.

The functional relationship of suicide rate to age for white males is quite sensitive to the level of general business activity. The regression coefficients used to approximate the suicide rate-age relationship for white males in individual years vary inversely with the level of general business activity. This sensitivity of functional relationships to the level of general business activity was not observed however for the suicide rate-age relationships of the other three race-sex subgroups.

Introduction

Presented in this paper is (1) a comparison of the functional relationships of suicide rate to age for the four race-sex subgroups, white male, white female, nonwhite male, and nonwhite female; (2) an examination of the changes in these relationships during the period 1914-1964; and (3) an examination of the hypothesis that these functional relationships are sensitive to the level of general business activity.

Suicide rates had been observed to vary with age for all of the race-sex subgroups considered in this study. It was therefore decided to concentrate the analyses on the functional relationship of suicide rate to age rather than suicide rates themselves. Suicide rates are considered as a function of age or as the dependent variable in the regression analyses with age as the independent variable. Simple linear regressions and the *b* values or regression coefficients were used in the analyses as a measure of the functional relationship between suicide rate and age, since changes in the slopes of these straight lines can be more easily interpreted and related to the actual phenomena than could changes in constants calculated for a second degree curve $(Y = a + bX + cX^2)$

This study differs from others in that it primarily focuses attention on these functional relationships and associates the changes in these relationships with the level of general business activity. Previous procedures have been to correlate suicide rates themselves with business activity rather than functional relationships.

Summary of Conclusions

The typical suicide rate-age relationship for white males is linear. White male suicide rates increase steadily with age and are considerably higher in every age group than corresponding rates for the other three race-sex subgroups. White female suicide rates increase steadily as age increases up to about age 45 to 54, then decrease as age increases in the older age groups. Nonwhite male suicide rates increase steadily with age up to about age 35. After 35 the rates increase at a lower rate as age increases. Nonwhite female suicide rates are considerably lower than corresponding rates for the other three race-sex subgroups. Rates for nonwhite females increase slightly as age increases until about age 35, then decrease slowly as age increases after 35. The suicide rate-age relationship for each of the four race-sex subgroups is adequately described for this study by a linear approximation.

The functional relationship of suicide rate to age for white males was observed to be quite sensitive to the level of general business activity. The slopes of the linear regression lines used to approximate the suicide rate-age relationship for white males in individual years vary inversely with the level of general business activity. This sensitivity of functional relationships to the level of general business activity was not observed for the suicide rate-age relationships of the other three race-sex subgroups.

Nature of Data

Simple linear regression lines $(Y = A + bX)$ fitted by method of least squares were used to approximate the functional relationship of suicide rate (Y) to age (X) for each year in each of the four race-sex subgroups. The slopes (regression coefficients or *b* values) of these linear approximations were arranged in a time series so that changes in slopes could be compared with the level of general business activity. The data cover suicide rates from ages 5 to 84 inclusive.[1] Ages 85 and over were not included because of the large sampling fluctuations due to small numbers in the population.[2]

The suicide rates were computed by the U.S. Department of Health, Education, and Welfare and represent the number of deaths by suicide per 100,000 population. A considerable number of these rates, especially in the nonwhite male and nonwhite female subgroups were based on fewer than 20 suicides and are subject to large sampling fluctuations.[3]

The category "White" includes those reported as Mexican and Puerto Rican, as well as those reported as "White." The category "Nonwhite" consists of persons reported as Negro, American Indian, Chinese, Japanese, and persons of mixed white-nonwhite races. Over 90% of this category is negro.

Average suicide rates per year were computed within each age group for each of the four race-sex subgroups. Unweighted averages were used because the typical suicide rate-age relationship was wanted, taking each year as equally important without regard to population size that year.

The indicated subgrouping by race and sex enabled an examination of the effect of each of these factors on the suicide rate-age relationship for each race-sex subgroup while holding the others constant.

Order of Analysis

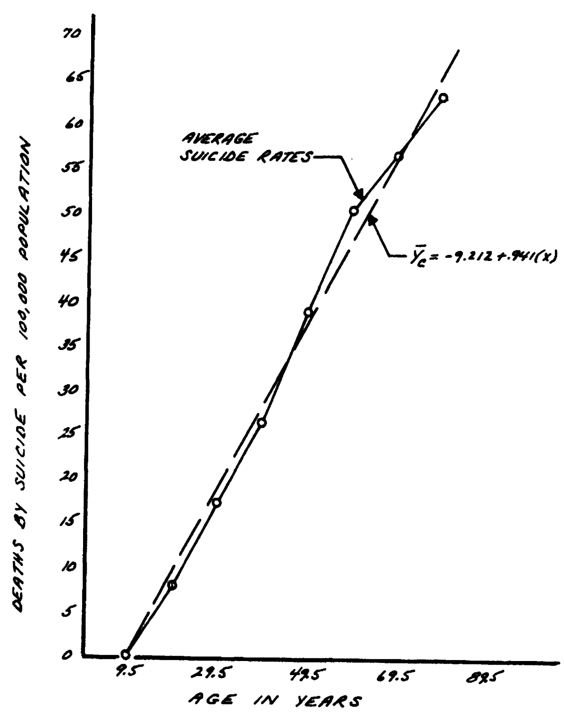
The suicide rate-age relationship for white males is examined first; second, the suicide rate-age relationship for white females; third, for nonwhite males and fourth, for nonwhite females.

White Male Suicide Rate-Age Relationship

Table I and Figure 1 show the mean suicide rates per year for white males classified by age. Table I and Figure 1 indicate that the mean white male suicide rates increase steadily with age.

Figure 1 indicates that the relationship of the average suicide rate to age for white males is closely approximated by linear regression. The slope (regression coefficient) of the fitted regression line for the average suicide rates to age relationship is plus 0.941.

FIGURE 1
AVERAGE SUICIDE RATES PER YEAR (1914-1964) OF WHITE MALES FOR U. S. POPULATION



A linear regression equation is also a good approximation for the individual years from 1914 to 1964 with the exception of a few years in which there is a slight leveling off in the two oldest age groups. The regression coefficients for individual years vary from a high in 1932 of 1.53 to a low in 1964 of 0.682

(Figure 2, Table 1). The nature of this changing slope is important. The changing slopes are brought about by changes in the suicide rates in the middle and older age groups, while the suicide rates in the younger age groups are less variable. The greatest variation of suicide rates appear from age 60 on with considerable variation between 30 and 60, but relatively little variation below 30 years. It is also apparent that when the regression coefficient increases the suicide rate increases in all age groups. The increase, however, is less in the younger age groups and greater in the middle and older groups respectively. The same is true when the regression coefficient decreases, there is a decline in the suicide rates in all age groups but the decline is greater in the middle and older groups. Thus, the fitted regression lines that represent the relationship of suicide rate to age for white males from 1914 to 1964 rotate about a relatively fixed point, the left end of the regression line.

The external factors that cause this change in slope of the regression line affect the suicide rates in all groups in the same direction. These effects are less, however, in the younger age groups than in the middle and older age groups.

Suicide rates decline during periods of increasing business activity and rise during periods of falling business activity. In prosperity, ambitions can be satisfied and upward mobility maintained. In depression economic strain is greater and suicide may become the only exit from an intolerable position.

It was indicated earlier that changes in the regression coefficient were primarily brought about by changes in the suicide rates in middle and older age groups. These very age groups that show the greatest variation in suicide rate are the age groups that are most affected by changes in business. The younger age groups are not so deeply involved in business activity. Their financial losses are less severe in a depression as are their expectations less in prosperity.

In the years from 1914 to 1921 the turning points of general business cycles as computed by the N.B.E.R. and the turning points of cycles in the time series of regression coefficients coincide reasonably well (Table 2, Figure 2).

TABLE 1

SUICIDE RATES OF WHITE MALES BY AGE FOR U.S. POPULATION, 1914-1964

Suicide Rate (suicides per 100,000 population)

