I

COMPUTER USES IN SOCIAL STATISTICS

Chairman, Joseph F. Daly, Bureau of the Census

Developments in Processing Statistics from Tax Returns - E. J. Engquist, Jr., Internal Revenue Service

Some Implications of Computer Processing of Economic Censuses and Surveys - Jack A. Scharff, Nathaniel Swersky, and Eugene L. Wendt, Bureau of the Census

Selected Computer Applications in Processing Data from a Continuing Health Household Interview Survey-Walt R. Simmons, Robert T. Little, and Eleanor L. Madigan, National Center for Health Statistics

DEVELOPMENTS IN PROCESSING STATISTICS FROM TAX RETURNS

E. J. Engquist, Jr., Internal Revenue Service

INTRODUCTION

The Treasury Department has developed a system of measuring the taxpayer's response to the complex of law, regulations, rulings, decisions, and advice that he receives, by compiling information from the tax returns as they are filed. A major portion of the quantitative information thus assembled is published annually in a "Statistics of Income" series of reports. This paper is principally concerned with a few aspects of one of these, namely, the one dealing with individual income taxes.1/

STANDARD COMPUTER APPLICATION

When the Internal Revenue Service entered into a computer co-ownership agreement with the Census Bureau in 1955, the usual advantages of a computer application were expected and subsequently obtained. Timeliness was improved through computer speed and by peak use of the larger resource; quality was improved by data-edit comparisons not feasible on a multipunch-card record but facilitated by the continuous tape record; and cost and time reductions were achieved by substitution of machine for manual operations, such as in applying sample weights.

Gains, of course, must be netted against the advent of new problems. New recruiting and training requirements, swing and split shift assignments, and tape file controls were among the many new factors requiring management attention. Technical problems of magnitude included the development of new class-file-sort theory to manipulate a large number of tax return units, each with a multifield record (equivalent to 7 or more punch cards), through many cross classifications.

NEW PRODUCTS FROM THE COMPUTER

The system revision that accompanied the application provided for new features. Extensive sample variability measurement was possible for the first time which permitted improvement in sample design. Tabulations of the number of occurrences, as well as amounts, provided new "frequency" data, an important product as well as a basis for evaluation of the dollar amounts. Calculation of ratios between items, on each tax return, permitted summarizing by ratio classes.

THE TAX MODEL

A more recent development is the use of the computer to construct and manipulate a "tax model." The purpose is to provide a quick, accurate method of calculating the effects of alternatives proposed in the individual income tax law provisions. Heretofore the effect of a proposed change was measured by making calculations based upon summary tabulations from the tax returns. These calculations were relatively accurate, depending upon the availability of information on the specific item affected and cross tabulations of independent or control variables, such as size of taxable income classes or tax status. For instance, it is not too difficult to make a calculation based on available summary data showing the effect of a change in the amount of exemption from $600 to $700. It becomes increasingly difficult, however, to measure the effect of two changes proposed simultaneously, such as an increase in the amount of exemption and a decrease in the tax rate.

The computer may be considered the key element in this development. The legislator, Government official, or analyst necessarily has antithetical objectives, a desire for maximum detail and accuracy opposed to his other need for speed and flexibility. With some compromises, it is true, the computer permits a reasonable achievement of all four desiderata in this application.2/

Input.--For this first attempt to develop and manipulate a tax model, the individual income tax returns filed in 1961, covering tax year 1960, are being utilized. Information from Forms 1040, 1040W, and 1040A is involved. The parameters and variables for input to the model are coded information as well as dollar amounts. A given item, such as a tax credit, may be manipulated, that is, assigned an arbitrary value in one problem but in another problem maintain its value as reported. A list of items making up the tax model record is contained in the Appendix. A short description of the definition of the data items is available in the "Statistics of Income" publications.

The tax returns from which the data are taken to produce the tax model record are a subsample of the sample of returns used in producing "Statistics of Income--1960, Individual Income Tax Returns."3/ Stratification, sampling rates, size of sample and relation of the tax model subsample to the Statistics of Income sample are shown in Table 1. The Statistics of Income sample is basically a stratified systematic sample. Since it is used in preparing an annual series of publications, it has been possible to optimize and balance the sample as well as to effect controls resulting from an extensive program of sample management. Results of sample research indicate that usable estimates of the items included in the tax model will be achieved.

To gain flexibility needed in the tax model, however, the size of the sample was reduced. This was achieved by assigning 10 random starting numbers between 1 and K for each sample code for each reel of tape in the Statistics of Income file. The returns to be selected for the tax model file consist of the returns that occur in sequence, beginning with the starting numbers and every K^{th} number thereafter. The K^{th} number for each sample class will be 70, 40, or 20, as shown in Table 1. In drawing the tax model sample, provision is thus made for classifying it, in a random manner, into 10 classes of equal size. Thus, for any characteristic being estimated there will be one estimate based on all 10 classes, and one for each of the 10 classes; the first is the estimate for the user, and the other 10 will provide an estimate of the sampling variability. It is expected that this replicate subsampling procedure will provide estimates of sampling variability on a timely and low-cost basis for the many characteristics that will be investigated.

The tax model input will be completed by the assumptions such as proposed changes in tax rates, tax brackets, tax credits, deductions, exemptions, changes in treatment of sources of income, and other factors. It is expected that these assumptions will be furnished from time to time in varying combinations, depending upon the nature of the problem to be solved.

The tape layout provides for a 60 12-digit word record, 57 of them representing items from the tax return. The first two of the 12-digit words are sorting keys consisting of coded information and the last 12-digit word is a "control" which represents the weight given that record.

In this particular computer application we expect about 2,000 records of this description to go on one reel of magnetic tape, using the 128 characters per inch tape density of Univac I. Since about 100,000 records will be pulled into the sample, the tax model file will consist of approximately 50 reels of tape.

Creation of the Tax Model File of Tape.--The tax model file of tape is a byproduct of the Statistics of Income material. After editing the returns and punching the cards for this publication, the cards are converted to a computer tape which contains many more items for many more returns. This tape is perfected through a data-edit process which also conventionalizes the record to a fixed length and item location, regardless of source document. Out of this process comes approximately 400 reels of tape of accepted records.

For each of these reels of tape there will be read into memory 80 random numbers and the K^{th} number. These random numbers will change for each subsequent reel. The 80 random numbers consist of a set of 10 random numbers for each of 8 sample codes. The computer is programmed to select and write on tape the record for the appropriate returns. At this stage there are 200 tapes produced for the tax model.

At the same time a "select" code is assigned so that all uses of the first random numbers in each sample code are assigned a "select" code 1. All uses of the second random numbers are assigned "select" code 2, etc. Select code 1 will thus contain a complete random sample of all Statistics of Income sample returns. Since 10 such subsamples will be created within the tax model subsample, an ability to produce valid low-cost estimates of variance is obtained.

The 200 reels of tape will be packed so as to produce a file of approximately 50. At this point the tape will be labeled and controlled in a tax model library file. They will be held there until the analyst supplies assumptions and requests manipulation. It is expected that the file will be completely established by June 1, 1962.

Manipulation of the Tax Model File.--The manipulation of the tax model file will consist of the creation of a new file with parameters changed in accordance with programmed instructions. The tax analyst will probably conceive of these changes in terms, relatively simple for him, such as a simultaneous modification of the number of exemptions allowed for blindness and a modification of the amount allowed for each exemption, regardless of kind. The systems analyst must, however, recognize that many other variables must be taken into account in recalculating tax liability in even this simple problem. Each variable in the tax model file must be reviewed to determine how it will be affected by the assumptions in the problem. For instance, the use of tax tables vary by marital status, and the application of tax credits will change upon the calculation of a different income tax liability.

The systems analyst will also devise subroutines. Some of these can already be foreseen. For instance, the several tax tables will be set up as a subroutine on tape, making use of table lookup instructions, so as to reduce the strain on internal machine memory. Tape outputs will be preserved to some extent because the output of one problem may become input for a subsequent set of assumptions.

It is hoped that the manipulation of the tax model file will be relatively quick and inexpensive after a library of subroutines and subfiles is established. It is expected that a systems analyst and programmer can prepare the system and machine instructions in about a week for most of the problems that may be proposed. Machine running time may average less than eight hours per problem. Thus, the total cost of systems analyst, programmer, and machine time for the solution of a problem is expected to average about $1,000.

Evaluation of the Tax Model File.--The analyst seeking an answer to a specific problem in tax administration has certain requirements that should be met by the tax model. The system should provide for a response to a majority of his problems. Responses to the assumptions that will be posed should be fast and accurate. To meet these requirements, this particular file was developed to contain a record of 57 items from the tax return. It is hoped that this length of record will answer most of the questions. At the same time, the record was kept short enough to be responsive to most of the requirements, although it will be impossible to answer questions the same day they are raised. It is currently felt that the accuracy requirement of the analyst may be least well satisfied. Most of his questions are likely to deal with relatively narrow distinctions between assumptions and require changes in parameters that occur less frequently and, therefore, are subject to high sampling variability.

The system that has been devised is obviously limited by the equipment and personnel assigned. While personnel is equally adept on each of Univac Model I and Univac 1105, the Model 1 was selected in spite of its slower internal speed because no conversion program is needed to inspect printouts of intermediate or final results.

Two steps will be taken to help the analyst evaluate the results he obtains from the tax model. The replicate sampling will permit relatively easy calculation of sampling variance. Table 2 shows, for selected adjusted gross income classes, the relative sampling variability at the 95 percent level for four items of information obtained from a subsample of 1959 returns of about the same size and characteristics as will be used in the tax model.

In addition, there will be prepared a relatively complete tabulation of items of information from the tax model subsample which will be compared and analyzed relative to similar information prepared from the larger sample for publication in Statistics of Income. Undoubtedly, because of sampling variability, universe estimates from the data in the tax model subsample, aside from any assumed changes in conditions, will vary from these in the Statistics of Income tabulations. It is necessary to differentiate differences due to the assumptions from those caused by sampling variability. Therefore, it is planned that each time an analyst proposes a set of assumptions there will be produced from the tax model a tabulation showing the information prior to the input of the assumptions as well as after recalculation. This should produce a direct measure of the effect of the assumptions. This measure of "change" due to the assumptions must then be translated to the estimate of the "level" by relating the change to information from the basic Statistics of Income tabulations.

A LONG-RANGE LOOK AHEAD

Other impacts of computers on tax statistics are in the making and will unfold. The development of a master file[4] of tax returns will permit innovations such as local area data, historical comparisons of the record of a single taxpayer, and the establishment of classifications of ratios as the class interval. It may be possible better to satisfy the research worker's need for comprehensive historical files and quick response to his analytical requirements without sacrificing timeliness or accuracy.[5] Such improvements will, of course, assist in the empirical quantifications provided by the computer in testing tax theory that previously was based on less exact hypotheses.

FOOTNOTES

1/ Statistics of Income--1959, Individual Income Tax Returns, is available from the Superintendent of Documents, U. S. Government Print-Office, Washington 25, D. C.

2/ The arrangements represent an agreement among the Office of Tax Analysis of the Treasury Department, the Statistics Division of the Internal Revenue Service, and the "Studies on Government Finance of the Brookings Institution." The last will arrange for computer time for appropriate summary runs of the file which may be desired by the research staff of that organization.

3/ For a description of this sample, see Statistics of Income--1959, Individual Income Tax Returns.

4/ William H. Smith, "Automatic Data Processing in the Internal Revenue Service," one of a series of papers discussed at the Harvard Conference on Taxation and Data Processing, April 13-15, 1961.

5/ Bertrand M. Harding, "The Future of Automatic Data Processing in the Internal Revenue Service," The Tax Executive, Vol. XII, No. 3, pp. 236-240, April 1960.

APPENDIX

ITEMS TO BE INCLUDED IN THE 1960 TAX MODEL TAPES

Statistical Codes

Adjusted Gross Income.--Twenty-six two-digit codes indicating adjusted gross income size classes.

Tax Status.--Four one-digit codes designating taxable returns with normal tax and surtax or with alternative tax and nontaxable returns with or without information.

Marital Status.--Five one-digit codes indicating joint return of husband and wife, separate return of husband and wife, head of household, surviving spouse, or single person.

Form of Deduction.--Two one-digit codes designating returns with itemized deductions or standard deductions.

Data Class.--Four one-digit codes indicating whether the return is to be included in certain programs such as the Audit Research Program, and whether certain special processing is required.

Form.--Three one-digit codes designating Individual Income Tax Forms 1040, 1040A, or 1040W.

Sample.--Eleven two-digit codes designating sampling strata, classified according to adjusted gross income.

File M.--A single alphabetical code to identify tax model tapes.

Exemptions:

Taxpayer.--Two one-digit codes indicating regular $600 exemptions claimed for husband and wife.

Age.--Two one-digit codes for taxpayer and spouse 65 or over at end of taxable year.

Blind.--Two one-digit codes for taxpayer and spouse blind at end of taxable year.

All Other.--Two-digit codes for number of dependent children and other persons, excluding spouse, claimed as exemptions.

Total Exemptions.--A two-digit code indicating total number of exemptions.

Taxable Income.--Thirty-nine two-digit codes for taxable income size classes.

District.--Sixty-eight two-digit codes for the Internal Revenue District Office in which return is filed.

Data Items1/

Excludable Sick Pay.--(page 1, line 6)

Salaries and Wages.--(page 1, line 7)

Farm and Nonfarm Profit or Loss.--(page 1, lines 8 and 9)

Adjusted Gross Income or Deficit.--(page 1, line 11)

Income Tax Before Credits.--(page 2, line 6 or 7, or page 1, line 12)

Dividends Received Credit.--(page 1, line 13 (a))

Retirement Income Credit.--(page 1, line 13 (b))

Income Tax After Credits.--(page 1, line 14)

Total Deductions.--(page 2, line 2, tax computation schedule)

Taxable Income.--(page 2, line 5, tax computation schedule)

Credit for Foreign Tax Paid.--(page 2, line 8 (a), tax computation schedule)

Credit for Tax Paid at Source.--(page 2, line 8 (b), tax computation schedule)

Throwback Tax Credit.--(from schedule attached by taxpayer)

Itemized Deductions: (page 2)

Total Contributions

Total Interest Paid

<u>Taxes</u>.--real estate; state income; state and local sales; total taxes.

<u>Medical Expenses, Allowable Deductions</u>.--("medical and dental expense," line 7)

<u>Other Deductions</u>.--casualty and theft; child care; alimony paid; employee business expense; entertainment expense; educational expense; total other deductions.

<u>Total Itemized Deductions</u>

<u>Dividend Exclusion</u>.--(page 3, Schedule A, line 3)

<u>Dividends in Adjusted Gross Income</u>.--(page 3, Schedule A, line 6)

<u>Interest Received</u>.--(page 3, Schedule B, total)

<u>Net Capital Gain or Loss</u>.--(page 3, Schedule D, line 1, or separate Schedule D, (I), column h, line 10 or 11)

<u>Other Property Gain or Loss</u>.--(page 3, Schedule D, line 2 or separate Schedule D, (II), column h, line 3)

<u>Pensions and Annuities--Taxable Portion:</u>

<u>Part I</u>.--(page 3, Schedule E, line 6)

<u>Part II</u>.--(page 3, Schedule E, line 5)

<u>Rents, Net Income or Loss</u>.--(page 3, Schedule G, computed by subtracting sum of columns 3, 4, and 5 from column 2 for rent only)

<u>Royalties, Net Income or Loss</u>.--(page 3, Schedule G, computed by subtracting sum of columns 3, 4, and 5 from column 2 for royalties only)

<u>Partnership Profit or Loss</u>.--(page 3, Schedule H, line 1)

<u>Estate or Trust Income or Loss</u>.--(page 3, Schedule H, line 2)

<u>Net Operating Loss Deduction</u>.--(page 3, Schedule H or attached schedules)

<u>Other Sources of Income or Loss</u>.--(page 3, Schedule H, line 3)

<u>Capital Gains and Losses</u>.--(from separate Schedule D):

<u>Capital Loss Carryover</u>.--(D (I), column h, line 3)

<u>Net Short-term Gain or Loss</u>.--(D (I), column h, line 4)

<u>Net Long-term Gain or Loss</u>.--(D (I), column h, line 7)

<u>One-half Excess Long-term Capital Gain</u>.--(D (I), column h, line 13)

<u>Balance for Partial Tax</u>.--(D (I), column h, line 14)

<u>Total Dividends Received</u>.--(computed by summing lines 2 and 5, page 3, Schedule A)

FOOTNOTE

1/ References in parentheses are to the location of data items on the Individual Income Tax Form 1040 for 1960 unless otherwise indicated.

Table 1.--Sampling Rates for Statistics of Income, 1960 Individual Income Tax Returns, and for the Tax Model 1/

Sample Class, Form, and Adjusted Gross Income	Number in universe	Statistics of Income Sample Plan		Tax Model Sample Plan		K[th] Number
		Rate	Number	Rate	Number	
	(1)	(2)	(3)	(4)	(5)	(6)
Forms 1040A	17,478,000	3/1000	52,434	1/4 x 3/1000	13,109	40
Forms 1040 and 1040W:						
AGI under $10,000:						
Nonbusiness	29,849,100	3/1000	89,547	1/4 x 3/1000	22,387	40
With Schedules C or F	8,963,300	1/100	89,633	1/7 x 1/100	12,805	70
AGI $10,000 under $50,000:						
Nonbusiness	3,302,300	3/100	99,069	1/7 x 3/100	14,153	70
With Schedules C or F	1,347,200	3/100	40,416	1/7 x 3/100	5,774	70
AGI $50,000 under $150,000:						
Nonbusiness	52,000	3/10	15,600	1/2 x 3/10	7,800	20
With Schedules C or F	56,400	5/10	28,200	1/2 x 5/10	14,100	20
AGI $150,000 and over:						
Nonbusiness	4,900	1/1	4,900	1/1	4,900	1
With Schedules C or F	5,200	1/1	5,200	1/1	5,200	1
Prior year delinquent returns:						
AGI under $50,000	387,200	1/100	3,872	1/7 x 1/100	553	70
AGI $50,000 and over	150	1/1	150	1/1	150	1
Total	61,445,750		429,021		100,931	-

1/ Based on "Sampling Plan for Tax Model for Tax Year 1960" dated May 4, 1961

Table 2.--Relative Sampling Variability at the 95 Percent Level for Selected Adjusted Gross Income Classes in TAX ANALYSIS OF INDIVIDUAL INCOME TAX RETURNS FILED FOR 1959 DURING 1960

Adjusted gross income class	Number of returns		Adjusted gross income		Taxable income		Income tax after credits	
	Number	Relative sampling error (Percent)	Amount (Thousand dollars)	Relative sampling error (Percent)	Amount (Thousand dollars)	Relative sampling error (Percent)	Amount (Thousand dollars)	Relative sampling error (Percent)
Grand total..............	60,275,050	0.03	[1]305,760,588	0.47	167,158,420	0.65	38,899,720	0.87
Taxable returns, total.........	47,534,171	0.39	288,480,915	0.51	167,011,593	0.65	38,899,720	0.87
$600 under $1,000............	1,304,610	5.33	1,082,458	5.35	186,993	6.22	37,467	6.23
$5,000 under $6,000..........	6,320,830	2.27	34,680,384	2.27	16,985,008	2.64	3,431,270	2.51
$100,000 under $150,000......	18,152	16.28	2,138,815	15.28	1,794,464	16.56	907,772	16.66
$1,000,000 or more...........	257	(2)	534,986	(2)	408,962	(2)	257,999	(2)
Nontaxable returns, total......	12,740,879	1.43	[1]17,279,673	3.78	146,827	16.00	-	-
No adjusted gross income.....	440,158	7.21	[3]1,404,972	37.13	-	-	-	-
Under $600...................	3,932,219	2.91	1,277,581	3.29	-	-	-	-
$600 under $1,000............	1,660,148	4.46	1,271,921	4.50	(4)	(4)	-	-
$5,000 or more...............	238,317	12.11	1,545,634	11.70	(4)	(4)	-	-

[1]Adjusted gross income less adjusted gross deficit.
[2]Not subject to sampling variability since the returns in these classes are sampled at a 100 percent rate.
[3]Adjusted gross deficit.
[4]Sample to small to yield reliable estimate.

Internal Revenue Service
Statistics Division
December 1961

SOME IMPLICATIONS OF COMPUTER PROCESSING OF ECONOMIC CENSUSES AND SURVEYS

Jack A. Scharff, Nathaniel Swersky, Eugene L. Wendt, Bureau of the Census

In 1951 a small group of individuals including the Chairman of this session, Dr. Daly, pioneered at the Bureau of the Census in the planning and programing of the Annual Survey of Manufactures to be run on a new piece of equipment called UNIVAC. They had no idea that from this small beginning there would be proposed such a wide range of computer applications as have been developed in the last 10 years. The consequences of the work of this group of people have had far reaching effects in many areas, not the least of which is the economic statistics surveys of the Census Bureau. The work that is now being done on the computer far exceeds the dreams of the people who were working on this one project ten years ago. It might be said that we are now celebrating an anniversary of the processing of economic statistics on the computer.

Why can it be so candidly stated that the original pioneers in programing had so little concept of the wide range of contemporary applications? The answer is found in that the implications and applications of computer processing were not fully understood at that time, and, in fact, currently, computer applications stretch as far as the imagination and desires of programers and management.

In Appendix A there is a partial list of the economic statistics projects that have been processed in whole or in part on the computer at the Bureau of the Census in the last few years. The size of this list is only a fair indication of the extent of computerization in the field of economic surveys. The estimates of work to be performed on the computer in the next few years is much greater as is indicated on this list, and would tax the capacity of existing Census computers without considering any other surveys or censuses which the Bureau will conduct. In order to overcome this shortage of computer hours there is under way at this time a feasibility study to determine how much additional capacity is required and which computer will best fulfill the requirements.

In order to more fully appreciate the work involved in the list of projects mentioned, one must consider some of the implications of this processing effort and of computer processing in general. Of paramount importance are the following:

1. Types of personnel involved in the staffing of a programing unit and their functions.

2. Programing methods utilized.

3. Systems analysis and data analysis as related to computer programs.

4. Interest of management in furthering objectives.

5. Scheduling of operations, programing and production by the computer.

<u>Staffing a Programing Unit</u>

There are three basic types of skills required within a successful programing unit. First is the executive or unit administrator who has the primary responsibility to receive all incoming requests for programing, to set up the necessary meetings with project sponsors in order to clarify and solidify specifications, to allocate the project to a responsible systems analyst or project planner, to establish a schedule for the completion of the project, to allocate test periods and computer priorities to insure the culmination of each job within the prescribed period, and to set programing policies, principles and techniques for the unit.

Next is the systems analyst or project planner who has the responsibility of evaluating the particular project, in order to determine the best method or computer system to be used in deriving final results and to prepare an operational flow-chart which will guide all programing efforts.

Last is the coding programer who works with the analyst to prepare and test the programs dictated by the operational flow-chart.

<u>Programing Methods for Computer Results</u>

One should never forget that implicit in the use of computers is the ultimate goal of obtaining results. A successful program is one which essentially fills the real <u>needs</u> of management within the desired time limits and as budgeted.

In order to attain this end some programing policies are implied. In the Economic Operations Division of the Bureau of the Census four basic programs are produced. The first three listed below are capable of producing almost immediate results, that is to say within a day and less than a week from inception to completion of project.

The first type of program and usually the easiest to write, is that which merely requires insertion of parameters into a generator which, after computer processing, provides a running program. Typical of this application are sorts and merges which are more or less standardized.

The next type utilizes a basic input/output framework program including subroutines for input/output, housekeeping and conversion. The actual processing steps required are coded manually and then entered into a provided storage area of the program.

The third type, different from the one above, utilizes as a basic source a previously completed program which accomplishes results similar to the current needs. Input data is tailored to fit the conditions and revisions are made in the program to serve the input.

Finally there is the full-fledged complex program required for high volume processing or for periodic surveys. Within this category of programs it has been found necessary to fully document every stage of operations so that modifications may be made in the future by persons other than the one who originated the program.

Within these four categories, during the course of the last six months, supposedly a slack period between economic censuses, there have been 500 programs written by fewer than 25 people.

Systems Analysis and Data Analysis as Related to Application to Computers

By far the most critical implication to successful processing is the preliminary work done prior to the program coding effort. This is the "leg work" in which the executive of the programing effort gathers together the requirements from management usually in the form of "specifications" and evaluates these specifications in concert with statisticians, methods determination experts, program analysts and the like, in order to determine how the particular program may best be resolved.

The duration of this initial effort is a variable depending upon the conciseness of specifications and the intricacy of the problem. The following summary of some actual applications is illustrative of some but not all of the implied considerations in computer processing.

General Effects of Computer Processing of Economic Statistics

There are a series of questions which can be asked when exploring whether a survey should be processed on the computer, but the most important one is, "What can I obtain from computer processing that I cannot get from another technique, whether in timing, or cost, or quality of work?". The first possible answer to this is consistency.

It is well recognized that one of the many advantages of the computer is the absolute consistency between the handling of one item and the next. Once the computer has been correctly instructed to do a job it performs its mission tirelessly and efficiently.

Consistency, however, is only a virtue if the results are consistently correct. The necessity to provide a system which represents a carefully developed plan with no logical loopholes creates the need for a different emphasis by statistical personnel than that required in a non-computer philosophy. The pre-test of plans is of greater importance; the need for investigation of all possible consequences of specifications for a survey is magnified; and the requirement for liaison between the statistician and the computer technician is mandatory.

Thus under a computer oriented system a person not only has to be proficient and understand the subject matter, but also should be able to translate this understanding into a format adaptable to the computer. Sometimes even the best plans cannot be translated into computer language because of what is called the lack of judgment of the computer. A person who is not only familiar with the subject matter but also with the computer can usually devise means of building a pseudo-judgment factor which can then be adapted for the computer. As an example, there are many surveys which use a consistency edit as one of the techniques for determining the reliability of input data. The computer makes decisions by applying a procedure to data and using as a base its recent experience with similar data, much the same way as a clerk would use his judgment.

Another of the results of computerization has been to transfer the work load from the clerk to the systems analyst. This phenomenon has had one advantageous effect in the Census Bureau where traditionally there have been periods of peak employment for comparatively short periods of time. It is now possible to conceive of a type of operation which will minimize the effects of this peak employment and will create a more stable working force. This requires a greater lead time to properly prepare for a major census or survey. A wholesale change can never be made without sufficient lead time to thoroughly check the results usually by processing on the computer a small portion of the data and reviewing the output to determine whether the work has been performed satisfactorily.

The computer, which is equally effective in accounting and statistical processing, can decide whether a particular set of conditions are acceptable and, if not, can substitute for the unacceptable data a closely related series. In the Foreign Trade Exports processing, in order to eliminate inconsistencies in the statistics every detail record is checked for unit price. Most of the items for each commodity are within the established range of acceptable prices. Previously, those items that fell outside the price range were reviewed clerically and were re-inserted in the statistics in the following month. These specifications produced statistics which were considered to be misleading because of the carry-over from one month to the next. In order to overcome this problem, the computer now derives a new quantity based on the characteristics of historic data. The value is left unchanged because in numerous studies it has been found to be one of the most reliable of data items on the record. After this imputation the detail is reasonably consistent, both internally and between months, however, this has had the effect of minimizing the legitimate, large differences which might occur in the universe.

Relationship of Computerization to Quantity of Output and Operation

There is a sign on the door of a programmer's office which may be attributed to an anonymous source: "Never begin a vast project with half vast ideas." This statement is particularly appropriate when applied to computer work. The amount of information and quality of operation that can be achieved with computers is only limited by the ability of the humans to conceive ideas. Certainly, it is true that projects as large as the surveys conducted by the Census Bureau deserve a considerable amount of preliminary thinking in order to make the use of the computer worthwhile. In addition to the miles of paper that can be generated by the computer, there is also the ability to utilize analytical tools which have previously been too costly to consider.

The Annual Survey of Manufactures, prior to 1959, was a probability sample designed to publish general statistics with considerable detail but with only limited coverage for product statistics or for statistics on small areas. The sample selected for the current panel of the Annual Survey of Manufactures increased the coverage sufficient to permit publication of data on product classes and local areas in detail not previously considered publishable because of the high variances.

Subsequent to the selection of the sample, a dual control was established on the Annual Survey of Manufactures panel. First, a mailing register was initiated to provide for printing the 1958 data on the report forms. Separately, using the same source material, a data register was established to be used for the computer editing and tabulating routines. This dual register has provided the means of controlling each file with an independent verification by matching the two files.

The mailing register of the Annual Survey of Manufactures is used for controlling and recording the receipt of correspondence and reports. Furthermore, it is a source of information to identify those companies to be included in a mail followup for delinquent reporting. This file contains sufficient information to facilitate selection of sub-samples from the Annual Survey of Manufactures universe; to study response patterns by selected categories; and to maintain a reference file of delinquent reporting characteristics.

The data register containing the prior year statistics for each company is used in the tabulation of the Annual Survey of Manufactures to control status changes and to provide a base for the editing of current data. After identifying and correcting duplicate and mispunched reports the prior year's register is used in the balancing and editing operations as a stabilizing influence to prevent the insertion of erroneous data. Likewise, for those plants for which no report has been received, it provides the base for imputation. The prior year's data register is also used in the tabulation program as a base year for estimates.

Change from Detail Analysis to Summary Analysis

Computers have provided the means by which it is possible to change from a review of data at the detail level to an intensive analysis of a small number of summaries. It is no longer necessary to construct a model within which the individual report is retained until the final stages of the summarization, but now it is feasible to assume that the raw data can influence the compilation only at the early stages and later can be subjugated to the effect of the cell changes. An excellent example of this method of non-reliance on detail may be found in some of the techniques utilized in the production of the 1959 County Business Patterns. For this publication an edit was made against historical data from the 1956 County Business Patterns and from the 1958 Censuses. If cell totals were found to be inconsistent with anticipated relationships to the historical data, the detail items for that cell were manually reviewed.

In the County Business Patterns tabulations many innovations were introduced or expanded. In the first place, the County Business Patterns were run on three types of computers -- IBM 705; Univac-I; and Univac 1105. The communications between the computers presented a number of problems. The resolution of these provided the experience needed to eliminate this problem for the future. This flexibility in computer hardware is the outgrowth of recent improvements but the ability to accomplish this is becoming more common.

For the 1959 County Business Patterns the computer was instructed to analyze and select a single set of codes and data from five files obtained from Census and Social Security. The processing of these files on the computer through a series of matches, edits, corrections and sorts required several hundred programs.

In order to facilitate the program writing there was developed for Univac-I an input/output framework generator which freed the programer from the necessity of debugging at least the housekeeping half of his program. After corrections were assembled, the more powerful 1105 computer was utilized to provide final tables. A single summary program provided, in three successive passes, county industry tables, state industry tables, and U. S. industry tables.

Capacity to Explore

One of the features which has been developed on the computer in the last few years has been that of experimenting with available data to determine what would be the effect of certain program changes. Foreign Trade statistics have been compiled on the computer for more than four years and these historical data have been frequently utilized in various studies to test theories. Many of the innovations which have been introduced into the Foreign Trade program have been the result of some of these studies.

The replacement of the logical approach utilizing quasi-accepted procedures, by an empirical approach limited only by the imagination and by the availability of data, has released surveys from tradition and allows them to operate within a freer framework. Objectives can then be tested based upon a new standard and an old data instead of the previous reliance on the opposite, namely, old standards and new data.

The limitation of cost can quite often be overcome in the computer by designing a program specifically for a given objective. Many illustrations of this could be easily cited. In the course of the 1958 Census of Manufactures and Minerals when the industry volume tables were being planned, instead of using a full record to process all of the tables it was thought to be more economical to produce a number of individual records each designed to fulfill the requirements of a specialized series of tables. For example, a twelve-digit record was devised which was intended for use in producing two tables of area by industry by employee size statistics. The entire Census of Manufactures universe was contained on two partial tapes, whereas it would ordinarily require 40 tapes. This same record was later used to produce the publication "Location of Manufacturing Plants." Similar economies made it possible to utilize the computer in areas which previously were not deemed efficient.

One of the many problems which exist in producing statistics from reports which are extensive in the relationship of one item to another is the difficulty of coordinating all of the interrelationships. In the 1958 Census of Manufactures it was decided that the entire record should be kept together until such time as the interrelationships had been completely edited. This necessitated making provisions for a maximum of 50 data fields of General Statistics, 250 Product items with 8 data fields for each item, 60 Material items with a maximum of 4 data fields for each item, 50 Special Inquiry items with a maximum of 24 data fields for each item, and a Fuels and Electric Energy item with 15 data fields required. The handling of these records during the course of industry coding and editing illustrates the potential to which the computer may be directed.

Minimization of Pre-Computer Processing of Input

In addition to quantity and quality of data there is a possibility now to extend the computer work both forwards in the process in order to take advantage of the raw data and backwards to deliver from the computer a finished product. The system which merely reflected a transfer of operations from punch card equipment to the computer is now being converted to one which examines the entire process from the beginning to end and allows the computer to be used freely as a tool within the system.

The use of check digits, alphabetic coding, analysis of complex interrelations, editing, nonsense balancing and the like has extended further the utilization of the computer. Also, the output of the computer can now be presented in such a form that with only minor manual editing the computer output may be published as produced. The first successful experiments of presenting data directly from the copy produced on the high speed printer indicated not only the feasibility of this approach but also the desirability and practicability.

The Interest of Management in Extending Objectives. The Dynamic Qualities of Computer Utilization.

It is particularly significant to observe the growth in interest by management in the use and functions of computer operation. Although enthusiasm may be present, there is also some trepidation in the transition from conventional punch card and clerical procedures to computer operations. Apparently a great number of procedures done clerically defy definition. Comments like "this is based on experience," or "we know intuitively that this is wrong," and "actual documents must be examined to resolve problems" are by constant questioning formalized into a computer system. Even so, the initial objectives are not too far extended from the conventional system.

However, once the particular program has been computerized and proves successful, there ensues a feed-back of ideas from management. Initial "fears" vanish and of paramount interest is the question of what else can the computer accomplish.

Illustrative of this feature of feed-back is the monthly Foreign Trade Program conducted by the Census Bureau. Approximately four years ago each of 400,000 commodity classification codes appearing on incoming documents was verified clerically and each unit of quantity was reviewed for consistency. In addition, approximately 8,000 items which were rejected by the computer for nonconformance to pricing criteria during an edit were manually reviewed and corrected for inclusion in the survey of the following month.

Today, because of management's questioning of clerical processes, approximately 10,000 items are manually coded per month and rejects for pricing are imputed by the computer eliminating clerical intervention and what is more important included in the current month's survey. Problems of this nature are resolved as they are recognized in concert with "statistical specialists" who analyze the situation and reduce the problem to a mathematical formula for computer resolution. In this instance the computer itself contributed in the resolution by evaluating data for an extended period in the past and introduced new levels for pricing criteria.

In the same program, historically, alphabetic information appearing on incoming documents were clerically encoded in accordance with two coding manuals. Effective in January 1962 this coding operation will be performed by the computer by use of 4 alphabetic characters from the document to express country and 5 characters to express foreign port. In essence these are the first characters of the country and the foreign port name. Today this function is successfully accomplished on 10% or 40,000 records. Once again it was the result of application of "specialists" cognizant of the problem, and the computer utilizing past data to analyze and assist in the preparation of a translation matrix. The implications for the future indicate the end is not in sight. Document "scanners" open a new vista for the computer.

Utilizing current and past data the computer has been developing statistical aides to be used as a guide in our current tariff dilemma and in the management of import quotas in the area of cotton and cotton textiles. These special surveys are a few of the illustrations of "feedback" of data produced by the computer to be utilized in a manner not initially contemplated.

<u>Scheduling of Operations, Programing and Production on the Computer</u>

Conceding that an efficient working programing organization exists, that the preparation of computer programs poses no problem; one is faced with the final hurdle of getting results.

In a unit there is more than one project entailing many computer programs being processed at the same time. Some order must be maintained; this implies the establishment of a unit whose sole responsibility is to maintain order. It is the responsibility of the head of this unit to evaluate the number of hours required, usually derived from experience or in consultation with the executive of the programing unit and to schedule computer time based on the priority needs of management.

Having scheduled time, work must be set up in advance and personnel made available at the computer to perform the necessary tasks. This is accomplished by following project operational flow charts and operating instructions previously prepared by the systems analyst and coding programer. Attendant also is the maintenance of records of computer time, controls to insure data has been processed properly and the like.

In conclusion, the solution of problems in existing surveys increases the potential of applying the successful techniques to newer surveys or to those not yet programed for the computer. The limitation of the computer a few years ago has diminished. Each year a different set of criteria is used to provide the basis for determining the range of computer applications. Automatic programing such as COBOL is a new feature which has been introduced and eventually should overcome some of the existing shortages of trained personnel. The acceptance of the computer as a powerful tool by educational institutions and businesses, as well as government has created the desire to further exploit this tool. Tomorrow we should be in a position to look back on today and sympathize with our present limitations.

APPENDIX A

ECONOMIC CENSUSES AND SURVEYS CONDUCTED ON THE COMPUTER AT THE BUREAU OF THE CENSUS

	COMPUTER HOURS	
	Used During Period July 1958-Nov. 1961 (Univac I & 1105 hours)	Estimated for Period Jan. 1962-Dec. 1966 (1105 hours)
Censuses of Business, Manufactures and Minerals		
Basic Censuses	18,000	22,500
Business Supplemental Program	800	
Industry Supplemental Program	300	
Concentration Statistics	170	
Central Business Districts	240	
Company Statistics	450	
Census of Governments	20	1,000
Current Business Program		
Monthly Accounts Receivable Survey	900	700
Monthly Retail Trade Report	1,800	1,300
Monthly Wholesale Trade Survey	800	600
Other Business Surveys	75	300
Business Trust Funds	650	-
Current Industry Program		
Annual Survey of Manufactures	2,700	2,200
Monthly Industry Survey	90	1,000
Annual Lumber Survey	120	1,000
Other Industry Surveys	60	5,500
Industry Trust Funds	650	-
Foreign Trade Program		
Exports	4,600	22,700
Imports	4,400	
Shipping	5,300	
Exports of Manufactured Products	100	
Other Foreign Trade	300	
Foreign Trade Trust Funds	1,600	
County Business Patterns	4,000	2,500
Construction Program	50	1,100
Miscellaneous	-	7,800

Figure 1. Condensed estimating equation, for any statistic x in the U. S. National Health Household Interview Survey.

NOTATION

Subscripts:

g = stratum
h = PSU
i = Segment
j = household
k = person
α = color-residence zone
β = age-sex-color class

Statistics:

x = a health characteristic; = 1 if present, equal 0 if not present.
y = current population value; = 1 for each individual person.
z = 1950 population value; = 1 for each individual person in 1950 Census.
$\cdot$ signifies summation over missing subscripts.

Number of Units:

m = number of sample segments
n = designed number of sample households
n' = responding number of sample households

(Subscripts indicate scope of inclusion where notation might otherwise be ambiguous.)

Weights:

u = original 1st stage design weight for individual person
v = original subsequent stages design weight for individual person
w = design subsampling weight for out-sized segments

(Subscripts are not shown in weights, but they vary along with the statistics which they multiply.)

Summations: All summations are over the universe unless otherwise indicated.

The Estimate of x is:

$$x' = \sum_{\beta} \frac{x'_{\cdot\beta}}{y'_{\cdot\beta}} \, y_{\cdot\beta} \quad , \quad \text{in which}$$

$$x'_{\cdot\beta} = \sum_{g}^{503} \sum_{h}^{1} \sum_{i}^{m} \sum_{j}^{n'_{ghi}} \sum_{k} \sum_{\alpha} \frac{n_{ghi}}{n'_{ghi}} \, \frac{\sum\limits_{g}^{503} \sum\limits_{h}^{1} \sum\limits_{i} \sum\limits_{j} \sum\limits_{k} \sum\limits_{\beta} u \, z_{ghijk\alpha\beta}}{\sum\limits_{g} \sum\limits_{h} \sum\limits_{i} \sum\limits_{j} \sum\limits_{k} \sum\limits_{\beta} z_{ghijk\alpha\beta}} \, u \, v \, w \, x_{ghijk\alpha\beta} \quad , \quad \text{and}$$

$$y'_{\cdot\beta} = \sum_{g}^{503} \sum_{h}^{1} \sum_{i}^{m} \sum_{j}^{n'_{ghi}} \sum_{k} \sum_{\alpha} \frac{n_{ghi}}{n'_{ghi}} \, \frac{\sum\limits_{g}^{503} \sum\limits_{h}^{1} \sum\limits_{i} \sum\limits_{j} \sum\limits_{k} \sum\limits_{\beta} u \, z_{ghijk\alpha\beta}}{\sum\limits_{g} \sum\limits_{h} \sum\limits_{i} \sum\limits_{j} \sum\limits_{k} \sum\limits_{\beta} z_{ghijk\alpha\beta}} \, u \, v \, w \, y_{ghijk\alpha\beta} \quad , \quad \text{and}$$

$y_{\cdot\beta}$ is the official independent control estimate for the β th age-sex-color group for the total U. S. civilian non-institutional population.

NOTE: The adjustment $\frac{y_{\cdot\beta}}{y'_{\cdot\beta}}$ is applied separately to each sample datum in the β th age-sex-color class would require calculation which was too costly in terms of dollars or time. Again the computer has made recent theoretical advances operationally feasible. At the risk of being guilty of huckstering, but in the effort to illustrate the point, we display a somewhat condensed version of the basic estimating equation for the Health Interview Survey.

Without belaboring the point, but with the reminder that this displayed estimating equation does not reflect all the detail of actual estimation, and noting that the complex estimating process is applied to each of the 130,000 persons interviewed each year, it is obvious that the computer permits an elaborateness in estimation which would rarely, if ever, be attempted by less powerful computational devices.

An important characteristic of computer estimation is that it permits differential weighting of original observations at very small cost. Thus, it invites the use of efficient survey designs which otherwise might be foreclosed.

An illustration of the kind of estimating steps taken is given here by a description in outline form of the last phase of the process, which is an adjustment of all data by ratio factors to bring them into agreement with independent official population estimates prepared by the Bureau of the Census. The algebra of this step is to multiply every datum in each of 60 age-sex-color cells by the ratio of a population control to the sample population estimate for that cell before adjustment. Two computer runs are required to accomplish this adjustment, controlled by machine instructions designed to (1) read all data records into the computer, isolate, classify, and tally each weighted person-record into the proper one of 60 age-sex-color classes and code the record with that class indicator; (2) introduce into computer memory the independent estimates for the 60 specified classes; (3) carry out the necessary arithmetic to calculate ratio factors and store these factors in a set pattern in memory for use in the next pass of the data records; (4) again read the data records, use the class code to select the proper adjustment factor stored in memory, and multiply the previous weight by that factor to produce a new weight; and (5) put this adjusted weight into all records associated with that person.

This phase of estimation is not a complex computer operation compared with several others in the HIS survey. However, the same adjustment on conventional punched-card equipment would develop into a complicated series of processing steps and likely require some supporting manual operations. Time and cost would be prohibitive whereas only four hours of computer running time for one year's data are required to complete this process.

Tabulation

An element of the estimation process, which is not explicitly evident from the equations, is that the computer is programmed to convert each observed person, condition, or hospital episode to an estimate of that part of the total universe which it represents. Hence, in principle, simple addition of the converted data for any class produces a population estimate for that class. This conceptually simple process is slightly more involved in operation for two reasons. The first of these may be thought of in general terms as a matter of sorting, classification, and rearrangement of records. Ordinarily tabulation cannot be accomplished by a simple automatic retrieval of data stored in the computer. It is necessary through one of several search and assembly procedures to collect cases which belong in the same cell, and then count them—or add their adjusted weights—to produce the desired table. The other special action which must be taken is adjustment of the data to the desired time scale. Basic weights in the estimation procedure are set at such a level that direct tabulation of the sample over 13 weeks of interviewing produces an average value of the relevant statistic for a 13-week period. When 52 weeks of interviews are tabulated, a multiplier of 1/4 must be introduced into the tabulation to preserve the feature of average-over-the year. For some statistics, such as number of physician visits, the reference period for the direct statistic is a 2-week interval. If total physician visits in a year are desired, and 52 weeks of interviews are tabulated, the correct multiplier is 6.5, because the basic weighting applied to the sum of 2-week experience for each of four quarters produces an estimate for eight weeks, which must then be multiplied by 6.5 to produce a 52-week aggregate. For other combinations of number of weeks of interview, length of reference period, and type of statistic other multipliers are appropriate. In all cases, the scaling is a simple matter, and great flexibility in tabulation is possible with a minimum of new programming.

Attention is called to a standard procedure used in NHS to facilitate programming for tabulation, although the process is not dependent on tabulation by computer. We refer to the regular practice in planning tabulations of preparing annotated dummy tables. These tables are prepared as early as possible in the entire collection process, preliminary versions usually being made before the questionnaire has been

SELECTED COMPUTER APPLICATIONS IN PROCESSING DATA FROM A CONTINUING HEALTH HOUSEHOLD INTERVIEW SURVEY

Walt R. Simmons, Robert T. Little, and Eleanor L. Madigan
National Center for Health Statistics

Introduction

In his Introduction to the Development of Economic Doctrine, Alexander Gray comments that in no other book on the subject are so many important topics omitted.[1] That thought is an appropriate caveat for the present paper. This report presents no new statistical theory or methodology of any consequence, nor does it offer any striking patterns of programming computers. It does not even mention such features as computer memory, capacity, speed, or other characteristics of the equipment. The sole purpose is to display for those practicing statisticians who may not have had much opportunity to use computers, some illustrations in which a computer has been used to facilitate the reduction of data to more useful information. Computer processing has accomplished some of these calculations more rapidly than would otherwise have been true, and in one or two instances has made feasible a method which otherwise might not have been undertaken. Some of the illustrations may suggest more elaborate techniques which the electronic computer makes possible at tolerable cost. The applications discussed were all carried out with a Univac I, but it is believed that, with relatively minor modifications, other smaller or larger computers might have been used.

Attention is invited to one very general kind of circumstance, before considering specific examples. It has become a commonplace thought that any, or at least very nearly any, logical process which can be described in sufficiently precise terms can be programmed for computer operation. Similarly, at least in theoretical terms, and given sufficient time, any process or computation which an electronic computer can accomplish, could be done by a person with much simpler tools. Thus any statistical processing or analysis which can be described in mathematical terms can be carried out on a computer. And there is no magic which the "Giant Brains" computers use, which constructively cannot be matched by ordinary mortals.

The U. S. National Health Survey

Illustrations offered in our remarks all come from the conduct of the household interview phase of the U. S. National Health Survey. The illustrations can be viewed in better perspective if leading features of the Health Interview Survey—which will be referred to as HIS—are recalled briefly. The HIS is a continuing survey which has as its purpose determination of the incidence and prevalence of illness, impairments, and injuries; the use of medical, dental and hospital facilities, and related health phenomena for the civilian noninstitutional population of the United States. The survey is the consequence of more than 10 years of exploratory activity culminating in legislation in 1956, which directed the Surgeon General of the Public Health Service to put into operation such a system. Working under detailed specifications of the Public Health Service, the Census Bureau prepared the original survey design and collects the data. Collection is by personal interview of people in their homes, mainly through a structured questionnaire. The schedule includes some 100 questions on demographic and socioeconomic characteristics of respondents and on their health status and experiences over recent time periods.

The sample design is such that each week's interviewing constitutes a representative sample for the nation. Hence, trends can be produced for high-incidence statistics, while combinations of samples for 13 weeks, or 52, or 104 weeks, or even longer periods, can produce data for items of lower frequency in the population. Most analysis thus far has rested on interviewing for 52 weeks. Over a year the sample includes approximately 130,000 persons in 38,000 households in 6,400 land area segments in 503 PSU's (counties or groups of counties). The sample design has several times been modified in the last four years, through joint consideration and decision by the National Health Survey and the Census Bureau. But it has, throughout its life, thus far, retained its original essential features. It is a multistage, highly stratified probability sample of land segments, and of the persons resident in those areas.

A minimum amount of coding and processing of questionnaires is carried out by usual manual methods. The data are then doc-sensed, and, via punch cards, placed on magnetic tape. Most further processing is accomplished on an electronic computer.

Types of Computer Processing Discussed

The illustrations of computer processing to which we call attention in this paper may be considered as coming from four classes or functions of statistical reduction of data: (a) Editing and Coding, (b) Estimation, (c) Tabulation, and (d) Special Analyses.

Editing and Coding

It has been found convenient to describe the survey record for a sample person in terms of the punch cards which are used to transcribe that record. These are a household card; a person card; a (health) condition card for each reported illness, injury, or impairment; and a hospital card for each episode. The purpose of the Editing and Coding is to clean the data and prepare them for estimation—they are to be converted from raw data into a more efficient input of higher quality. This is a general process which is as old as the concept of tabulating statistics, but only in recent years has there been both full recognition of the fundamental importance of the editing steps, and the capacity to carry out extensive editing at tolerable costs. Computers have made the execution possible and have sharpened awareness of the value of particular editing steps.

One series of runs or passes through the computer cleans the deck of intolerable data, edits fields to code range and to item consistency, and transfers household and person characteristics to appropriate person-, condition-, or hospital-record. In the same series, experiences from the condition- and hospital-records are summed and entered into the proper person-record for later use.

Among the editing processes is a programming technique, which is a type of indirect addressing, and which is used to recode each health condition and hospital episode from the detail of the International Classification of Diseases (ICD) to a broader classification of 280 groups—and then still further to groups of 38, 30, and 11 classes. Concurrently, with this recoding a total of 206 subroutines are used to edit data on diagnosis and on hospital procedure for internal consistency and plausibility.

This is accomplished by the computer examining the ICD code for validity. If found valid the number itself is reshaped by a series of mathematical manipulations into an instruction used to address memory. The generated instruction directs the computer to examine the content of the appropriate memory positions. The numerical structure of the three digit number found determines whether it is a recode of the ICD and is put into the record, or whether through additional modification becomes a transfer instruction that causes a branch to the appropriate subroutine for the prescribed edits.

To illustrate, suppose the condition ICD 611 (prostatitis) was reported. The computer would use the ICD number 611 to generate an instruction that would cause it to branch and examine the data at address 453. An inspection of the 3-digit number found at this address would determine if this number is to be used to recode ICD to one of the categories in the 280-group recode or used to cause the computer to branch to a routine that would examine the sex, age, and type of condition fields.

Since prostatitis is a disease that is chronic and can occur only in males, age 25 or over, in this case the 3-digit number or word will direct that a sequence of age, sex, and chronicity tests be made. If the case meets the tests, it will be recoded 184 as a prostate disease. For the other possibilities it will be recoded into more general categories.

At the end of each pass the computer types on the console printer the number and kind of edit changes that have been made. A supplementary listing identifies the person for which a change was made, and thus permits a human review of the original schedules and consequent corrective action, whenever the volume of editing changes of a particular kind appears to warrant such action. At the current level of editing, the volume of computer changes in data is quite modest, the annual rate being slightly over 1,000 changes for something like 13,000,000 pieces of data. Of the changes, about 250 are modified diagnoses.

Estimation

Not so many years ago estimation from probability samples was restricted largely to inflation of sample data by the reciprocal of the sampling fraction. This was followed by methods which used the simpler forms of stratification and ratio estimation. Estimation continued to be—for large-scale operations—a relatively uncomplicated procedure, partly because theory was perhaps unimaginative, but partly too, because more complex estimators

Figure 2.

Fiscal 1961
Quarterly

PREVALENCE OF TOTAL CHRONIC CONDITIONS CLASSIFIED BY VARIOUS MEASURES OF SEVERITY AND DAYS OF RESTRICTED ACTIVITY, BED DISABILITY AND WORK OR SCHOOL LOSS ASSOCIATED WITH THESE CHRONIC CONDITIONS ACCORDING TO SEX AND AGE

Sex and Age	Number of Chronic Conditions						Number of Days		
	Total	Total Medically Attended	With bed days in the year	With Restricted Activity Days in 2 Weeks	With Bed Disability Days in 2 Weeks	With Days Lost From Work or School in 2 Weeks	Of Restricted Activity	Of Bed Disability	Lost from Work or School
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Both Sexes									
Under 15									
15-24									
15-16									
17-24									
25-44									
45-54									
55-64									
65-74									
75+									
Male									
Ditto Age									
Female									
Ditto Age									

1. Include all chronic condition cards with "1" in column (aa)
2. Medically attended - column C code 1
3. Restricted Activity Col. G
   Bed Days Col. H
   Work & School Loss Col. I & J
4. Cols. 1-6 x basic weight
   Cols. 7-9 x 6.5 weight

$$S_2 \times A_7 \left[(T + SP_1 + DBY + DRW + DBW + DWSW)_c + (DRW + DBW + DWSW)_d \right]$$

designed. A typical illustration of such a table is shown here as figure 2. Those symbols which are not self-explanatory in the table are code terms which are a shorthand description of NHS tables.

In 1961 approximately 775,000 separate cells of tabulated data were assembled for this project with a total running time of about 700 hours.

Special Analyses

Clearly there is no limit to the possible special analyses which might be undertaken when one has many thousands of items of data on tape and the power of a computer with which to relate some of those data to others. The most important point we could make here probably is just that: the variety of possible analysis is restricted only by one's imagination, and that extensive actual analysis becomes feasible because the computer can make millions of comparisons and computations on individual items in a matter of minutes.

We shall restrict our discussion to a single illustration. This concerns a matter on which at NHS we have expended a good bit of energy, but feel we have just scratched the surface of the work that needs to be done.

Among the questions in the interview is a series which asks whether the person had a hospital episode within the past year, and, indirectly through admission date and duration of stay, whether and when a hospital discharge occurred. Regular processing of these data produces an estimate of hospital discharges experienced over the year previous to interview by the population living at the time of interview. Except for discharges within that period for persons no longer alive at the end of the period, this estimate ought to be equivalent to total number of discharges from hospitals annually. For appropriate definitions of hospitals and of discharges, no very precise count of this latter figure is known from any source outside the NHS. Informed speculation did lead to a preliminary guess, however, that the NHS tabulated data on number of discharges showed a total which likely was 15-20 percent smaller than the unknown true count would be. A pilot study of hospital experience of decedents in their last year of life[2] suggested that decedents might account for something of the order of one half of these 15-20 percentage points. This left an unexplained deficiency possibly in the neighborhood of 7-12 percent. Several plausible hypotheses might "explain" the difference. One of these postulates a recall-loss or memory decay which increases with distance between the occurrence of an event and the time of interview. This notion has a certain intuitive plausibility, and has received attention from a considerable number of people, but published findings are quite limited. (See e.g., references 3, 4, and 5.) A more precise statement of this hypothesis would recognize recall-loss as the resultant not of pure memory failure but of a complex of factors including such elements as differential evaluations over time, financial impacts, and a variety of motivations.

For NHS data, it is possible to calculate for each reported discharge the interval between discharge and date of interview. A full description of this process requires several pages, and is not necessary to the present account. It is sufficient to call attention to several key features: (1) All dates involved, some of which were originally expressed in days, some in weeks, and some in months, must be translated to a common scale to permit direct arithmetic operations; (2) persons which had reported hospitalization, none of which upon edit was in the year previous to interview, should be excluded; (3) episodes of persons who were still in the hospital on the Sunday night before interview should be excluded, since they did not constitute discharges; (4) provision must be made for treating instances of "information not available" for each item of original entry; (5) several classes of case can be distinguished on the basis of relationships among admission dates, length of stay, and interview date; since no single method of computation is best for all classes, a routine was introduced which chose the exact form of calculation for discharge date according to the class of the case. For those interested, an outline of the procedure is given in figure 3.

This matter of interval between discharge and interview was given attention because of the following hypothesis. It ought to be true that, within sampling error, approximately equal numbers of hospital discharges occur in the first month prior to the interview date, in the second month before interview date, the third month, or during the k^{th} month before interview for k, say, not greater than 20. If the NHS data could be tabulated to show the number of hospital discharges reported by length of interval between discharge and interview—called henceforth, the J-interval—the relationship between the a priori hypothesis of level experience and reported data could be analyzed.

The topic is an intriguing one. This paper might easily have devoted its full allotment of space, and more, to discussion of associated theory. But that is not the function of the present forum. Accordingly, we shall treat the subject in terms of a very simple model. Our purpose is not to argue necessarily that this is a best choice of model, but to illustrate how the computer aids in solving the problem of this model.

Assume that for a specified population there are $\bar{D}$ hospital discharges in a unit time period, say a lunar month or 4-week period. $\bar{D}$ is assumed to be constant. Let q_x be the probability that a discharged person dies x lunar months after the discharge—for this analysis, the lunar month is considered an indivisible time unit; q_o is the probability that the dischargee is dead at the time of discharge. The quantity r_J is the rate of reporting discharges by living persons for a specific J-interval—i.e., J lunar months between discharge and interview. The ratio $r_j = 1$ represents perfect reporting; r_j less than unity represents net underreporting; r_j greater than unity represents net overreporting. With this formulation, the reported number of hospital discharges for a given J interval is:

$$D_J = \bar{D} \left(1 - \sum_{x=o}^{J} q_x\right) r_J .$$

The quantity in parentheses is the probability that a dischargee is still alive J months after discharge. This quantity is necessarily equal to or less than unity and is monotonically decreasing for increasing values of J. If the recall-loss hypothesis holds in its simplest form, r_J is also monotonically decreasing, and at $J = 0$ is equal to unity. Under these assumptions, D_o is equal to the number of discharges of live persons in a month, and D_J is a monotonically decreasing function of J. The 2nd order power series approximation to the D_J function might be written

$$D'_J = a_o + a_1 J + a_2 J^2$$

and the estimated level monthly volume of discharges as

$$\bar{D}' = \frac{D'_o}{1 - q_o} .$$

Using the methods sketched earlier, the computer calculated from observed data, the quantities D'_J for each value of J, for J = 1 to 13. (Actually the first computer calculations were made for J in terms of days; they were assembled into lunar month terms manually, but it would have been a simple matter to have converted to lunar months on the computer.) Calculations were made separately for each of

Figure 3. Outline of Calculation for Interval between Discharge and Date of Interview

Part A. Initial Data and Notation

X: = 1 if case has 1 + days in hospital in previous 12 months

= 0 otherwise

Y: = 1 if still in hospital on previous Sunday night before Census date

= 0 otherwise, including D.K.

A': = month and year of admission

A: = A' converted to a coded day, with June 1956 coded zero; July 1956 coded 1, etc. In essence, the code assumes that all admissions occur on the 15^{th} or 16^{th} of the month. Thus an admission of August 1957 is coded A = 365 + 31 + 15 = 411 [Coding done by table-look-up process.]

C': Week of interview

C: C' translated into day code. Table assumes interviewing takes place on Tuesday, which is median day. [Also by table-look-up process.]

D: "12 months before interview week" is calculated in computer as C minus 367

E: "Number of hospital days in previous 12 months," is reported on questionnaire

F: Derived Date of Discharge (See below)

J: Derived interval between discharge and interview date: J = C - F

Part B. Computer Operations

(a) Select all persons for which X = 1, and base subsequent work on hospital cards for these persons

(b) Translate date of admission A' into coded date A

(c) Translate date of interview C' into coded date C

(d) Test Y: If Y = 1, set F = C - 500, J = 500

Y = 0, skip to (e)

(e) Calculate D = C - 367

(f) Test A:D: If A ≤ D, D + E = F

If A > D, A + E = F,

[But if "still in hospital" is unknown, or if A' or E is unknown, then F = C - 600, and J = 600]

(g) J = C - F

(h) Put J and C into record.

2 years' data. Derived values of D'_J were obtained by fitting the 2nd order curves to observed data by usual least squares methods. In these calculations, the value $q_o = 0.035$ was used, which is consistent with other evidence on proportion of dischargees who are dead on discharge.[2, 6]

Of only incidental relevance to the present account, but of possible interest to readers are the following results. In both years the 2nd order curve gave very good fit. In 1958 variance about the regression curve was but 9 percent as large as variance between months, and in 1959 variance about the regression curve was only 7 percent of between-months variance. Using 13 $\bar{D}'$ as the "adjusted" estimate of total annual discharges, the process implied that the total deficiency in reported discharges—including the omission of discharges of persons not living at time of interview as a "deficiency"—is 15 percent for both 1958 and 1959 data. (A preliminary calculation for 1960 data yields a figure of 17 percent.) Thus these results are not inconsistent with the speculations regarding totals using external data, and may well be more precise.

The value of the computer calculation as contrasted with other possible modes of calculation is emphasized when one looks beyond these first global results. Both a priori judgment, and other exploration[7] suggest that the shape of the curve $D'_J = f(J)$ might vary for differing age-sex classes, or length-of-stay groups, or other categories, The computer programming makes it a simple matter to tabulate D_J values for any age-sex, length-of-stay, or other subclass which is identifiable in the case record. This is not the place for discussion of substantive results, but we do note that values have been computed for a considerable number of such subclasses, and are being studied. They apparently contain a good bit of information; for example, the rate of decline or dropoff in reported discharges with increasing J-interval is sharper for hospital stays of very short duration than for stays of longer duration—a tendency that might have been expected.

Other avenues are immediately suggested by these results, including multivariate analysis of the initial figures, and a considerable variety of schemes for adjustment of reported data in the estimation process. One step into these areas has been taken by the NHS in its official publications. Study of the shape of the D'_J curves has led to the view that data on discharges reported to have occurred within 6 months of the date of interview are considerably superior to those reported for the 6-month period which lies in the range of 6 to 12 months prior to interview—although it does not necessarily follow that a 6-month reference period in the questionnaire would be an improvement over a 12-month reference period. In line with this view, hospital estimates for 1959 and 1960 are being based only on discharges reported for the most recent 6 months of the 12-month reference period in the questionnaire. Computer processing of basic data has made this change a low-cost by-product of the J-analysis operation. The use of the 6-month period produces an estimate of total number of discharges which from 1959 is 8 percent larger than the 1-year reference period would yield, and for 1960 is 10 percent larger. Quantitatively this procedure would appear to have removed something over one half of the total deficiency. In conjunction with data on discharges for decedents, the process might remove quantitatively almost the entire deficiency. Whether the resultant figure is completely unbiased may not be certain.

We shall say no more here about this particular problem. It was introduced only to illustrate the flexibility and adaptability of computer processing to nonroutine types of analysis. We do wish to reemphasize our belief that the long-range benefits of computer applications in social statistics will be found not so much in reduced costs or in increased speed, but rather in the twin characteristics of methodological strength and flexibility which make feasible a more intensive reduction of data and more perceptive analysis.

REFERENCES

1. Alexander Gray: *Introduction to the Development of Economic Doctrine.* Longmans, Green & Co., London, 1934.

2. U.S. National Health Survey. *Hospital Utilization in the Last Year of Life.* Health Statistics, Series D-3. PHS Publication No. 584-D3. Public Health Service. Washington, D. C., January 1961.

3. Percy G. Gray: The Memory Factor in Social Surveys. *Journal of American Statistical Association.* Vol. 50, pp. 344-363, Washington, D. C., 1955.

4. Robert Ferber: On the Reliability of Responses Secured in Sample Surveys. *Journal of American Statistical Association.* Vol. 50, pp. 788-810, Washington, D. C., 1955.

5. R. K. Som: *On Recall Lapse in Demographic Studies.* International Population Conference (Proceedings, p. 50). Vienna, 1959.

6. Guide Issue. *Journal of the American Hospital Association.* Chicago, 1957.

7. U.S. National Health Survey. *Reporting of Hospitalization in the Health Interview Survey.* Health Statistics, Series D-4. PHS Publication No. 584-D4. Public Health Service. Washington, D. C., May 1961.

II

PANEL DISCUSSION - REPORT OF THE TECHNICAL COMMITTEE ON BROADCAST RATINGS

Chairman, Andrew Stevenson, Committee on Interstate and Foreign Commerce, U. S. House of Representatives

Panel:
- William G. Madow, Stanford Research Institite
- Raymond J. Jessen, C-E-I-R, Inc.
- Bud Sherak, Kenyon & Eckhardt, Inc.
- Hyman H. Goldin, Federal Communications Commission
- Jay Eliasberg, CBS Television Network Research

SOME REMARKS* ON THE REPORT OF THE TECHNICAL COMMITTEE** ON BROADCAST RATINGS ENTITLED "EVALUATION OF STATISTICAL METHODS USED IN OBTAINING BROADCAST RATINGS"

By William G. Madow, Stanford Research Institute

1. Introduction

On March 11, 1960, the Hon. Oren Harris, Chairman of the then existing Special Subcommittee on Legislative Oversight of the Committee on Interstate and Foreign Commerce and also Chairman of the latter Committee, in a letter to Dr. Morris H. Hansen, then President of the American Statistical Association, included the following paragraphs that summarize very well the motivation, objectives, and conditions of work of the Technical Committee on Broadcast Ratings organized by the American Statistical Association in response to Congressman Harris' letter:

"...the Subcommittee recently has had occasion to consider the existing statutes and regulations, or absence of them, applicable to the character of the programs which licensed radio and television stations are broadcasting over the public airways. It appears from the testimony that the choice of the kind of program broadcast over networks during prime viewing hours has often been predicated upon public acceptance of preference as indicated by certain 'ratings' ascribed to programs by certain 'rating services.'

As it is clear that the determination of any such ratings must be derived from statistical procedures involving sample surveys, our Committee has requested you to arrange for an examination and evaluation of the statistical methods used by the principal rating services. It is my understanding that you have taken this request up with your Council and that they, and you, recognizing the public interest and professional responsibility involved, have agreed to designate a group of scientists which would make an independent study for us. I understand that the group would act in its own capacity, being free to present its findings without your prior review or that of our Committee." (Report, p. 1)

*Prepared for a panel discussion of the report at a meeting of the American Statistical Association, December 27, 1961.

**The Technical Committee on Broadcast Ratings, a Committee of the American Statistical Association, consisted of William G. Madow, Chairman, Stanford Research Institute; Herbert H. Hyman, Columbia University; and Raymond J. Jessen, General Analysis Corporation (now CEIR, Inc.); assisted by Paul B. Sheatsley, National Opinion Research Center, and Charles R. Wright, University of California, Los Angeles.
The report of the Technical Committee on Broadcast Ratings was published as House Report No. 193 of the 87th Congress, First Session, entitled, "Evaluation of Statistical Methods Used in Obtaining Broadcast Ratings," A Report of the Committee on Interstate and Foreign Commerce, Oren Harris, Chairman. All references are to the report.

The Technical Committee on Broadcast Ratings came into existence late in March, 1960 and submitted its report early in March 1961 to the American Statistical Association, which in turn transmitted the report to Mr. Harris. The Technical Committee, its task having been completed, ceased to exist in June 1961.

Seven rating services were considered: American Research Bureau, Inc.; C. H. Hooper, Inc.; A. C. Nielsen Co.; The Pulse, Inc; Sindlinger and Co., Inc; Trendex, Inc.; and Videodex, Inc. Without exception, they were helpful and cooperative.

Although the report is also concerned with radio ratings this paper will, for simplicity, consider primarily television ratings. This paper is intended to summarize my views on some aspects of the report.

2. The Setting of the Study

The quotation from Congressman Harris' letter makes it clear that while the desire to have an examination and evaluation of the statistical methods used by the rating services in estimating the ratings was motivated by problems of programming policies, the task of the Technical Committee included neither the examination nor the evaluation of programming policies, nor was the Technical Committee asked to estimate the ratings that would be received if different programs were available. Our responsibility was solely to examine and evaluate the statistical methods used by rating services in estimating the ratings. No limitation was placed on the extent of our examination of statistical methods.

As remarked above, the rating services cooperated with us fully throughout the study. I am sure they often wished that we might merely use material that had been given others and not press them further. However, they did cooperate. Whatever omissions there may be in our knowledge of their statistical procedures are, I am sure, due to our not inquiring rather than to their not being willing to provide the information.