Year	AGE - 5-14	15-24	25-34	35-44	45-54	55-64	65-74	75-84	Regression Coefficient
1914	0.3	12.6	27.0	35.3	52.1	64.1	60.6	59.4	.941
1915	0.3	11.7	26.2	35.8	50.6	68.2	66.8	63.1	1.030
1916	0.3	10.8	22.1	31.1	41.8	59.6	63.2	65.4	1.010
1917	-	-	-	-	-	-	-	-	-
1918	0.2	9.5	24.1	29.0	33.3	45.2	49.8	53.9	.777
1919	0.3	7.7	19.9	27.1	31.7	43.3	49.4	59.7	.839
1920	0.2	7.5	15.7	22.7	30.2	38.3	46.6	52.2	.761
1921	0.4	8.5	19.4	29.9	42.5	52.2	56.6	60.7	.929
1922	0.3	7.4	17.8	27.4	40.0	49.3	59.7	60.1	.947
1923	0.3	6.3	16.9	27.3	38.4	48.0	55.3	65.2	.967
1924	0.2	7.7	17.1	28.8	40.5	52.3	59.0	58.9	.944
1925	0.3	7.4	17.0	27.3	40.3	51.1	63.0	61.5	.989
1926	0.3	7.7	16.8	29.6	42.0	53.0	64.7	67.7	1.060
1927	0.3	8.3	18.4	30.0	45.2	56.8	66.3	68.5	1.080
1928	0.2	8.1	19.3	30.1	48.2	58.6	72.1	77.0	1.200
1929	0.3	8.9	19.2	31.8	46.7	58.1	72.6	76.8	1.180
1930	0.3	9.6	21.6	36.2	55.7	70.6	75.3	76.1	1.230
1931	0.3	9.2	21.8	36.4	61.2	81.4	88.6	80.1	1.390
1932	0.3	9.4	23.0	36.5	62.1	85.9	92.2	88.5	1.530
1933	0.2	9.1	20.4	33.2	56.3	74.7	84.1	88.2	1.420
1934	0.3	9.6	21.3	31.0	48.1	63.6	74.8	78.1	1.220
1935	0.3	9.5	20.7	29.0	47.0	58.5	63.6	75.8	1.120
1936	0.4	9.1	21.0	30.7	44.9	57.3	61.6	73.0	1.070
1937	0.5	9.5	21.5	32.2	48.5	59.2	61.6	73.8	1.090
1938	0.3	9.4	22.2	33.1	50.1	63.6	61.9	68.5	1.060
1939	0.5	8.5	19.8	29.3	44.8	58.2	60.2	70.4	1.060
1940	0.4	8.8	19.9	30.1	44.1	58.8	58.2	65.9	1.010
1941	0.3	7.8	18.3	26.4	35.8	49.2	55.2	65.7	.958
1942	0.4	7.2	17.4	25.3	33.0	46.8	49.9	61.1	.883
1943	0.4	6.5	13.5	19.9	26.2	37.1	44.0	60.2	.822
1944	0.4	6.2	14.0	20.1	25.8	33.8	40.0	57.4	.762
1945	0.4	7.7	16.7	23.3	29.0	38.6	44.7	58.4	.797
1946	0.5	7.7	14.7	23.4	32.4	42.8	45.7	56.4	.811
1947	0.5	6.8	13.8	23.3	32.6	45.5	48.3	67.6	.941
1948	0.4	6.8	13.3	21.8	33.3	44.7	48.8	58.1	.865
1949	0.4	6.9	13.3	24.0	33.6	46.6	51.2	59.6	.897
1950	0.3	6.6	13.8	22.4	34.1	45.9	53.2	61.9	.919
1951	0.3	6.8	12.8	20.5	30.3	39.7	50.3	53.8	.813
1952	0.2	6.9	12.6	19.6	29.8	38.4	47.2	56.7	.815
1953	0.4	7.0	13.0	20.5	31.0	38.4	47.5	55.9	.807
1954	0.2	6.9	13.5	19.9	33.0	41.6	46.2	52.6	.787
1955	0.2	6.2	12.7	19.8	31.7	42.8	45.6	54.7	.810
1956	0.3	6.4	12.9	19.1	30.2	41.5	46.4	58.1	.835
1957	0.4	6.5	12.7	20.1	30.3	38.1	43.5	55.3	.781
1958	0.4	7.6	14.3	22.1	34.0	41.9	45.8	58.3	.823
1959	0.5	7.9	14.4	21.5	32.8	41.8	45.0	56.6	.800
1960	0.5	8.6	14.9	21.9	33.7	40.2	42.0	55.7	.763
1961	0.4	7.9	14.7	22.3	32.8	39.7	39.9	53.2	.733
1962	0.5	8.7	16.5	22.8	33.4	40.5	41.3	57.0	.763
1963	0.5	9.2	16.9	23.6	33.0	40.2	40.3	52.5	.713
1964	0.5	9.3	17.0	22.2	31.8	38.5	39.1	50.5	.682
$\overline{Y}$ =	.346	8.158	17.556	26.534	38.998	50.484	55.778	63.316	
σ_y =	.096	1.391	3.734	5.205	9.073	11.936	12.720	9.246	

$$\overline{Y}_c = -9.212 + .941\ (x)$$

TABLE 2

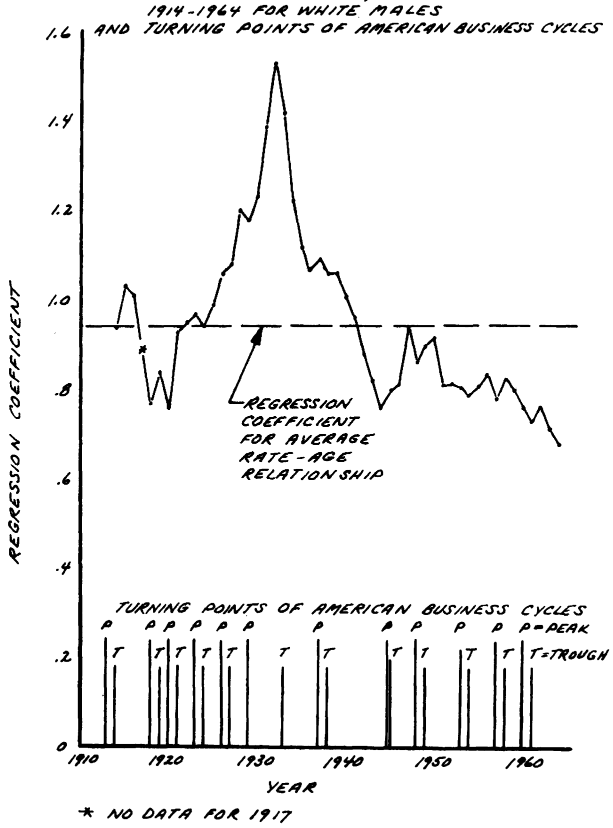
TURNING POINTS OF AMERICAN BUSINESS CYCLES AND OF TIME SERIES OF REGRESSION COEFFICIENTS FOR WHITE MALE SUICIDE RATE TO AGE RELATIONSHIP,

1914 to 1921

Business Cycles[4]		Time Series of Regression Coefficients	
1914	Trough	1915	Peak
1918	Peak	1918	Trough
1919	Trough	1919	Peak
1920	Peak	1920	Trough
1921	Trough	1921	Peak

In the period from 1921 to 1929 the general level of business activity and the regression coefficients increased. In this period of heightened economic activity, the regression coefficients increased rapidly. Why did the regression coefficients rise during this prosperous period? A closer look at the nature of the boom shows that it was not a unified rise in prosperity for all

FIGURE 2
TIME SERIES OF ANNUAL REGRESSION COEFFICIENTS OF SUICIDE RATES AND AGE, U.S. POPULATION 1914-1964 FOR WHITE MALES AND TURNING POINTS OF AMERICAN BUSINESS CYCLES



segments of the economy. For example, there were many bank failures in this period. Nearly 1,000 banks failed in 1926 alone and over 600 in every year from 1923 to 1929. Business failures also increased considerably.[5] Agriculture did not share in the boom to the same degree as did the rest of the economy. There was a heavy increase in farm mortgage debt and foreclosures of farm property were high throughout the twenties. Farm prices and income collapsed in 1920-21 when European agriculture output returned to its prewar level. By 1925 it had made a satisfactory recovery but there was little expansion after 1925.[6]

The fact that prosperity was not shared by all segments of the economy during the twenties is perhaps the most important factor causing the regression coefficients to increase. While some individuals were enjoying the unprecedented boom, others were losing their savings, business, and farms. Again it was the middle and older age groups that were affected most severely. Suicide rates increased in these age groups causing the regression coefficients to increase. After the collapse of business in 1929 regression coefficients increased to a peak of 1.53 in 1932. Business activity reached its trough in 1932 and slowly started to improve again.

After 1932 regression coefficients declined and business activity slowly began to increase. Perhaps an important reason for the decline in regression coefficients was the result of programs undertaken by the federal and state governments. Such programs as social security, unemployment compensation, F.D.I.C., and public works projects greatly reduced the economic strains and especially the feeling of uncertainty present in the middle and older age groups.

The regression coefficients reached a low point, .762 in 1944, and began to rise in 1945 as the war ended and business was faced with the problem of reconversion back to peace time production. The variations in regression coefficients during the post war years and the turning points in general business activity do not coincide nearly as well as in the period from 1914 to 1921 (Table 3). There is evidence of a relationship between the two series,however, and the continued prosperity is reflected by a decline in regression coefficients especially since 1950.

In conclusion, it has been shown that the functional relationship of suicide rate to age for white males varies considerably between 1914 to 1964 and this variation is closely associated with the level of business activity. This is not to imply that changes in this relationship are completely explained by changes in the level of business activity. There are undoubtedly a multiplicity of factors which affect the relationship. Nevertheless, the association between the suicide rate-age relationship for white males and peaks and troughs of general business activity is quite remarkable.

TABLE 3

TURNING POINTS OF AMERICAN BUSINESS CYCLES AND OF TIME SERIES OF REGRESSION COEFFICIENTS FOR WHITE MALE SUICIDE RATE TO AGE RELATIONSHIP, 1945 to 1964

Business Cycles[7]		Time Series of Regression Coefficients	
1945	Peak	1944	Trough
1945	Trough	1947	Peak
1948	Peak	1948	Trough
1949	Trough	1950	Peak
1953	Peak	1954	Trough
1954	Trough	1956	Peak
1957	Peak	1957	Trough
1958	Trough	1958	Peak
1960	Peak	1961	Trough
1961	Trough		

White Female Suicide Rate-Age Relationship

Table 4 and Figure 3 shows the average suicide rates per year for white females classified by age. White female suicide rates increase steadily as age increases up to about age 45 to 54 then decline in the last two age groups (Figure 3). A simple linear regression equation $(Y = A + bX)$ was used to approximate the functional relationship of suicide rate (Y) to age (X). Although the actual relationship is curvilinear, the linear approximation used is accurate enough for comparing the changes in the relationship from 1914 to 1964. The slope of the fitted regression line approximating the average suicide rates to age relationship is .130.

White female suicide rates increase steadily as age increases in the younger age groups, level off in the middle age groups, and decline in the older age groups in every year from 1914 to 1964. The ages when the highest rates appear vary between 45-64. It is interesting to note that the period of menopause (usually between ages 45-50) occurs at the time when suicide rates are highest. This period of menopause, which is a time of nervous tension and frustration for women, may be the primary factor causing white female suicide rates to be highest during the period.

The decline of suicide rates in the last two age groups suggests that the white female is better able to adjust to old age than is her male counterpart. This is perhaps because old age for white females is not associated with the same loss of prestige or integration in society.