There was one aspect of this study, however, that the Technical Committee felt to be important and for which we turned, not primarily to the rating services, but to those who are their actual or potential clients. For what purposes are ratings used and what ways are the ratings used?

To evaluate ratings by some absolute criterion such as, for example, a standard error of less than so and so taking into account all possible sources of error, might be desirable but it might also be that relations among ratings suggest conclusions even when standard errors are not available. Requirements on data are relative to their uses just as requirements on any communication system are relative to the uses to be made of that system. In practice, the uses of data depend on much more than the data themselves--and the requirements on the accuracy of data are naturally affected by both the importance of the decisions they influence, and the extent to which they influence the decisions. I am glad to say that, despite occasional grumbling, we received a great deal of cooperation from those sponsors, advertising agencies, networks, and broadcasting companies to whom we wrote about the uses, and I am sure that had we had more time and been able to accept some of the invitations we received to visit with those who replied, we could have obtained more information.

The Technical Committee was not asked to investigate, evaluate, or compare the rating services with one another. In the study, we tried to avoid making comparisons of the rating services. I should like to caution against using parts of the report to compare the services with one another. We did not write or review the report in such a way that the statements we made about different rating services are comparable, except perhaps for the populations covered in rating surveys. Each rating service was considered within its own framework. Thus, a service that attempted little might have less criticism in the report than a service that undertook to and did accomplish more, but in so doing provided more occasions for criticism. This should not be taken to imply that the rating services are, in our judgment, equally good sources of estimates of ratings.

I should also like to stress, as we did in the report, that there may be errors in some of the details we give concerning the practices of the services. While the errors are not important for our evaluation, they obviously may lead to erroneous conclusions if used to judge or compare rating services.

When we began this study, the Committee had hoped to submit its report in draft form to the rating services for their comments before the report became final. For various reasons, including the fact that neither the American Statistical Association nor the Congressional Committee was to review the report prior to publication, it was deemed best at the end of the study to issue the report without comments in advance from the rating services.

The report would certainly have benefited from being submitted to others, including the rating services, for comment. We did ask the rating services for comments after the report was published. Those that we received did not disagree with our conclusions although, as intimated above, there were some errors in details and there is some feeling that the report may appear to have been more critical of one service than another at various points. On the whole, the comments have not been unfavorable.

It is worth noting that the rating services are well aware of the defects in their surveys and that they seem to inform their clients. My impression is that they face severe pressures in the timing and costs of their surveys, and it is these pressures rather than lack of interest or knowledge of improvements that lead to the existence of the defects to which we call attention. Also, my impression is that many of their clients are well informed of the possible defects of rating surveys, and use them with these in mind. Again, I want to caution against any implication that all rating surveys are equally good or bad.

In the following remarks I have selected some of the major aspects of the report for comment. Further information is given in Dr. Jessen's paper. Details will be found in the report.

3. A Summary View of the Uses of the Ratings in Relation to the Statistical Methods Used by the Rating Services

To obtain material on uses, we wrote to some sponsors, some advertising agencies, some broadcasting stations, including chains, and to the three major networks. No sample was selected since we were primarily concerned with whether a picture of uses emerged that could be synthesized.

Some parts of the industry claim that ratings are not the major factor in their decisions but only a portion of the evidence they use, some parts claim that ratings or cost per thousand are a major factor if not the sole factor, in decisions affecting them. Both statements are correct; they apply to different parts of the industry. But they can be effectively combined for our purposes in the statement: Other factors being equal, the program having highest rating, or lowest cost per thousand, is preferred. The other factors considered vary from none to some that are possibly more important in certain cases than the ratings themselves.

Omitting non-numerical factors in preference, the other numerical variables most often cited besides ratings, cost per thousand, circulation, share of the audience, coverage, and so on were the composition of the audience by sex, age, income, geographic location, and other demographic variables. Many users stressed the importance of trends as well as level of ratings*.

Another way of looking at the ratings is that the organization that makes many small decisions, for example, an organization that buys many spot announcements, will be likely to base the decision primarily on the major readily available data such as ratings and cost per thousand. If the organization deals with smaller markets, it cannot obtain trends very easily; the surveys aren't made often enough.

The organization that makes large decisions spends more on each and may bring in many factors such as, for example, balance of programs, strategy, image-product compatibility, and the results of special surveys. It does not rely so heavily on ratings, and cost per thousand, and trends in them as the organization that makes

*We shall use the word "ratings" to refer to the many different measures cited above.

many small decisions. But no organization ignores ratings, cost per thousand, and the other data cited above.

To summarize, ratings, cost per thousand, and related data including audience composition and trends in these data, are important to all parts of the broadcast industry. The more costly or rewarding the decision, the more that is spent on auxiliary information that reduces dependence on the ratings. Organizations with greater resources, knowing the possible faults in ratings and either depending on large smaples or on many smaller samples for the economic consequences of their decisions, treat the ratings as one part of the information on which they base their decisions. Organizations such as stations or advertisers in small communities have smaller resources and are more dependent on the results of a single rating survey--and therefore on the statistical methods used in a single local survey rather than on those used in national samples or on the average effects of the statistical methods used in many local surveys.

I should like to stress that I am not speaking just about a large enough sample versus a small sample, but about the greater dependence of the small user on a single chance event. Rating surveys are not made weekly or monthly in smaller market areas.

Finally, although some users doubtless use technical statistical, or operations research, or mathematical-model approaches to decision making to some extent, no such methods were mentioned by the users.

4. Over-all Evaluation and Recommendations

The Technical Committee found many details and a few fundamental matters to criticize concerning the statistical methods used by the rating services and the descriptions of methods and quality of the ratings published by the rating services.

Let me first consider two major policy issues that are often raised concerning the ratings and that are directly related to the statistical methods used by the rating services. These issues require an over-all evaluation of the effects of the various criticisms that we made of the statistical methods used by the rating services--and it is to over-all evaluations and recommendations that this section is devoted.

I should like to present my own interpretation of some of our conclusions.

For the big policy question--whether the ratings are sufficiently accurate to reflect the preferences of the audiences nationally for cultural as compared with non-cultural programs from among the programs available to the audiences --the ratings are sufficiently accurate, when, despite the different methods used, they are in agreement, as they seem to be on this issue. If anything, the incomplete coverage of the population is likely to increase the proportion of households claiming to prefer cultural programs, except that at least some services do not report the audience of educational television stations.

A second important statistical issue is whether the ratings are sufficiently accurate for comparing programs of the same type with one another, or stations in the same market area with one another. Ratings of so many programs are published, and the ratings of so many stations are compared, that we can be certain that even though many comparisons of programs and stations are correct, many are incorrect. Many apparent differences of ratings, many rankings of ratings, and many apparent trends in ratings are, in fact, just results that could occur by chance. Where samples are small, as in small market areas, or where ratings are close to one another there is greater likelihood of error.

"Error" tables, including both sampling and non-sampling errors, are needed for the comparison of stations and programs. The Technical Committee felt that the rating services were not publishing or otherwise making available sufficient information to permit satisfactory use of their "error" tables, nor are their "error" tables sufficiently extensive. In the report, we listed various recommendations to remedy what we considered to be inadequacies in this area. The inadequacies were primarily in the omission of estimates of non-sampling errors, and the omission of any treatment of many comparisons rather than individual estimates. To rectify these inadequacies will probably require some research. In addition, the rating services should publish more educational material and guides for the use of ratings, particularly as measures of trends.

In connection with both of the above conclusions, we should like to point out that the statistical defects of the rating surveys are likely to have a much more serious effect on the so-called qualitative information, namely, age, sex, size of family, income level, and other demographic characteristics, than on the ratings themselves. In view of the frequent statements of increasing reliance on these qualitative characteristics we feel it important to call attention to this danger.

These are two "big" conclusions from the Technical Committee's report. Now, what did the Technical Committee propose be done that would improve the situation?

The Technical Committee felt that even though the ratings as currently made were useful for many purposes, there was much that could be done to improve the ratings and the understanding of how to use ratings of individual programs and stations. Also, it seemed clear that in view of the frequent changes in the kinds of ratings issued and the new problems and uses of ratings that would occur, the interest in the quality of ratings should be permanent.

One source of improvement in the ratings is competition among the rating services themselves. But we felt that improvements might be more rapidly made if certain recommendations were adopted. Three major recommendations were made. These are given in the report in greater detail, but are restated here:

(a) The rating services should increase the information they furnish on how they estimate the ratings and related data and on measures of the quality of the ratings and related data, and they should regularly publish such information. Obviously, as the rating services increase the information on quality, their clients will have greater means of understanding the quality, improving the uses made of ratings, and achieving the increases in quality required by those uses. "Error" tables should be constructed for the demographic characteristics of ratings as well as for the ratings themselves.

(b) The rating services should not only do research, but should publish research on estimating and using ratings. The services now do methodological research, and they make some research available to their clients, but it would improve the quality of ratings and their uses if methodological research were published.

(c) The broadcast industry should develop an industry-wide Office of Research Methodology--either to conduct or support a program of research in making and using audience measurements of quantitative and qualitative types.

My own summary is that we thought the ratings on the whole are useful but could and should be improved, and we suggested some approaches that should result in improved ratings. How the individual rating services respond to our recommendations will depend on the pressures on them and on the economics of the situation. I think we suggested how the broadcast industry can proceed so that it can expect steady improvement in the estimates, uses, and understanding of the ratings.

Let me now turn to some of our specific criticisms.

5. The Statistical Methods Used in Estimating Ratings

Different definitions are used by the rating services in estimating ratings. None is necessarily better than the others and the variability resulting from differences in definitions is probably not a major cause of differences among ratings.

However defined, a rating of a program is usually the percentage of households with television sets in a specified area who were part of the audience of that program.

In putting such a general definition into practice many variations occur that affect the measurement.

The area might be the United States or a metropolitan area, or a market area, or an area served by television stations on a coaxial cable permitting the simultaneous broadcast of the programs.

Being in the audience may mean that the set was tuned to the program, or that at least one person in the household was viewing or listening for at least so many minutes. The audience may be the average audience say, per minute, or the total audience in the time period.

Ratings may be based on households and persons satisfying socioeconomic criteria, such as being in an income class or being teen-agers, as well as on all households or persons.

For purposes of the present comments, it will be desirable to omit consideration of the many different definitions of ratings and of related measures such as share of the audience and coverage. It is important, however, to recognize that differences in concept, area, and population surveys and in measuring instruments all contribute to the variability of estimates in addition to sampling.

Our primary criticism of rating surveys is in the difference between population to which ratings are applied and the populations from samples of which the services actually obtain information.

We do not have data with which to estimate the biases in ratings and related data resulting from this cause. But, it is probably important and should be estimated, at least from time to time. These biases should be discussed or approximated in connection with "error" tables.

We found that all the methods in use at the time we made our study about a year ago, resulted in effectively taking samples from between 50 and 70 percent of the population.

If diaries or meters are used, then there are difficulties in obtaining cooperation, and not all of those who agree to these methods actually produce usable data.

If telephones are used, then only about three-fourths of the households have access to a telephone and the proportions with access to a telephone differ considerably in various geographic and socioeconomic categories. Furthermore, some households with access to a telephone are not listed in telephone books. Telephone calls to toll-call areas may be omitted, or call-back telephone calls to such areas may not be made.

When personal interviews are made, there may be no call-backs possible without being involved in a recall period of more than 24 hours.

Now the audience characteristics of those who provide data in these surveys may well differ from those who do not. We know that for some of these methods their demographic characteristics do differ (Report, Ch. 4), and we believe that such differences exist to some extent for all methods.

While the requirements of rapid reporting may lead the rating services to continue to sample from these pseudo-populations, they should at least from time to time make surveys to determine the effects of their use.

All the rating services are aware of probability sampling methods and all use them--or at least use methods that should be equivalent to probability sampling--but I felt that they tended to cut corners more than desirable. In particular, there is some tendency to omit "error tables" or to use "error tables" that do not correspond to the sampling methods and estimation equations. Not all rating services are equally to be criticized. At least one has made great efforts not only to use probability sampling but also to have its "error tables" correspond to the procedures it uses. But there is much for the rating services as a whole to clean up here--even though, as mentioned above, we feel that the ratings are adequate for many of the uses to which they are put.

Another major criticism of the rating services is in the lack of objective evidence of quality control of the process of obtaining information. Respondents, enumerators, and data processors all make errors. They make them often, and the errors may be serious. While the cost of estimating the biases and variability due to such errors may be too great for a single survey, the rating services are continuously collecting data, and are using methods subject to error. However carefully they are organized, there is no substitute for the planned study and reporting of biases and variability arising from human or machine errors.

6. Summary

It is certainly unfair to some rating services to generalize about all of them. And obviously only a small part of the report has been discussed here.

My own summary is something like the following:

If you don't like the major programming policies, don't blame the ratings. If anything, I would expect the preference picture to be worse for cultural programs than the ratings show.

If you think that the rating services should use larger national samples, well, any sample is better when larger if no loss in quality occurs, but it will not have much effect either on preferences of types of programs or on major programs. Current national sample sizes, although not too large, are not too small.

If you think that the sometimes erratic behavior of local market area rating surveys is evidence of unethical behavior on the part of the rating services, well, we didn't investigate them. But even without any unethical behavior, there could be large errors in the data from some of those surveys--errors that might balance out over many stations or areas, or over time.

If you think that the above statements imply that we approve the statistical methods used by the rating surveys, please note that we think we have made some severe criticisms and some recommendations that we feel the industry can and should take into account.

Like most of us, the rating services have been doing a job good enough for many needs, but they could do better. We have tried to suggest steps that would ensure their doing so.

THE RATING COMMITTEE REPORT: A PANEL STATEMENT

Raymond J. Jessen*, C-E-I-R, Inc.

Fortunately, or perhaps unfortunately, much of what I have to say on Broadcast Ratings is contained, together with those of my colleagues, in the report of our committee which was submitted to the American Statistical Association last January and published by the U. S. Congress as House Report #193. Today, in order to help get a discussion going for this panel I shall make some remarks bearing on the following 4 points:

1. A brief summary of the complaints on the broadcast ratings and our reaction to them

2. The Public Interest versus the Advertiser Interest in Ratings

3. Requiring clear Labelling versus setting Standards

4. Need for methodological Research

I. Summary of Complaints on Ratings and our Reactions to them.

Sifting through the complaints that one hears about the ratings the following three rather broad reasons seem to have some bearing on how or why the complaint may have arisen.

(1) The ratings are not really what they pretend to be,

(2) The ratings are what they pretend to be, but are inaccurate and

(3) The ratings should be something more than what they pretend to be.

In the first group the complaints appear to be concerned with the matter of coverage, both that of time and space. Some rating services use or have used a specified period such as a week, a/ for example, during which their data were collected but in the report may imply or declare that the rating applies to some longer period of time such as a month or a quarter. If a broadcaster has exclusive knowledge of when a rating period is scheduled he has opportunity to offer special attractions during that period (known as "hypoing") to make a better showing than he would otherwise. But even if all broadcasters know when the rating period occurs then disputes may arise as to the "representativeness" of whatever period is selected. However this practice of basing ratings on short periods when longer periods are presumed to be represented appears to be used much less now than several years ago - and therefore this problem is disappearing.

In addition to the complaints which arise because of restricting the time period covered by ratings, are those that arise because of restrictions on the area or the segments of the population covered. Some ratings for example fail to include such things as: certain outlying portions of a market or broadcast area, persons not-at-home at the time interviewers call, certain foreign speaking homes etc. Usually these omissions are for economic reasons but the usefulness of the results is so seriously impaired thereby that the economy may well be false.

Several of the services could improve the understanding of the appropriate period of time and to what universe for which their results were applicable. The use of a "survey week" to represent the average goings-on during a month not only leads to fruitless arguments on how well it represents some longer time period but is wasteful from a sampling point of view. Confusion on just what universe is being dealt with is also not difficult to eliminate if the service will clearly present it in his report. However the effects of the omission of homes because they are too difficult to reach because of distance, language, etc., is difficult to assess without considerably more evidence than is now available.

The segments of homes omitted when relevant by nearly all the services are those in which a respondent is not-at-home at time of call, those not listed in a phone book, and those which refuse to cooperate. Regardless of method used, these groups appear to average about 40% of all eligible homes - a rather large group on which to have no viewing or listening information. In interview surveys, this group could be reduced substantially if call-backs were made; but in diary and surveys the non cooperators are a difficult problem and no direct approach seems feasible.

The second group of complaints seem to arise not because the ratings are erroneously conceived but because they are inaccurate. Examples of these complaints are: "hypoing" (where calculated misdeeds are attributed to the broadcaster), fraudulent results (where calculated misdeeds are attributed to the rating service), technical faults and weaknesses attributable to the ignorance or carelessness of the rating services and to sampling variation, which might be blamed on the fates as well as on the rating services. Included in this class are such charges as, "small samples cannot properly represent the diversity of human behavior in this county", "scientific sampling may be all right for determining such things as the number of refrigerators or garden tractors but dealing with people is another matter", "I know of interviewers who just sit in their homes and fill in the answers!", "Rater Q gave a rating of 10.2 to a program that didn't exist!", etc.

Even if its purpose is crystal clear, there are a number of inaccuracies in almost any survey. These inaccuracies or deviations from the true are generally classified by survey statisticians into the two types: (i) sampling deviations and

*Member of the Technical Committee on Broadcast Ratings

a/ Trendex uses 1 week per month. Pulse dropped the single week period about three years ago. Hooper uses 1st week in month. Sindlinger is continuous.

(ii) non-sampling deviations. The likely sizes of deviations due to sampling can be determined if the type of sample is appropriate (such as random, for example) and its size is known. Some of the rating services provide the means for making this determination, some don't. In the case where clusters are used at some stage in the sampling some of the services fail to provide the user with information on how this feature affects sampling precision. There are some instances when the effect of clusters must be quite large and therefore seriously misleading if only the number of homes sampled is given.

Non-sampling deviations arise from errors in the original data, failure to obtain data from all designated homes in the sample, processing errors, errors inherent in the method of measurement, etc. These deviations in ratings are a more serious problem than those of sampling because they can not be made determinable through correct statistical design and generally require special studies for their detection and measurement. Although the raters presented some studies on this problem, they were of little help for judging the effects of such deviations on the accuracy of the ratings.

In the third group the complaints arise because the complainant seems to expect something more than that which the trade ordinarily understands to be a rating. The complainant takes a dim view of the utility of "nose-counts" and demands ratings that measure the "quality" of the audience, the satisfactions it receives, its advertising or commercial effectiveness, etc.; or perhaps the program's effect on the morals, culture and well-being of the people so that "public interest" can be more adequately evaluated. Also it is felt by some that the needs and desires of those <u>not</u> in the audience might also be included in some kind of "rating".

I am inclined to be sympathetic with these complaints, and agree that there is a need for statistical measures other than "nose-counts" so we can determine the nature and feelings of the audience as well as its size. However the ratings in their present form can and do serve useful functions. But this does not mean we should not be looking for something more.

II. The Public versus the Advertiser Interest in Ratings

There is a basic difference in the needs for ratings among the different interest groups in and out of the broadcast industry. Advertisers are mostly interested in some simple measure of the effectiveness of broadcasting as a means for selling their products. A program is a good program and a station or a network is a good station or a good network if it does an efficient job of stimulating the advertiser's sales. Since advertisers usually use a number of different media at the same time and since the advertising effectiveness of any particular campaign may be indirect as well as direct, it is very difficult to determine the sales effects of any broadcast program; or even for any part of the firm's overall advertising program. In the absence of these unavailable measures of effectiveness the advertiser depends on the use of audience measures such as size, type, etc. - measures of characteristics which he believes are related somehow to his actual or potential sales. Thus broadcast ratings give the advertiser a means for evaluating the expected effectiveness of television, newspapers, magazines, etc. for his advertising dollar. They may not be very good for this purpose but they <u>are</u> available. When the advertiser feels the ratings are not doing a good job of measuring advertising effectiveness he may be inclined to complain about their inaccuracies and other inadequacies.

On the other hand there is a group which has little interest in the matter of advertising effectiveness but rather is concerned with what it regards as the "public interest". People in this group are generally interested in how "moral", or "educational" or "cultural" a program might be. Obviously the standard ratings do not measure these things, nor is it very obvious how they <u>can</u> be measured. I personally feel that the problem of finding suitable measures should not be ignored and I would like to see the behavioral and social scientists devote more effort to its solution.

III. Standards versus Labelling.

A fixed set of standards for ratings has a strong appeal to many people in the industry - particularly those who must make many decisions using many different ratings from the different rating services. The simplicity and economy of some industry standards would certainly reduce the burden of evaluating the different ratings as to accuracy and comparability and facilitate communication and understanding among the various persons involved. For example, if everyone used the same "adjustment factor" to translate rating X into rating Y a lot of time and argument might be saved. The industry can do much on its own to bring about some standardization, and this would be facilitated if an industry-wide office of methodology is established.

But there are certain kinds of standards that I feel ought to be avoided, such as for example, setting standard sample size requirements. Since samples of different designs may differ in precision for a given number of homes contacted, a size requirement in terms of a fixed number of homes may encourage the use of inefficient sample designs. Setting size requirements in terms of <u>random sample equivalence</u> may avoid this difficulty, but of course at a cost of some loss in simplicity and understanding of what is happening.

On the other hand if certain accuracy requirements are established, rather than sample size, then there is a danger that they will be too high for some purposes and perhaps too low for others. The rating business is fairly technical and will perhaps become much more so, but it seems inevitable that many concepts and practices in the rating business will become more standardized, I hope that paralleling this development will be an increased use of technically trained people to cope with the more sophisticated, more useful and more efficient concepts which also seem almost sure to come.

In order to avoid some of the undesirable "side effects" of standards, I suggest a more vigorous attempt to adopt the practice of "clear labelling". By this I mean that if rating services would clearly state what their ratings mean statistically, such as the proper projectable universe, the sampling error, the non-sampling error (as well as it can be determined) and of course a statement somewhere of the procedures used, we would have less pressure for standards. By labelling fully and clearly, good practices may have a fighting chance to win over bad practices. And new a more useful approaches will have a better opportunity to find a possible market. Also technical development and technical understanding would be speeded in an industry where progress has been too slow.

IV. Need for Methodological Research.

In the earlier parts of this statement I tried to elaborate on the idea that many of the criticisms of the ratings arise because of lack of understanding of one sort or another. Some are due to failure to understand what the ratings are purported to be. Others are due to a lack of understanding of what the various errors in ratings might be in the practical problem at hand rather than in some very specific and perhaps dramatic instance. Still other criticisms are really appeals for ratings which are fundamentally different than the current ones and perhaps would have only a limited market. And of course a lot of criticism is essentially warfare between the protagonist and antagonist of the different methods used and the houses that produce them.

An industry-wide Office of Methodological Research would do several desirable things. It would be an excellent place to bring about an understanding of the underlying common elements in the rating business, which may be overlooked otherwise. This it could do by what would be essentially an educational program. By preparing and issuing reports on procedures used by each service and the supportable declarations that each gives on the advantages of his procedures. In addition - and this would be the primary function of the office - it would itself carry out or would sponsor research in problems common to the industry. These might be problems of setting standards, of clear labelling, of determining non-sampling errors, of determining suitable concepts for measurement, etc. It would be used otherwise, but I hope it would be used to encourage progress in a very important industry - what with the coming of the computer age and its appetite for more and more data - an industry which will become much more important in the future.

REMARKS ON THE REPORT OF THE AMERICAN STATISTICAL ASSOCIATION TECHNICAL COMMITTEE ON BROADCAST RATINGS

Bud Sherak, Kenyon & Eckhardt, Inc.*

Each of the discussants on this afternoon's panel has been invited to comment on the report of the Technical Committee from the viewpoint of the organization he represents. Therefore, I think it is appropriate if I preface my comments with a few remarks on how and why advertising agencies are concerned with broadcast ratings.

One of the important services of an advertising agency is to develop media plans for its clients. In broadcast media planning the advertising agency screens availabilities, evaluates them according to the clients' needs and interests, and makes specific recommendations. Then, with the approval of its clients, the advertising agency negotiates with the stations and networks and places the order to buy - or cancel, as the case may be.

Years ago, many of the key decisions involving broadcast media were based primarily on judgment. But today this is almost unthinkable. While judgment still plays a large role in decision making, it is almost always based upon and supported by facts about the media. The stakes and risks involved in media decisions are now too great to be left to judgment alone.

Frequently, ratings from the rating services are the only objective data we have to appraise the broadcast media. They are the basic source of our information on the size and characteristics of the audiences of different programs and on the amount of tuning or viewing at different time periods and in different parts of the country.

Ratings are almost indispensable to our work -- and, to some extent, so are the rating services. I think if the rating services were abolished today the broadcast industry would have to revive them tomorrow. The alternative would be to generate our own ratings. And the ratings we would develop and use would not be nearly as accurate as those currently supplied by the rating services. We couldn't afford the expense.

Even with the rating services to defray expenses among their subscribers, the cost of ratings information is extremely high. Just this past year, Kenyon & Eckhardt alone spent over $75,000 for broadcast ratings. And I am sure that this is comparatively small relative to the amount spent on ratings by the networks and stations, whose very livelihood depends on what ratings may indicate about their programs.

In these introductory remarks I have indicated that we are heavily committed to ratings. However, I wish to make it clear that we don't look upon ratings as the ultimate source of answers to broadcast media problems. Actually, for many important media decisions, ratings are inadequate. And there is a growing movement among media research practitioners to supplement ratings with other pertinent information about the broadcast media.

Essentially, what we are trying to do is to develop value functions for all media and media units. This would permit us to answer such questions as: What influence does the program have on the effectiveness of the commercial message? Independent of the size and characteristics of the audience, and all other things being equal, is it better for an advertiser to reach a prospect through a commercial in a variety show or in a situation comedy -- or in a western? What is the value of program loyalty? If listeners or viewers enjoy and appreciate a program, to what extent do they reward the advertiser?

Most of these questions are now answered on the basis of judgment. And in many instances, media planners find this to be entirely satisfactory. However, there are comparable questions on the effectiveness of various media units, that are now answered by judgment, which media planners think should be answered by research.

A television advertiser has a choice of many different media units. In network programming he can use 2-minute commercials in TV Specials, 1-minute commercials in prime-time participations or alternate week sponsorships, back-to-back 30-second commercials, or minute daytime participations.

He can also buy television spots using 1-minute commercials in late night participations or in syndicated films in marginal time, 20 or 30-second station breaks in prime time, 8 to 10-second I.D.s in prime time, or 1-minute daytime participations and station breaks.

The critical question here is what is the relative selling power of these media units?

This is an extremely difficult problem because the value of a media unit depends on many things, such as the nature of the advertised product, its marketing problems and objectives, and the intrinsic selling power of the commercials. Ratings alone can't answer the question.

*Now with Market Facts, Inc. -- New York

There are also questions of cross media comparisons. For example, what are the relative values to an advertiser of a 1-minute commercial in an evening network television program compared to a page advertisement in a given magazine or newspaper? Here again, ratings by themselves are only of little help to us.

I regret to say that most of the research I have seen on this problem has been inadequate also. Usually, the research is based on measuring the extent to which commercials are recalled rather than estimating the effects they produce in building acceptance and sales of the advertised product.

Some companies are attempting to use simulation and linear programming methods to help optimize the allocation of the advertising budget. They use conventional ratings information and estimate the values of media impressions by judgment and consensus.

This more or less sums up where we stand in media research. There is much that we know about media, and much more that we have to know. We also know where our major efforts in research will have to be made in the future.

With this as background, I would like to turn now to the report of the Technical Committee.

From the correspondence in the Preface of the Report we see that the American Statistical Association was consulted because of a concern by the House Subcommittee on Legislative Oversight that the choice of programs available to the public is being influenced by ratings. However, the House Subcommittee did not ask for an investigation of how ratings are used in programming decisions nor did it ask for an opinion on whether the use of ratings was in the public interest. Instead, its assignment was for the American Statistical Association "to arrange for an examination and evaluation of the statistical methods used by the principal rating services".

In my judgment, this assignment was not adequate for the problem that created it.

Whether or not the use of ratings is in the public's interest is essentially independent of whether ratings are judged to be accurate. Furthermore, even if a passing grade on accuracy is a necessary condition for ratings to be acceptable it is not sufficient.

To be sure, most criticisms of ratings usually begin with questions about their accuracy -- specifically, on the size and adequacy of the sample. The criticism cited in the report that ratings assume "if one Republican dentist from North Dakota is listening to Bob Hope, then all are listening" is very familiar to us. In the same vein, we have also heard TV personalities argue that they get fan mail from more people each week than Nielsen has in its entire panel.

But times are changing and people associated with the broadcast industry are accepting the idea that ratings may be fairly good estimates of audience size. However, many still question whether ratings are adequate measures of a program's performance -- even from the point of view of the advertiser.

The argument advanced is that a show can have a poor rating but might still have such loyal viewers that it does a better than average job of selling the advertiser's products. This is more difficult to refute than some of the unwarranted criticisms of the sampling methods used by the rating services. In fact, the argument may be correct sometimes.

I wish the Technical Committee had been given the broader assignment of exploring the entire subject of the use of ratings. This would have been much more valuable to the users of ratings than a study of the statistical methods of the rating services.

Long before the committee's report was published, I think most of the major users of ratings had already concluded that the audience data from the principal rating services were sufficiently accurate for their purposes. The report merely confirmed their judgment.

This is not meant to be critical of the work of the Technical Committee; it is only a criticism of the limitations on their assignment. Actually, considering the complex task of evaluating several different survey methods, each of which is made up of many complex operations, I think the committee did a superb job and should be highly commended for its work. Also, with only a few minor reservations, I agree with the committee's conclusions and support its recommendations.

I was especially glad to see a recommendation that the broadcast industry develop an Office of Research Methodology. I don't know to what extent the rating services would be willing to cooperate with such an organization, but it makes a great deal of sense from a users point of view. Certainly it will make the user feel more secure in accepting the research findings.

It is no secret that methodological research reported by the rating services often seems to prove what the rating services would want it to prove. And I suspect that many users of ratings simply throw up their hands at most of it. This may be grossly unfair to the rating service which may have conducted the research primarily to learn how to improve its own service.

The Technical Committee was apparently aware of this problem, and commented that the methodological research conducted by the rating services "suffers from the defects that it has been made by the services most directly concerned". But at the same time, in Recommendations 4 and 5, the committee urges the rating services to perform methodological research to estimate the size of the sampling biases and response errors in their rating surveys. This seems to be inconsistent with the committee's earlier views on the sponsorship of methodological research.

Under Recommendation 4 there is also a suggestion that the rating services try to increase representation of the general population in their rating surveys to a minimum of at least 80%. The committee estimates that the field yield in rating surveys now ranges from 52% to 62% among the principal services.

To increase this yield by 20 to 30 percentage points seems to me to require a major overhauling of the rating services and probably would increase the cost of ratings to subscribers very appreciably. I would seriously question the wisdom of this additional expense. If we now judge ratings to be sufficiently accurate for our needs, we should look to invest additional research funds elsewhere -- specifically in research that can provide estimates of the relative values of reaching prospects with commercials in different programs and different media units. The same comment would apply to Recommendation 9, though I am in sympathy with the need for improving rating surveys in small marketing areas.

Recommendation 8 on the use of an audience per dollar measure also makes sense to me -- not only because it is unbiased, unlike cost per thousand, but because it makes the return on the media investment much easier to visualize. I think the idea, say, of a program delivering 300 homes per dollar is much more readily conceptualized than the idea of a program costing $3.33 for every 1,000 homes delivered.

Also, the homes per dollar measure will simplify our media arithmetic. It is additive among the various subgroups in the audience so that media researchers can simply multiply the audience size of each group by the reciprocal of the cost of the program. With the cost per thousand measure, the denominator changes with each calculation so that the job of computing this index for each subgroup requires a separate division.

I found the analysis of the methods used by the rating services to be extremely helpful. But I was struck by what seemed to be a small, yet important, omission in this section of the report.

The committee reviewed the sampling methods of the rating services with considerable thoroughness. It also devoted a full chapter to the interviewing procedures. But I missed a discussion and critique of the questionnaires used by the rating services --- particularly the diaries.

Apart from sampling considerations, our greatest concern about diaries is that they may not be administered methodically and accurately. There is also a difference of opinion in the broadcast industry on optimum diary design.

In filling out the Nielsen diary the viewer merely has to indicate when the set is on and off, enter the station call letters or channel number for the time periods in which he viewed, and list the number of men, women and children viewing. This has been criticized on the grounds that when people come to fill out the diaries they are more likely to be sure of the program they viewed than the station to which the set was tuned.

In contrast, the ARB diary requires the viewer to write the name of the program as well as enter the station call letters. But he must also list the starting and ending times whenever the channel is changed and report the number of men, women, and children paying attention to the program. This has been criticized on the grounds that it requires too much work for the diary keeper and may result in excessive reporting errors.

It would have been instructive to have heard the views of the Technical Committee on this subject.

In summary, I think that the committee did an excellent job of carrying out the assignment of evaluating the statistical methods used by the rating services. I would have preferred the assignment to have been broader in scope, to include technical problems of the application of ratings and other broadcast measurement. But perhaps these are subjects for another report.

The finding that the methods of the rating services are reasonably adequate for the purposes they serve is important to the broadcast industry. It should answer most of the unfounded criticisms of the accuracy of ratings. It would not satisfy all the critics of ratings -- and, in fact, it should not. For ratings by themselves are not adequate for many broadcast research problems.

The conclusions of the committee are basically sound, and both the broadcast industry and the rating services would profit by following the committee's recommendations. I would caution only against over-investing in ratings to improve their quality. There are new areas in broadcast research now developing which could benefit more from our support.

PANEL DISCUSSION OF THE REPORT OF THE TECHNICAL COMMITTEE ON BROADCAST RATINGS

Dr. Hyman H. Goldin, Chief, Research and Education Division
Federal Communications Commission 1/

I should like to clarify at the outset the interest of the Commission in broadcast ratings. We don't subscribe to any service. Our interest in this area is basically different from that of networks, stations, or agencies.

We are interested in the subject of ratings as it bears on our general responsibility to license and regulate stations in the public interest. In carrying out this mandate, we are required to keep ourselves informed on the structure and functioning of the broadcast industry. This is necessary both as the basis for formulating our own general policies and practices, and to provide Congress with expert information. Specifically in the matter of ratings, it was the Subcommittee on Communications of the Senate Interstate and Foreign Commerce Committee that initiated a public hearing on the subject in June 1958. Subsequently, the Special Subcommittee on Legislative Oversight of the House Committee on Interstate and Foreign Commerce, in March 1960, called on the American Statistical Association to undertake a technical evaluation of ratings. The Commission, as the expert government agency in the field of broadcasting, must be prepared, if called upon, to provide Congress with its independent views on this subject.

Aside from its responsibility to Congress, the Commission has an interest in ratings insofar as it bears on the output of the broadcasting industry. Let me make clear that the Commission does not censor individual programs. This is specifically excluded by the Communications Act. We are concerned with programming service in a more general way: The extent of program balance (for example, as between entertainment and other types of programs), the degree of diversity in program choices, and the forces that promote or hinder overall program improvement in the public interest.

Ratings obviously play an important role in the programming decisions of broadcast stations and networks. We know that, by and large, a high-cultural program does not regularly make the top 10 or 20 Nielsen or Arbitron. We don't blame the surveys for this result: It reflects a highly skewed distribution in public education, tastes, and TV usage. While I recognize that most of the time stations for economic reasons must serve the majority audience, they have an obligation also to serve minority tastes. If an Eisenhower interview, for example, reaches 7 and 1/2 million people, as compared with 25 million reached by each of two alternative network programs, we don't have the same adverse reaction as the network programming the Eisenhower interview. There are few alternative single organs of communication which could attract 7 and 1/2 million people.

Currently, we are conducting an inquiry into the general area of program selection, particularly at the national level. This proceeding has been going on since October 1958, and we have had testimony from writers, program producers, networks, advertising agencies, talent, and other segments of the broadcasting industry. Various witnesses have referred to the role of ratings in the selection or rejection of network programs. In a further hearing, beginning January 23, 1962, we have specifically requested the networks to discuss in detail their use of rating information and the effect of such use on the ultimate composition of the network schedules.

In addition to our broad affirmative responsibility to improve programming service in the public interest, we also have a responsibility to root out abuses and undesirable practices reflecting unfavorably on the character of broadcast licensees. This is a responsibility which, in some aspects, we share with other governmental bodies. Currently, we are cooperating with the Federal Trade Commission which is inquiring into certain alleged abuses involving station claims as to rating results as well as to the objectivity of certain rating operations.

I hope that this brief statement clarifies our interest in broadcast ratings.

I am very happy that Congressman Harris undertook to enlist the support of the ASA in obtaining a technical evaluation of the rating services. This is a subject fraught with emotional reactions. Typically, the word "ratings" in governmental circles bring a negative response in a spectrum ranging from those who are not persuaded that ratings have any validity to those who are convinced that they are the devil's work and the source of all evils in broadcasting. A great value to me personally of the ASA report is that now I have an authoritative reference source that I can cite. This may not convince non-believers but at least it shifts the burden of proof. It helps greatly to be able to say that a group of recognized and uncommitted statistical experts satisfied themselves that the national rating services are, by and large, valid for the purposes for which they are normally used

1/ The views expressed are those of the author and not necessarily of the F.C.C.

The Report has another value. After disposing of irrelevant and uninformed criticisms, it brings into focus very real problems which exist in this area. Certainly I agree that considerably more research into methodology is required; that more needs to be known about the uses to which the data are put; and that considerably more information should be published as to the size of sample, the sampling error, extent of non-responses, and adjustments, if any, in the responses. We agree also that an area in which considerable improvement is needed is in local ratings, particularly in smaller markets. Most important, I feel, is the need for an Office of Methodological Research, properly staffed and properly financed.

I should like to underline some areas which are mentioned in the Report but not fully developed. There is considerable dissatisfaction in the radio industry with respect to ratings. As the Report indicates, the fractionalization of the audience resulting from the great number of radio stations and the difficulties of accurately portraying out-of-home listening have created very great technical problems in economically and effectively measuring the audience for a particular radio station. In addition, the intense competition among stations resulting from the great expansion in the number of stations has put a premium on being first in the market. I have heard from a number of stations that despite the technical shortcomings of the radio ratings (resulting from low ratings, small samples, multiple sets, and out-of-home listening) advertising agencies place great emphasis on such ratings in buying national spot time. The stations feel very strongly a need for improvement in radio ratings. Ardent, free-enterprise broadcasters, who resent the Commission's attempts to obtain more detailed data on programming in the public interest, have suggested that the Commission itself undertake ratings or at least establish standards of good survey practice.

We have concern in assuring ourselves that the ratings are "honestly" drawn and "honestly" used. But our basic concern is rather with research which will throw light on the uses of broadcasting to serve long-range citizen needs. As one recent publication in the field stated:

> One of the discouraging aspects of present conditions in American broadcasting is that the industry is conducting almost no research into the long-range, fundamental influences of the media. In the twentieth century our missiles are guided, but our radio-television transmissions have no comparable sense of direction. Large sums are being spent to measure audiences and to discover how best to sell products to them, but there is almost no careful study of the overall impact of the media on our society. Network executives have speculated publicly about the desirability of such an investigation, but so far it has had a very low priority. Of course the presuppositions and design of such a study would be of tremendous importance. A superficial job would be worse than none because it would be dangerously misleading.

The Commission has no funds for conducting such research. It is conceivable, however, that the Office of Methodological Research properly conceived and properly financed could play an important role in this connection.

Broadcast communications act on the individual as only one of many stimuli, their specific effects are not readily discernible, and may differ in influence substantially as between the long-run and the short-run, and as related to various demographic factors. I am hopeful that the broadcast industry in its professional capacity will some day launch a full-fledged, long-range, perhaps ASA-certified study on the effects of broadcast media. Meanwhile, however, the Office of Methodological Research when established might concern itself with this question: How can the regular continuing audience surveys be used to provide more detailed and reliable information on the selection of programs by various subgroups within the general population, consistent with continuing commercial patronage of the surveys? For example, can audience composition data by age, sex, income, occupation, and education be refined further to provide meaningful profiles of the viewing patterns of various identifiable groups in the community (viz. community leaders, etc.)?

I am sure that a number of other lines of inquiry will present themselves to an Office of Methodological Research. This office might well be part of a larger center for broadcasting research. I am heartened by the expressed interest of leaders in the broadcasting industry to take a more positive role in promoting and stimulating and financing research directed not only to the media's commercial effectiveness but its role in furthering the public interest. I recognize that substantial sums of money are involved but I am hopeful that the broadcast industry, including networks, stations, advertisers, and agencies will take a broad and realistic view of their long-term enlightened self-interest. It is not inconceivable that such research may help to improve the industry's output in the long run.

Discussion

Jay Eliasberg, Director of CBS Television Network Research

Dr. Stevenson, members of the panel, ladies and gentlemen:

Before I make any comments on the Report of the Technical Committee on Broadcast Ratings, I would like to congratulate everyone connected with it. I think it was a wonderful job.

A friend of mine named Larry Deckinger was chairman of another committee that was supposed to do something about radio and television ratings. I'd like to borrow a remark Larry made in prefacing a progress report on the work of that Committee. The remark seems especially appropriate at the Christmas season. Larry stood up, looked around and then said, "When the good Lord wanted to save the world, He sent His only begotten Son. He did not send a committee."

Several of us on this panel worked on the committee that Larry headed -- it was called the Radio-Television Ratings Review Committee of the Advertising Research Foundation. It strikes me that the ASA Committee whose work we are here to discuss today had three advantages over the ARF Committee. First, I think that they were helped to some extent by having the awesome power of the House Interstate and Foreign Commerce Committee behind them. The other two advantages, however, were probably more important. Bill Madow's Committee had only three members; ours had nineteen. According to Parkinson's Law those are odds of better than six to one in favor of the smaller ASA group. And last but not least, the ASA Committee had to do its job, or felt that it had to, in a year; the ARF Committee has been in existence for almost ten years. We seem unwilling either to finish our work or to disband.

It is particularly in the light of my work on that Committee that I want to offer my heartfelt congratulations to Bill and his associates for the job they did. It was a wonderful job, and I hope that fact will be kept in mind as I make a few further comments about the report.

The most important criticism I have to make about the report is that I think it failed to make its point as clearly and definitely as it should have.

I base my feelings about what the major point of the report should have been on the letter to Morris Hansen from Oren Harris, Chairman of the House Interstate and Foreign Commerce Committee. That letter said, in part, that "It appears from the testimony that the choice of the kind of programs broadcast over networks during prime viewing hours has often been predicated upon public acceptance or preference as indicated by certain 'ratings' ascribed to programs by certain 'rating services'. As it is clear that the determination of any such ratings must be derived from statistical procedures involving sample surveys, our committee has requested you to arrange for an examination and evaluation of the statistical methods used by the principal rating services." You'll notice that Chairman Harris' letter talks only about network programs in prime viewing hours. The quotation from Chairman Harris' letter might be paraphrased as asking whether national ratings in prime nighttime hours can be accepted as reasonably reflecting the viewing habits of the public. I am somewhat a party at interest, of course, but I think that the answer to that question is, "Yes." The Legislative Oversight Committee of the House Interstate and Foreign Commerce Committee apparently agreed, since its press release on the report we are discussing today said, "To the extent that network programming policies are based on the ratings, it is unlikely that technical improvements in methods of estimating the ratings will lead to any support for a change in programming policies." And I think, that with extremely careful reading of the ASA Committee's report, anyone can come away with a similar impression. However, the closest I can come to finding a direct answer to the question in the report itself is on page sixteen where the Committee says, "The committee judges that the differences among the methods used by the various rating services are such, and the quality of the rating surveys is such, that when they agree the chances are very good that at least the policy decisions will reflect the listening and viewing behavior of the population of the United States."

I do not mean in any sense to depreciate the work that the Committee did with respect to other aspects of ratings and their uses. However, I would rather have seen the report devote less time to local ratings and to the various business decisions which hang at least partially on ratings, and to have devoted the additional space and time to an unequivocal answer to the basic question which unfortunately still plagues us to some extent. I would have been happier, in other words, if the next time my dinner partner said to me, as she almost invariably does, "You don't believe those television ratings, do you?", I could pull the ASA report out of my pocket and say, "Here. Please read this."

More seriously, I think the report would have been much better if it had concentrated more on the chief problem -- network ratings. Discussions of other matters -- local ratings, for example -- should not have been allowed to become confused with the major subject of the report.

But since the report did go into other matters, I would like to comment on one of them. Conclusion number eight of the report says that, "The type of measure called 'dollars per thousand (audience units)' is biased and has a large variance, especially if the rating and effective sample size are small." I've added those words 'audience units' to replace three asterisks in the report. The report goes on to point out that this is because such a measurement involves a constant (the program cost) in the numerator of the fraction and a random variable (the estimate of the number of audience units -- usually homes) in the denominator. The Committee then recommends the use of units per hundred dollars, instead of cost per thousand units. I agree fully with the Committee's recommendation. I would much rather we talked about a measurement which went up as it became better, as units per hundred dollars does in contrast to cost-per-thousand. Also, the work of research people in our industry would be simplified if we could make three divisions to get X women per hundred dollars, Y men per hundred dollars, and Z children per hundred dollars, and then two additions to get X plus Y adults per hundred dollars and X plus Y plus Z people per hundred dollars instead of having to make five divisions.

But I must object to the report's use of the word "biased" with respect to "cost per thousand." I am aware that technically the word is correctly used, in the sense that the expected value of the estimate of a program's cost per thousand differs from the actual value of the program's cost per thousand. However, I think I should point out that, to many of us laymen, the word "bias" is like a red flag. It says to us, in essence, "Stay away. Don't use this number or you are likely to make an incorrect decision." Since the ASA report came out, I have been trying to construct an example, no matter how hypothetical, where use of a program's estimated cost per thousand units would lead me to a different course of action than would use of the same program's units per hundred dollars. I have not been able to find such an example. I doubt that one can be constructed. For this reason, I don't think that the word "biased" should have been used in this connection, since the ASA report was bound to fall into the hands of laymen and, indeed, was perhaps primarily intended for them.

I might make two other points in this connection. The first is that the "bias" - in the technical sense of the word now - is extremely small - at least for national programs. In such cases, cost per thousand estimates are based on audience estimates, which in turn are based on samples of more than a thousand homes. One would have to search hard to find a program for which the estimate of cost-per-thousand-homes has a relative bias of more than 1%. I suspect that it may be impossible to find a program for which it is more than 5%.

The second additional comment is that this "audience units per hundred dollars" conclusion and recommendation of the report assume that cost figures are fixed and are precisely known. Unfortunately, the cost figures used in cost per thousand estimates are more often than not estimates. And, even more unfortunately, they are frequently less accurate estimates than are our estimates of audience.

One further comment, if I may. The report suggests the formation of an Office of Methodological Research. I am not so sure. There are certainly many rating methodological questions which have not been answered to date. However, one might infer from the report that the only reason these haven't been answered is that there has been no organization to answer them -- or to see that they were answered. This, I think, would be an erroneous conclusion. I think that there are two chief reasons that most of these methodological questions have not been answered. First, in the present state of the art of research, some of them can't be answered. And second, and perhaps most important, many of the questions would cost a tremendous amount of money to answer. If I may refer again to the ARF Ratings Review Committee, that Committee suggested an extended program of research on ratings. The program would have cost a great deal. From the dim mists of time, a figure of more than a half a million dollars seems to occur to me. As I recall it, that figure stopped everyone when it was first mentioned, and I have no reason to think that the reaction would be much different today.

In closing, may I once again congratulate Bill and his associates for the fine job they did.

Thank you.

III

INTERNATIONAL STATISTICS

Chairman, William R. Leonard, United Nations Statistical Office

Change of Manufacturing Output - H. S. Booker, London School of Economics

Techniques of Sampling and Statistical Evaluation for Census Work-Eli S. Marks, Case Institute of Technology

Uses of Electronic Computers in Processing Censuses-James L. McPherson, Bureau of the Census

Discussion - A. J. Jaffe, Bureau of Applied Social Research, Columbia University

CHANGE OF MANUFACTURING OUTPUT*

H.S. Booker, London School of Economics

There are many things that can make international comparisons misleading. Some are due to the fact that quantitative data are collected in different ways and with different coverages in the various countries concerned and that, despite precautions, accuracy is never perfect. Some are due to differences in conditions in the various countries which mean that if the statistics are in fact correct, it may still be misleading to draw inferences from them without having a lot of other ancilliary information. This may arise from correct calculations of average density of population per square kilometre where necessary auxiliary information would be about the climatic conditions and fertility of the soil or the presence of minerals. Or it could arise from correct information of the proportion of government expenditure on social services where necessary auxiliary information would be on the ratio of government expenditure to national income and the nature of government expenditure not on social services (e.g. on armaments and on economic development).

The purpose of this paper is to show how misleading, in some respects, index numbers can be and quite irrespectively of any inaccuracies in the data from which they are composed. The experiment is made with the current world series of index numbers of manufacturing production based on 1953. In the experiment use is made of other quantitative information all published by the United Nations, of net manufacturing output, of population and of persons engaged in manufacturing in various countries. No doubt these figures are more satisfactory for measuring change in individual countries than for making comparisons between countries. It is felt, however, that the picture presented, though it may be wrong in detail, is reasonably accurate in the aggregate and well worth attempting. Some of the conclusions are quite striking and arise, first from the use of one year as base, and next from the fact that index numbers only indicate proportionate change, and not absolute change. The first requires some consideration of the representative nature of the year 1953 and an acknowledgement that what happened between 1953 and 1960 should also be studied in relation to what happened before 1953. The second requires that the proportionate changes should be related to the size of the factor being measured and, perhaps, also to the possibilities of further change. In a country with almost no manufacturing an increase of 100% may mean very little; in a country with only a small proportion of its labour force in manufacturing an increase in manufacturing labour force and hence output may be more easily undertaken than in a country with already a large proportion of its population in manufacturing; in a country with its manpower inefficiently employed it might be easier to increase by a large proportion the output per person employed than in a country already efficient. In a country which is already efficient further growth in output per person may be dependent upon completely new knowledge and inventions as well as new capital; in a country not so efficient it may mean only the adoption of types of organization and methods already commonly applied elsewhere with some new capital.

In the paragraphs that follow an attempt is made to give numerical values to the factors which have been mentioned above so that they can be isolated. Unfortunately information that seems even remotely useful for this purpose is only available for a limited number of countries. Thus the centralized economies such as Russia and mainland China are excluded, there is nothing for African countries and the further the analysis is taken the fewer are the countries covered so that in the last tables there is nothing apart from information on North America, European countries, Argentina and Japan. Population estimates are available for all countries, and weights proportional to net output in manufacturing for 50 countries. Annual manufacturing output indices are available for 32 countries and indices of employment in manufacturing for 27 countries but not always co-incident with the output indices.

The current indices of manufacturing output, based on 1953, present a well known picture of slow growth in Northern North America (United States and Canada), substantial growth in Europe and outstanding growth in Eastern and Southeastern Asia dominated by Japan. Thus in 1960 with 1953=100, the index for North America was 118, for Europe 158 and for Eastern and Southeastern Asia 249. Latin America showed a record comparable with that of Europe with the index at 164.

It should be recorded that for North America the choice of 1953 as the base year is unfortunate. Manufacturing output in that year was 5% higher than the average for the three years 1952 to 1954. Even when allowance is made for this the relationships remain essentially the same. If the average rate of production in the years 1952 to 1954 is made equal to 100, instead of the rate for 1953 alone, the index for North America in 1960 is raised from 118 to 124, still well below the index of 158 for Europe.

The situation looks very different when the indices, as in table 1, are recalculated on an earlier base (using 1948 net outputs as weights to 1950 and 1953 net outputs subsequently). With 1938=100 the growth recorded in North America, Eastern and Southeastern Asia and Latin America are all of the same order with average rates of increase, compounded, of about 5% per year. In Europe the average rate of increase was less than 4% per year so this measure shows Europe still lagging materially, despite recent rapid growth. With this new base the relative stagnation of North American production in recent years looks like a reaction from a very rapid rate of growth between 1938 and 1950. The recent rapid rate of growth in Europe and Eastern and Southeastern Asia seems to be a catching up again after a stagnation associated with the war years. In Latin America growth continued at a relatively constant rate throughout the whole period.

The remainder of the analysis relates to

*I have to thank Miss Camille LeLong for help in preparing the tables of the UN Statistical Office

individual countries for the period 1953 to 1960.

There are various ways in which the degree of industrialization of a country can be measured. A typical way is to estimate the proportion of the economically active population which is engaged in manufacturing. That method ignores the efficiency with which those occupied in industry are employed and also ignores the fact that different proportions of the population may be economically active in different countries. Another method is to measure the proportion of the national income which is produced by manufacturing industry but that ignores the fact that some countries produce more per person than others irrespective of the distribution of economic activities. The first method used here is to divide the net output from manufacturing in 1953 by the total population in that year, and to relate these net outputs per person of total population to the result obtained for the United States. The resulting values are called index numbers of industrialization. This method has the advantage that good estimates of population are available for all countries and for many there is reasonable information of net outputs. It has the disadvantage that net outputs are compared by using current exchange rates which may not be realistic for this purpose. They are, however, more likely to be comparable for measuring outputs of material goods, which are likely to enter into international trade, than for total national income which includes a large volume of services. Another disadvantage is that the output of very small manufacturing concerns tends not to be recorded in censuses of production and in the less developed countries this output, as a proportion of total manufactured output, is likely to be higher than in the highly developed countries. This may not be a disadvantage if industrialization is looked upon as the development of production in factories.

The second method is more restricted in scope, it is concerned with production and employment in manufacturing only, shows how much a change in production is associated with the change in employment and how much with a change in output per person employed. Finally it attempts to measure the output per person employed in 1953 and 1960 in absolute terms and relate it to the average output of a worker in the United States in 1953. This might be called an index of efficiency of employment.

The results of the first calculation for 1953 are shown for 49 countries in table 2; the summary at the end for the different areas covers only the individual countries shown in the table. Manufacturing output per head of the total population in the United States and Canada (combined) was more than twice that for Australia and New Zealand, three times as much as the average for Europe, fifteen times as much as for Latin America and nearly fifty times as much as the average for the Asian countries recorded. Japan, the most highly industrialized country in Asia had an industrialization index, in 1953, only one-thirteenth of the value for the United States, and some of the countries had an index less than one-hundredth of the United States value.

Since 1953 these values have been changing, mainly under the influence of increased manufacturing production but partly also under the influence of population growth. In the seven years to 1960 population growth was under 10% in European countries and Japan and above 10% in most other countries thus helping to reduce the industrialization of other countries including the United States and Canada related to that of Europe and Japan irrespective of the growth in manufacturing. As manufacturing was growing more quickly in European countries and Japan than in North America the two factors were, not absolutely but comparatively speaking, working for them in the same direction.

From the first two columns of table 4 it can be seen that for most countries with records, the increase in population is quite overshadowed by the much greater increases in manufacturing production and it can be said generally that population growth has almost insignificant effect. Exceptions relate to Canada, United States and Argentina where in each country, both increases were of about the same proportion so that the industrialization indexes were almost unchanged. The only other population growth of special note is the large one of 29% for the Philippines which had the effect of modifying substantially the 90% increase in manufacturing production.

The results of dividing net output of manufacturing industry by total population in both 1953 and 1960 and relating the quotient to that obtained for the United States are shown in table 3 for 27 countries for which information is available. The change from 1953 to 1960 is shown both in absolute terms and as percentages. The two most highly industrialized countries in 1953 as measured by this index (United States and Canada), show almost no change to 1960, and only modest gains when comparison is made with 1952-4 instead of 1953. The industrialized European countries show substantial growth in manufacturing output per head expressed both absolutely and proportionately. West Berlin is, perhaps, a special case and the most striking other advances are by Yugoslavia and Japan especially when expressed as proportions; their absolute advances, however, are exceeded by several European countries, including the United Kingdom, Germany and France. The rapid proportionate growth of Pakistan is very small in absolute terms. In general even the proportionate increases for countries with an industrialization index of less than 200 in 1953 are quite modest when compared with the increases in the European countries already highly industrialized. There is little evidence here that the less developed areas are catching up in manufacturing output per head of total population, but there is evidence that the United States and Canada are lagging.

A new factor can now be introduced into the analysis if the analysis is reduced to cover only 20 countries and that new factor is the use of index numbers of employment in manufacturing industry. By use of these index numbers as well as the index numbers of manufacturing output it is possible to divide the increase in manufacturing output into two factors, that associated with an increase in employment and that associated with an increase in output per person employed. For example with 1953=100 the index of production for Germany was 183 in 1960 and the index of employment in manufacturing was 137. Dividing 183 by 137 gives 1.34 and the index of production per person employed is therefore 134. This information is given in table 4. In 16 of the 20 countries the increase in output per person was the more important, the exceptions being Yugoslavia, Germany, Hungary and Denmark. Amongst the 16 where output per person was more important than the number of persons employed there were still some countries where the increase in employment has had a big influence and this is especially so in Japan and Austria. A regression line of percentage increase in manufacturing production (Z) on percentage increase in employment in manufacturing (E) gives the equation $Z = 1.94E + 31.3$. This equation suggests a tendency for output to have increased by rather over 30% (31.3) in the seven years irrespective of any change in employment and a further increase of nearly 2% (1.94) for every increase of 1% in the numbers employed. There seems clear evidence of increasing returns from manufacturing; in general the countries with the greatest increases in employment also had the greatest increases in output per person. This effect is not simply the short term effect associated with recovery from depression when there was under-employed labour and capital. In relation to this regression line Japan, France, Italy, Poland and the Philippines have done well and Denmark, the United Kingdom, Ireland, Canada, the United States and Sweden badly.

A final table (table 5) relates to only 14 countries. It attempts to measure, for 1953 and 1960, the output per person engaged in manufacturing. It is more tentative than the earlier tables because of the possibility of errors in the various series having a cumulative effect. In particular it depends upon the 1953 net outputs used as weights in the index numbers of production being consistent with the figures of the numbers of persons engaged in manufacturing industry as well as the reliability of the output and employment indices. For some countries the numbers engaged in manufacturing industries as recorded in population censuses are markedly higher than the values used in the calculations, the differences presumably being accounted for by the home workers, self employed persons and small employers and their employees whose output tends to be omitted from censuses of production and index numbers of production.

The greatest interest in this table 5 relates to the absolute and proportionate growths of output per person in manufacturing. Generally, small percentage growths since 1953 are associated with high outputs per person in 1953. The countries are arranged in this table in order of their efficiencies of employment in manufacturing in 1953 the most efficient at the top and if the relationship were consistent the percentage increases in the last column would rise consistently and there are striking exceptions. The correlation however is not very strong though the regression line gives the equation $Z = -0.0385X + 56.5$ where Z = percentage increase in output per person from 1953 to 1960 and X = output per person in 1953 (U.S.=1000). For Japan, France and Italy the percentage increases were respectively 32, 22 and 17 above the regression line and for Argentina, the United Kingdom and Belgium they were respectively 25, 20 and 15 below.

When consideration is given to the absolute growths of manufacturing output per person employed in manufacturing the increase for the United States is seen to be the second highest in the table, suggesting that though manufacturing is growing less than in most other countries, its efficiency in terms of output per person is still improving rapidly. This may be partly due to the increased use of capital and partly to the more efficient employment of labour. The enormous percentage growth for Japan represents quite a modest increase in absolute terms, 9 of the 14 countries showing a greater absolute growth. The absolute growth for Argentina is by far the smallest for the countries shown, but for Belgium and the United Kingdom, both industrial countries of long standing, the growth is less than the average in absolute as well as percentage terms. The absolute growths for France, Italy and Canada though below that for the United States are very satisfactory and that for Yugoslavia outstanding. In Germany too the growth is good, especially as capital also had to be found for a greatly expanded labour force in manufacturing referred to previously.

The apparent lagging of the United States in manufactured output since 1953 is now seen as a stagnation of total output associated with a very substantial gain in efficiency of production, the very rapid rise in Japan is only in relation to a low level in 1953 and in absolute terms is only about one-half of the growth in output per person in manufacturing in the United States.

Table 1 — World index of industrial production

	1960		Rate of increase per year			
	1953=100	1938=100	1938 to 1950	1950 to 1956	1956 to 1960	Whole Period
Northern North America	118	337	7.1%	4.6%	2.2%	5.5%
Latin America	164	328	5.0%	5.4%	6.6%	5.1%
E and SE Asia	249	319	-0.9%	12.4%	13.2%	4.8%
Europe	158	226	1.7%	6.4%	5.6%	3.7%
World[1]	140	280	4.2%	5.7%	4.5%	4.7%

[1] Excluding centrally planned economies such as Russia and mainland China.

Source: U.N. Monthly Bulletin of Statistics, August 1960 and August 1961.

Table 2 — Indices of Industrialization

Manufacturing output per head of total population, 1953

(United States = 1000)

Country	America and Oceania	Europe	Asia
United States	1000 (950)		
Canada	714 (692)		
United Kingdom		620	
Norway		549 (556)	
Switzerland		547	
Sweden		501	
New Zealand	462		
Germany, Fed. Rep.		438 (441)	
Australia	437		
Denmark		423 (431)	
West Berlin		367	
Belgium		339	
Netherlands		313 (316)	
Austria		310 (324)	
France		291 (301)	
Italy		203	
Finland		198	
Uruguay	147		
Ireland		139 (136)	
Argentina	133 (137)		
Yugoslavia		120	
Spain		108	
Venezuela	101		
Japan			75 (73)
Mexico	71		
Portugal		71	
Chile	69		
Greece		66	
Colombia	48		
Malaya and Singapore			45

Table 2 (continued)

Country	America and Oceania	Europe	Asia
Peru	43		
Brazil	41		
Costa Rica	35		
Turkey		32	
Ecuador	30		
El Salvador	30		
Paraguay	29		
Guatemala	27		
Formosa			18
Philippines			16
Honduras	14		
India			13 (12)
Thailand			10
Nicaragua	8		
Ceylon			8
Korea			8
Pakistan			6
Burma			6
Indonesia			4
Summary			
Northern N. America	976		
Latin America	62		
E and SE Asia			20
Europe		310	
Oceania	441		

Note: Figures in brackets denote values when 1952-54 instead of 1953 is taken as base.

Sources: Relative net outputs in 1953 from the weights in the world index of manufacturing production, see Supplement to the U.N. Monthly Bulletin of Statistics, 1959 and U.N. Statistical Yearbook, 1960.

Population figures from U.N. Monthly Bulletin of Statistics, August 1961.

Table 3 Changes in industrialization, 1953 to 1960
Manufacturing output per head of total population (United States, 1953=1000)

Country	Index of industrialization		Change from 1953 to 1960	
	1953	1960	Absolute	Percentage
United States	1000 (950)	1039	39 (89)	3.9 (9.4)
Canada	714 (692)	702	-12 (10)	-1.7 (1.4)
United Kingdom	620	785	165	26.6
Norway	549	741	192	35.0
Germany, Fed. Rep.	438	738	300	68.5
Denmark	423	557	134	31.6
West Berlin	367	919	552	150.3
Belgium	339	442	103	30.4
Netherlands	313	455	142	45.3
Austria	310	524	214	69.1

Table 3 (continued)

Country	Index of industrialization		Change from 1953 to 1960	
	1953	1960	Absolute	Percentage
France	291	485	194	66.7
Italy	203	354	151	74.4
Finland	198	283	85	43.1
Ireland	139	174	35	24.9
Argentina	133	131	-2	-1.7
Yugoslavia	120	280	160	133.1
Venezuela	101	171	70	69.5
Japan	75	195	120	160.0
Mexico	71	102	31	44.2
Portugal	71	111	40	56.0
Chile	69	69	0	0.0
Greece	66	107	41	61.5
Guatemala	27	31	4	15.5
Formosa	18	27	9	51.1
Philippines	16	23	7	46.8
India	13	18	5	37.8
Pakistan	6	14	8	132.5

Note: Figures in brackets take 1952-54 as base instead of 1953.

Sources: As table 2 with information of growth of manufacturing output from U.N. Monthly Bulletin of Statistics, September 1961, table 7.

Table 4 Factors affecting Industrialization

Country	Increase in Population 1953 to 1960	Increase 1953 to 1960 in		
		Manufacturing Production	Manufacturing Employment	Manufacturing Production per person
	%	A%	B%	C%
Japan	8	180	54	82
Yugoslavia	9	155	74	47
Poland	13	97*	20	64
Philippines	29	90	23	55
Germany	9	83	37	34
Italy	4	81	12	62
France	7	78	6	68
Austria	2	72	25	38
Netherlands	9	59	12	42
Hungary	4	55	27	22
Finland	8	54	8	42
Norway	7	44	5	37
Denmark	5	38	19	16
Belgium	4	36	5	30
Sweden	4	34*	6	26
United Kingdom	4	31	9	20
Ireland	-4	20	1	19
Canada	20	18	-3	22
United States	13	17	-5	23

Table 4 (continued)

Country	Increase in Population 1953 to 1960	Increase 1953 to 1960 in		
		Manufacturing Production	Manufacturing Employment	Manufacturing Production per person
	%	A%	B%	C%
Argentina	14	12	-10	24

*General index of production.

Sources: As in previous tables with index numbers of employment in manufacturing from U.N. Monthly Bulletin of Statistics, September 1961, table 4B.

$$A = 1.94B + 31.3 \qquad r = 0.88$$

$$C = 0.42B + 31.8 \qquad r = 0.45$$

Table 5 Indices of Efficiency of Employment in Manufacturing (United States, 1953=1000)

Country	Output per person 1/		Increase 1953 to 1960	
	1953 (X)	1960	Absolute (Y)	Percentage (Z)
United States	1000	1232	232	23
Canada	851	1035	184	22
Norway	743	1019	276	37
Sweden	484	612	128	26
Germany, Fed. Rep.	450	601	151	34
United Kingdom	433	520	87	20
Yugoslavia	431	632	201	47
Netherlands	305	433	128	42
Italy	300	485	185	62
Belgium	283	366	83	30
France	272	457	185	68
Austria	271	373	102	38
Argentina	180	224	44	24
Japan	148	269	121	82

1/ In manufacturing industry.

Sources: As in previous tables with information as to numbers engaged in manufacturing from U.N. Statistical Yearbook, 1960, table 66. Some of the values for 1953 obtained by interpolation of figures for nearby years.

$$Y = 0.17X + 75 \qquad r = 0.68$$

$$Z = -0.0385X + 56.5 \qquad r = -0.52$$

TECHNIQUES OF SAMPLING AND STATISTICAL EVALUATION FOR CENSUS WORK

Eli S. Marks, Case Institute of Technology

The past 20 years have witnessed a virtual revolution in techniques of census-taking and tabulation. On the mechanical side, we have the extensive use of electronic data-processing machines by a variety of countries and the use of mark-sensing devices in the United States and Canada to eliminate the expense and error associated with punching and verification of cards. While these developments have very important implications for census work, the equipment involved is rather expensive and not readily available in industrially undeveloped countries. While some of the smaller, non-industrialized countries may look to electronic equipment for the solution of their census problems, most such countries must continue to rely for many years to come upon hand and punch-card tabulation techniques.

The restriction of available mechanical devices to those of 20 or 30 years ago does not, unfortunately, mean a restriction of census problems to those of 20 or 30 years ago. For one thing, almost all of the smaller countries have experienced a considerable population growth - in the case of the newer countries, one could appropriately consider the rate of growth infinite! More important, however, is the increased demands placed upon a census. Public and private agencies in almost all countries have become aware of the importance of census data as a basis for action and the demands for more, for more timely, and for more complex, statistics have increased exponentially while the available personnel and facilities have increased (at best) arithmetically.