Retirement does not demand as severe an adjustment for white females as it does for white males because the pattern of life for the female is not so sharply changed as for the male after

TABLE 4

SUICIDE RATES OF WHITE FEMALES BY AGE FOR U.S. POPULATION, 1914-1964

Year	Age - 5-14	15-24	25-34	35-44	45-54	55-64	65-74	75-84	Regression Coefficient
	Suicide Rate (suicides per 100,000 population)								
1914	.2	9.3	10.3	11.5	12.6	15.5	11.7	9.5	.112
1915	.3	8.1	10.4	11.6	13.3	14.0	14.0	10.8	.142
1916	.2	7.2	9.3	9.7	12.4	12.9	11.5	8.5	.111
1917	-	-	-	-	-	-	-	-	-
1918	.3	5.8	9.1	10.1	11.5	13.4	11.2	10.9	.138
1919	.2	6.1	8.5	10.5	12.9	13.4	11.9	10.6	.142
1920	.2	5.5	8.2	10.0	11.7	11.1	9.9	8.6	.109
1921	.2	5.7	8.0	9.2	11.5	11.8	10.2	9.3	.119
1922	.1	6.0	7.6	9.6	11.4	11.9	10.7	10.7	.134
1923	.2	5.7	8.1	9.4	11.3	12.9	9.5	10.9	.132
1924	.2	5.4	7.6	9.5	11.1	11.9	10.1	8.0	.111
1925	.2	5.0	8.4	9.9	12.3	11.3	10.7	6.9	.103
1926	.2	6.1	8.6	10.2	12.6	12.3	10.7	10.3	.128
1927	.2	5.3	8.8	10.6	12.4	12.4	10.9	9.4	.125
1928	.1	6.1	8.8	10.3	12.5	11.9	11.8	8.2	.116
1929	.1	6.0	9.3	10.8	12.9	14.1	12.3	10.7	.146
1930	.1	6.4	9.8	11.4	13.2	14.5	12.7	10.0	.140
1931	.1	5.9	9.9	11.8	14.2	15.2	12.3	10.5	.148
1932	.1	5.9	9.9	11.1	13.9	15.0	14.7	9.0	.149
1933	.1	5.9	9.0	11.3	12.9	14.6	12.1	9.1	.134
1934	.1	5.7	9.2	11.1	13.8	13.3	10.9	9.9	.131
1935	.2	5.4	9.4	10.9	13.0	13.8	11.9	7.9	.122
1936	.2	5.0	9.8	10.9	13.3	13.4	11.6	9.0	.129
1937	.1	5.1	9.0	11.6	14.6	14.4	11.5	9.8	.143
1938	0	4.6	8.5	11.7	14.3	14.0	11.4	9.0	.139
1939	.1	3.9	7.6	11.0	14.3	13.7	12.4	7.5	.139
1940	.1	3.9	8.6	11.5	14.0	13.1	12.9	9.0	.147
1941	.2	4.1	8.1	10.3	12.8	12.5	10.7	8.3	.126
1942	.1	3.2	7.1	9.3	11.8	12.0	9.7	12.2	.161
1943	.2	3.2	6.1	8.4	10.8	12.4	11.2	9.1	.148
1944	.1	3.3	6.1	9.0	10.5	11.5	9.6	10.4	.146
1945	.1	3.0	6.8	9.5	12.0	11.4	10.9	9.9	.149
1946	.1	3.4	6.3	9.4	11.7	11.9	10.3	8.7	.137
1947	.1	2.7	6.0	8.9	11.8	10.9	11.5	8.8	.147
1948	.1	3.1	5.5	9.0	10.7	10.6	9.4	9.7	.139
1949	.1	2.6	5.1	8.2	10.4	11.5	10.0	7.6	.133
1950	.1	2.7	5.2	8.2	10.5	10.7	10.6	8.4	.139
1951	.1	2.3	5.0	7.7	9.8	9.2	9.1	8.3	.126
1952	.1	2.1	5.0	6.9	9.4	9.2	9.0	7.0	.117
1953	.1	2.3	4.8	6.5	9.0	9.3	8.3	7.5	.116
1954	.1	1.8	4.4	6.8	8.2	9.5	8.7	6.3	.113
1955	.1	2.0	4.9	6.6	10.3	10.4	9.4	8.3	.136
1956	.1	2.0	4.7	6.7	9.6	10.5	9.4	6.8	.124
1957	.1	1.8	5.0	7.1	8.8	10.1	8.3	6.2	.110
1958	.1	2.4	5.9	7.1	10.4	10.0	9.6	6.6	.116
1959	0	2.1	5.7	7.5	9.5	10.6	9.8	7.6	.129
1960	.1	2.3	5.8	8.1	10.9	10.9	8.8	9.2	.136
1961	.1	2.3	6.1	8.3	10.8	10.4	9.1	7.6	.121
1962	.1	2.9	7.3	9.5	12.4	11.1	8.8	7.6	.115
1963	.1	3.1	7.5	10.9	12.9	11.6	9.5	7.7	.119
1964	.1	2.9	7.3	10.9	12.5	10.9	10.4	6.7	.114
$\overline{Y}$ =	.132	4.292	7.468	9.560	11.828	12.098	10.672	8.810	
σ_y =	.061	1.809	1.748	1.564	1.556	1.649	1.411	1.387	

$$\overline{Y}_c = 2.343 + .130\ (x)$$

FIGURE 3
AVERAGE SUICIDE RATES PER YEAR
(1914-1964) OF WHITE FEMALES FOR
U. S. POPULATION

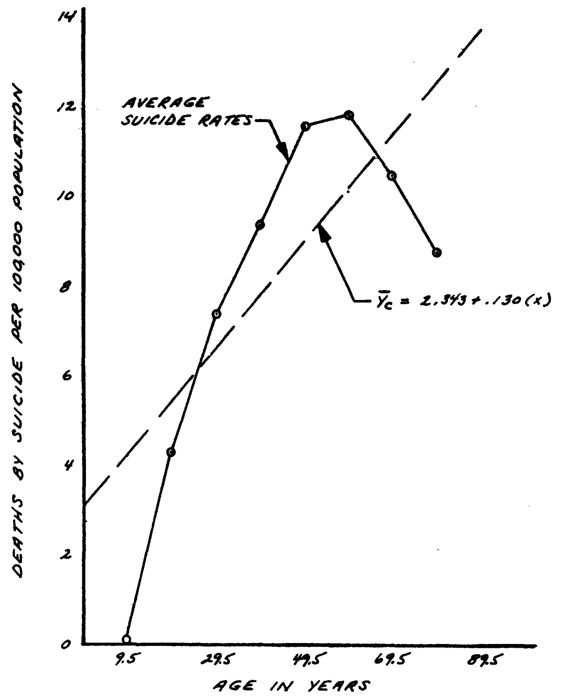
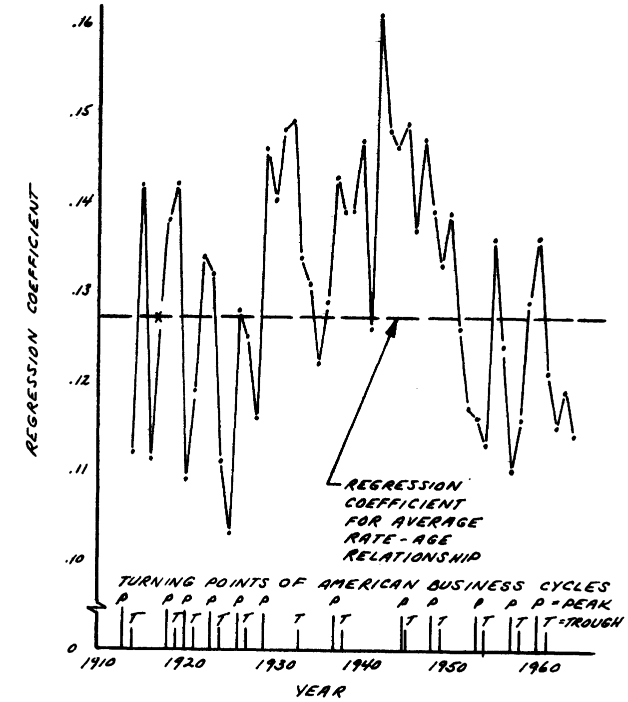


FIGURE 4
TIME SERIES OF ANNUAL REGRESSION COEFFICIENTS
OF SUICIDE RATES AND AGE, U.S. POPULATION
1914-1964 FOR WHITE FEMALES
AND TURNING POINTS OF AMERICAN BUSINESS CYCLES



* NO DATA FOR 1917

about age 65. The white female's role and functions in the family remain relatively unchanged as compared with the period before age 65.

White female suicide rates are influenced somewhat by the general level of business activity. They tend to rise during periods of depression and fall during periods of prosperity. The slopes of the regression lines representing the relationship of suicide rate to age for individual years vary only from .103 in 1925 to .161 in 1942 and there is no close association between the fluctuations in slopes and the fluctuations in the level of general business activity (Figure 4). This low degree of variability in regression coefficients is partly due to using a linear approximation when the actual relationship is curvilinear. The most important reason, however, for the low degree of variability in regression coefficients is because of the relative stability of suicide rates of white females.

In conclusion, the functional relationship of suicide rate to age for white females is not linear, but linear regression lines have been used as approximations to the true relationships. There is little change in the regression coefficients of these lines during the period studied.

Nonwhite Male Suicide Rate-Age Relationship

The average suicide rates per year for nonwhite males increase steadily as age increases in the first three age groups (5-14, 15-24, 25-34) Table 5 Figure 5). After 35 years of age the average rates increase at a lower rate as age increases and in the last three age groups (55-64, 65-74, 75-84) the rates are almost constant (14.2, 13.9, 14.0). A straight line regression equation was used as an approximation to the true relationship of average suicide rates to age. The slope of this fitted regression line is .174. The relationship of suicide rate to age varies for individual years. Suicide rates declined in the last age group (75-84) in 24 of the years studied and increased in 22 of the years studied. There is no apparent pattern to these occurrences nor do they coincide with periods of prosperity or depression. The differences observed are probably due to the fact that some of the suicide rates, especially in the last age group, were based on fewer than 20 suicides.

A tentative explanation for the reduced rate of increase of suicide rates in the middle age groups and lack of any definite directional movement of the rates for the last three age groups is that vertical socio-economic mobility for many nonwhite males exists only within the bounds of their race. This limit on upward mobility results in ambitions aimed at more modest positions and therefore suicide rates do not increase steadily as age increases because disappointments of not achieving desired goals are not as severe. Another explanation may be that frustrations associated with an unattained goal may be directed outward toward the dominant class rather than inward resulting in suicide.

FIGURE 5
AVERAGE SUICIDE RATES PER YEAR (1914-1964) OF NON-WHITE MALES FOR U.S. POPULATION

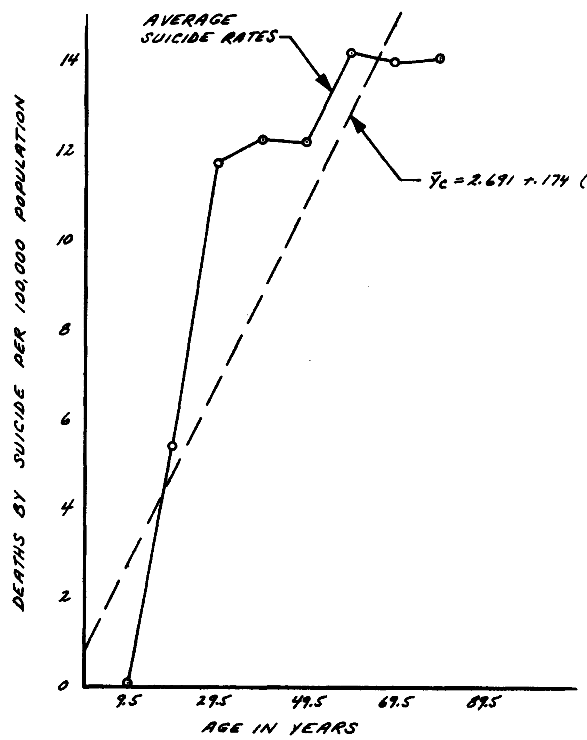
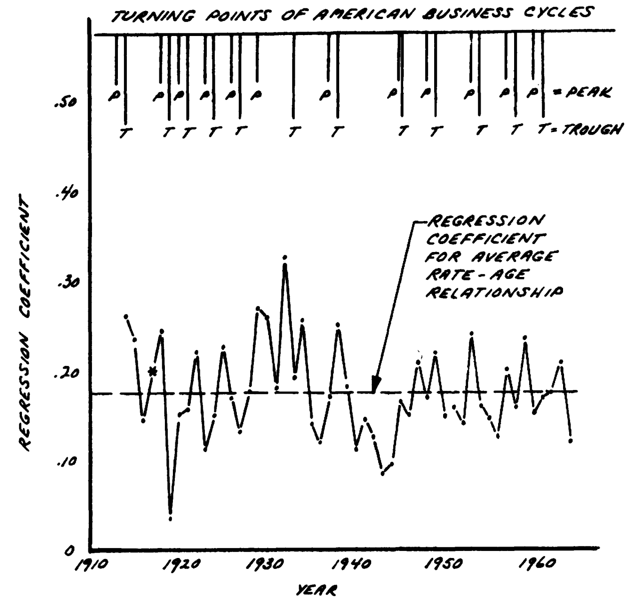


FIGURE 6
TIME SERIES OF ANNUAL REGRESSION COEFFICIENTS OF SUICIDE RATES AND AGE, U.S. POPULATION 1914-1964 FOR NON-WHITE MALES AND TURNING POINTS OF AMERICAN BUSINESS CYCLES



* NO DATA FOR 1917

The regression coefficients of the fitted regression lines for individual years vary from .036 in 1919 to .325 in 1932 (Table 5). It is difficult to associate changes in suicide rates for nonwhite males with changes in the level of general business activity. In periods of increasing or decreasing business activity no clear association can be made with the suicide rates in all age groups. Some of this lack of sensitivity to the level of general business activity can be attributed to suicide rates in many age groups being based on a small number of suicides. The conclusion,however, that nonwhite male suicide rates are only moderately sensitive to the level of general business activity seems reasonable. Perhaps the most important reason for this low degree of sensitivity of nonwhite male suicide rates to business activity is due to their minor role in the family. In many cases the nonwhite male is not the head of the family and often is only a transient visitor. He assumes no obligation to support the family in many cases, nor is he even an integrated member of the family. Therefore he is not severely affected by changes in the level of general business activity.

The time series of regression coefficients for nonwhite males show no close association with the level of general business activity (Figure 6) the fluctuations observed are perhaps the result of suicide rates being based on small numbers of suicides in the older age groups.

To summarize, suicide rates for nonwhite males do not increase steadily as age increases but level off after about 35 years of age. Suicide rates of nonwhite males are only moderately sensitive to changes in the level of general business activity and the fluctuations in the time series of regression coefficients show no close association with the level of general business activity.

Nonwhite Female Suicide Rate-Age Relationship

The average suicide rates per year for nonwhite females increase with age in the first three age groups (5-14, 15-24, 25-34) but then decrease as age increases in the next three age groups (35-44, 45-54, 55-64), and lack any definite directional movement in the last two age groups (Figure 7 and Table 6). The functional relationship of average suicide rates to age for nonwhite females is approximated by a linear regression line. The slope of the line which describes the average suicide rates-age relationship is almost 0,(.008). This is important for it suggests that age has no effect on the average suicide rate for nonwhite females.

There is considerable variation of suicide rates in individual years from the typical pattern. These variations are probably due to suicide rates in the last four age groups being based on less than 20 suicides, such rates being subject to

TABLE 5

SUICIDE RATES OF NON-WHITE MALES BY AGE FOR U.S. POPULATION, 1914-1964

Suicide Rate (suicides per 100,000 population)

Year	AGE - 5-14	15-24	25-34	35-44	45-54	55-64	65-74	75-84	Regression Coefficient
1914	0.4	8.4	26.9	23.6	13.9	29.7	26.3	19.1	.262
1915	0.0	7.2	24.0	20.2	18.7	19.5	23.1	18.6	.235
1916	0.0	4.8	15.5	15.4	12.6	11.2	14.9	12.2	.144
1917	-	-	-	-	-	-	-	-	-
1918	0.2	6.7	14.8	15.7	8.6	16.1	23.8	17.6	.244
1919	0.0	5.2	11.8	10.9	8.8	9.8	8.0	3.6	.036
1920	0.2	3.7	8.5	8.3	10.7	10.0	12.6	10.7	.150
1921	0.1	6.3	10.8	12.8	10.3	10.3	19.0	10.4	.157
1922	0.1	4.6	9.2	14.1	10.7	9.9	13.2	20.5	.221
1923	0.0	3.5	7.5	11.9	13.2	10.4	12.0	5.7	.111
1924	0.2	4.1	9.6	13.6	10.2	13.8	11.8	11.1	.148
1925	0.2	4.8	9.7	12.0	8.2	9.1	16.6	19.3	.224
1926	0.2	3.8	9.7	10.3	12.1	14.0	14.5	10.8	.171
1927	0.2	4.3	12.8	12.8	12.4	12.6	16.5	7.6	.133
1928	0.0	5.2	10.8	14.0	14.2	15.8	9.9	15.6	.177
1929	0.2	5.7	11.6	11.9	16.6	13.8	17.5	22.3	.270
1930	0.1	4.4	11.9	13.4	14.3	16.5	23.0	15.6	.259
1931	0.1	4.6	12.7	14.2	14.6	20.0	14.9	11.0	.180
1932	0.2	5.9	11.9	16.1	16.8	24.9	16.3	25.8	.325
1933	0.1	5.9	11.2	16.3	17.5	13.8	17.4	13.7	.193
1934	0.1	6.3	11.9	11.3	13.8	16.7	10.8	25.0	.256
1935	0.2	5.1	10.7	13.0	14.8	17.8	10.8	9.4	.139
1936	0.4	5.5	9.2	13.2	10.6	17.8	11.2	7.3	.119
1937	0.4	4.4	10.7	13.7	13.2	15.9	13.0	12.5	.171
1938	0.3	5.5	9.8	12.7	14.0	17.2	9.0	24.4	.251
1939	0.0	3.9	10.1	11.8	13.1	12.0	9.3	16.9	.182
1940	0.4	5.1	11.5	10.6	14.8	12.6	13.5	6.5	.110
1941	0.1	5.1	10.3	10.5	12.5	11.7	11.3	12.1	.145
1942	0.2	5.0	9.3	9.7	8.0	14.5	9.2	10.1	.125
1943	0.4	4.9	7.8	6.2	6.9	10.3	6.5	8.2	.084
1944	0.1	3.9	7.6	7.2	8.0	11.1	6.7	7.7	.094
1945	0.0	5.1	9.1	8.8	9.1	9.1	12.2	14.6	.165
1946	0.2	5.0	9.4	8.8	9.9	12.8	10.7	12.6	.152
1947	0.3	5.0	8.9	11.1	9.1	13.8	14.3	16.5	.208
1948	0.0	5.0	9.1	12.2	14.1	10.9	14.5	12.6	.171
1949	0.1	4.6	9.9	11.3	12.0	16.4	15.8	15.3	.219
1950	0.1	5.3	10.1	11.3	11.7	16.8	15.0	7.9	.147
1951	0.3	5.2	10.7	10.8	11.4	12.0	10.1	15.1	.158
1952	0.3	4.5	10.7	8.9	9.9	14.3	11.0	10.9	.141
1953	0.1	4.2	10.1	10.3	12.6	12.0	12.2	21.9	.239
1954	0.1	5.5	12.8	10.6	11.4	12.2	13.6	13.4	.158
1955	0.1	6.5	9.6	9.6	10.5	12.7	10.8	12.8	.144
1956	0.1	5.8	10.8	10.5	9.8	12.2	10.0	11.5	.124
1957	0.0	5.6	13.0	10.6	11.8	11.9	17.4	15.6	.198
1958	0.2	5.4	12.6	12.3	12.8	15.3	14.2	11.5	.157
1959	0.2	6.6	14.4	11.0	13.9	14.8	14.9	22.9	.244
1960	0.1	5.3	12.9	13.5	12.8	16.9	12.6	11.3	.150
1961	0.1	7.6	16.3	11.5	14.0	15.0	13.2	16.6	.169
1962	0.1	7.5	12.8	12.8	12.4	14.6	16.8	13.6	.174
1963	0.3	7.5	15.9	14.9	13.6	12.7	18.2	19.0	.207
1964	0.1	8.0	16.2	12.8	11.8	12.3	13.4	12.5	.120
$\overline{Y}$ =	0.158	5.380	11.702	12.220	12.174	14.150	13.870	13.988	
σ_y =	0.119	1.133	3.539	2.935	2.537	3.808	4.180	5.159	

$$\overline{Y}_c = 2.691 + .174\ (x)$$

large sampling fluctuations.[8]

Suicide rates for nonwhite females tend to rise during periods of depression and fall during periods of prosperity, but due to the large fluctuations in rates it is difficult to observe this relationship.