Along with the increased demand for statistics, there has been of recent years an increasing sophistication with respect to the pitfalls of inaccurate statistics. In general, this sophistication has not kept pace with the increasing demand and it is not unusual for the hard-pressed census statistician to be confronted with requests for "figures - any figures, deliver them yesterday; and never mind these fool questions about how accurate they should be; we must, of course, have perfect accuracy!" In spite of this type of behavior, most "producers" of census statistics, and an increasing number of users of census statistics, are conscious of both the unattainability and the superfluity of "perfect accuracy." We may yearn for the "good old days" when a census was accurate by definition (and official decree) but most of us are willing to face up to the realities and are unwilling to deliver statistics of unknown but dubious accuracy, regardless of the pressures involved.

In the situation of increasing demand for both quality and quantity of statistics and of sharply limited resources, the producer of census data must turn more and more to modern techniques of sampling and evaluation -- sampling to meet those demands which can be met and evaluation (if for no other reason) to counter the "unmeetable" demands! Even the large and industrially developed countries are relying more and more on these techniques and for the smaller undeveloped countries they are indispensable. Furthermore, even those countries which can afford the latest model of digital computer that high-pressure salesmanship can supply, are discovering (or will discover) that electronic gadgets are a supplement to, rather than a substitute for, human ingenuity and sound statistical methods.

In the search for solutions to the problems of demand, other countries have looked, of course, to the United States. Fortunately, in the area of statistical developments, the United States has retained its pioneering vigor and can offer more than mechanical devices and public relations "gimmicks." In particular, the work of the United States Census Bureau in the area of sampling and statistical evaluation, has provided a model which can profitably be studied even by countries which may be more advanced in other statistical fields.

The application of one country's methods to another country's problems is, however, neither simple nor painless. Mistakes have been made in a too-literal following of United States instruction manuals and statistical texts. In the light of these mistakes, the experience with applying United States developed sampling and evaluation techniques to the Chilean censuses of 1960 should prove illuminating - both for its indication of what to do and its indication of what not to do. My own experience with the Chilean censuses related primarily to two problems and my remarks are, therefore, directed primarily at these two problems.

The two problems with which I was concerned in Chile were (a) the development of a measure (or of measures) of the completeness of the census coverage and (b) the selection of a sample to provide rapid and (reasonably) accurate advance tabulations of the census data. The latter problem is a particularly difficult one for all but the handful of countries which can afford the newest and fastest electronic data-processing equipment. And some of these latter are discovering in 1961 and 1962 what the United States discovered in 1950 and 1951 -- that having high-speed equipment and having the ability to use equipment at high-speeds are not the same thing.

Since a paper on the evaluation of the Chilean censuses of 1960 is to appear in the near future, I shall confine my remarks to the problem of sampling for advance tabulations. The general lessons to be learned from this problem are applicable also to the problem of evaluation although there are, of course, techniques specific to each problem.

The problem of selecting a sample for advance census tabulations is -- in its design aspects -- an extremely simple one. A complete sampling frame exists and the sampling theory is straightforward and well-developed, since concern is entirely with estimating the results of the census and not with determining some "true value" which may be undefined (or even undefinable). But even during the stage of sample design, both the Chileans and I were (fortunately) aware of the problems of carrying-out a sample design. I say, "fortunately," because other countries have tried to use samples designed purely from a sampling cost and variance standpoint without considering the costs and biases of carrying-through on the sample design.

The Chileans had had an experience with the problems of carrying through a sample design in connection with the sample for advance tabulation of their 1952 censuses of Population and Housing. There a well-designed sample came to grief in execution and ended with a bad bias, whose correction required about two years and the services of a top United States Census Bureau sampling expert.

The problem in 1952 was that the clerks selecting the sample (a 2% household sample selected systematically for the whole country) tended to avoid the larger households, substituting the schedule immediately preceding or following. This type of bias is somewhat hard to detect without special control measures and was, in fact, not detected until the sample for the entire country was selected and punched and the preliminary totals compared with hand counts made previously.

The experience of the 1952 census was recent enough to provide a vivid object lesson. My own indoctrination prior to going to Chile had included a first hand briefing on the incident and many of the 1960 personnel of the Chilean statistical office had been there also in 1952. The head of the processing section remembered quite vividly the disorganization of his own work attendant upon the belated discovery of the bias and was determined to avoid a similar fiasco in the 1960 censuses. Quite independently of me, he hit upon one leg of what I think of as the tripod on which a sound sample must rest -- simplicity -- the other two being good design and good control. Unfortunately, in stressing this leg of the tripod he sawed off the "good design leg" although his basic analysis of the situation was sound.

The processing section head's proposal was that we sample entire enumeration districts (or, in the local terminology, enumeration zones). Here the problem of keeping track of what was selected and of avoiding substitutions and other biases would be greatly reduced in magnitude. However, although the Chilean enumeration districts average about one-tenth the size of United States enumeration districts (i.e., average about 20 households), the intraclass correlation is quite large and, on some characteristics, might mean that an enumeration district sample would have to be 5 to 10 times as large (in terms of number of persons) as a sample of households in order to attain the same accuracy. For that difference in sample size we could buy a very adequate third leg -- good control -- to stabilize our tripod of good design and simplicity.

The preceding statement about sample sizes is deceptively easy to make now but the difficulties of arriving at this conclusion illustrate the problems of applying for the first time in a country, techniques which are so "obvious" to "experts" that they never bother to explain them in detail! On my first visit to Chile, in March-May of 1960, the question of sampling enumeration zones or households was raised (as noted above). Before trying to answer this question, I naturally wanted to examine the variances and sample sizes involved. At this point I discovered that there were no enumeration zone totals available from the previous census and there were no distributions of households or families by size or other characteristics. We laid out, therefore, a program of preparing distributions of enumeration zones and households by size and numbers of males and females for four provinces (the major administrative subdivisions of the country).

Since these data had not previously been entered on punch cards and since the punching and tabulation facilities of the statistical office of Chile were overburdened already with other work (current surveys and census pre-test tabulations), the tabulation of the desired distributions was set up as a hand tabulation, with two of the office's limited number of clerks assigned to the job.

It developed that some of the materials for parts of two of the four selected provinces could not be located and these areas had to be omitted, curtailing further the already limited information which would be available for making a critical decision. In addition, I failed to point out the "obvious" fact that "open-end" intervals could introduce serious errors into a variance computation if they contained any appreciable number of cases. I also forgot the obvious fact that clerks must have tabulation intervals specified in advance and cannot be expected to notify the statistician in charge that there is a large number of cases in an open interval. By the time I discovered the problem, it was too late to do anything except try to make allowances for the range of possible error.

Eventually the decision was made to sacrifice some simplicity to good sample design and to use a sample of schedules. The requirements for sampling indicated different sampling rates would be desirable for different provinces since separate estimates by province were to be prepared and the provinces vary considerably in population. To balance these complexities, it was decided to use a systematic sample with separate starts in each comuna (equivalent to United States county). The system was to number consecutively all schedules of the comuna and then sample those schedules with (randomly) prespecified ending numbers -- the number of "endings" varying from 2 to 18. The plural random starts, although adopted primarily in the interests of simplicity and ease of sample control, do have the advantage of giving a design which permits unbiased estimates of sampling variance.

All of the above is extremely simple. The formulae for anticipated sampling error and required sample sizes appear in any sampling text and selection of suitable "endings" requires only a table of random numbers -- although we did have to resort to the Inter-American Statistical institute's center at the University of Chile to obtain the latter! The important job was that of "sample control." This is a dull, routine task. It involves no new and exiting discoveries, no fascinating mathematics or other theory. It involves, for example, spelling-out the meaning of "number consecutively all schedules of the comuna" in full (and dull) detail, so that not even the most unskilled and lazy clerk can find an excuse for biasing the numbering -- "continue the numbering from one zone to the next within the comuna; start the first zone with 000-001; do not omit or duplicate any number, etc."

Sample control involves such trivia as checking that the highest number assigned agrees with an independent hand-count of the schedules and instructions on the procedure to be followed if it does not -- "look for omitted and duplicated numbers, assign duplicates the lowest unassigned number, assign omitted numbers to the last schedule(s) cancelling the number(s) previously assigned to it (them)," etc. These are "idiot" instructions but are unfortunately, necessary and cannot be replaced by the assumption that the use of discretion by the clerks will be equally satisfactory even though this assumption may sometimes be true.

In addition to the checks on numbering and the simplification of the sampling process, the sample control included a comuna-by-comuna check to assure that number of schedules, number of households (different from number of schedules since some households had two schedules), and number of persons were within sampling error bounds of the expected values. For the number of schedules and number of households, binomial sampling error formulae were applicable but the variance of numbers of persons had to be estimated. This estimate was obtained from the 1952 size of household distribution mentioned above. The controls described were implemented by a form on which population and sample figures were entered and which provided formulae for computing the sampling error "tolerance limits" and space for entering the results of each step in the computation.

All of this may seem like "idiot" procedure. Unfortunately, experience demonstrates conclusively its necessity. I understand that a very similar sampling for the Argentinian census ran into difficulties due to omission of some of these controls. Furthermore, the comuna-by-comuna checks in Chile pointed-up a defect in the sample design. This defect had to do with the sampling of institutions. The first schedule of an institution was the regular Population and Housing schedule with space for twelve persons. Continuation sheets, however, had space for listing 66 persons. In the design of the sample we considered drawing a sample of persons within institutions instead of a sample of schedules but rejected this alternative in order to keep the sampling procedure simple and uniform.

The sample checks for the first province completed, indicated several comunas where failure of the sample to satisfy the estimated "tolerance limits" could be traced to the large variance introduced by sampling whole schedules in institutions. Steps should have been taken immediately to reduce this variance by modifying the sample design. Unfortunately, there were delays in communication so that remedial action was delayed until it was too late to remedy more than a part of the problem. However, the mechanism for detecting trouble was present and this is the first half of avoiding trouble!

Another problem in the use of United States sampling techniques in the Chilean context centered in the estimation of sampling errors. Before leaving Chile I spelled-out the procedure for computing these estimated sampling errors. The defects in my assumptions did not become painfully apparent until seven or eight months later when people tried to follow my instructions. Then I received a letter which stated (among other things): "On calculating the sampling errors for housing, I have no problem. For population errors, difficulty arises because the process is a very long one, people-and-time-consuming, and I am very short in both respects...We have no tabulator with summary punch."

Again I had applied assumptions based on my United States experience in a situation where those assumptions were inaccurate. I had assumed that a job which was easy in the United States would present no difficulties elsewhere, forgetting that a "trivial" job can become gigantic if one lacks equipment, personnel, and money.

The conclusion to be drawn from the above experiences is not that United States or Canadian census experience is inapplicable to other countries. On the contrary, I found a great many unanticipated similarities to United States experience in the Chilean setting -- including a hot controversy raging on my arrival about the utility of the inquiry about "condition of dwelling unit" which made me feel right at home! Furthermore, almost all the principles of sample and survey design that I learned in the United States were directly applicable to the Chilean situation. What are not directly applicable are the specific procedures developed in the context of a highly mechanized economy. A brief consideration of the realities of the Chilean situation indicated that there was a solution to the problem of calculating sampling errors (from United States experience) which was applicable to Chile. What is needed is then the use of advanced techniques but with the necessary adaptations to local conditions. Above all, we need good design and simplicity and good control in the United States, in Chile, or in Ghana.

USES OF ELECTRONIC COMPUTERS IN PROCESSING CENSUSES

James L. McPherson, Bureau of the Census

Brazil, Canada, France,West Germany, India, Israel, Italy, Japan, Mexico, Norway, Sweden, the United Kingdom, the United States, Venezuela and Yugoslavia are countries which either are now, or soon will be, using electronic data processing equipment installed in their own territory for Census purposes. In addition, Finland and the United Arab Republics expect such equipment in neighbor countries to process some of their census materials. This we, at the U.S. Bureau of the Census, believe is an exhaustive list. We know there are installations of electronic computers in other nations. But we are reasonably sure that only the countries enumerated above are, at present, using or planning to use electronic equipment to process census data.

Russia

The U.S.S.R. is perhaps the major example of a country well equipped with powerful electronic computing equipment which has not assigned the processing of census data to such equipment. A group of United States computer engineers visited Russia in 1959. They visited several scientific laboratories and saw electronic computers in operation, new computers being built, and still newer computers being designed. It was their general impression that the state of the art in Russia has kept pace with western world development with the exception of the use of magnetic tape as an input-output medium. When a group of Russian engineers visited the United States several months before the U.S. engineers made a return visit, the Russians were quite impressed with our use of magnetic tape. There is evidence Russia has progressed significantly in this area. For example, we recently learned that at least one group is actively working on a mathematical model of the whole Russian economy a la Leontieff. This requires manipulation of matrices of high order for which magnetic tape is almost essential. We believe their problem has been one of producing tape with adequate magnetic quality rather than one of building transports to move the tape rapidly while it is being read from or written on.

However, despite the availability of electronic data processing equipment, the Russian population Census of 1959 was tabulated with punched card equipment. A short undated report entitled "Employment of Computing Machines for the Processing of Statistical Data in the U.S.S.R." describes the processing of the "1959 All-Union Population Census--an operation of great importance and involving a large volume of work."

This report indicates that a decentralized organization accomplished the basic processing. Evidently 209 million "perfocards" were punched and tabulated at 57 statistical boards throughout the union. These computing stations prepared summary cards which were forwarded to the Central Census Computing Machine Station of the Central Statistical Board. There were about 10 million of these summary cards which were run through sorters and tabulators to prepare district, city, regional, territorial and republican totals. Photo offset reproduction of tabulator output was used for publication of results.

The report claims that "modernization of computing equipment made for considerable reduction in labour expenditure in processing census data (by approximately one third as compared to the preceding census)." However, the kinds of equipment modification described in the report relate to "relays" and "sorting of perfocards" which suggest that the 1959 Russian Census was processed on multi-column sorter-tabulators similar to IBM 101 equipment rather than with electronic data processing machinery.

The speed with which the work was done is impressive particularly in view of our impression that only manual and punched card techniques were employed. Preliminary results based on apparently manual "sorting of data according to territorial units and packaging of census sheets in lots" were published on May 10, three months after the enumeration. A brief report on the first part of the final results based on punched card tabulations was published February 4, 1960, less than one year after enumeration. And "A third report, giving population distribution according to means of subsistence, social groups, economic branches and professions, and educational levels of mental and manual works, was published in December 1960."

Canada and the United States

The United States processed a comparatively small part of the 1950 Censuses of Population and Housing with Univac I. Almost all of the tabulating for the 1954 and 1958 Economic Censuses was accomplished with electronic computers. Tabulation of the 1960 Census of Population and Housing is being performed entirely with electronic data processors.

Electronic equipment contributed significantly to an important "first" in the U.S. history of Population censuses. For every one of the seventeen censuses of the United States preceding the 1960 Census, the counts for the apportionment of seats in the House of Representatives among the several States were "hand counts." The 1960 counts for apportionment purposes were produced by the electronic data processing equipment at the U.S. Bureau of the Census.

The Dominion Bureau of Statistics of Canada has installed electronic equipment to process the 1961 Census of Population, Housing and Agriculture. Effectively all of the tabulations of these censuses will be accomplished with this installation. (The session on "The Methods of the 1961 Census of Canada" at these meetings is scheduled to include a paper on Canadian tabulation plans.)

Both Canada and the United States aspire to reduce processing time through the use of electronic equipment. This is but one of several objectives but it has been considered (by Canada and U.S.) to be an extremely important one. Processing time can be defined in many ways. However, from the point of view of the general user of census results, the elapsed time between collection of the data and final publication of the results is the time statisticians spent processing the information and is, therefore, in his mind at least, the "processing time."

A part of the processing time involved in any census is the time spent in input preparation. Here, "input preparation" means conversion of the information recorded on a document by an enumerator or a respondent to a medium of some kind--punched cards, punched paper tape, magnetic tape--amenable to processing through mechanical, electro-mechanical, or electronic equipment. In many types of censuses, some clerical operations such as coding and editing of the information on the enumeration document precedes the preparation of the tabulation medium. But even when the only work which precedes input preparation is of a routine housekeeping nature, the use of mechanized tabulation facilities must await the creation of a suitable input medium.

The Canadian census officials were, to the best of our knowledge, the first to tackle the problem of speeding up the preparation of input. Manually operated card punching machines introduces significant delays in a census processing operation. As early as the 1951 Census of Canada, the DBS statisticians pioneered the use of document sensing equipment. This was an extension of the IBM mark sense machine to accommodate a form over twice the size of the conventional punch card. The Canadian enumerator recorded the answers to the census questions on these mark sense documents and machines, rather than key punch operators, created the punched cards which were then processed through sorters, tabulators and other conventional punched card processing machinery. This was a successful and satisfactory solution to the input preparation problem and the use of document sensing schedules was repeated in the 1956 Census of Canada.

In the mid-1950's we at the U.S. Bureau of the Census sought the advice and assistance of our National Bureau of Standards in connection with the input preparation problem. We described the Canadian use of document sensing to NBS engineers. We indicated that, in our opinion, there were at least two restrictions the document sensing approach imposed that we hoped might be eliminated. One of these was the rigidly specified size of the document which we believed was too small to accommodate the amount of information we wanted to collect per unit of enumeration. The other was the requirement that a specific writing instrument--either an electrographic pencil or a special fountain pen with electrographic ink--had to be used.

Our FOSDIC equipment resulted from our collaboration with NBS. This equipment permits a wide range of document sizes so we can make the dimensions of the form suit the content of the census. It does not require a special writing instrument, although we recommend a medium soft pencil. The input to Fosdic is microfilm images of census schedules. The output is magnetic tape ready for processing by our electronic data processing equipment.

For their 1961 census, the Canadians have similar equipment which they call electronic document reading equipment. The output is magnetic tape ready for their computers. The input is the document itself rather than a microfilm picture.

In both countries the objective has been rapidly to create an input medium for the electronic data processors with a view to speeding up the census processing time. We Censusites in Canada and the U.S. attach great importance to early publication of results and try to act accordingly. There are perhaps one or more other nations which do not share our concern with early publication and there certainly are other countries behaving in a manner which holds little or no promise of significantly reducing the time between enumeration and publication regardless of how rapidly electronic computers may accomplish the tabulation phase of a census activity.

In addition to facilitating a speed-up of the whole census process, electronic equipment can contribute in other important ways to census work. More complex cross classifications of the data than could be economically managed with punched card equipment become comparatively easy with electronic computers. The uniformity with which internal inconsistencies in the basic information can be detected--and in many census type investigations automatically corrected--is several orders of magnitude better with electronic computers than it was with earlier combination clerical-punched card processors. Definitions of derived descriptors of the data can be quite complex thus making them more useful when electronic machines are used to compute them. For example, the U.S. 1960 Census reports on family income will include, for hundreds of geographic areas, the median income of families consisting of man, wife and two children under 18 years of age. In earlier censuses we were able to show only over-all median family income with no control on family composition. For many purposes area to area comparisons will be much more meaningful when the median income relates to a "standard" family.

These are but a few illustrations of ways in which electronic computers can be and are being used to improve the quality of census data. All countries which are using or expect to be using electronic equipment to process their censuses expect to realize this kind of benefit.

Continental Europe

Every one of the six countries on the continent using or planning to use electronic computers for census work contemplates the use of IBM computers. Norway, where a British computer--the Deuce will perform some of the tabulating for the Census of Fisheries, is the only country using any equipment not supplied by IBM.

A convenient classification of electronic computers has three categories--small, medium, large. Since we are statisticians at this meeting, none of you will have difficulty pairing high, low and medium costs with their appropriate mates.

Small electronic computers generally have small, slow memories and very slow input-output facilities. We know of no country where small computers are used for census work.

Medium computers usually have small but reasonably fast memories; they almost certainly have punched card input-output facilities and may have magnetic tape transports associated with them. The IBM 650's and the 1401 in Italy and the 1401 in Norway are medium computers.

Large computers have large fast memories, magnetic tape input-output facilities and may or may not have direct punched card input-output equipment on line. The IBM 705's for France and Yugoslavia and the 7070's for West Germany and Sweden are large computers. Both Sweden and West Germany use an IBM 1401 as a peripheral device. Typically, in a 1401-7070 installation, the 1401 is used to prepare input for a 7070 by converting data from punched cards to magnetic tape and to accept magnetic tape output from the 7070 for printing on hard copy.

Italy is the only country not only on the continent but anywhere in the world that plans to emulate Canada and the U.S. by using electronic equipment to prepare magnetic tape without the intermediate medium of punched cards. Plans for the Italian population census contemplate use of the same kind of document to magnetic tape machinery as the Canadians are using. Norway, Sweden, France, West Germany and Yugoslavia are not underdeveloped countries in the general sense of this term and they certainly have on their census staffs personnel sophisticated in statistical work particularly insofar as the tools of the census trade are concerned. Without arguing the merits of either side of the case, we can point out that Italy, Canada and U.S. consider it important to employ electronics to accelerate input preparation while Norway, Sweden, France, Germany and Yugoslavia apparently place less emphasis on exploiting electronics for this purpose.

The United Kingdom

The U.K. is using an IBM 705 to process their 1961 Population census. A British machine the Deuce IIA is used for their annual census of agricultural holdings. Another British machine --the National-Elleott 405-- is used for their censuses of production and distribution.

The U.K. census personnel are allied with the continentals (Italy excepted) in the matter of input preparation. That is, they are evidently satisfied to prepare input by manually punching cards. In fact, the plans for processing the 1961 Population Census of the U.K. indicate our British colleagues are rather relaxed about processing in general. One gathers the impression that the philosophy is to use electronic equipment because it is more reliable than alternative machinery and will therefore contribute to producing a higher quality product. There does not, however, appear to be any particular desire in the U.K. to exploit to speeds of electronic equipment for the sake of getting the job done faster.

Latin America

Mexico and Venezuela have medium size computer--Remington Rand SS90's and Brazil has a large computer--the Remington Rand 1105 (which is the same equipment as the United States Census uses).

We understand the Mexicans have just about completed tabulation of their 1960 Census of Population with their SS90 and are currently tabulating the 1960 Census of Agriculture on this equipment. According to the manufacturer, the Mexican computer is not equipped with magnetic tape input-output equipment. The input has been punched cards and output is either punched summary cards or printed copy.

The Venezuelans have used, and are now using, consultants from the United States. After a visit in the fall of 1959 your author recommended early aquisition of tabulating facilities by the Venezuelans for their (then) forthcoming census. In his view it was more important to decide what machinery would be used than to deliberate about the relative merits of alternative equipments. Not until more than a year later was a decision to install an SS90 reached. We understand, from a recently returned U. S. consultant, that the equipment is now installed and ready to begin processing the 1961 Census of Venezuela. Our informant tells us, however, that card punching has not yet started. Evidently the Venezuelans have not yet been persuaded to overlap the various operations involved in processing a census. Thus they concentrated first on getting a hand count of population. Almost all personnel resources were devoted to this task which was completed only a month or two ago. The next step will be the selection of a sample (designed by a consultant from the U.S.) of the household schedules. Presumably not until this sample has been selected and any necessary coding has been completed will the preparation of the input medium begin. Whether and how much electronic computers will contribute to processing the Venezuelan Census remains to be seen.

Brazil embarked on an ambitious attempt to emulate U.S. techniques and procedures for processing their census. An early action in this program was the aquisition of a Remington Rand 1105 computer. We are familiar with many of the problems which attended delivery and installation of this machinery in Brazil. The recent political upheaval in Brazil has slowed information about the use of this equipment to a trickle and what little information we do get relates to technical and engineering matters concerning the equipment rather than plans for its use in connection with censuses.

Other Countries

We know there is, or soon will be, an IBM 705 in Japan for Census use and a Russian built "Ural" computer in India. This is about all we know about these two countries. We must be candid and admit we haven't endeavored to learn more.

Israel originally planned to use punched card handling equipment for all their census tabulations. Recently they decided to use an IBM 1401 at the Mechanization Center in Jerusalem to do some of the work connected with inflating some sample questions and to compute (from data tabulated using electro-mechanical equipment) certain percentages, ratios, averages and medians for each of 800 settlements for which they plan to publish these derived data.

EDITING

Editing of census information involves not only the identification of missing or internally contradictory information but also action of some kind to correct the deficiency. Electronic computers can be depended upon consistently, thoroughly and reliably to apply whatever rules the user specifys to identify those observations which require corrective action of some kind before the data are good enough to tabulate.

Views with respect to just how the requisite corrective action should be accomplished seem to vary widely. At one extreme we would place the United Kingdom. As we understand them, the U.K. plans for the population census contemplates using the electronic computers only to identify cases where information is missing or self contradictory and refer these to humans for corrective action.

At the opposite extreme we would place the Norwegians and Swedes who hold the view that not only can the "untabulatable" observations be recognized by the computer but that the computer can be directed to impute satisfactory and consistent answers by reference to complex mathematical models which can be thought of as multidimensional function generators. These function generators would be unusual since the several terms of the function would sometimes be independent variables--the satisfactory answers to census questions--and sometimes dependent variables --the missing or inconsistent answers to census questions for which satisfactory answers are to be imputed.

Our strong view at the U.S. Census on this question is a two-fold one. First we use our computers to both identify need for corrective action and to impute "tabulatable" answers where necessary. Second, and of major importance, we believe any workable method for imputation can be used and consequently it need not be particularly complex.

The important consideration which leads to our position on this question relates to the basic quality of the information we have collected. If we fail to get answers to a significantly large number of questions or if we get internally inconsistent answers for a large proportion of the population enumerated we cannot expect to make "good" information out of "bad" regardless of how fancy an imputation model we might invent. We can keep track in the computer of the number of times we impute answers to questions. Our philosophy is that when it is only infrequently necessary to impute answers, then any imputation technique can do little violence to the subsequent tabulations. When imputation is necessary for a large proportion of the observations we face a serious problem. It is obvious that we have been unsuccessful in our attempt to collect the information. Whether we should (a) suppress publication, (b) publish what we were able to obtain, (c) embark on a reenumeration activity, or (d) take some other kind of action are decisions that must be made in this kind of a situation. One thing we certainly would not plan to do in such an unhappy event would be to expect our electronic computer to choose an appropriate course of action.

Discussion

A. J. Jaffe, Bureau of Applied Social Research, Columbia University

The most important element, it seems to me, involves the purpose for which a sample is drawn or tabulations made, or statistics otherwise collected or processed. Once the questions are answered, "What will you do with the statistical findings? What action will you take?" we have a reasonable basis for deciding what type of sample to draw and how accurate it should be, or what type of tabulation equipment to use.

Mr. Marks' tripod is insufficient, it seems to me, as a basis for a sound sample. In addition to simplicity, good design, and good control, the purpose to be served by the statistics must be borne in mind. In this connection it seems to me that most purposes for which statistics will be used, especially in lesser developed countries, require knowledge of approximate levels, rather than highly exact information. For example, information that the ratio of female children under age five to women aged 15-44 is 100 children per 100 mothers establishes the fact that the birth rate is very high. Whether the exact ratio is 95 or 105 is irrelevant for any action purpose. Therefore, a sample should be accurate enough to establish the fact of a high birth rate, but need not be accurate to within one percent, or less, of the presumed "correct and true" ratio of children to women.

If you accept this emphasis on purpose to be served by the statistics, then, it seems to me, simplicity is of very great importance, especially in lesser developed areas where there may be shortages of qualified clerical and statistical workers. Simple procedures which can be followed even by semi-literate clerks and control procedures which a poorly educated supervisor can follow, become of over riding importance.

I emphasize simplicity even at the expense of "good sample design" if necessary. The design need be good enough only to provide a reasonably good estimate of the level of the statistic commensurate with the purposes for which the statistics are to be used. If the sample is so designed as to provide the best possible estimate of the complete universe from which it was drawn, and as a result the procedures required to handle it are not as simple as they might otherwise be, that design is not good, in my opinion.

I apply the same reasoning to the use of computers and other ultra-modern tabulation equipment, as reported in Mr. MacPherson's paper. If one must tabulate the census for all the 100 or 200 million or more people in a country, it is certainly desirable to have the most modern tabulation equipment. However, if the uses to which the statistical results are to be put, can be served by tabulating information for as few as one-half million people, then less emphasis and importance need be attached to the nature of the tabulation equipment. One-half million cases can be tabulated with semi-modern equipment, if need be, (and taking into consideration the skills of the operating personnel available and the funds available for purchasing equipment). I submit that in very many cases, in the majority of cases, the final purposes to be served by the statistics, can be had by tabulating a small sample of all the cases. Therefore, I am not unduly worried by the nature of the tabulation equipment which might be utilized.

In short, I am emphasizing the necessity to think through problems and their possible solutions. I do not believe that one can depend on "tried and true text book" methods, nor on inanimate machinery, to do the basic thinking necessary. I admit, however, that if homo sapiens will not think, then automatic procedures - learned or operated by rote - are preferable to no procedures.

As for Mr. Booker's paper, I am glad to see that he updated a few of the statistics presented in the United Nations document Patterns of Industrial Growth, 1938-1958, prepared by Mr. A. Aidenoff (1960). Mr. Booker's emphasis on the fact that measuring change in terms of percentage increase, or decrease, often exaggerates the true situation, is excellent; one must examine the absolute values (of income, employment, etc.) and absolute changes in such values before one can determine how much "progress" is actually being made in a country.

For a fuller and more complete description of what has transpired in the fields of production and employment during the past two decades, this discussant recommends the volume, Patterns of Industrial Growth.

IV

EDUCATIONAL STATISTICS

Chairman, Francis G. Cornell, Educational Research Services, Inc.

SOME STATISTICAL FINDINGS FROM NATIONWIDE TEACHER POLLING

Glen Robinson, National Education Association
Chester H. McCall, Jr., Booz-Allen Applied Research, Inc.

Introduction

Data collection and processing are becoming increasingly important in the field of education. To be of maximum value, such data must be obtained rapidly, accurately, and on a nationwide basis.

During the past two and one-half years, the Research Division of the National Education Association has been engaged in a project to develop and test data collection and analysis procedures to meet the rising demands in the field of education. The purpose of this paper is to describe some of the results of this project.

Background of Project

During the early 1950's, there were few timely and accurate nationwide statistics that pertained to teacher and school principal populations. Those that were reported represented either compilations of statewide figures or compilations of statewide estimates. There were practically no studies based on probability samples drawn from all the school districts in the nation. One reason for this was the lack of appropriate listings of all the school districts.

A new era was opened in nationwide educational sampling when the 43,500 school districts in the nation enumerated in the 1957 Census of Governments were placed on magnetic tape by the Governments Division of the U. S. Bureau of the Census. For the first time probability sampling of the nation's school districts could be achieved at a reasonable cost.

Meanwhile, new research techniques in probability sampling had been developing in areas outside education, and the NEA Research Division staff believed that the application of these techniques to the Division's work could result in greater efficiency, reliability, and speed. Consequently, the staff undertook an intensive project to apply and test the best of these new sampling and polling techniques in the field of educational research. The study was considered to be of prime importance to the future work of the Research Division.

Since the extensive work involved in a project of this scope had to be carried on concurrently with the Division's regular operation, special assistance was needed for the project. The firm of Booz-Allen Applied Research, Inc. was retained as technical consultants.

The project had three primary purposes:

1. To study and evaluate the sampling procedures then being used by the NEA Research Division
2. To examine, develop, and test recent sampling and polling techniques which could be properly applied to the types of studies made by the Division
3. To develop within the Division a continuous training program in the use of new techniques.

To date the following nationwide surveys have been made as a part of this project:

1. Three polls of teachers' opinions
2. Two studies of teacher mobility and turnover
3. One study of the status of the American teacher
4. One poll of elementary-school principals' opinions

All seven studies were mailed surveys. These studies were carefully designed and co-ordinated so as to yield a maximum of data relating to sampling problems such as the effect of nonresponse bias, the effect of sampling error, the time required for response, and the effectiveness of certain techniques in improving response rates. These data have been placed on punch cards for detailed analysis. Such analyses are presently under way.

General Procedures

The design used in the studies was a two-stage cluster sample. The first stage consisted of a stratified random sample of school systems (clusters) drawn from the 43,500 school districts enumerated by the Bureau of the Census in its 1957 Census of Governments. The first-stage sample was selected for the NEA Research Division according to the Division's specifications by the Bureau of the Census.

Letters were sent to the superintendents in each of the sample systems requesting a copy of their directory listing the names and school addresses of teachers and principals in their school system.

The second stage consisted of a random sample of teachers or principals drawn from these school districts. Samples were selected so that all teachers in the nation had the same probability of being selected. Hence, the samples were self-weighting.

The survey instruments were mailed directly to the sample of teachers or school principals. These questionnaires were accompanied by a personally addressed and typed letter, signed by the Director of the Research Division, explaining the nature of the survey and the importance of the teacher's response. The general pattern of follow-up procedures used consisted of a personal letter sent by air mail to nonrespondents at the end of two weeks. Telegrams were sent to nonrespondents at the end of four and five weeks. The mean net response rate (useable questionnaires returned) for the seven studies was 95.7 percent. Some persons did not respond because they were critically ill, had died, had moved without leaving a forwarding address, or had written saying they would not take part in the poll. When this group was added to the net response group, a gross response rate of 97.7 percent was obtained for the seven studies.

Survey instruments were pre-tested on samples of teachers drawn from school systems not selected in the first-stage sample.

Some of the statistical findings resulting from these procedures will be described later in this paper.

General Results and Implications of Project

Although much work remains to be done in analyzing the data received and in making improvements in the sampling procedures, a number of benefits are already resulting from the sampling project. These are:

1. Improved accuracy in data collection and analysis by the NEA Research Division. Since these samples were drawn in accordance with probability theory, it is possible to report the results with a stated degree of accuracy and confidence. The high response rates have greatly reduced the possible nonresponse bias. Analysis of data is yielding more knowledge of nonresponse groups.

2. Development and use of better sampling frames. Although the first stage of the samples for the seven studies was drawn for the NEA Research Division by the Bureau of the Census from 1957 lists, the Division has since purchased from the Census Bureau a complete listing of all school systems as of the school year 1959-60. This more recent information is on IBM cards and is presently being used by the Division for sampling purposes.