One can only postulate reasons why nonwhite female suicide rates decrease as age increases after about 35. An important factor seems to be the important role of the nonwhite female in the family. The following pattern appears to be quite typical in nonwhite family life. Due to a high illegitimacy rate and large number of working mothers among nonwhites, a female becomes the head of the family. The mother must work to support her family, leaving the grandmother to take care of the children and the grandmother becomes the power-holder in the family. She raises the children and takes care of the home. It is the grandmother's probability of suicide that is reduced because of her important role in the family. Her interests and time are taken up by her many duties and she performs a necessary function in the family. This unique role of the nonwhite grandmother reduces her probability of suicide. She is supported by the other members of the family and her function of raising the children gives her a feeling of satisfaction and usefulness. It is important to note that this role of the grandmother is not reduced in cases where the male member of the family, or father, lives at home. In most cases the wife still works leaving the children with the grandmother. It is believed that this unique family relationship had its origin during the period of negro slavery in the south.

Figure 7 indicates that there is no tendency for nonwhite female suicide rates to increase during the period of menopause. The important role of the nonwhite female in the family during this period is perhaps often sufficient to counteract the feeling of nervous tension and frustration associated with menopause.

The slopes of the regression lines fitted to the annual data vary from -.037 in 1914 to .071 in 1936 (Table 6). There is no close association between the time series of regression coefficients and the level of general business activity (Figure 8). The fluctuations in regression coefficients are primarily the result of suicide rates in the last four age groups being based on small numbers of suicides.

In conclusion, it has been observed that nonwhite female suicide rates typically increase only in the first three age groups, and after 35 the rates decrease as age increases. No close association between the time series of regression coefficients and level of business activity was observed nor were suicide rates very sensitive to the level of general business activity.

Conclusions

Figure 9 indicates that the mean white male suicide rates are considerably higher in each age group than the mean rates for the other three race-sex subgroups. This pattern is true for every year from 1914 to 1964, annual white male suicide rates are higher in each age group than the corresponding rates in the other three race-sex subgroups for every year from 1914 to 1964.

The question arises: Why is suicide more prevalent among white males than among the other three race-sex subgroups in this study? Emile Durkheim attributes the high suicide rates of white males to their dominant place in our society.[9] Durkheim found that dominant classes are more prone to suicide than subordinate classes. The dominant class possesses a higher mobility aspiration than subordinate classes. The subordinate status reduces the probability of suicide because of the restraint enforced by the dominant class.

> ....the horizon (mobility aspiration) of the lower classes is limited by those above them, and for the same reason their desires are more modest. Those who have only empty space above them (unlimited mobility aspirations) are almost invariably lost in it, if no force restrains them.[10]

This freedom of upward mobility of white males due to their dominant position in our society places great strains on them. It allows them to desire positions far beyond their capabilities. If they fail to achieve these goals, they may be unable to accept defeat. This important theoretical view is best explained by the following quotation from *The Sickness Unto Death* by Soren Kierkegaard.

FIGURE 7
AVERAGE SUICIDE RATES PER YEAR (1914-1964) OF NON-WHITE FEMALES FOR U.S. POPULATION

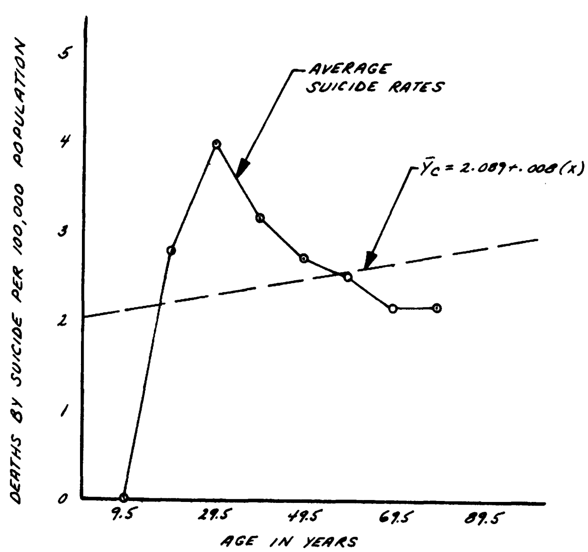


TABLE 6

SUICIDE RATES OF NON-WHITE FEMALES BY AGE FOR U.S. POPULATION, 1914-1964

Suicide Rate (suicides per 100,000 population)

	AGE - 5-14	15-24	25-34	35-44	45-54	55-64	65-74	75-84	Regression Coefficient
1914	0.0	7.8	8.9	5.6	2.8	1.8	6.7	0.0	-.037
1915	0.0	5.4	10.4	8.2	5.5	5.3	0.0	8.8	.019
1916	0.0	3.8	3.6	2.2	3.1	1.2	2.2	0.0	-.017
1917	-	-	-	-	-	-	-	-	-
1918	0.0	3.8	4.1	3.4	2.7	2.7	8.1	4.3	.056
1919	0.0	2.5	4.6	3.2	1.8	0.8	2.7	0.0	-.015
1920	0.1	3.7	2.9	2.3	3.3	1.5	1.4	3.6	.012
1921	0.0	3.7	4.9	2.9	1.8	2.1	1.3	0.0	-.026
1922	0.2	3.1	3.3	3.4	3.6	2.4	0.0	0.0	-.024
1923	0.0	2.5	3.6	3.6	1.5	1.1	0.0	2.9	-.002
1924	0.1	3.2	4.9	3.0	2.9	2.2	1.1	0.0	-.023
1925	0.0	2.3	3.9	3.2	3.4	1.6	2.2	0.0	-.009
1926	0.2	2.8	5.1	3.9	4.3	3.1	1.1	5.3	.026
1927	0.1	2.8	4.8	4.3	3.2	2.4	0.0	0.0	-.028
1928	0.2	2.9	5.3	3.2	2.3	3.4	5.4	2.2	.024
1929	0.0	3.3	4.2	3.6	2.6	3.4	1.7	2.2	.004
1930	0.1	3.6	3.8	3.7	2.9	2.9	3.4	2.1	.011
1931	0.0	3.5	5.2	3.0	2.1	4.5	2.4	4.2	.025
1932	0.2	3.7	4.5	4.7	2.6	4.0	1.5	0.0	-.020
1933	0.1	4.0	3.9	3.5	3.4	2.9	1.3	1.9	-.006
1934	0.1	4.3	5.1	4.2	3.2	2.8	2.5	3.6	.008
1935	0.1	3.6	6.3	4.0	1.2	2.0	1.2	5.3	.010
1936	0.3	3.0	4.9	2.7	2.8	2.3	2.2	10.4	.071
1937	0.1	3.4	4.7	3.0	3.8	2.9	1.6	1.7	.004
1938	0.1	2.9	5.3	3.1	2.9	4.1	3.0	0.0	-.005
1939	0.1	2.7	3.2	2.4	3.7	1.8	2.3	4.8	.033
1940	0.0	3.3	3.5	3.0	3.2	1.4	2.5	1.5	.000
1941	0.1	2.5	2.6	1.8	2.9	2.3	1.9	2.9	.020
1942	0.1	2.8	4.0	2.9	2.1	1.3	0.9	0.0	-.023
1943	0.1	1.6	1.7	2.7	1.4	1.9	0.9	3.9	.027
1944	0.1	2.0	2.1	2.0	1.8	1.2	2.5	2.5	.018
1945	0.0	2.2	2.8	2.3	1.5	1.1	0.4	2.4	.002
1946	0.0	2.3	2.8	2.8	3.4	2.2	1.6	0.0	-.006
1947	0.1	2.6	3.1	2.5	1.7	1.5	0.8	0.0	-.019
1948	0.1	1.8	3.1	2.4	2.2	0.8	1.5	0.0	-.011
1949	0.2	1.7	2.5	2.6	2.1	2.2	1.8	2.0	.014
1950	0.1	1.7	2.8	2.2	4.0	1.2	2.5	2.9	.025
1951	-	2.4	2.6	2.7	2.3	2.9	2.3	1.8	.015
1952	0.1	1.3	2.7	2.5	1.6	2.2	0.3	0.9	-.002
1953	0.2	2.0	2.1	1.6	1.8	1.9	2.2	0.8	.006
1954	-	1.5	2.8	2.0	2.8	2.9	2.6	0.8	.015
1955	-	1.8	2.8	1.7	2.8	2.6	2.5	1.5	.017
1956	0.1	1.2	2.6	3.0	3.0	3.3	2.7	1.4	.022
1957	-	1.4	2.5	2.3	1.3	4.0	1.9	2.1	.025
1958	0.0	1.7	3.6	2.8	3.1	2.9	2.4	3.3	.030
1959	0.1	2.2	3.7	2.7	2.9	4.3	3.0	1.9	.022
1960	0.0	1.5	3.5	3.7	3.2	3.4	3.8	4.2	.048
1961	0.2	2.0	3.5	3.9	2.7	3.2	1.9	1.6	.009
1962	0.0	3.0	4.6	3.9	3.3	3.9	2.0	3.2	.018
1963	0.2	2.6	5.0	4.0	2.7	2.3	5.0	1.5	.014
1964	0.0	2.0	4.7	4.1	3.4	4.0	3.2	2.8	.027
$\overline{Y}$ =	0.078	2.788	3.982	3.168	2.732	2.522	2.168	2.194	
σ_y =	0.011	.160	.220	.152	.119	.147	.219	.310	

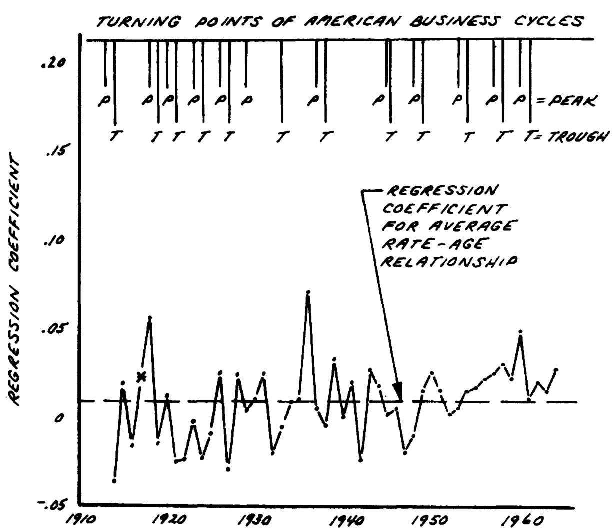
$$\overline{Y}_c = 2.089 + .008\ (x)$$

> Thus when the ambitious man whose watchword was "Either Caesar or nothing" does not become Caesar, he is in despair thereat. But this signifies something else, namely, that precisely because he did not become Caesar he now cannot endure to be himself. So properly he is not in despair over the fact that he did not become Caesar, but he is in despair over himself for the fact that he did not become Caesar.[11]

The four race-sex subgroups show three similar relationships of suicide rate to age. The slope of the regression line representing the functional relationship of average suicide rates to age for white males is plus .941, indicating on the average almost an increase of 1. suicide per 100,000 population for every 1 year increase in age. The slopes of the regression lines representing the functional relationship of average suicide rates to age for nonwhite males and white females are similar, .174 and .130, respectively. Thus suicide rates increase at a much lower rate as age increases for white females and nonwhite males than for white males. The slope of the regression line representing the functional relationship of average suicide rate to age for nonwhite females is .0087. This indicates that using this regression line we would predict almost no increase in rate as age increases.

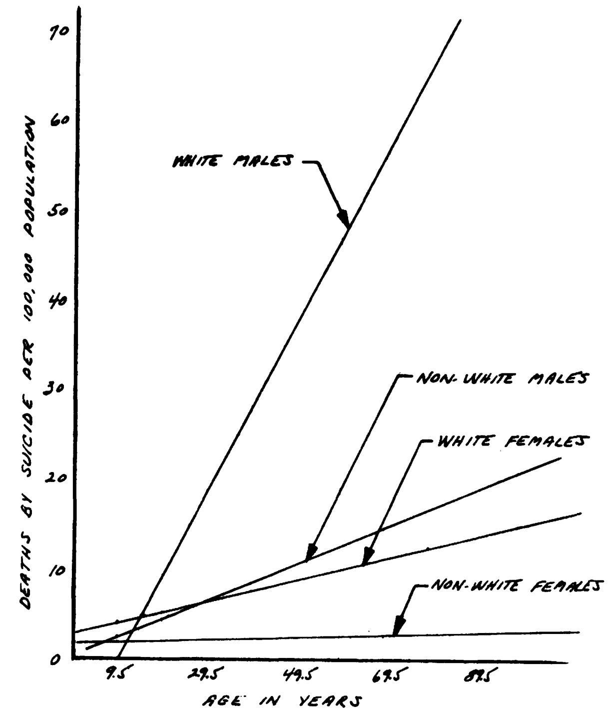
It was also seen that the suicide rates of white males were extremely sensitive to the level of general business activity and that the time series of regression coefficients were closely associated with the level of general business activity. This extreme sensitivity and close association with the level of general business activity was not observed for the other three race-sex subgroups.

FIGURE 8
TIME SERIES OF ANNUAL REGRESSION COEFFICIENTS OF SUICIDE RATES AND AGE, U.S. POPULATION 1914-1964 FOR NON-WHITE FEMALES AND TURNING POINTS OF AMERICAN BUSINESS CYCLES



* NO DATA FOR 1917

FIGURE 9
AVERAGE SUICIDE RATE-AGE RELATIONSHIP FOR WHITE MALES, WHITE FEMALES, NON-WHITE MALES AND NON-WHITE FEMALES



[1] This age span is broken down into eight age groups, 5 to and including 14, 15 to and including 24, etc. Assuming ages were recorded as of nearest birthday the mid-points were taken to be 9.5, 19.5, 29.5, etc.

[2] *Suicide*, U. S. Department of Health, Education and Welfare, Vital Statistics - Special Reports, Vol. 43, No. 30, August 22, 1956 (Washington, 25, D. C.), p. 464.

[3] *Ibid*

[4] R. A. Gordon, *Business Fluctuations*, New York: Harper and Brothers Publishers, 1952, p. 216.

[5] C. R. Whittlesey, *Principles and Practices of Money and Banking*, New York: The MacMillan Co., 1954, pp. 359-61.

[6] Gordon, *op. cit.*, p. 379.

[7] Business cycle peaks and troughs from 1945-1948, R. A. Gordon, *op cit.*, p. 216. Peaks and troughs from 1948-1961,

The Postwar Cycles, William B. Franklin, National Industrial Conference Board, Inc. New York 22, N.Y.

[8] *Suicide*, U. S. Department of Health, Education and Welfare, Vital Statistics - Special Reports, *op. cit.*, p. 476.

[9] Emile Durkheim, *La Suicide*, translated by J. A. Spaulding and G. Simpson, Glencoe, Ill: *The Free Press*, 1951, pp. 246-58.

[10] Durkheim, *op. cit.*, p. 257.

[11] Soren Kierkegaard, *The Sickness Unto Death*, Garden City, N.Y.: Doubleday Anchor Books, 1954, p.152.

References:

Durkheim, Emile, *La Suicide*, Translated by J. A. Spaulding and G. Simpson, Glencoe, Ill: The Free Press, 1951.

Gordon, R. A., *Business Fluctuations*, New York: Harper and Brothers Publishers, 1952.

Henry, A. F. and J. F. Short, Jr., *Suicide and Homicide*, Glencoe, Ill: The Free Press, 1954.

Kierkegaard, Soren, *The Sickness Unto Death*, Garden City, N. Y.: Doubleday Anchor Books, 1954.

U. S. Department of Health, Education and Welfare, *Suicide*, Vital Statistics - Special Reports, Vol. 43, No. 30, August 22, 1954, Washington 25, D.C.

U.S. Department of Health, Education and Welfare, *Suicide in the United States 1950-1964*, National Center for Health Statistics Series 20 Number 5, August 1967, Washington, D. C.

Whittlesey, C. R., *Principles and Practices of Money and Banking*, New York: The MacMillan Co., 1954.

SOME GUIDES TO INTERPRETATION OF THE FIGURES ON SCHOOL ENROLLMENT AMONG AMERICANS OVERSEAS IN THE 1960 CENSUS OF POPULATION*

Harley M. Upchurch

Human Resources Research Office, The George Washington University

It is generally known that, since the end of World War II, increasingly large numbers of Americans have been sojourning in other lands. Until very recently, however, even such basic information about this "overseas American population" as its size and geographical distribution has been hard to come by. Reasonably complete and reliable data on its other characteristics have been even more difficult to assemble if--in fact--they could be had at all. Then, in 1964, the Bureau of the Census published a special report entitled Americans Overseas. To quote this source:

> Selected groups of Americans living overseas have been counted in the decennial censuses since 1900. Because of the increased number of Americans now overseas, the 1960 enumeration was extended to cover all types of Americans residing abroad, particularly those designated here as "other citizens," i.e. civilians other than Federal employees, their dependents, or crews of merchant vessels...
>
> Furthermore, the content of the questionnaire was expanded in order to obtain additional information on the demographic and economic characteristics of Americans abroad.[1]

It seems likely that as the existence of these data becomes more commonly known they will be put to a wide variety of uses. If they are collected on a continuing basis they may well come to represent a major breakthrough in U.S. Census taking.

On the other hand, any complete enumeration of Americans abroad must necessarily be "conducted under operational conditions very different from those used...in the United States,"[2] and this one was a pioneering effort. Undoubtedly, then, there is much to be learned (both from the experience of collecting, processing, and presenting the data as well as from attempts to interpret it) which will be of value the next time such a census is taken. This paper represents one user's endeavor to contribute something on that score. Hopefully, it will also serve to alert other researchers to some of the problems they may encounter in working with the data. If, in passing, it also stimulates some interest in the enrollment status of American school children overseas, so much the better.

Over the past year I have had several occasions to make use of the "overseas Americans" data and in particular those which are relevant to the educational status of kindergarten, elementary, and high school aged children. Specifically, I have been attempting to compare the extent to which children of Members of the Armed Forces, Federal Civilian Employees, and Other Citizens overseas are enrolled in school relative to one another and to children in the U.S. population as a whole. Because (at this writing) all of the evidence is not in, I shall not say anything about the conclusions which seem to be taking form. Instead, I will discuss some of the considerations which have had to be taken into account in the analysis to date. Mixed in with these comments are some thoughts about why the educational characteristics of children overseas might be given special emphasis in future enumerations as well as some suggestions concerning changes which might be made in the presentation of future tabulations.

To a large extent, the analysis which motivated these comments was, for its part, inspired by the somewhat startling picture which emerges when one inspects the figures on school enrollment appearing in the Americans Overseas report and compares them with the enrollment status of children in the U.S. (Table 1).

*The research reported in this paper was performed by HumRRO Division No. 7 (Language and Area Training), Alexandria, Va., under Department of the Army contract with The George Washington University. The contents of this paper do not necessarily reflect the official opinion of the Department of the Army.

[1]U.S. Bureau of the Census. U.S. Census of Population: 1960. Selected Area Reports. Americans Overseas. Final Report PC(3)-1C. U.S. Government Printing Office, Washington, D.C., 1964, p. VII.

[2]Ibid., p. VIII.