3. More timely reporting of data. As a result of the project, it is now possible to report survey data within the same school year in which the data were collected. These surveys have demonstrated that with proper sampling and follow-up procedures, response rates in excess of 90 percent can be achieved within two months of initial mailing date.

4. Reduced unit cost. A substantial reduction in the unit cost of surveys has been achieved by the project. An appraisal of the full extent of this accomplishment is forthcoming. A work-sample survey was conducted by the Research Division staff for a six-month period to estimate the unit cost of conducting surveys by the new and old procedures. These data are scheduled for analysis in the near future.

5. Implications for future operation. The sampling project has initiated a new period in improved data collection and analysis within the NEA Research Division. Improved sampling techniques will facilitate the use of survey designs not previously feasible. Improved sampling techniques have also indicated that further improvements can be achieved through the use of automated data processing. Action is being taken to secure computer equipment (of the IBM 1620 class) in the immediate future.

Some Statistical Considerations

The previous sections of this paper have presented a brief description and background of the present sample survey project within the NEA Research Division. This portion of the paper will emphasize the statistical and related aspects of the program.

The test vehicle chosen to examine the usefulness of a two-stage sampling procedure was the Teacher Opinion Poll. These polls consisted of approximately 10 questions designed to stimulate interest and hence elicit a high response. Eleven status questions were selected to test the consistency of the survey findings with external estimates and the reproducibility of survey findings from one year to another. The first Teacher Opinion Poll was mailed February 1960. The following section describes the method by which the sample was selected.

Teacher Opinion Poll Design

The first problem encountered in selecting a design for sampling teachers was the lack of a complete listing of all public-school teachers. It was necessary, therefore, to examine other possible procedures. Since the Bureau of the Census maintained a listing of all school systems in the United States on magnetic tape, it was believed that a suitable design would be the selection of school systems by some random device and then, having selected these systems, to obtain a sample of individual teachers. This design was selected out of necessity rather than desired optimum sampling. The two-stage plan was executed in the following manner:

1. School systems were stratified into eight strata according to the number of teachers in the system as of October 1956.

2. A sample of 31 systems was selected from each of the eight strata. From these systems a random selection of teachers was then obtained. The teachers selected constituted the sample for the teacher opinion polls and other studies of teacher populations.

With only limited information available on the important characteristics of the population stratified by this method, it was felt that a first estimate of an appropriate sample size would be to assume the required sampling for a purely random selection from among all teachers in the United States. This method was utilized to ascertain the required sample size. It is recognized that minimum accuracy requirements may not have been met in some cases because of the effect of the inter-cluster contribution to the variance of the estimates. As the information from these studies is analyzed, it will be possible to revise the sampling procedure and bring it more in line with optimum allocation techniques.

In 1960, two teacher samples were selected from the same 250 systems. These are referred to as Teacher Opinion Poll No. 1 and Teacher Opinion Poll No. 2. In 1961, a new set of systems was selected from the Bureau of Census records. Teacher Opinion Poll No. 3, The Status of the American Teacher, and The Elementary-School Principal Opinion Poll were all selected from the 1961 sample of systems. In 1961, the respondents to Teacher Opinion Polls 1 and 2 were mailed a questionnaire designed to measure teacher mobility and turnover. These latter two studies were made in order to assess the amount of change occurring in teacher sampling frames within a one-year period.

The design used in the Teacher Opinion Polls has been criticized by some persons as not being the most efficient design. We are aware of this possible weakness and would point out that the major reason for selecting the sample in this way was the restriction imposed on the project by the lack of an appropriate listing of all teachers. As the data from the studies of the past two years are more thoroughly analyzed, more appropriate statistical designs will be introduced where feasible.

Response Rates

One of the criticisms which is frequently directed at those who use mailed survey techniques is the generally poor response. The project has overcome this deficiency through careful planning, efficient control of mailing lists, appropriate processing of incoming questionnaires, and use of effective follow-up procedures. Remarkable response rates were achieved in seven surveys conducted during 1960 and 1961. Table 1 shows the response rates for each of the seven surveys, the mailing date, the cut-off date for accepting replies, and the sample size. It is interesting to note that the lowest net response rate was approximately 93 percent. While we have run no statistical tests on this, we find little reason to believe that within reasonable limits the length of a questionnaire necessarily influences the tendency to respond. It is possible that the time of mailing within the school year influences response as much as the length of the questionnaire.

TABLE 1.--MAIL AND RESPONSE INFORMATION FOR SEVEN SAMPLE SURVEY STUDIES, 1960 AND 1961

Sample survey	Length of questionnaire	Date mailed	Date of last reply	Sample size	Response rate Net[a/]	Response rate Gross[b/]
1	2	3	4	5	6	7
Teacher Opinion Poll 1	2 pages	23 February 1960	20 April 1960	1149	.968	.981
Teacher Opinion Poll 2	2 pages	29 March 1960	24 May 1960	1147	.949	.982
Teacher Opinion Poll 3	2 pages	3 February 1961	7 April 1961	1633	.966	.975
Teacher Mobility and Turnover 1	1 page	27 February 1961	13 June 1961	1112	.950	.989
Teacher Mobility and Turnover 2	1 page	29 March 1961	14 June 1961	1089	.965	.983
Status of American Teacher	4 pages	27 March 1961	27 June 1961	2104	.926	.949
Elementary-School Principal Opinion	3 pages	17 February 1961	18 May 1961	721	.976	.978

a/ Net response rate includes only those persons who returned useable questionnaires. Average net response rate for seven studies was .957.

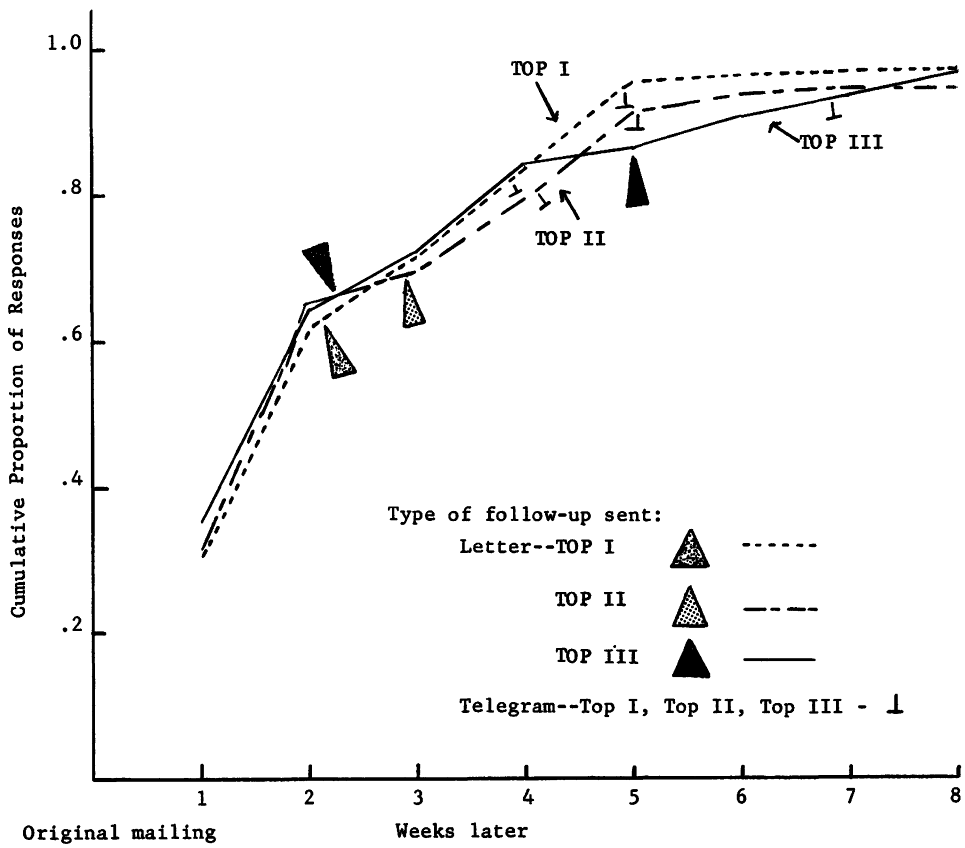
b/ Gross response rate includes persons who returned useable questionnaires plus those persons whose reason for nonresponse could be determined--deceased, critically ill, moved without leaving forwarding address, or wrote saying they refused to answer questionnaire. Average gross response rate for seven studies was .977.

A system of questionnaire numbering was used to maintain records and to identify nonrespondents.

In conducting the studies, special effort was made to glean as much information as possible about the populations studied and their response patterns. For example, Figure I shows the cumulative response curves for the three Teacher Opinion Polls along with the mailing dates of the follow-ups. The striking agreement among these three curves suggests that this pattern might be expected when similar follow-up procedures are used with the same type of population and questionnaires. On the other hand, the effect of deviation from the general follow-up procedure, described previously, is clearly evident in the chart.

There was also need for research on the tendency to respond by members of various subgroups within the teacher population. For example, little was known about the response patterns of men teachers versus women teachers or elementary-school teachers versus secondary-school teachers. In

FIGURE I--RESPONSE PATTERNS FOR THREE NATIONWIDE TEACHER OPINION POLLS, 1960 AND 1961



Cumulative response rates[a]			
Week	TOP I	TOP II	TOP III
1	.301	.315	.351
2	.614	.648	.640
3	.711	.694	.722
4	.841	.797	.842
5	.953	.912	.868
6	.963	.940	.909
7	.968	.949	.933
8	.968	.949	.966

a/ Response rates plotted are cumulative net response rates which include only those persons returning useable questionnaires.

each of the three Teacher Opinion Polls the date on which the questionnaire was mailed, the date the questionnaire was received, and the postmark date on the return envelope were noted on the returned questionnaires. This information yielded approximate estimates of the length of time the questionnaire was actually in the respondent's hands; these data were punched into the IBM card for each questionnaire. Some interesting observations were made on the basis of a preliminary analysis of the time-in-respondent's-hands data. This analysis indicated that there was a highly significant difference (probability less than .001) between the response curves for men and women teachers in both elementary and secondary schools. Men were found to respond more rapidly.

Although not a part of the seven studies, another interesting experiment, conducted in 1961, pertained to factors affecting response rates. This was done in a study of small high schools in the United States. Instruments were sent to both the principals and the teachers in a sample of small high schools. Half the questionnaires to each group were sent in stamped envelopes. The other half were sent in metered envelopes. The follow-up procedures were identical for both groups. Utilization of the Kolmogoroff-Smirnov test demonstrated that in both groups the response distribution was significantly different at the .001 level, with the persons receiving stamped envelopes answering in greater proportion. The inference here is that the use of postage stamps created either a more personal impression or an impression of importance that stimulated greater response. Further study is needed to substantiate this.

<u>Selected Estimates</u>

Several choices in the selection of statistical estimates were available for analyzing the two-stage sample design used in the Teacher Opinion Polls. An excellent reference for such estimates is a book by Sukhatme.[1/] Of the various estimates which were available, it was believed than an unbiased estimator would be the most appropriate. For this reason the design was set up so that the probability of being selected in the sample was identical for all teachers in the United States. This procedure, in effect, yielded a self-weighting sample.

An estimate for the mean (in the general case) of the i^{th} stratum when k_i clusters were subsampled is:

$$\bar{x}_i = \frac{1}{N_i} \times \frac{K_i}{k_i} \sum_{j=1}^{k_i} N_{ij}\bar{x}_{ij} \qquad (1)$$

where $$\bar{x}_{ij} = \sum_{t=1}^{n_{ij}} x_{ijt}/n_{ij}$$

The variance for these sample estimates, however, is much more complicated. An estimate of the variance of the sample mean in the i^{th} stratum, for example, is given by the following expression:

$$\hat{\sigma}^2_{\bar{x}} = \left\{\frac{1}{k_i} - \frac{1}{K_i}\right\} s^2_{bi} + \frac{1}{k_i K_i} \sum_{j=1}^{k_i} \left\{ \left(\frac{N_{ij}}{\bar{N}_i}\right)^2 \left(\frac{1}{n_{ij}} - \frac{1}{N_{ij}}\right) s^2_{ij} \right\} \qquad (2)$$

where $$s^2_{ij} = \sum_{t=1}^{n_{ij}} \frac{(x_{ijt} - \bar{x}_{ij})^2}{n_{ij} - 1} \qquad (3)$$

$$s^2_{bi} = \frac{1}{k_i - 1} \sum_{j=1}^{k_i} \left\{ \frac{N_{ij}}{\bar{N}_i} \bar{x}_{ij} - \bar{x}_i \right\}^2 \qquad (4)$$

and $$\bar{N}_i = \sum_{j=1}^{K_i} \frac{N_{ij}}{K_i} \qquad (5)$$

The terms in the equation are as follows:

- K number of strata in population.
- K_i number of clusters in i^{th} stratum.
- k_i number of clusters selected from i^{th} stratum. (note that $k_i \leq K_i$).
- N_{ij} number of observations in j^{th} cluster in i^{th} stratum.
- N_i number of observations in i^{th} stratum.
- N total population size; $N = \sum_{i=1}^{K} N_i$.
- N'_i total number of observations in clusters which are being sampled in i^{th} stratum.
- n_{ij} number of observations in sample from j^{th} cluster in i^{th} stratum (note that n_{ij} may equal zero).
- n_i number of observations in total sample from i^{th} stratum;

$$n_i = \sum_{j=1}^{k_i} n_{ij}.$$

1/ Sukhatme, Pandurang. <u>Sampling Theory of Surveys with Applications</u>. Ames: Iowa State College Press, 1958.

n total sample size; $n = \sum_{i=1}^{K} n_i.$

X_{ijt} the t^{th} sampled observation in the j^{th} sampled cluster in the i^{th} stratum.

Since $\bar{N}_i$, the average number of observations per cluster in the i^{th} stratum, is not available in this particular equation, it is estimated by the following equation:

$$\hat{\bar{N}}_i = \bar{n}_i = \sum_{j=1}^{k_i} \frac{N_{ij}}{k_i} \qquad (6)$$

We recognize that our use of this term may be subject to some error.

It should be noted that S_{ij}^2 is the variance within the j^{th} sampled cluster in the i^{th} stratum. S_{bi}^2 is the weighted variance among the sampled cluster means in the i^{th} stratum. In addition, the value of K_i is not known for the two years under study. This figure has been estimated by the projection and interpolation of data for the school years 1956-57 and 1959-60. We are of the opinion that the errors in these estimates do not contribute significantly to the amount of error in the estimates of the variance of sample means.

An estimate for the mean in the total population is given by the expression

$$\bar{\bar{X}} = \sum_{i=1}^{K} \frac{N_i}{N} \bar{X}_i$$

$$= \sum_{i=1}^{K} \frac{K_i}{k_i} \sum_{j=1}^{k_i} \frac{N_{ij}}{N} \bar{X}_{ij}. \qquad (7)$$

While an estimate for the variance of $\bar{\bar{X}}$ is then seen to be

$$\hat{\sigma}^2_{\bar{\bar{X}}} = \sum_{i=1}^{K} \left\{\frac{N_i}{N}\right\}^2 \hat{\sigma}^2_{\bar{X}_i} \qquad (8)$$

The statistics which have been tabulated for presentation in this paper are the following mean characteristics for the classroom-teacher population: years of teaching experience and age of teachers. Table 2 summarizes the estimates of these parameters for 1960 and 1961 as found in Teacher Opinion Polls 1 and 3. In addition, included in Table 2 are the estimates of the standard errors of these characteristics from equation (8) and also under the assumption of a purely unrestricted random sample, as computed from equation (9), namely,

$$\hat{\sigma}^2_{\bar{\bar{X}}} = \frac{\sum\sum\sum\left(X_{ijt} - \bar{\bar{X}}\right)^2}{n(n-1)} \qquad (9)$$

Although the estimates of the standard error, as derived from equation (8), indicate variances greater than we feel are desirable, they provide a good base from which to make adjustments in the sample design. From a detailed analysis of the strata and cluster data used in computing these estimates, it should be possible to improve the sample design and adjust the sample size so as to reduce these variances. As a consequence of the observed standard errors, it appears worthwhile to consider the use of estimates which will have smaller variation; for example, an estimator which will provide for a reduction in the magnitude of S_{bi}^2. One possible estimator would yield

$$S_{bi}^2 = \frac{1}{k_i - 1} \sum_{j=1}^{k_i} \left(\frac{N_{ij}}{\bar{N}_i}\right)^2 (\bar{X}_{ij} - \bar{X}_i)^2. \qquad (10)$$

Unfortunately, for this presentation time did not permit an evaluation of other possible estimates and standard errors.

We would welcome any suggestions or comments that might be helpful in improving either the sample design or the method of assessing sampling variability.

TABLE 2.--NATIONWIDE ESTIMATES OF SELECTED CLASSROOM-TEACHER POPULATION CHARACTERISTICS AND THE STANDARD ERRORS OF THESE CHARACTERISTICS

Characteristics	Survey[a/]	Estimate of mean	Standard error	
			Computed by Equation (8)[b/]	Computed by Equation (9)[c/]
1	2	3	4	5
Teaching experience	TOP 1	14.16 years	.58	.35
Teaching experience	TOP 3	13.74 years	.66	.28
Age	TOP 1	40.76 years	1.31	.37
Age	TOP 3	40.35 years	1.85	.32

a/ Teacher Opinion Poll 1 was mailed February 1960; Teacher Opinion Poll 3 was mailed February 1961.

b/ Equation (8) provides an estimate of the variance of the sample mean which takes into account the amount of variance derived from the use of a cluster sample. These numbers are based upon a preliminary machine run and are subject to revision.

c/ Equation (9) provides an estimate of the variance of the sample mean which assumes an unrestricted random sample.

ON ESTIMATION OF SCHOOL POPULATION

Philip M. Hauser and Evelyn M. Kitagawa
University of Chicago

Never before in the history of the United States has the need for estimating school population been greater than at the present time. Never before has such estimation been more difficult. For the forces determining school population in a given locality have never before assumed the complexity, the scale, or the variability of those now at work on the American scene.

More specifically, efforts to estimate school population today must, among other things, reckon with great population and community changes, and changes also, in proportions of children enrolled in school. These changes include among the population factors: resurgent national growth; increased concentration in Standard Metropolitan Statistical Areas; accelerating decentralization within metropolitan areas; great waves of in-migrant peoples of diverse race, cultural background and school enrollment practices; and rapidly changing age structure. Simultaneously, almost unprecedented community changes are under way, including changes in patterns and densities of land use arising from urban renewal programs, public housing projects, federal and local highway construction, increased utilization of physical planning techniques, and ever increasing traffic flows. Furthermore, the problem of estimating school population for specific local educational purposes is not made easier by the differential rates of expansion of public, private and parochial schools, respectively, and by the increasing need to replace obsolescent or obsolete school structures even while grappling with problems of rapid growth.

The interplay of these factors is by no means uniform and it requires a separate evaluation city by city, community by community and school district by school district to estimate school population. Most American communities are faced with problems of rapid growth; but some also face actual decline, not only in total population but, also, in school population. In some areas, trends may be stable enough for reasonably accurate and useful projection; in others, the erratic character or the magnitude of population and community changes may defy all efforts and all available techniques to predict.

Efforts to estimate school population must necessarily take into account each of the changes of the type to which reference has been made and their interaction as they affect particular metropolitan areas, cities, or subdivisions thereof. In general, the larger the area for which the estimate is being made, the more accurate is it likely to be. The smaller the area, the greater is the likelihood of error. Moreover, any number of special circumstances may affect small areas, particularly individual school districts, including matters of policy of various local administrative units.

Let us turn to a consideration of the types of methods that are available for estimating school population in different types of situations, with the national trends in mind. Because of the difficulties inherent in attempting to predict the actual course of future events--especially births and migration--that affect the size of school enrollment at future dates, the term "projection" rather than "prediction" has been attached to most figures computed for this purpose. For the same reason, short-term projections are subject to less error than long-term projections. In fact, if the projection period extends no more than five years into the future, all of the children who will be enrolled in school at that time are already born and there is no need to estimate future births, and one potential source of error is eliminated. In general, it can be ventured that local projections beyond ten years into the future are subject to such a wide range of error as to be of little use for planning purposes.

In general, three approaches have been used by persons or agencies to estimate future school populations: (1) Establish grade progression ratios from one grade to the next in successive years in the past, and apply these ratios to the numbers currently enrolled in each grade to obtain estimates of enrollment for successive years in the future; statistics on births for appropriate years are used to estimate kindergarten and first grade enrollment in this procedure. (2) Obtain projections of the population of school age for desired future dates, estimate enrollment rates for the same dates and apply them to the population projections. (3) Obtain projections of dwelling units for desired future dates, and estimate school enrollment from projected ratios of population and students to dwelling units. The use of grade progression ratios is less complicated and time-consuming than the cohort-survival method of obtaining population projections, and it usually should give as accurate results providing reliable statistics on school enrollment are available for a series of consecutive recent years. The third method is particularly applicable to new or rapidly expanding communities, where only rough approximations to future enrollments can be expected. The three approaches to school enrollment projections are discussed below, for the elementary and high school grades.

I. GRADE PROGRESSION METHOD

The procedure to be followed to obtain projections of school enrollment by this method can be described in three steps.

Step 1. Compute grade progression ratios for each grade of elementary and high school, from data on school enrollment for successive years in the area for which the projection is to be made. For example, dividing the number of students enrolled in the third grade in 1951 by the number enrolled in the second grade in 1950 gives the grade progression ratio for grades 2 to 3 from 1950 to 1951. Similar ratios are computed for the other elementary and high school grades, using data for the same years. If there has been no migration into the area during the two-year period, the ratios will be less than unity (assuming enrollment data for the two years are comparable and accurate), since there will be some losses due to mortality or drop-outs from one grade to the next. Grade progression ratios in excess of unity usually indicate migration into the area (although the ratio for kindergarten to grade one may be an exception, and there may be other special cases). For this reason, it is preferable to compute a series of grade progression ratios for a span of the last five or ten years, or longer if the basic data are available. This gives some indication of the trend in the ratios over time, and permits extrapolation of ratios from the observed time series into the future. Before grade progression ratios are utilized in this fashion, care should be taken to ascertain: (1) that the enrollment data are compiled for the same area and on a comparable basis during the time period covered, preferably a count of enrollment as of October of each year; (2) that the figures are reasonably complete and accurate; (3) that enrollment in parochial and other private schools are added to public school statistics before the ratios are computed. If the necessary enrollment statistics are not available for parochial or other private schools, ratios may have to be based on statistics for public schools only, but in this case the proportion of enrollment in public schools on the projection date must be estimated if total enrollment projections are to be obtained (see step 3).

Step 2. Estimate grade progression ratios for the projection period. For example, if projections up to October 1963 are the objective and if the latest enrollment statistics are for October 1958, it is necessary to estimate grade progression ratios for 1958-59, 1959-60, 1960-61, 1961-62, and 1962-63. These five ratios for each grade sequence--from grade two to grade three, from grade three to grade four, etc.--may be estimated by extrapolating from the time series of observed ratios for that grade sequence. For instance, if in step 1 ratios have been computed from enrollment statistics for 1950 through 1958, the eight ratios for progression from grade two to grade three in successive years--1950-51, 1951-52, etc. to 1957-58--may be plotted on a graph. If the 1950-58 statistics are comparable and reliable, changes in the ratios from one year to the next during this period should be due primarily to migration, at least in the elementary grades where drop-outs are very low. Hence, for each elementary grade, the 1959-1963 grade progression ratios may be extrapolated from the observed 1950-58 ratios by postulating an increase in the ratio after 1959 if it is assumed that migration will be higher in 1959-63 than in 1950-58, or by postulating a decrease in the ratio if it is assumed migration will be lower in 1959-63. Extrapolation of grade progression ratios for high school grades may be more difficult. Job opportunities and the military draft influence drop-out rates at the high school level and, especially for boys, migration may not be the major influence on changes in grade progression ratios over time.[1/] As a result, ratios for these grades should be extrapolated with assumptions that take into account these two factors as well as migration.

Step 3. Apply the estimated grade progression ratios to current enrollment figures to obtain projections of future enrollment. Projections of school enrollment in each grade on October 1959 are obtained by multiplying the October 1958 enrollment in the preceding grade by the estimated 1958-59 grade progression ratio between the two grades. For example, the projected enrollment in grade five on October 1959 is computed by multiplying the observed October 1958 enrollment in grade four by the estimated 1958-59 grade progression ratio for grades four to five.

Similarly, projections of enrollment in each grade on October 1960 are computed by multiplying the estimated October 1959 enrollment in the preceding grade by the estimated 1959-60 grade progression ratio between the two grades. Hence, projected 1960 enrollment in grade six is obtained by multiplying the projected 1959 enrollment in grade five by the estimated 1959-60 grade progression ratio for grades five to six.

In this way enrollment in each grade is estimated for each year from the current date to the last projection date. If the observed grade progression ratios computed in step 1 are based on enrollment statistics for public schools only, an analysis of recent trends in the proportion of elementary and high school enrollment in parochial and other private schools can be used to estimate the proportion expected to be in such schools on the projection dates. This proportion can be used to inflate the projections obtained for public school enrollment using the grade progression ratios computed for public schools.

Special procedure for kindergarten, grades one and two. Factors such as the popularity and availability of kindergarten facilities, and changes in the tendency to retard first grade pupils, greatly influence trends in grade progression ratios for these grades. As a result, enrollment projections for kindergarten and grade one may be based on extensions of observed trends of the relationship between births and enrollment in these grades, and the progression pattern for grades one to two can be adjusted on the basis of expected school policy regarding retardation.[2/] If ratios of kindergarten enrollment in a given year to the number of births five years earlier, and ratios of first grade enrollment to births six years earlier, are to provide the basis for making projections of future enrollment in these two grades, it is necessary not only to compute these ratios for a number of recent years and to estimate corresponding ratios for the projection dates, but if the projection period extends more than five years into the future, it is also necessary to estimate the number of births in future years. For example, if the projection period extends ten years into the future, births for the first five of those years must be estimated in order to project kindergarten and first grade enrollments for the last five years of the period.

II. ENROLLMENT RATES APPLIED TO POPULATION PROJECTIONS

While the procedure to be followed in making enrollment projections by the grade progression method is more or less clear-cut and can be illustrated by one example, this is not true of the second approach. There are several different methods for making projections of future populations, and the appropriate method in a given situation depends on a number of factors, including the length of the projection period, the kind and quality of available data, the size and location of the community or area for which the projection is to be made, and the rate of net migration to the area.

The problem is further complicated by the fact that if several years have elapsed since the last census of the area, it is advisable first to make estimates of the current population, by age, before making projections of the population as of a specified future date. Because methods for estimating the current population of an area utilize one or more series of current data that are "symptomatic" of population change in the area since the last census (for example, annual statistics on births, deaths, school enrollment, residential construction, electric meters, telephones, etc.) projections made from current population estimates based on such data are probably subject to less error than projections made directly from the recorded population of the last census. Thus, the person who wishes to make projections of school enrollment at some future date by first making population projections for the date, and then applying enrollment rates to these projections, not only has to choose a projection method appropriate to his situation but, if a number of years have elapsed since the last census, he must also select (from a variety of methods) an appropriate procedure for making current population estimates.

It is not our purpose here to describe in detail the various methods that have been used to make estimates of current population, or the techniques for obtaining projections of future population. References to available source materials on both subjects are cited below, together with brief summaries of the approach in several commonly-used methods. These sections are followed by a description of the procedure for converting projections of school-age population to projections of school enrollment.

Estimates of current population. The United States Bureau of the Census has published the results of a survey of the sources and types of local population estimates prepared by state and city agencies (Current Population Reports, Series P-25, No. 178, June 27, 1958). The report includes brief explanations of the various methods for making current population estimates, and cites references to more detailed source materials that can be used as guides for computation purposes. It also gives some evaluation of the accuracy of different methods. The table below indicates how often each method was used by state agencies to make county estimates, and by city agencies to make city estimates.

Method used	Number of state agencies	Number of city agencies
Agencies reporting, total	62	32
Migration & natural increase	28	7
Bureau of Census, Method I	3	
Bureau of Census, Method II	15	
Combination with other method	5	
Other	5	
Composite method	5	2
Natural increase alone	5	2
Censal ratio	7	17*
Simple form	5	
Complex form	2	
Proration	6	-
Arithmetic extrapolation	4	2
Other	7	2

*Principally the dwelling unit method

The brief explanations given for these methods in the Census Bureau report are repeated below:

Migration-and-natural-increase methods are those in which the components of population change (i.e., natural increase and net migration) are estimated separately. In Method I of the Bureau of the Census, the net migration rate for a given area is estimated, on the basis of school enrollment or school census data, as the difference between the percentage change in the population of school age for the area and the corresponding change for the United States.[3/] In Method II of the Bureau of the Census, net migration is estimated, using school enrollment or school census data, from the difference between the actual population of elementary school age and the population expected on the basis of the 1950 Census and births, and from current data on the variation of migration rates by age.[4/] In the other migration-and-natural-increase methods, net migration is estimated in various other ways, e.g., by the use of data on school enrollment for successive school years and grades (grade-progression method), by the use of estimates of net migration for previous periods, etc.

The composite method makes use of several series of "indicator" data--births, deaths, school enrollment, etc.--to estimate the size of the various age segments of the population to which these basic indicators are most applicable. Summing the estimates for separate age groups yields an estimate of the total population. In one form of the composite method--the age-specific death rate method--deaths from all causes or from selected causes are used as an indicator series for all or most age groups.

The natural increase method involves merely adding postcensal natural increase (births minus deaths) to the census figure. It assumes, therefore, that postcensal net migration equals zero.

The simple form of the censal ratio method involves (1) computation of ratios of population to a single symptomatic element (school enrollment, births, deaths, etc.) at the last census date for each county, and (2) application of these ratios to the corresponding postcensal symptomatic element to obtain postcensal estimates of county population. Sometimes, these county estimates are adjusted to make them add to an independent estimate for the state as a whole. In the complex form of the censal ratio method, specific allowance may be made for the postcensal change in the ratio of population to symptomatic data or two or more simple ratio estimates may be averaged. The vital rates method, for example, averages two estimates based respectively on birth and death statistics, and allows for the postcensal change in the birth and death rates. In the dwelling unit methods, data on building permits issued or data on electric, gas or water meter connections are used to measure postcensal changes in the number of dwelling units, and assumptions are made regarding postcensal changes in the number of persons per occupied dwelling unit and vacancy rates.[5]

The proration method involves commonly the distribution of the postcensal state total on the basis of current "symptomatic" data such as school data, births, and deaths. This procedure implicitly assumes that the ratio of population to the symptomatic item is the same for all areas in the state. The state estimate may also be prorated on the basis of local population at the last census.

In arithmetic extrapolation, it is assumed that the yearly amount of population change in an area in the postcensal period equals the average yearly amount of change in the area in a recent past period, usually the most recent intercensal period. In geometric extrapolation, the average yearly rate of change is assumed to remain the same as in the past period.

Most of the methods described in the preceding paragraphs yield estimates of the total population currently resident in an area. However, in order to obtain projections of the future population by the cohort-survival method it is also necessary to know the age composition of the present population. In this respect, the composite method of estimating current population has the advantage that the age composition is obtained as part of the estimating procedure. When other methods are used, the age composition may have to be estimated by additional assumptions after the total population estimate has been obtained. For example, the estimated current age distribution for the United States as a whole, which is published annually by the Census Bureau, might be used as a guide or it may be adjusted on the basis of the relationship between the age distribution of the local area and the age distribution of the United States at the last census.

Projections of future population. The three most commonly used methods for making projections of the future population of small areas are the ratio method, the cohort-survival method, and a simpler type of component method in which separate allowance is made for births, deaths, and migration but without taking age into account. Other methods include graphic extrapolation and mathematical extrapolation, and generalized types of projections, for example of dwelling units.[6]

The cohort-survival method involves separate projections of the three components of population change--births, deaths and migration--with age taken into account. That is, after careful analysis of past trends, and on the basis of general assumptions concerning future economic, military and health conditions, projections are made of age-specific mortality rates, age-specific fertility rates, and age-specific migration rates throughout the projection period. These age-specific rates are then applied to the estimated age distribution of the current population (or the recorded age distribution at the last census if only a short time has elapsed) to obtain population projections by age for the desired dates. Often, several alternative assumptions for the components of population change are projected, and a range of population projections rather than a single projection is the end product.

The procedure used by the Census Bureau for its latest projections of the United States population, by age and sex, for 1960 to 1980, is described in their Current Population Reports, Series P-25, No. 187 (November 10, 1958). This report discusses in detail the fertility and mortality assumptions used by the Bureau, and provides useful guides for persons who may wish to make similar assumptions for local areas. Along this line, the Census Bureau has also made projections of the population of each state to 1970, using fertility and mortality assumptions derived by multiplying the projected birth rates for the nation (that is, the rates used to make the United States projections) by a factor expressing the expected ratio of the state's birth rate to the nation's (see Current Population Reports, Series P-25, No. 160, August 9, 1957). In computing these factors it was assumed that the state differences in fertility would disappear in 50 years, and the expected ratio for a given state on a specified projection date was the result of an interpolation between the observed ratio for the state in 1950-55 and an assumed ratio of unity 50 years later. Mortality assumptions for states were derived from United States life tables for whites and nonwhites, using the racial composition of each state to obtain a weighted average of age-specific white and nonwhite mortality rates. A similar approach could be used by local areas to derive fertility and mortality assumptions for the cohort-survival method. However, the migration assumptions for a local area will usually require special analysis of the local situation. A method for estimating the net migration rate in local areas since the last census is described in the Census Bureau's Current Population Reports, Series P-25, No. 133 (March 16, 1956), pages 4-9. This net migration rate may provide some basis for extrapolating migration rates in the future.

A more complete description of the computation procedure for the cohort-survival method, after the fertility, mortality and migration assumptions have been determined, is given in an earlier Census publication, Current Population Reports, Series P-25, No. 43 (August 10, 1950), "Illustrative Projections of the Population of the United States, 1950 to 1960."

The ratio method provides a simpler approach to the problem of making population projections. Briefly, this method derives projections of the future population of an area on the basis of a series of observed historical ratios between the population of the given area and the population of a larger area for which projections are already available. For example, in 1952 the Census Bureau used the ratio method to make projections of the population of the 48 states to 1960 as follows:

(1) projections for the nine geographic divisions of the United States were first obtained by extrapolating ratios of the division's total population to the national total on the basis of observed ratios computed from the 1920 to 1950 censuses, and then applying these extrapolated ratios to already available population projections for the United States; (2) projections for the 48 states were prepared next, by extrapolating the ratio of each state's total population to the appropriate division total on the basis of observed ratios for 1920 to 1950, and then applying the extrapolated ratio to the population projection for the appropriate division. The assumptions used in making the extrapolations are described in the Census Bureau's Current Population Reports, Series P-25, No. 56 (January 27, 1952).

When the ratio method is used to make population projections for a city, several successive stages may be used to obtain the final projections, for example, to the two stages described above for obtaining projections for states might be added a third stage in which the historical series of ratios of the city's population to the appropriate state's population is extrapolated and applied to the projected population of the state. Such a procedure was used by the Philadelphia Planning Commission to project the population of the Philadelphia-Camden Industrial Area to the year 2000. In this case, the projection was carried through the following four stages of a ratio-estimation process, using the already available total United States projections as the starting point: (1) United States urban population projections were obtained by extrapolating the historical ratios of the U.S. urban to the U.S. total population and multiplying the extrapolated ratio by the total population projections of the U.S.; (2) the Northeastern Industrial Region's urban population was then projected by extrapolating ratios of its urban population to the U.S. urban population, and applying them to the U.S. urban population projections; (3) the Philadelphia-Camden Industrial Area's urban population was next projected by extrapolating from ratios of its urban population to the urban population of the Northeastern Industrial Region, and applying them to the projected urban population of this Region; (4) lastly, projections for the total population of the Philadelphia-Camden Area were obtained by extrapolating ratios of its total population to its urban population and applying them to the projected urban population of the Philadelphia-Camden Area.[7/] The combination of stages to be used to make projections for a particular community will depend on the available data, and some investigation into ratios that are likely to produce reasonable results.

While the ratio method is less complicated and less time-consuming than the cohort-survival method, it does assume that reliable population projections are already available for a larger political or geographic area, and that a time series of historical ratios of the population of the area for which the projection is desired to this larger population can be computed from past censuses (or other materials). One disadvantage of this method, for purposes of making school enrollment projections, is that it yields projections of the total population only; additional assumptions and computations are required to estimate the age composition of the population on the projection dates, and this composition is necessary to obtain projections of school enrollment.

The simpler type of components approach mentioned earlier differs from the cohort-survival method in that the mortality, fertility and migration assumptions applied to the current population estimates (or the last census count) do not take account of age; rather, general rates for these three components are applied to total population figures, and the resulting population projections are for the total population. These projections, like those obtained by the ratio method, require additional work to estimate the age composition of the projected population.

Projections of school enrollment. Once population projections for persons of school age have been obtained, projections of school enrollment on the same dates may be computed by applying enrollment rates--representing the proportion of persons of a given age who are enrolled in school--to appropriate age groups of the population. The enrollment rates selected for this purpose may be the same as those observed at the time of the last census, or they may be extrapolations of past trends in enrollment rates for the particular area, or for an area presumed to have a similar population. Generally, enrollment rates at the elementary level--where school attendance is compulsory--are so high that rates for the last census can usually be used. Rates at the high school level are not so uniform from one geographic area to another, nor so stable from one time to another, so that some allowance may be made for possible changes in these rates if it seems advisable in the local situation.

The particular form in which school enrollment rates should be computed depends on the detail of the enrollment projections desired, and also on the amount of age detail available in the population projections to which the enrollment rates are to be applied. For example, if enrollment projections by single grades are desired, grade-specific enrollment rates for each year of age should be computed and multiplied by the projected population in each year of age. So detailed a projection is probably warranted only if the projection is for a short span of years--no more than five years, perhaps--and if reasonably reliable projections of the school age population by single years of age have been obtained. In this connection, it should be noted that reliable short-term projections by single years of age can be computed for many large well-established communities but are probably unreasonable for new communities or those currently experiencing a spurt of rapid population growth.

When enrollment projections are desired for longer periods of time, or when population projections are not available by single years of age, projections of total elementary school enrollment and total high school enrollment may be obtained as follows: (1) compute an enrollment rate expressing the ratio of elementary enrollment in grades 1-8 to the number of persons 6-13 years of age, from statistics for the last census; (2) compute another enrollment rate expressing the ratio of high school enrollment to the number of persons 14-17 years old at the time of the last census; (3) if desired, adjust the high school enrollment rate on the basis of assumptions as to future trends in job opportunities, the military draft, or other factors which affect the rate of drop-outs at the high school level; (4) multiply the elementary enrollment rate by the estimated population 6-13 years of age on the projection date, and multiply the adjusted high school enrollment rate by the estimated population 14-17 years old on the projection date. The resulting products are the estimated elementary school enrollment and the estimated high school enrollment on the projection date. If it is desired to include kindergarten enrollment in the elementary school projection, the ratio of grades kindergarten-8 to persons 5-13 years old may be used in the first step and multiplied by the projected population 5-13 in the last. However, as mentioned earlier, kindergarten enrollment rates are not as stable as those for other elementary grades.

III. PROJECTIONS OF DWELLING UNITS, POPULATION AND PUPIL-POPULATION RATIOS

The third approach to projection of school enrollments is perhaps the only feasible one for rapidly expanding communities which are in their early period of growth and which still have relatively large parcels of unused land available for residential development. In this approach, a local survey is undertaken to obtain some of the following kinds of information: the composition of households living in various types of dwelling places at the present time; the type of housing recently constructed in the area; the types of families moving into the new housing, including the number of persons per family and the number of elementary and high school students per family; the amount of land available for residential construction and the type of construction likely to be undertaken on this land (zoning ordinances may help here but are not the only determining factors); and any other information that seems relevant to the future development of the particular area.

The information obtained in the local survey should provide a basis for (1) estimating the rate of land utilization and the rate of construction of dwelling units, which can then be converted into projections of numbers of dwelling units on specified future dates; (2) estimating the population per dwelling unit and pupil-population ratios--separate ratios should be obtained for elementary and high school pupils--which when applied to the projected number of dwelling units will give the projections of total population and school enrollment. (In the second step, estimates of pupils per dwelling unit may be made directly and applied to the projections of dwelling units without the intermediate step of estimating total population, but this will cut out the total population estimate which may be of interest in itself.)

This approach to school enrollment projection was used in studies, made by the Committee on Field Services of the University of Chicago Department of Education, of school enrollment in two suburban communities of the Chicago Metropolitan Area.[8/]

PREDICTING SCHOOL NEEDS IN DIFFERENT TYPES OF COMMUNITIES

The best method for predicting school needs in a particular community or school district depends to a great extent on the characteristics of the community itself, including its age, size, location and the rate of net migration to the area. Of course, the kinds and quality of data available must also be taken into account. Below are suggestions for approaching the problem in five different types of communities, the first four representing communities in various stages of development and growth, and the last one a school district situation which may present special problems. Needless to say, in a particular situation methods other than those suggested for the general type of community may prove to be more useful.

Exploding community. Most of the areas falling in this category are in the suburban part or outlying fringe of large metropolitan areas. Most of them are either new communities or areas of rapidly expanding residential development in older communities with large parcels of vacant land. Projection of school needs for communities of this type is exceedingly difficult without a special survey of the local area; the third approach to school enrollment projection described in the preceding section of this report is recommended for these communities; in fact, it was discussed in terms of its applicability to this type of situation.

Older community still experiencing growth. The grade progression method of making enrollment projections may be used for these communities, providing: (1) a recent time series of grade progression ratios is available, based on comparable statistics of enrollment, preferably as of October of each year; and (2) this time series of ratios covers a sufficiently long period to give some indication of fluctuations in the ratios due to migration into the area, thus permitting adjustment of the ratios on the basis of the level of migration that is assumed for the projection period. If the necessary data for the grade progression ratio approach are not available, the second approach is suggested, that is, computing projections of population by age, and applying enrollment rates to these population projections.

Stable, well-established community. The grade progression method is recommended for these communities, providing the kinds of data specified for this approach in the preceding paragraph are available. If the data are not available, again the second approach is recommended.

Declining community. There is less need to project school enrollment in these areas since present facilities are more likely to be sufficient. However, the changes in age structure of the population as a result of the extended period of high birth rates since the last war, may make it worthwhile to project enrollments; either the first or second approach described in the preceding section of this report may be used for this purpose.

Rural and consolidated school district. The grade progression method is recommended for this situation. If some difficulty is experienced in obtaining the required birth statistics for projection of kindergarten and first grade enrollments by this method--the available counts of births may not correspond to school district boundaries--a special analysis of the birth records, available in the local office where they are recorded and filed, may be necessary, that is, the births may be allocated to school districts using the residence of the mother that appears on each certificate. Before computing grade progression ratios for a consolidated school district, the enrollment statistics for all of the individual schools in the district should be added; if the boundaries of the consolidated district changed during the time period for which the ratios are computed, the ratios should be adjusted accordingly.

When the grade progression method cannot be used because of deficiencies in the enrollment statistics, some difficulty may be experienced in trying to use the second approach, because the population projections and the estimated enrollment rates needed in the second approach require Census statistics for the area as a starting point. If the school district boundaries do not correspond to the geographic areas tabulated in the census statistics the latter will have to be adjusted for this discrepancy before the method can be used; often there is little basis for such an adjustment.

CONCLUDING OBSERVATIONS

Few communities in the United States during the coming years can avoid the task of estimating school population. In the nation as a whole, the need to expand elementary school facilities as a result of the postwar baby boom became most acute during 1959, the year in which practically all elementary school children were postwar babies. In the years which lie immediately ahead, to 1964, high schools will be absorbing the postwar baby crop.

In many communities, however, and especially in rapidly developing suburbia, elementary as well as high school facilities will prove inadequate for some time to come. In the materials which have been presented, some of the major forces producing needs for additional schools have been indicated and techniques for estimating school enrollment have been described. Moreover some indication has been given of the techniques best adapted for estimating school population in different types of local situations. These materials should be helpful to agencies and persons faced with the task of setting policy, providing school facilities and administering school programs. The social statistician, who has played an active role in the development of methods of estimation has already been called upon and may be expected to be called upon with increasing frequency for the important social engineering task of estimating future school population.

FOOTNOTES

1/See California Department of Finance, Financial Research Section, Projections of Public School Enrollment in California to 1960 and 1965 (Sacramento, April 1954), p. xii. This report discusses in some detail the influence of migration on changes in grade progression ratios over time.

2/This approach was used in the California projections cited in footnote 1, although details of the procedure are not elaborated in this report.

3/For a more detailed explanation and illustrative example, see U.S. Bureau of the Census, Current Population Reports, Series P-25, No. 20, May 6, 1949.

4/For a more detailed explanation and illustrative example, see U.S. Bureau of the Census, Current Population Reports, Series P-25, No. 133, March 16, 1956. Series P-25, No. 165, November 4, 1957, describes a recent amendment in one of the steps of the method.

5/For additional explanation, see Current Population Reports, Series P-25, No. 156, April 30, 1957, p. 2.

6/For a discussion of the various methods and citation of relevant source materials, see Jacob S. Siegel, "Forecasting the Population of Small Areas," Land Economics, February, 1953, pp. 72-87.

7/Population Estimates 1950-2000, Philadelphia-Camden Area (Philadelphia City Planning Commission, April 1948).

8/A Study of the Problems Related to Future Expansion of the Glenbrook High School, Northbrook, Illinois (Board of Education, Northfield Township High School District 225, December, 1957); A Study of the Problems Related to Future Expansion of the New Trier Township High School, Winnetka, Illinois (Board of Education, New Trier Township High School District 203, December, 1957).

Discussion

John Folger - Florida State University

Dr. Hauser has suggested some conditions under which the grade-survival method of projecting enrollment would be appropriate and some others where the use of measures of residential growth would be appropriate. Experience with projections in a metropolitan area (Atlanta) indicates that a combination of these two methods is often desirable.

Enrollment projections are desired by local school systems for two main purposes. First are the projections for the entire system which are used to plan for personnel, future revenues required, and similar purposes. Second is the planning of specific locations for school plants. For this latter purpose, projections based on residential growth are the major means of pin-pointing school needs in specific areas. These projections can be cumulated for the entire school district to get the total enrollment which can be checked against projections -- made for the entire area by the grade-survival method. Projections made by the grade-survival method are likely to be inappropriate for small areas of a metropolitan district. Most parts of large cities are changing rapidly in size, composition, and in other ways which make any kind of trend line subject to a wide margin of error. Therefore, grade-survival projections for areas of cities small enough to plan school location will usually be less useful than methods that relate enrollment to expected future dwelling units.

A dwelling-unit projection which ascertains the maximum number of dwelling units in an area when it is completely built up is very useful for planning specific school location. The rate at which an area is approaching its maximum school enrollment size is not as important for some kinds of school planning as the total size it is likely to reach.

When the dwelling-unit approach is combined with the grade-survival or ratio-to-population methods for the entire area, projections of maximum accuracy should be achieved. (There is little empirical evidence to indicate the superiority of one approach over another.)

Discussion

Louis H. Conger, Jr., U. S. Office of Education

The best method depends on the data available and the nature of the projected figures desired. When all is said and done, it is still a matter of judgment and experience, since alternative methods are nearly always possible. At the school district level, enrollment projections are usually wanted for planning school building programs; thus detail by area is needed. In this context, birth and grade enrollment data are almost always more ample than population data. If the data permit, it is desirable to combine private school enrollment from an area with that of public schools in making the projections, in order to stabilize the grade retention rates. The projected enrollment can then be split into public and private components by specific assumptions.

It is my observation that much of the work in making projections is getting your basic data in good order -- checking it carefully for accuracy and consistency. Failure at this point is not easy for the consumer to detect, but the effect on the projections may be disastrous.

In the case of long-range projections for rapidly-growing areas, ridiculous overestimates may arise from the compound interest effect of grade retention rates. A land-use analysis of the maximum population at the saturation point is one way to detect this; another way is to project enrollment for a larger area and compare it with the sum of the projections made for the individual parts.

V

THE METHODS OF THE 1961 CENSUS OF CANADA

Chairman, Walter E. Duffett, Dominion Bureau of Statistics

SOME PROBLEMS ENCOUNTERED IN THE PLANNING AND TAKING OF THE 1961 CENSUS OF CANADA

O.A. Lemieux, Dominion Bureau of Statistics

The organization which it is necessary to put in motion for the planning and taking of a census is always a large one. A large number of operations must be undertaken which may appear unimportant to persons who have had no experience in this kind of work, but which are very important if we are to be ready to start enumeration on the day set by law for it. In the preparation of the 1961 Canadian Census the situation was complicated by a series of circumstances which were unprecedented in our experience and which had to be met as we went along.

The date set by law for the taking of the Canadian Decennial Census is June 1 of the year ending in "one". In this case the date was June 1, 1961.

Shortly after the compilations of the 1956 Quinquennial Census were under way, the Dominion Bureau of Statistics began to study the possibility of acquiring an Electronic Computer to compile the 1961 Census. The 1951 and 1956 Censuses had been done with the use of mark-sensing equipment; i.e., the enumerators made their marks on a card which was put through the conventional mark-sense punching machine. This system had proven very satisfactory and time saving in both censuses. In conjunction with an Electronic Computer, it was important to devise a reader which would be able to use the documents mark-sensed by enumerators and transfer the information directly on to a magnetic tape which then could be fed into the computer. The return to the slow and bothersome process of punching cards would have seemed to be a retrograde step.

We talked to the representatives of firms which manufacture various types of computers always with the idea that a reader was necessary which would transfer the information directly from the document to the tape so as to do our job as efficiently as possible. Finally, the International Business Machines Company undertook to develop such a reader. Thus, a reader was eventually developed and put in operation sometime in 1960. In 1958 it was decided that we would acquire an I.B.M. 705 to which a 1401 was added sometime later. These were installed ready for use during the second half of 1960.

Programming. - The next very important problem was the acquisition of programmers to start writing the various programmes for the runs required for the census. At that time, programmers were not very plentiful in Canada and it became quite a task to acquire the necessary staff to programme the various runs and test the programmes in time to be ready when the returns of the census began arriving from the field. It was necessary to employ a large number of young persons who had had no previous experience and to train them. Some of them, after they had obtained some knowledge and experience, left us for more lucrative positions in business and industry, and the turn-over in programmers was quite high.

Another very important fact was that not one of the subject-matter people in the Census Division, from the Director down to the most junior statistician, knew anything about a computer. The majority of them had never even seen one. None of us knew what a computer could do, what it could not do, how it functioned, etc., etc. We had heard stories of the marvelous things it could do and, I fear, some of these were a trifle exaggerated. It was, therefore, necessary to give our subject-matter specialists some knowledge of the computer so that they would have an idea of what the machine could do. This training had to be done at a time when our statisticians should have been exceedingly busy doing something else.

Questionnaires. - Now that we had a reader, it was necessary to plan a questionnaire that would fit it. Of course, the size of questionnaire to be used had been decided before building the reader, but the arrangement of questions on this questionnaire had to be settled. This had to be arranged in such a way that a very ordinary enumerator would be able to mark it successfully under field enumeration conditions.

Such a questionnaire was devised and then had to be tested under field conditions.

Determination of questions. - After the physical shape of the questionnaire had been settled, it became necessary to determine what questions would be asked. There, we had been told of the huge capacity of the machine and people started to overwork their brains to find out new questions that should be asked in the census. As you know, some census questions are essential and must be asked at every census. These are questions like sex, age, marital status, relationship to head of household, occupation, etc. Then the Dominion Bureau of Statistics works in co-operation with the Statistical and Population Commissions of the United Nations, with the Food and Agriculture Organization, with UNESCO and the Inter-American Statistical Institute. A number of questions have to be asked to meet the requirements of these organizations for purposes of inter-continental and world-wide comparability.

Then there is the question of making an effort to meet the requirements of our users in business, trade, social work, etc., etc., both in Canada and the United States, because, as you may know, a large proportion of our clients are from the United States. A large sample of our users was queried by mail as to what questions would be most useful in the census. We, of course, reserved the right to refuse to accept questions after it was felt that the questionnaire had reached a size beyond which we could not go.

A very large number of questions were suggested by users of questions which had a great deal of merit. There are three things to remember, however, and these may be summarized as follows:

1. It is of the utmost importance to determine what is the maximum number of questions that can be asked without jeopardizing the reliability of the statistics obtained. Past experience has proven that there is a definite limit to the number of questions that we can expect the enumerators to ask and the people to answer. We were of the opinion that the questionnaire for the 1961 Census should not be more complicated nor much longer than the one used for the 1951 Census. It should be noted that the prolongation of the interview time may bring it to a point where fatigue and annoyance on the part of the respondent affect the quality of the answers.

   Another serious consideration, and one not so generally recognized, is the deterioration in the quality of all census results arising when an unduly heavy load is imposed upon an army of quickly recruited and often inadequately trained enumerators. Each question added to the schedule contributes its share to the volume of instructions which need to be assimilated in advance of enumeration and the cumulative effect of a large number of additional questions may be to spread the training time over so many subjects that the enumerator is not properly equipped to get reliable information on any one of them.

2. The questions asked must be such that they will produce statistics which are of general interest. We cannot afford to ask a question on the census just to serve the purposes of one individual or one organization.

3. There is no point in asking questions that, we know in advance, a large proportion of respondents (largely housewives) cannot answer.

Test Census. - On June 1, 1959, a Test Census was held in two areas; one English-speaking area in Ontario and a French-speaking one in Quebec. In Canada, both English and French are official and questionnaires must be provided in both languages. The purpose of the test was to try out the new type of document to see whether or not enumerators could make marks, under enumeration conditions, that would be recognized by the reader; to test the questions to find out whether or not they were readily understood by respondents; to discover whether all questions asked were apt to produce usable statistics; to test the proposed tabulation on the computer; and, finally, to assist in determining rates of pay to enumerators.

The test questionnaire contained 22 questions to be answered on a 100-per cent basis. These questions covered personal characteristics, occupations, employment and salaries and wages. In addition, there were 15 questions asked of a 20-per cent sample of the population. Five of those referred to family size, one to migration and nine to income from all sources.

Analysis of the Test Census. - The results of the test showed that the reader could very well read the markings of an ordinary enumerator. It also showed that a few questions which had been urged upon us would not produce usable statistics and that others would have to be rephrased in order to be understood by both the enumerators and the respondents.

Publicity around the Test Census. - The questionnaires used in the Test Census were given to the press and this started a campaign, which at one time was violent enough, against a number of proposed questions from the census. The questions which were objected to were those on family size, on income and on ethnic origin.

Family size - One newspaper started the news that the Dominion Bureau of Statistics proposed to ask of every woman the number of illegitimate children she had borne. This started a string of letters of protest and the subject was even mentioned on the floor of the House of Commons. Of course, this was not true, the Bureau had never proposed such a question.

Income - In this case the newspaper protests were that the Government was getting too nosey, that this information was private and none of the business of the Government and that the information had already been supplied under oath to the Income Tax Department. Then there were protestations about the fact that people would have to give this information to an enumerator living next door or just a small distance down the street.

Ethnic origin - This question in the Canadian Census has been a source of worry to a large proportion of the population for some time. In order to appreciate the problem it is necessary to go into some preliminary details. Now, what do we mean by the ethnic origin of the Canadian population? How is it determined? The concept is at least indefinite if not to say nebulous. However, whatever the statistical value of the question may be, a large proportion of the population feels that it has considerable administrative use in our country which is officially bi-cultural.

Because of adverse publicity in previous censuses, certain groups, namely, the Canadians of French, Jewish and Ukrainian origins were concerned lest this question be dropped from the 1961 Census. Representatives of some of these groups waited upon the Government and were assured that the question on ethnic origin would be asked in the 1961 Census in the same way as it had been asked in 1951.

On the basis of this assurance the question was included in the Test Census with exactly the same instructions as had been given in 1951.

Preparation of the final population questionnaires. - The criticisms levelled at the size of family and income questions, in particular, caused us to prepare two questionnaires, one for the questions which were asked on a 100-per cent basis and the other for the questions asked on a sample basis. These questions were on migration, size of family and income. This questionnaire was made in the form of an envelope. It was left by the enumerator to the respondent for him to fill, seal and return to the enumerator at a later date, and this envelope would be opened only in the Dominion Bureau of Statistics thus ensuring confidentiality in so far as the enumerator was concerned.

The questionnaires were finalized but in doing so it was decided, on the question of ethnic origin to add the words "Canadian" and "American" as acceptable ethnic origins. It is obvious that from the very beginning this question referred to non-North American origins and whatever the value of the question may be this addition tended to destroy the purpose of the question.

Printing. - The printing of the questionnaire presented another problem. The printing demanded a very high degree of precision so that the various positions could be read by the reader. It was a slow process. Documents had to be checked at very close intervals to ensure that they did not exceed the permissible tolerance which is of the order of one one-hundredth of an inch between dots. The printing took several months and all necessary materials were packed in 32,000 boxes ready to be shipped to individual enumerators across the country.

Protests of various organizations. - When the questionnaires became available, it was seen that a change had been made in the questionnaire by the addition of "Canadian" and "American" as acceptable ethnic origins, and, as has already been said, this failed to recognize the purpose of the question and threatened to render it useless. A certain number of organizations and a sizeable portion of the press began a series of protests, claiming that the intent of the question had been changed in spite of the statement of the Government, and that the question was no longer asked as it had been asked in the 1951 Census. Another section of the press and other organizations and individuals favoured the addition of these two terms and a long and protracted polemic followed. The Members of Parliament were assailed from all sides.

Finally, it was decided to revert to the original question and remove these two terms. This meant the reprinting of some 11,000,000 questionnaires and a change in the instructions to enumerators. The decision for reprinting was not taken till sometime in March, - 32,000 boxes had to be unpacked and repacked with the new documents. There was great pressure for time and this created confusion and errors which did not come to light until the enumerators were in the field.

This paper is not for the purpose of discussing the pros and cons of a question on ethnic origin in the census but since this turned out to be one of our leading problems in the 1961 Census it seems advisable, in order to understand the question, to say a few words about it.

What is ethnic origin and how did it get included in the Canadian Census in the first place? At best it is a vague concept which is exceedingly difficult to define in terms that will be understood by ordinary enumerators and by respondents who have been in Canada for many generations.

The Encyclopaedia of the Social Sciences, Macmillan, 1931, has six pages of definitions and explanations and after reading them carefully one remains with the idea that while ethnic groups exist they are exceedingly difficult to define in terms that an ordinary person will understand.

In essence, the Encyclopaedia defines ethnic communities or groups as groups bound together by common ties of race, nationality or culture, living together within an alien civilization and remaining culturally distinct. They may occupy a position of self-sufficient isolation or they may have extensive dealings with the surrounding population while retaining a separate identity. In its strict meaning the word "ethnic" denotes race; but when applied to communities in the above sense it is loosely used, in the absence of any other comprehensive term, to cover the more general concept of culture.

Then it goes on to explain how these communities and groups have developed and it gives examples such as the Czechs, Slovaks, Poles, Jews, etc.; of the Canadians of French origin it says: "In Canada the presence of an established French population which antedated the English induced a tendency of tolerance toward other groups, their cultural autonomy being in a measure acknowledged through permission to retain their language for some official and educational purposes."

Then the Encyclopaedia goes on to explain that "the chief basis for cohesiveness is race. Physical differences which cannot be changed or concealed set one group apart from another and mark any man who seeks to leave his community (or group) and become part of the surrounding culture."

"Where no marked racial differences exist, cultural difference forms the basic cohesive bond and appearance merely supports cultural barriers."

"Supplementary to race and nationality the strongest reinforcing factors are language and religion, both of which are apt to be essential parts of the national culture complex. The retention of the national tongue is often the principal aim of those who seek to prevent an ethnic group from losing its identity while the loss of that language is taken as a measure of amalgamation."

Origin of the question in the Canadian Census. - As far as I was able to ascertain, this question appeared for the first time in a census of Nova Scotia taken in 1767. In this census the population was divided into English, Irish, Scotch, Americans, Germans and Acadians. More than half of the population was reported as "American", a term which has not been acceptable as an ethnic origin between 1871 and 1951.

In 1842, in Upper Canada, and in 1844, in Lower Canada, the birthplace table separates the Canadian-born into "French Canadians" and "English Canadians". In 1848, and in 1851 and 1861, again the Canadian-born were divided into "French origin" and "not French origin" in both Upper and Lower Canada. It is of interest to note that at that time the population was composed as follows: of the Canadian-born, 35 per cent were of French origin and 65 per cent of British and other origins; 23 per cent of the population was foreign-born and of this number only 16 per cent were born outside of British possessions.

Naturally enough the ethnic origin question in the census started by being for the purpose of counting the Canadian of French origin separately from the others.

In 1871, there was a change and two tables were produced, one on the birthplace of the population and the other on origin. In the introduction to the published volume, the then director of the census wrote as follows: "The subject matter of Table III (ethnic origin) is a new feature of our census statistics. None of the former censuses of the various provinces had it, except in so far as the French origin was concerned, in the former Province of Canada. What is given in previous returns under the head of "origin" was simply the enumeration of the people by their place of birth. But a moment's reflection shows at once that these two subjects are as different as they are important."

Now, how different were they? By racial origin, in the 1871 Census the British origins were 60.5 per cent of the total population, the French 31.0 per cent and the others were 8.4 per cent. By birthplace the percentages are British 62.1, French 31.0 and others 6.9 per cent. So, as it turns out, the difference in this census was not great.

The instructions given to the enumerators on this question were as follows: "Origin is to be scrupulously entered as given by the person questioned in the manner shown in the specimen schedule by the words English, Irish, Scotch, African, Indian, German, etc."

The enumerators and respondents must have understood what it meant since the question was answered and only 0.2 per cent reported "not stated". It could be that we could get better answers nowadays if we wrote less instructions.

Of course in 1871, as the previously quoted figures indicate, the problem was rather simple.

The question was repeated at every census since 1871, with the exception of 1891 when only Canadians of French origin were counted.

Since 1871, the instructions to enumerators have been amplified but I am not too sure that we have improved them very greatly.

The 1961 instructions to enumerators read as follows: "It is important to distinguish carefully between "citizenship" or "nationality" on the one hand and "ethnic" or "cultural" group on the other. "Ethnic" or "cultural" group refers to the group from which the person is descended; citizenship (nationality) refers to the country to which the person owes allegiance. Canadian citizens belong to many ethnic or cultural groups - English, French, Irish, Jewish, Scottish, Ukrainian, etc.

For census purposes a person's ethnic or cultural group is traced through his father. For example, if a person's father is German and his mother Norwegian, the entry will be "German".

If the respondent does not understand the question as worded on the questionnaire you will ask the language spoken by him on arrival if he is an immigrant, or by his ancestor on the male side on first coming to this continent. For example, if the person replies that his ancestor on the male side spoke French when he came to this continent, you will record "French". However, if the respondent should reply "English" or "Gaelic" to this question, you must make further inquiries to determine whether the person is English, Irish, Scottish, or Welsh." (Parenthesis: This seems to establish language as the only determinant of one's ethnic origin.)

Then the instructions continue: "If the respondent does not understand the question as worded on the questionnaire or you cannot establish the ethnic or cultural group through the language of the ancestors", you will ask: "Is your ethnic or cultural group on the male side English, French, Jewish, Negro, North American Indian, Norwegian, Scottish, Ukrainian, etc.?"

Then starting in 1951 an innovation was made as follows: "Since the question refers to the time when the person or his ancestors came to this continent, the answer should refer to the ethnic groups or cultures of the Old World. However, if, in spite of this explanation, the person insists that his ethnic or cultural group is "Canadian" or "U.S.A.", enter his reply in the write-in space."

This, in 1951, was the thin edge of the wedge which led to our problems in 1961.

As you see, many words were added to instructions to enumerators but I doubt very much if the concept was made any clearer.

Through the years, particularly after the first World War, a body of opinion has been formed opposing the question on the grounds that it was no longer meaningful because of intermarriage between groups and because of the fact that some people claim that after several generations in Canada they did not know to what cultural group their ancestors belonged on first coming to Canada and besides, they did not care, and wished to call themselves plain "Canadians".

Reasons for asking this question in Canada. - There are many reasons for asking this question in the Canadian Census. These may be summarized somewhat as follows:

1. As we have already mentioned, the original purpose of the question was to count Canadians of French origin separately from the others. Under the Canadian Constitution they have special rights. In order to preserve and exercise these rights, they should be able to count themselves periodically, assess their own position in order to determine where they stand. For example, if in a certain area a fairly substantial percentage of Canadians of French origin no longer speak French, the knowledge of this fact allows the French community to try to do something about it.

2. Among all groups, this is a measure of the degree of assimilation into one or the other of the two official cultures of Canada. Since the beginning of the century, Canada has received a very large number of immigrants from a very large number of countries. Some of these groups integrate themselves more readily than others with the Canadian community.

3. In the future, when Canada has fully developed, sociologists and anthropologists will have a better measure of what went into the building up of the Canadian nation than any other country which was formed before statistics were available.

4. For groups other than the French and the Anglo-Saxon, it allows them to assess their contribution to the building up of the Canadian nation. Canada, unlike her neighbour to the south, has not adopted the concept of the "melting pot". On the contrary, immigrants from foreign lands have been encouraged to cherish the traditions which they brought from their homeland and to use them to add to the Canadian culture.

This has been an unduly long parenthesis but I thought it advisable to give these details because it is with this question of ethnic origin that our most serious problems in the 1961 Census have arisen.

Field organization for the 1961 Census. - In 1951 and 1956, the field organization of the census was, in many ways, similar to that of the United States. We had one Commissioner (called Supervisor in the United States) for each Electoral District. Under his direction were a number of Field Supervisors (called Crew Leaders in the United States). The number of Field Supervisors in each Electoral District varied with the area and population of the District. Thus in 1951, we had 263 Commissioners and somewhat in excess of 1,000 Field Supervisors.

In order to remove one level of supervision in 1961, the Electoral Districts were divided into Census Districts; the number of Census Districts in an Electoral District again varying with the area and the population. Each Census District was headed by a Commissioner who had under his supervision from 25 to 30 enumerators in urban districts and some 15 to 20 in rural areas. Each Commissioner was responsible for the census in his Census District independently of other Commissioners in the Electoral District. Each one dealt directly with the Dominion Bureau of Statistics or with the Regional Office of the area where he was located. This made it possible for him to receive his instructions directly from headquarters and permitted to have a much more intimate contact with his enumerators. We ended with 1,350 Commissioners for the 265 Electoral Districts.

In addition, a group of employees (50 in number) were added to strengthen the Regional Office staff so that they would be in a position to give more attention to individual Commissioners than had been possible in previous censuses.

Enumeration areas. - In past censuses, enumeration areas had been delineated to contain from 1,500 to 2,000 population in urban areas and 800 to 1,200 in rural areas. In the hope of speeding up enumeration and because it was felt that Commissioners would have more time to devote to individual enumerators, it was decided to make enumeration areas that would average somewhere around 120 households. It was hoped that such an enumeration area would be covered in two weeks in urban areas and three weeks in rural ones. It has to be kept in mind that a Census of Agriculture was taken in conjunction with the Censuses of Population and Housing. Our census is taken as of June 1, and it is important that enumeration be completed before it gets complicated by the holiday season when families go away to summer cottages, extended vacations in the United States, or Europe, etc.

Appointment of Commissioners. - As in the United States, Commissioners are not appointed by the Civil Service Commission but are recommended to the Minister of Trade and Commerce, under whose jurisdiction the Dominion Bureau of Statistics comes, by the Members of Parliament. When these nominations are made, the candidates are interviewed by members of the staff of the Dominion Bureau of Statistics and accepted if satisfactory.