Table 1

NUMBER OF AMERICANS, AT HOME AND ABROAD, AGED 5-13 AND 14-17, AND PERCENT ENROLLED IN SCHOOL BY POPULATION TYPE AND SEX AS REPORTED BY THE U.S. CENSUS OF 1960

Age and Sex	At Home	Abroad: Total	Military[A]	Government[B]	Other Citizens[C]
5-13					
Both Sexes	32,727,246	177,093	129,096	13,671	34,326
Percent enrolled	89.5	47.4	43.4	57.1	58.6
Males	16,648,338	90,843	66,189	6,911	17,743
Percent enrolled	89.4	47.7	43.6	57.6	58.9
Females	16,078,908	86,250	62,907	6,760	16,583
Percent enrolled	89.6	47.1	43.2	56.5	58.2
14-17					
Both Sexes	11,260,157	27,967	18,191	2,715	7,061
Percent enrolled	87.4	62.9	57.7	73.0	72.5
Males	5,744,349	13,382	8,448	1,401	3,533
Percent enrolled	87.8	66.9	62.8	75.1	73.6
Females	5,515,808	14,585	9,743	1,314	3,528
Percent enrolled	87.1	59.2	53.2	70.9	71.5

Sources: U.S. Bureau of the Census. U.S. Census of Population: 1960. Selected Area Reports. Americans Overseas. Final Report PC(3)-1C Table 2 (Washington: U.S. Government Printing Office, 1964): and U.S. Census of Population: 1960. United States Summary. Detailed Characteristics. Final Report PC(1)-1D, Table 166 (Washington: U.S. Government Printing Office, 1963)

A. Dependents of Members of the Armed Forces.
B. Dependents of Federal Civilian Employees.
C. Other U.S. citizens living abroad.

It can be seen that in 1960, roughly 90 percent of the age group 5-13 "at home" were reported as enrolled while the same was true for only about 47 percent of those overseas. Furthermore, children of Members of the Armed Forces seemed to have been considerably less likely to be enrolled (by about 14 percent) than their fellows in the other two overseas population types. As regards persons aged 14-17, about 87 percent were counted as enrolled at home while the corresponding figure for the overseas population as a whole was approximately 63 percent. Once again, dependents in the Military population appear to have been markedly "under-enrolled" (by about 15 percent) relative to those in the Government and Other Citizens types. However, perhaps the most remarkable feature of Table 1 is the extent to which female dependents of Members of the Armed Forces appear to be missing from the ranks of the enrolled at the high school level. Only about 53 percent of them were reported as attending school.

It has often been pointed out that one of the major concerns of American parents abroad is to see that their children continue in school and, more especially, to see that they receive the kind of education which will serve them in good stead once they return to the U.S.[3] This circumstance has been accompanied by the establishment of a large number of "American type" schools in other lands. Many of them (about 325 in 1965) belong to an overseas school system which is operated by the Department of Defense (DOD). Other evidence of our government's concern for the education of American children in foreign lands is the fact that, in 1966, some 166 privately operated American schools overseas received financial assistance from the Department of State and the Agency for International Development.

Considerations such as these cause the picture which emerges from Table 1 to seem all the more remarkable. This, in turn, suggests that a closer scrutiny of the context in which they originally appeared (i.e. the overall report) may reveal some factors which should be held in mind for purposes of interpretation. A number of these are discussed below.

Under-enumeration

In evaluating its data, the census points out that the decentralized and far flung nature of the procedures employed prevented the utilization of review and quality control operations which are standard in the U.S. One of the results may have been a certain amount of under-enumeration. This is especially true for the Other Citizens because they were asked to cooperate on an entirely voluntary basis. Whether or not an incomplete count could have systematically biased the reporting on school enrollment is a moot point but nevertheless it should be held in mind.

The Definition of Enrollment

The question which was used to elicit information on school enrollment overseas differed somewhat from the one used in the U.S. This last reads as follows:

> P16. Has he attended regular school or college at any time since February 1, 1960?
>
> If he has attended only nursery school, business or trade school, or adult education classes, check "No"
>
> Yes... No...
>
> P17. Is it a public school or a private school?
>
> Public school..........
> Private or
> parochial school.......

Enumerators were instructed to regard public or private kindergartens and accredited correspondence courses as "regular" schooling which, in essence, was defined as:

> that which may advance a person toward an elementary school certificate or high school diploma, or a college, university, or professional degree.[4]

The corresponding question employed overseas read as follows:

> Is [this person] now enrolled in a school, college or university?[5]
>
> (If this person is taking correspondence courses (given by USAFI, a university, etc.) for high school, college, or university credit, check "Yes".)

It should also be remembered that the overseas questionnaires were wholly "self-completed" i.e. without the intervention of trained enumerators.

A comparison of the criteria involved suggests slightly differing definitions of "school enrollment" which, to some extent, could be accounting for the observed under-enrollment of American children abroad relative to their compatriots at home. In framing the item used overseas care was taken to make explicit mention of accredited correspondence courses at the high school level or above. Perhaps it would have been better to specify elementary school correspondence courses as well. The Calvert School, headquartered in Baltimore, Maryland, provides a program of home instruction which covers grades 1 through 8. Its courses are approved by the Department of Education of the State of Maryland and the school is a member

[3] This point is emphasized in many of the writings on Americans overseas. To mention only a few: Ruth Hill Useem, "The American Family in India," The Annals of the American Academy of Political and Social Science, vol. 368, (November, 1966),pp. 132-145. Harlan Cleveland, "The Pretty Americans," Harper's Magazine, (March, 1959), pp. 31-36). Frederick L. Redefer, "The Care and Feeding of Provincials," Saturday Review, vol. 43 (October 22, 1960), pp. 13-14, 39-40. George W. McCown, "U.S. Foreign Service Dependents Schools," School Life, vol. 45, (November 1962), pp. 22-24.

[4] U.S. Bureau of the Census. U.S. Census of Population: 1960. Detailed Characteristics. United States Summary. Final Report PC(1)-1D. U.S. Government Printing Office, Washington, D.C., 1963, p. XVI.

[5] Americans Overseas, op. cit., p. XIX.

of the Educational Records Bureau. The Armed Forces encourage parents who have elementary school aged children with them, at locations where no schools are available, to make use of this program and provide financial assistance for doing so. In 1964 approximately 3,000 overseas American children (including dependents of both civilian and military persons) were enrolled in the Calvert School program.

Secondly, the question used in the U.S. indicates that a child should be reported as enrolled if he has attended school at any time since February 1, 1960. No such provision is made in the question asked of the overseas population. From information collected by the census concerning the date at which respondents arrived abroad it can be conservatively estimated that more than 11,000 persons aged 5-17 had arrived between January 1 and April 1(the date to which the census figures apply). Conceivably a significant number of them had arrived in mid-semester and had not enrolled in school as yet; either because they had not had time, were waiting for a new semester to begin, or for some other reason. Furthermore, many of these "temporarily unenrolled" persons may have attended school in the U.S. after the February 1 cut off date and hence would have been counted as enrolled by the definition used "at home."

In the United States the census collects and publishes information on the kind of school attended (i.e. public or private). Recognizing that it can not possibly cater to the whims of all of its users, I would like, nevertheless, to suggest that (in the future) consideration be given to gathering and tabulating data on the kinds of schools in which overseas Americans are enrolled also. It might be interesting to know, for example, what proportion of them are following curricula similar to those offered in the United States as opposed to purely "host country" ones.

Broadly speaking, American type schools overseas can be divided into four large categories 1) church affiliated schools, 2) international schools, 3) company operated schools, and 4) those operated by the DOD.[6] Again it might be interesting to know the proportion of American children overseas which each of these types enroll.

The student body of DOD schools is primarily composed of children from Military families. However, children from the Government and (much less frequently) Other Citizens types are also represented. The extent to which children of the three population types were enrolled in DOD schools is a question which I would have liked to have been able to answer with the census data.

Finally, a number of DOD schools have facilities for boarding children for 5, 6, or 7 days a week and I would also have liked to learn what proportion of the enrolled were attending school away from their family.

Non-reporting

Apart from the apparent under-enrollment which the "raw data" reveal, one of their most striking features is the exceptionally high rate of non-reporting which they show. In the United States persons for whom the question was left unanswered were considered to be enrolled if they were 5 through 17 years of age. If they were 18 or older

Table 2

PERCENT OF NON-REPORTING ABROAD ON ENROLLMENT AND YEARS OF SCHOOL COMPLETED BY POPULATION TYPE AND AGE GROUP: 1960

	Population Type and Age Group							
	Total		Military		Government		Other Citizens	
	5-13	14-17	5-13	14-17	5-13	14-17	5-13	14-17
Non-reporting on:								
Enrollment	19.2	9.6[A]	18.7	10.0	15.6	6.9	22.5	B
Years of School Completed	2.9	1.9	2.5	1.5	2.1	1.1	4.8	3.1

Source: U.S. Bureau of the Census, U.S. Census of Population: 1960. Selected Area Reports. Americans Overseas. Final Report PC(3)-1C, Tables 2, 9, 11, and 13. (Washington: U.S. Government Printing Office, 1964).

A. Excludes Other Citizens.
B. The number of Other Citizens in the 14-17 age group who did not respond to the question on enrollment is not given by the Census.

[6]For a brief discussion of the four categories see Ruth Dunbar, "American Schools Overseas," NEA Journal, vol. 50, May, 1961, pp. 18-21.

they were treated as "not-enrolled." However, prior to "allocation for non-response" about 4 percent of the age group 5-13 at home were unreported on enrollment.[7] The comparable figure overseas is approximately 19 percent (see Table 2). Although non-reporting abroad falls off considerably in the 14-17 age group it remains fairly high (about 9 percent).

The amount of non-reporting on enrollment stands out even more sharply when it is contrasted with the percentages of missing responses to the question on grade of school completed. Finally, there appear to be pronounced differences in the completeness of the reporting from population type to population type.

It seems probable that there is no single cause for the overall pattern, but I would like to hazard a guess about one of the factors involved. It has to do with the way in which the relevant questions were structured and the order in which they were asked (see Figure 1). Perhaps the parents of many of the 5 and 6 year olds who had never attended school (and were not currently enrolled) checked the space labeled "none" on question 11 and skipped question 12 altogether because they thought it was not applicable. If such were the case it could partially account for the observed differences in non-reporting, both between the two questions and between the two age groups. If one had access to the raw data, this hunch could be checked simply by cross-tabulating responses to the two questions. If an exceptionally high frequency appeared in the cell corresponding to the answer "none" on grade of school completed and "no answer" on enrollment, the explanation tentatively offered here would be strongly supported.

In any case, non-reporting is something a researcher must keep well in mind when dealing with the data as they are given in the census report.

Figure 1

QUESTIONS EMPLOYED IN THE OVERSEAS QUESTIONNAIRE (FORM 60PH-15) TO COLLECT INFORMATION ON SCHOOL ENROLLMENT AND HIGHEST GRADE OF SCHOOL COMPLETED

11. WHAT IS THE HIGHEST GRADE (OR YEAR) OF SCHOOL THIS PERSON HAS EVER ATTENDED

Check one box. If the grade (or year) was in a school outside the U.S., check the box that stands for that grade(or year).

None ___ Kindergarten ___

Elementary school(grade)	1 ___	2 ___	3 ___	4 ___	5 ___	6 ___	7 ___	8 ___	
High or secondary school (year)	1 ___	2 ___	3 ___	4 ___					
College or university(year)	1 ___	2 ___	3 ___	4 ___	5 ___	6 or more ___			

- -

DID HE FINISH THE HIGHEST GRADE (OR YEAR) HE ATTENDED?

Finish this grade (or year)?

Yes - ___ No - ___

12. IS HE NOW ENROLLED IN A SCHOOL, COLLEGE, OR UNIVERSITY?

If this person is taking correspondence courses (given by USAFI, a university, etc.) for high school, college, or university credit, check "YES"

Yes, enrolled in school, college, or university - ___

No, not enrolled - - - - ___

Source: U. S. Bureau of the Census. U.S. Census of Population: 1960. Selected Area Reports. Americans Overseas. Final Report PC(3)-1C. pp. XVIII, XIX. (U.S. Government Printing Office, Washington, D. C., 1964).

[7] U.S. Bureau of the Census. U.S. Census of Population: 1960 Detailed Characteristics. United States Summary. Final Report PC(1)-1D. U.S. Government Printing Office, Washington, D.C., 1963. Appendix, p. 1-810.

Age Categories Employed

The foregoing remarks lead to the question of whether or not more appropriate breakdowns by age could have been used in tabulating the data. Perhaps the most important consideration on that score is the usefulness of grouping 5 year olds with persons 6 through 13. Enrollment figures by single years of age and individual grades are available for the "at home" population.[8] They show that about 45 percent of the 5 year olds were reported as enrolled with roughly 38 percent in kindergarten and some 7 percent in the first grade. There is some evidence that 5 year olds in general, and dependents of Members of the Armed Forces in particular, have less opportunity to attend school overseas than at home. DOD schools do not incorporate kindergartens and according to the report of a study conducted in 1962 although:

> The services are to be commended for the efforts they have made to organize kindergartens on a tuition basis....this is a limited and unsatisfactory response to a need that exists for all (sic) young children.[9]

By mentally interpolating the figures given in Table 3 one can see that the Military population probably contains a much larger proportion of 5 year olds than do either of the other two population types overseas or the population as a whole at home. In addition roughly 74.1 percent of all Americans aged 5-9 overseas are military dependents. These circumstances could 1) account for the fact that underenrollment among 5-13 year olds appears to be greatest in the Military population type, and 2) explain a large portion of the observed difference on enrollment between the total "overseas" and "at home" populations. Since an unknown (but perhaps large) proportion of the overseas dependents of Federal Civilian Employees rely upon DOD schooling for their children the lack of kindergartens could be depressing enrollment in that population type as well. Finally, conversations held with civilians who have lived abroad give me the impression that privately operated American schools overseas sometimes have kindergartens but, even so, they are less common than in the U.S.

The foregoing suggests that in the future, if figures on school enrollment could not be given by single years of age it might be more meaningful

Table 3

NUMBER OF CHILDREN 14 YEARS OF AGE OR LESS, AT HOME AND ABROAD, BY POPULATION TYPE AND AGE GROUP: 1960[A]

Age Group	At Home	Abroad Total	Military	Government	Other Citizens
Total	55,796,970	365,710	282,437	23,457	59,816
Less than 5 years of age	20,321,864	180,664	148,346	8,986	23,332
5 through 9 years of age	18,659,141	114,788	85,079	8,396	21,313
10 through 14 years of age	16,815,965	70,258	49,012	6,075	15,171

[A]See Table 1 for source and definition of Population Type.

[8]E.g. U.S. Summary, op. cit., Table 167.

[9]Department of Defense, "Overseas Dependent Schools: Recommendations for Improvement," (processed). A report of the Survey Committee appointed by Dr. Edward T. Katzenbach, Jr., Deputy Assistant Secretary of Defense (Education), December, 1962, p. 25.

at least to show 5 year olds separately.

In general, the age groups used in tabulations dealing with dependents of Members of the Armed Forces were similar to those dealing with dependents of Federal Civilian Employees. Other Citizens, however, were frequently categorized somewhat differently by age. For example, persons 14-17 in the first two population types were treated as an individual group in most tables dealing with enrollment. By contrast 14 through 17 year olds were often included in a 14-24 year age group where Other Citizens were concerned. This inconsistency made it impossible to use a uniform procedure in adjusting the figures to allow for such things as marital status.

The proportion of married females (14.7 percent) in the age group 14-17 for military dependents is very large in comparison with the proportion among dependents of Federal Civilian Employees (.3 percent). On the grounds that wives are less likely to enroll in school[10] and this might be distorting our analysis somewhat it was decided to exclude married females of "high school age" (i.e. 14-17) from our calculations. Because of the age break-down employed, this was not possible for Other Citizens.

In practice the handicap was not very great when marital status was being looked at because it seems certain that the number of wives among Other Citizens who were from 14 to 17 years of age is very small. However, I have chosen the example because of its bearing on the exceptionally low enrollment observed among females aged 14-17 in the Military population overseas. If wives are excluded from the calculations, the percentage of enrolled females is raised from 53.2 to 62.3 and becomes about on a par with the corresponding figures for males (62.8 percent).

[10]Only single persons are permitted to enroll in DOD schools.

SUMMARY AND CONCLUSION

Because of the increasing number of Americans residing abroad the 1960 census was extended to include the enumeration of all Americans Overseas. The resultant figures are a unique and valuable source of information on this little studied segment of our population. However, both because this was a pioneering effort and because special procedures were necessitated there is probably much to be learned which will be of value in the future. This paper is one users attempt to contribute something on that score. In it are discussed some of the difficulties which he encountered in attempting to make use of the data on school enrollment for persons 5-17 years of age overseas. In the main, these difficulties revolve around the questions employed, an unusual amount of non-reporting, and the age groups which were used in presenting the figures.

MINUTES OF ANNUAL MEETING OF THE SOCIAL STATISTICS SECTION

Washington, D.C., December 28, 1967

The meeting was opened at 7:50 a.m. by Jacob Feldman, Chairman. The Chairman announced that, on the basis of the 1967 elections, the officers for 1968 were to be:

Chairman	John D. Durand
Chairman-Elect	Henry S. Shryock
1st Vice Chairman	Elijah L. White
2nd Vice Chairman	Eleanor B. Sheldon
Secretary	Regina Loewenstein
Representative on Board of Directors	Margaret E. Martin
Council Representative	Leslie Kish
Editor of Proceedings	Edwin D. Goldfield

Implications of the Publications Committee's recent decision to divide the Journal into two parts were discussed. Each issue of the Journal will contain a "Theory and Methods" and an "Applications" section. This innovation resulted from expressions of discontent over the relative paucity of articles with an applied focus in the Journal during the past several years. Now it should be evident to potential contributors that a theoretical breakthrough is not a sine qua non of a JASA article. However, it was not felt that this change of image alone would induce a sufficient supply of good applied, expository, and historical articles. The Social Statistics Section still has the responsibility of stimulating the production of appropriate articles and having them submitted to the Journal. Among the ideas which came out of the discussion were:

1) The papers appearing in the Proceedings could be scrutinized for potential contribution to the Journal. Authors of such papers should be encouraged to make whatever revisions might be desirable for formal publication.

2) Certain sessions at the annual meetings could be designed with the objective of eliciting papers suitable for publication in the "Applications" section.

3) Members of the Section should be on the lookout for relevant papers while attending other conferences and while reviewing the work of colleagues or students.

It was suggested that the incoming Chairman appoint a committee which would thoroughly examine the papers appearing in the Proceedings during the past several years and which would also explore other potential sources. Members were urged to submit their papers to the Journal. Ideas as to how to discover and inspire papers should be submitted to the Chairman.

There was considerable discussion of the desirability of splitting the JASA into two separate journals, following the precedent of the Journal of the Royal Statistical Society. The opinion was expressed that such a separation would be harmful to both theoretical and applied statisticians within the Association. There is already little enough contact between the two groups without erecting an additional barrier to communication.

The program for the 1968 annual meeting was discussed. Members were asked to submit ideas for sessions to the Program Chairman.

A report was made concerning problems arising from the joint sponsorship of sessions at the annual meetings. A question was raised as to the way the overall Program Committee has counted such sessions against Section quotas. The incoming Section Chairman and Program Chairman were instructed to explore the matter further with the various interested parties. The Section officers have the authority to resolve the issue however they see fit.

The incoming Chairman was entreated to consider arranging the Section's 1968 business meeting for a time period later in the day than breakfast.

1967 Officers of the Social Statistics Section

Chairman	Jacob J. Feldman
Chairman-Elect	John D. Durand
1st Vice Chairman	Henry S. Shryock
2nd Vice Chairman	Elijah L. White
Secretary	Philip C. Sagi
Representative on Board of Directors	Margaret E. Martin
Council Representative	John K. Folger
Editor of Proceedings	Edwin D. Goldfield