As early as June 30, 1960, the Minister of Trade and Commerce wrote a letter to each Member of Parliament asking for recommendations. These recommendations were to be in the hands of the Minister by October 1, 1960. The intention was that between the 1st of October and the end of February, the representatives of the Dominion Bureau of Statistics would have time to interview and select the best candidates so that all appointments would be made by March 1st. This would give adequate time to explain their assignments to them and give them the necessary time to do the preliminary work before enumeration began on June 1.

Unfortunately, for many reasons, the most important of which have already been mentioned, some of the acceptable nominations did not come to hand until sometime in May. This created problems which caused confusion all through enumeration.

Because of the planning and in spite of these delays, over 52 per cent of the enumeration areas had been returned to the Regional Offices by the end of June as compared to 12.2 per cent in 1951. In all our enthusiasm, we had not anticipated such rapid returns and consequently we had not asked the Civil Service Commission to supply us with processing staff in the Regional Offices soon enough. As a result, our Regional Offices were literally swamped with returns long before they could have trained staff to handle them. This created a considerable amount of confusion in our Regional Offices and delayed the returns to Ottawa to be processed.

In spite of all this, the first preliminary reports of population were published on July 28, 1961, giving the preliminary population figures for about 1,000 municipalities. This was some weeks in advance of 1951. The last preliminary report giving the population of municipalities was published on October 6, again earlier than in 1951.

Postal check. - Prior to their 1960 Census, the U.S. Bureau of the Census had experimented with a postal check to assure the completeness of the census in areas covered by house-to-house mail delivery. They did not follow through with it. Like good neighbours we took over the idea and carried out the experiment for them.

Each enumerator after having enumerated a household made a card with the address of the household. These cards were turned over to the postal authorities, they sorted them and then gave us a list of the addresses that they had on their records and for which no card had been given to them. These addresses were checked and when they had been missed they were enumerated and added to the records. The result was that a population of over 38,000 was added to the census at a cost of about $30,000. It also gives the census authorities the satisfaction of knowing that all addresses in urban areas are accounted for and this provides an excellent argument when city authorities begin to argue that we have missed a large number of people in their municipality.

Processing in Regional Offices. - As in 1951, acceptability checks, coding, payment of enumerators, was carried on in our eight Regional Offices scattered across the country. This work was done in record time, the whole process being completed by the end of October.

When the returns were sent to Ottawa they were, in theory at least, ready to be forwarded to the reader for transfer to tape.

When we were preparing the census, we were not certain that all enumerators' markings would be of good enough quality to be picked up by the reader. Consequently, we had provided for a fair percentage of rejects caused by the failure of the machine to pick up the marks made by the enumerators. The results were way beyond our expectations. The machine read the marks and instead of the anticipated percentage reject, the actual was much lower.

Preliminary counts of population. - As already stated, a preliminary bulletin published on October 5 completed the publication of the preliminary population counts of all organized municipalities in Canada.

From these figures it is possible to get a glimpse of how the population has grown and distributed itself during the last five and ten years.

First, if we look at our two cities of more than 500,000 population, namely, Montreal and Toronto, during the period 1956-1961, the population of the city of Montreal proper increased by 4.1 per cent. During the preceding five years the increase had been 8.6 per cent and for the ten-year period the increase was 13.0 per cent.

In the city of Toronto proper, on the other hand, the population decreased 1.6 per cent between 1956 and 1961, 1.2 per cent during 1951-1956 and a decrease of 2.7 per cent during the decade 1951-1961.

Now, if we look at the Metropolitan Area of Montreal, apart from the city proper, we find that the population increased 42.2 per cent during 1956-61, 41 per cent during 1951-56 and 100.5 per cent during the decade.

The Metropolitan Area of Toronto increased 36.7 per cent during 1956-61, 57.3 per cent during 1951-56 and 115.1 per cent during the decade.

Between 1951 and 1961, these two cities and their Metropolitan Areas have accounted for 28.2 per cent of the anticipated increase in the Canadian population.

Now, if we take a look at the cities of a population between 100,000 and 499,999 which are not situated within the Metropolitan Areas of the two preceding cities we have a picture which is not dissimilar to the preceding one.

The cities themselves, there are ten of them, increased 14.8 per cent between 1951-56, 10.4 per cent between 1956-61 and 26.8 per cent between 1951 and 1961. The rate of increase has slowed down considerably between 1956 and 1961.

Now, when we take them with their Metropolitan Areas, the situation is pretty much the same as in the case of Montreal and Toronto.

Between 1951 and 1956, the Metropolitan Areas (other than Montreal and Toronto) increased 20.9 per cent, between 1956 and 1961, 18 per cent and between 1951 and 1961, 42.6 per cent. Between 1951 and 1961, these ten Metropolitan Areas accounted for 30 per cent of the total population increase of Canada. If to this we add the 28.2 per cent accounted for by Montreal and Toronto, the 12 municipalities and their Metropolitan Areas accounted for 53.2 per cent of the population increase of Canada.

Now, a look at the cities of 30,000 to 99,999 which are not located within the above-mentioned Metropolitan Areas. There are thirty of them. Between 1951 and 1956, they increased 18.6 per cent, between 1956 and 1961, 20.7 per cent and between 1951 and 1961, 43 per cent. These cities accounted for 11 per cent of the population increase since 1951. The percentage increase will be greater when account is taken of the urbanized areas which have developed on their periphery. Were these taken into account as they will be when the compilations are more advanced, their contribution would be somewhere around 15 per cent or more, 42 municipalities then are responsible for about 75 per cent of the total population increase between 1951 and 1961.

This leaves only about 25 per cent of the increase to be distributed among the remaining 4,800 odd municipalities.

From these preliminary counts, a few obvious observations can be made:

1. The large cities are becoming larger.

2. The intermediate size cities are also becoming larger.

3. The smaller cities and villages are remaining fairly stationary.

4. The rural areas have decreased considerably.

5. The movement from central cities to the periphery has continued between 1956 and 1961.

To sum up, I think I can say, without fear of being called too boastful, that the 1961 Census operation has been quite successful in spite of the difficulties which arose at the outset. There now remains to be seen what sort of a job our computer will do in producing tabulations.

A COMPARISON OF SOME OF THE CENSUS CONCEPTS USED IN CANADA AND THE UNITED STATES(1)

J. T. Marshall, Dominion Bureau of Statistics

The subject of this paper is not only a difficult one from many points of view, but a somewhat hazardous one as well. It presupposes that the author has sufficient intimate knowledge of the conceptual problems of the census of a neighbouring country as well as his own. Anyone familiar with the vast range of subjects covered by the modern-day census will appreciate that there are many fundamental reasons for, and purposes served by, the underlying concepts which govern the census content of a given country. Not the least significant factor, by any means, is the way in which these concepts have evolved over the years in one particular country as compared to another. Thus, it becomes a difficult task to attempt a discourse on this subject for one's own country. To this is added a considerable element of risk when one attempts to make comparisons with one's neighbour. But I am among friends, and I am sure that these risks are worth the taking, if in so doing it enlarges, even in a small way, our perspective of the statistical comparability of census information between the two countries.

One would have to say at the outset, of course, that there are far more similarities than differences in the concepts employed in the United States and Canadian censuses. This produces the happy situation of a large body of census statistics which knows no international boundary line. Differences which do exist vary greatly as to their cause and effect. There are some, for example, which on the surface appear to be at variance, but which on closer examination, indicate a slightly different approach to obtain the same information. Thus, the United States Census may have found that asking "date of birth" is the preferable way of obtaining the age distribution of the United States population, while the Canadian Census may have found that from their experiences to date they have preferred to use "age at last birthday". The arguments for and against, as in most all conceptual problems, are numerous and complex. In this particular instance, for example, they involve such considerations as memory bias on the part of persons answering this question for themselves or for other members of their household, the use and desirability of conversion tables, the methods being used to enumerate, process, and tabulate the data, and so on.

For some concepts, it is doubtful whether it is necessary, or even desirable, that the two countries must be exactly identical in their definitional approach. There are others, however, involving more fundamental differences, where comparability becomes difficult, and in fact misleading if one is not aware of the nature and extent of these differences. It is to some of these more outstanding cases that attention should possibly be directed, and even if, despite the best intentions of statisticians on both sides of the border, some of them may never be able to be reconciled, it is important that users of census statistics be made aware of them.

The Concept of a "dwelling" or "housing" Unit

Such awareness might prevent the recurrence of what is now, to us, the famous "telephone" incident. Statistics of the 1950 United States Census and the 1951 Canadian Census indicated that a significantly higher percentage of dwellings were equipped with telephones in Canada than in the United States. In the ensuing rush by Canadian statisticians to determine the cause, it was found that the United States showed a higher proportion of "one-and two-person" households than Canada. Seemingly minor differences in the definition of a dwelling unit for marginal cases, relating principally to such factors as the cooking and preparation of meals and the degree of structural separateness, produced a noticeable variance as to whether persons should be classed as lodgers in a boarding-house room or as living in a separate dwelling unit. The installation of a telephone was obviously not a criterion, and the Canadian definition, or at least its interpretation in the field was more restricted than in the United States. Thus, it can be seen that even in one of the most basic concepts of the census, there were elements of non-comparability even though the definitions, from outward appearances, were closely parallel.

Percentage of households by number of persons

Persons	1950 Census of U.S.	1951 Census of Canada
1	9.3	7.5
2	28.1	21.0
3	22.8	20.3
4	18.4	19.0
5	10.4	12.9
6+	11.0	19.3

This difference in dwelling concepts will carry over to comparisons of the results of the 1960 and 1961 Censuses of the United States and Canada. The United States Census defines a "housing unit" as a group of rooms or a single room if it is occupied by or intended for occupancy as separate living quarters, i.e., when the

(1) This paper was prepared jointly with Douglas Ralston; Ray Davy; Alice Wood; Mabel Waddel and Bob Ellis of the Census Division, Dominion Bureau of Statistics.

occupants do not live and eat with any other persons in the structure, and when there is either (1) direct access from the outside or through a common hall, or (2) a kitchen or cooking equipment for the exclusive use of the occupants. The Canadian definition for the 1961 Census places more emphasis on structural separateness, and states that to be classed as a "dwelling unit", living quarters must be structurally separate and must have a private entrance either from outside or from a common hall or stairway inside the building. The entrance must be one that can be used without passing through anyone else's living quarters. Instructions to Canadian census enumerators emphasized the fact that rooms on the second or third floor of what was originally a single house, for example, should be classified as a separate dwelling unit, or units, only if there had been some structural change to separate them from other living quarters in the building.

Whether the one or the other of these concepts represents the better definition could be argued "ad infinitum". Canadian statisticians might argue that their definition is closer to what one generally considers as a separate dwelling unit in terms of single houses, apartments and suites, etc., as distinct from boarding-house rooms wherein persons may or may not cook their own meals. They would be the first to agree, however, that what they set out as their concept in theory, is not always obtained in practice. In fact, the lack of uniform application by enumerators in the treatment of marginal cases seems a valid criticism of the Canadian concept.

The concept of a "family unit"

Another basic concept which gives rise to misunderstanding when the figures of the United States and Canada are compared, is that concerned with the definition of a "family unit". There is no pretense here at comparability in concept, and each country has been going its separate way for a number of censuses as to the group of individuals within a household which it defines as a "family". For either country now to adopt the other's concept would involve impairment of comparability within its own series. Nor am I sure that there is any strong desire on the part of either country to admit that the other employs a more useful definition.

Basically, the Canadian definition of a family considers only the husband and wife (with or without unmarried sons or daughters), or one parent with one or more unmarried son or daughter, as a family. Thus it can be seen that under the Canadian definition, a family always consists of two or more persons in a parent-child or husband-wife relationship. The United States definition, on the other hand, is broader in its concept, in that a family comprises all persons related by blood, marriage or adoption, within the household, of which there must be two or more such persons. Thus, for example, when a married daughter and her husband live in her parents' home there would be two families according to the Canadian definition, but only one under the United States concept. Any combination of brothers, sisters or other related persons comprising a household in the United States would constitute a family, but would simply be a group of two or more related persons in Canada.

Thus, we in Canada tend to think of our family definition as using the "immediate family" approach, and the United States as employing an "economic family" approach. Some indication of the comparability of statistics based on these two definitions can be observed from figures derived from the 1956 Census of Canada. A special tabulation was made of "economic families", using the same definition as the United States concept of a family and then compared with statistics based on the "immediate family" concept of the Canadian census. The resulting figures showed a total of 3,640,000 families using the United States definition, and 3,710,000, or 2 per cent more, using the Canadian definition. This small difference of 70,000 more families under the "immediate family" approach, however, does not tell the whole story. Some 16 per cent of all families were affected by the change in definition, while for the remaining 84 per cent, the same group of persons comprised the family under either definition. Further, the average size of the "economic" family was closer to the average size of household than it was to the average size of the "immediate" family.

Item	"Immediate" family concept(1)	"Economic" family concept(2)
1. Number of families	3,711,500	3,642,500
2. Persons in families	14,077,200	14,594,500
3. Persons in families as percentage of population	88.2	91.4
4. Average persons per family	3.8	4.0

(1) Corresponding to Canadian definition of a family.
(2) Corresponding to United States definition of a family.

Differences in the composition of families under the two definitions were most pronounced, as would be expected, for persons of adult ages. Of the 517,000 persons who were non-family members under the "immediate family" definition but became family members under the "economic family" approach, almost 88 per cent of these were 25 years of age or over. This means that the "economic family" definition has the effect of bringing into family status a relatively large group of adults such as widowed fathers, mothers, mothers-in-law, as well as a fair number of brothers and sisters, and other relatives. While this may sound as though the more extended definition includes, in the words of Gilbert and

Sullivan, "and his sisters and his cousins and his aunts", it has certain definite advantages, one of which I might comment on.

The value of this concept for classifying such census statistics as family income, for example, is undeniable as it relates more closely to the idea of the family spending unit than the "immediate family" definition. The 1961 Census of Canada, recognizing this value, plans to provide a number of such tabulations on the basis of "economic families". However, as in the past, the main tabulations on families and their composition will again be based on the "immediate family" definition. Housing authorities in Canada, for example, have expressed preference for this concept in regard to estimates of housing supply and demand. Some social analysts appear to prefer the more limited and precise concept for certain fields of research. How valid, or how important these reasons are, I am not prepared to say.

The fact, however, that both definitions are capable of being broken down into various sub classifications using such groupings as "primary" and "secondary" families, "sub-families", "related" and "non-related" families, etc., provide the analyst with some means of comparison providing he is fully aware of the meaning and uses of these terms in each country. It would be beyond the scope of this paper to go into these numerous facets of the family definition. Suffice it to say, that anyone making use of the census family statistics of both countries in any sort of comparison must, for the present at any rate, proceed with all due caution in respect to the independent viewpoints of two neighbours on this subject.

Metropolitan and Urbanized Areas

A geographic concept which has assumed ever-increasing importance in more recent years is that defining the "metropolitan area" of a major city. The United States has devoted a great deal of attention to this concept and has developed over the past few decades a standard set of uniform criteria upon which to base the establishment of metropolitan area boundaries. As yet in Canada we do not employ some of the criteria which give more preciseness and uniformity to the United States definition. However, we do make use of some of the more basic criteria such as population densities, proportions of farm population and distances from the central city, to define the metropolitan area limits of a given city.

Possibly the most basic difference in this concept as between the two countries relates to the inclusion in the United States of complete counties in the fringe areas when they satisfy the given criteria, whereas in Canada this is done on the basis of minor civil divisions. The main reason for this in Canada is that the county divisions are generally larger administrative units than their counterparts in the United States. If we were to adopt the county as the basic unit about 15 out of our 17 metropolitan areas would be unduly exaggerated.

That Canada is gradually moving in the direction of the United States in regard to this concept, however, is evidenced by the fact that for the first time in the 1961 Census, we have tried to define the truly "urbanized areas" based on urban densities, as is done in the United States, within the framework of the metropolitan boundaries. Since the Canadian concept of a metropolitan area, however, is at present somewhat more restricted than in the United States, chiefly through the use of minor civil divisions as building blocks rather than counties, the resulting difference between the "urbanized area" of a given city and the "metropolitan area" of the same city will not be too significant in Canada's 1961 Census. However, we are in a transitional stage for the 1961 Census, and we hope to take a much closer look at the present metropolitan area boundaries on the basis of the 1961 results, with the view to possible enlargement and unification of their boundaries on a less restricted concept of a city's outgrowth.

Thus, for the 1961 Census, the "urbanized area" within the boundaries of Canada's metropolitan areas will be used to give a more precise delineation between rural and urban, but we are planning to issue statistics on population and housing characteristics only on the "metropolitan area" basis, due to the relatively close correspondence which presently exists between the two sets of boundaries. This is evident when it is considered that the Canadian definition of a metropolitan area would naturally include any complete minor civil division in which a significant portion of that civil division would also be part of the "urbanized area". On a more extended definition of a metropolitan area, as in the United States, this would not be the case.

Ethnic or racial groups, and languages

A concept which has been accompanied by much controversy in recent years, in Canadian censuses at any rate, is that concerned with the measurement of population in terms of ethnic or cultural groups. In Canada, a question on this subject has been included in every census, but one, since 1871. It is difficult to comment on the criteria for this inquiry as compared, for example, to the question on "Colour or race" of the United States census. The basic difference, of course, is that in addition to trying to measure the numbers and characteristics of such non-white ethnic groups as "Negro", "North American Indian", etc., the Canadian census attempts to classify the total population into major ethnic categories such as French, German, Italian, Jewish, Ukrainian, and so forth. One of the principal reasons for such an inquiry in Canada is to meet the many requests for this type of information from within the various ethnic groups themselves.

Also, census information is one means of measuring the extent to which various ethnic groups have retained some of their ethnic and cultural identity, as for example, by comparisons of ethnic group statistics with those of mother tongue (which are largely determined on a language criterion). Thus, the Canadian Census defines mother tongue as "the language first learned in childhood and still understood". The United States, on the other hand, asks mother tongue only of persons born outside of the United States using the concept of "language spoken in the home before coming to the United States".

The United States inquiry on "birthplace of parents" provides information on national or geographic origin in the case of the foreign-born and first generation Americans of foreign parentage. However, the United States concept could not deal adequately, in Canada, with the measurement of the French ethnic group, for example, with its long history on the Canadian scene, particularly insofar as measuring its gains or losses.

What the future holds for the Canadian inquiry on ethnic groups is difficult to predict. Despite its numerous imperfections, one wonders what sort of substitute information might be forthcoming in the way of diverse estimates of all sorts from many quarters, were the census not to continue to try to provide for the many interested groups and agencies in Canada at least as factual and uniform estimates as possible on this complicated subject.

The economically-active population

In the complex subject-matter area of measuring the economically-active population, where seemingly arbitrary rules have to be set up to define what we mean by such terms as "labour force", "unemployment", "class of worker", and numerous others, one might expect serious problems of diversity in the conceptual approach adopted by Canada and the United States in their current censuses. However, such is not the case, and differences for the most part are concerned more with the amount of detail and supplementary information obtained by each country. Possibly the most serious element of non-comparability in the basic labour data is not conceptual in nature at all, but rather stems from the fact that the measurement of labour force characteristics in the United States census relates to a week in April, while in Canada it is a week in June. The intervening period is a season of change, as for example, in the primary industries from forestry operations, trapping, and the like, to farming and construction activities. It is a period of improvement in unemployment rates, of new workers such as students entering the labour force on a part-time or full-time basis, so that Canada's Census in this respect at least should be on a more optimistic note than that of the United States.

A few other differences might be mentioned briefly in this general area. The United States, in addition to obtaining information on the number of persons currently at work, unemployed or with a job but not at work, also included in the 1960 Census a question on the most recent year in which persons who were not currently working did any work. The 1961 Census of Canada obtained data on the number of persons with a current job, or were unemployed or who had a job anytime during the preceding 12 months. Thus, the Canadian Census will not have a corresponding count to that of the United States of persons not in the current labour force who did any work prior to the 12 months preceding the census. A further enlargement of the scope of the United States criteria as compared to Canada provided for questions to obtain the job description of all persons who had worked anytime during the past 10 years. In Canada, we did not obtain the job description of persons unless they had been economically active at some time during the 12 months preceding the Census.

In the field of income statistics, Canada has followed the lead of the United States, not without some trepidation on our part, in attempting to obtain for the first time statistics on personal income from all sources, rather than income from salaries and wages only, as in previous censuses of Canada. In the 1960 United States Census, the basis of inquiry for collecting income data was the calendar year (1959) whereas in the recent Canadian census it was the census year (June 1960-May 1961), although the amount for the calendar year (1960) was accepted if the exact amount for the past 12 months was not known or could not be estimated. In this instance, as in several others I could mention, the April 1st United States census date was clearly superior to Canada's reference date in June.

Housing concepts

Turning for a moment to the Census of Housing, one finds again a high degree of correspondence in the concepts of Canada and the United States. Such differences as there are, for the most part are due to differences in terminology and the amount of detail obtained. Thus, for example, the United States provides very precise criteria by which to determine condition of dwellings. Although the terms describing condition differ, both countries will produce results for three degrees of housing quality, viz., in the United States "sound, deteriorating and dilapidated"; in Canada, "good condition, in need of minor repair and in need of major repair". Much greater detail is obtained with regard to vacant dwellings in the United States than in Canada. Again, while the concept of toilet and bath facilities, including exclusive and shared use, is essentially the same in both countries, the United States goes farther than Canada and counts the number of bathrooms and partial bathrooms in each housing unit.

Certain variations occur in questions regarding household accessories. Thus, while both countries inquire as to homes equipped with home freezers, television sets, and passenger automobiles, Canada includes also a question on refrigerators, while the United States includes questions regarding radios, telephones, air conditioning units, clothes dryers and washing machines. The Canadian Housing Committee decided that radios and telephones were no longer as indicative of standards of living in Canada as certain other criteria. Data on clothes dryers and washing machines are obtained from Special Surveys. However, air conditioning units are not yet sufficiently prevalent in Canadian homes to warrant their inclusion.

Concepts regarding tenure, value, and rents are basically alike. As in the United States, we in Canada have attempted to obtain a figure representing gross rent. The items and basis of payment for additional services included in the gross rent are the same in both countries. In addition to the amounts paid for these services (water, gas, electricity and fuel), Canada included a question to ascertain whether or not cash rent included a refrigerator, cooking stove, furniture or garage.

The Census of Agriculture

The United States and Canadian Censuses of Agriculture are very similar with respect to concepts, questionnaire content, and types of tabulations, although there are a few notable differences. The United States Agriculture Census, for example, was taken in the fall of 1959, as a separate census operation, while in Canada it was taken in June, 1961, in conjunction with the Population and Housing Censuses. With the heavy work-loads of decennial censuses, it will be a matter of prime consideration in the planning of future censuses of Canada as to whether the present arrangement will continue, although there are a number of advantages to the combined operation which we would not like to forfeit. A further difference is that a substantial section of the 1959 United States questionnaire was enumerated on a 20 per cent sample, whereas in the 1961 Census of Canada sampling techniques for agriculture were abandoned as a result of the persistent demand that all information be tabulated for small areas.

A change in the concept of a "farm" was made for both the 1959 United States Census and the 1961 Canadian Census. Briefly, this change introduced sales of agricultural products as a criterion in place of production alone. Consequently, some holdings which produced only for home use are no longer classified as farms for census purposes. The United States definition specifies as farms those places of less than 10 acres with agricultural sales of $250 or more for the year, or places of 10 acres or over with agricultural sales of $50 or more. The 1961 Census of Canada defines a farm as a holding of one acre or over with agricultural sales of $50 or more. From this it can be observed that the main difference in the "farm" concept is that the United States would include a holding of less than one acre as a farm if sales were at least $250, whereas Canada did not include as farms any holdings of less than one acre. On the other hand, Canada would include holdings of one to 10 acres if agricultural sales amounted to $50, rather than $250 as in the United States.

One other related concept which might be mentioned briefly in the field of Agriculture relates to the "commercial farm" definition, for which important agricultural data are provided in both countries. The definition of a "commercial farm" for the 1961 Census of Canada is simply a farm with sales of agricultural products for the year of $1,200 or more. A more involved concept was used in the 1959 Census of Agriculture in the United States, as follows:

(a) Farms with sales of agricultural products for the year of $2,500 or more;
(b) Farms with sales of agricultural products for the year of $50 to $2,499 provided the farm operator was under 65 years of age, and (1) he did not work off the farm 100 or more days, (2) the income that he and members of his household received from non-farm sources was less than the total value of farm products sold.

The farm definition proper should not contain too great an element of non-comparability, since in Canada only 5 per cent of farms in 1956 were 10 acres or less in size at the point where the sales criteria differ. However, the difference in the "commercial farm" concept between the two countries could be of greater significance, although the extent of this would be difficult to estimate at this time.

The problems of reconciliation

Since it is not possible, of course, to compare in this paper the full range of census concepts, I have attached a summary statement which attempts to outline in a very brief way the main differences in some of the more significant definitional areas. There is some danger to this, of course, in that such a summarization may tend to over-simplify and minimize these differences. In certain cases too, they give the illusory appearance that it would be extremely easy to bring the concept of one country or the other into line in order to obtain complete comparability.

One has only to observe the first item on the list, however, to see that this is not so. The first item has to do with the concepts which underlie the counting of population at their usual place of residence. Both the United States and Canada basically take a "de jure" census and their residence rules appear to show a difference only in respect to the counting of university students. Since one country counts such persons basically where they are attending university, and the other where their "home" addresses are, it would seem a simple matter for one or the other country to change its rule in this regard. On looking closer, one wonders how many university students might be missed in the United States Census if on April 1 they tried to enumerate these at their usual homes or how much fun the Canadian Census would have in enumerating them at their University on a June 1 census date.

One might say that the solution is first to make the census dates coincide and then decide on a uniform rule. We in Canada are envious of the April 1 date used in the United States being farther removed from the vacation season than June 1, but our enumerators in the more rural types of areas report enough difficulties as it is from late spring weather problems which can be encountered in our country even in early June.

Each question has its own set of problems, and one can readily see the difficulties which the Statistical Commission of the United Nations is up against when it tries to encourage uniformity in census concepts among many vastly dissimilar countries, instead of only two whose general characteristics, geography, and so on, are in many ways so much alike. I feel sure that in the case of the United States and Canada a number of the present differences in our census concepts will disappear over the years. As census experience builds up in regard to given sets of criteria, the better ones gradually become apparent and replace the less useful ones.

We hope in Canada that we will be sufficiently intelligent to take advantage of the great pioneering work that has been done in these fields by our neighbour to the south. In fact, I am sure that you can tell from the many similarities in our fundamental approaches to most census concepts that it has been achieved through the meetings of the joint committees where a spirit of mutual respect and cooperation has existed between the two census organizations for the past 10 years. We feel there is every assurance that this will continue in the future.

Summary statement of some of the more significant differences in

census concepts employed in the United States and Canada

Concept	Canada	United States	Remarks
POPULATION			
1. The "de jure" population	Basically similar to U.S. residence rules, except that university students are counted at their "home" address.	University students are counted as residents of the communities in which they are residing while attending college.	Differences likely grew out of use of April 1 census date in U.S. as compared to June 1 in Canada.
2. Rural and urban	Urban defined as all centres of 1,000 and over, plus the "urbanized" fringes of cities of 10,000 and over.	Urban defined as all centres of 2,500 and over, plus "urbanized" fringes, as well as certain localities included by special rules.	Main difference is in the use of the 1,000 or 2,500 minimum. The 1,000-2,500 range amounted to 8 p.c. of the total urban in the 1956 Census of Canada.
3. Metropolitan Areas	Consist mainly of complete minor civil divisions below the county level based on urban densities, proportions of farm population, etc.	Consist mainly of complete counties on the basis of a standard set of criteria to ensure uniformity among cities.	Criteria not as refined in Canada as in the U.S., and present boundaries closer to "urbanized area" concept.
4. Farm population	All persons living on a farm as defined by Census of Agriculture, (i.e., one acre in size and $50 sales, minimum).	Same except that "farm" defined as follows: Under 10 acres: Sales of $250 or more 10 acres or over: Sales of $50 or more.	In the 1956 Census of Canada, 5 p.c. of Canadian farms were of 10 acres or less.
5. Dwelling (or housing) Unit (See also Items 16-26 re Housing Concepts	To be classed as a dwelling, living quarters must be structurally separate and must have a private entrance either from outside or from a common hall, lobby, or stairway inside the building. Such entrance must be one that can be used without passing through anyone else's living quarters.	A room or group of rooms occupied as separate living quarters, i.e., when occupants do not live and eat with other persons in the structure, and when there is either (1) direct access from the outside or through a common hall, or (2) a kitchen or cooking equipment for the exclusive use of the occupants.	Greater emphasis on structural separateness in the Canadian definition, which in practice leads to some lack of uniformity in the treatment of marginal cases. U.S. housing unit definition is based more on living arrangements rather than structural separateness.
6. Family Unit	Consists of a husband and wife (with or without children who have never married) or one parent with one or more children never married, living together in same household.	Consists of two or more persons living in the same household who are related to each other by blood, marriage, or adoption; all persons living in one household who are related to each other are regarded as one family.	Both definitions employ several sub-group classifications. Average size of Canadian family using U.S. definition was 4.0 in the 1956 Census, compared to 3.8 using Canadian definition of "immediate family".
7. Age distribution	Age in completed years at last birthday prior to the census date.	Same except that question asked in terms of "month and year of birth".	Different method of asking question should not affect materially the comparability of age statistics between the two countries.

Concept	Canada	United States	Remarks
8. Marital Status	Four categories (viz., single, married, widowed, and divorced). The "married" category includes persons separated for any reason unless a divorce has been obtained.	Five categories asked for (viz., single, married, married but separated, widowed, and divorced).	Previous attempt in 1941 Census of Canada to obtain a count of "separated" not successful. Alternative use made of family data on "married" (both present)", and "married (spouse absent)".
9. Ethnic or racial groups	Origin question asked as: "To what ethnic or cultural group did you or your ancestor (on the male side) belong on coming to this continent? Instructions to enumerators provided certain rules for determination, principal among which was the language spoken at time of arrival.	U.S. question on "Colour or race" somewhat comparable for non-white groups. Some ethnic data for other groups on first-generation basis obtained from questions on country of birth of parents.	Questions appear to be designed for different purposes to meet specific internal needs for information of the type collected by each country.
10. Mother Tongue	Defined as "the language first learned in childhood and still understood".	Question restricted to foreign-born population, and defined as "the language spoken in the person's home before coming to the United States".	Uses of this question in each country related to previous item on ethnic or racial groups, and account for different approach.
11. Education	Two questions asked: (1) Highest grade or year of schooling ever attended? (2) Attended school or university since last September?	Basically similar concepts, except that two additional questions asked: (1) Completed the highest grade attended? (2) If attending school, whether a public or private school?	Enrolments by type of school obtained in Canada by Education Division, D.B.S. Also Canada did not attempt the more difficult question on grade completions, particularly in view of census date.
12. Employment Status	Concepts basically similar to those of U.S. regarding persons with a job, unemployed, etc. Labour force questions asked of persons 15 years of age and over to tie in with standard age groupings, rather than 14 and over as in the U.S. Change-over in Canada also influenced by relatively small numbers in labour force at this age level.	U.S. questions provided for separate category of persons "With a job but not at work". Also, year last worked will be available for persons not in current labour force who worked at any previous period. Job descriptions obtained for all persons who worked within past 10 years, as compared to past 12 months for Canada.	Comparability of basic data on the economically active population likely affected more by difference in census reference date (viz., April in U.S., and June in Canada) than by any conceptual differences.
13. Hours worked per week	Based on the <u>usual</u> number of hours worked each week.	Based on the <u>actual</u> number of hours worked in the given census week.	Canada used "usual hours" in the expectation that it will be more representative when related to weeks worked during the year.
14. Weeks worked per year	Refers to weeks worked for wages or salary only, for all persons who were economically active anytime in the 12 months preceding the census date.	Refers to weeks worked in the preceding calendar year for all persons who did any work in that period.	Weeks worked is used in Canada mainly to relate to hours of work and incomes from wages and salary.

Concept	Canada	United States	Remarks
15. Income from all sources	Wages and salaries obtained from population universe. Earnings of persons in non-farm households from self-employment and other forms of income obtained from 20 p.c. sample basis. All questions based on 12 months preceding the June 1 census date.	Income inquiries on 25 p.c. sample basis, referring to calendar year 1959. Income from sources other than earnings obtained through one question, as compared to seven separate questions on Canadian questionnaire.	Calendar-year figures in Canada would relate to a period commencing some 18 months prior to the census enumeration. Provision made for 1960 calendar-year figure if person could not estimate income for immediately preceding 12-month period.
HOUSING			
16. Collective dwelling or group quarters	Institutions, camps, hotels, and lodging houses of 10 or more persons.	Living arrangements for institutional inmates and for groups of 5 or more persons unrelated to the head of the household.	Main difference is that the U.S. classifies all dwellings with 5 or more unrelated persons as "group quarters".
17. Vacant dwelling	A vacant dwelling is one suitable for occupancy but unoccupied at Census date (unless occupants only temporarily absent).	A housing unit is vacant if no persons are living in it at the time and occupants are not just temporarily absent.	No difference in the basic concept but U.S. categorizes types of vacancy as "vacant, under construction", "being converted", "used for non-residential purposes", "unfit for human habitation", "abandoned", etc. Canada distinguishes only as to whether vacant dwellings were occupied before, and their type.
18. Dwellings under construction	A dwelling is considered under construction from the time the foundation is begun until ready for occupancy.	Included only as part of "vacant" count if all exterior windows and doors are installed and usable floors are in place, but not occupied; otherwise not enumerated.	See Item 17, above.
19. Structure	Not defined.	A structure is a separate building that either has open space on all four sides or is separated from other structures by dividing walls that extend from ground to roof.	By emphasizing the "structural separateness" of dwellings, it was hoped, in Canada, to avoid the need for defining "structure" in 1961. (See concept below for "type of dwelling")
20. Type of dwelling	single detached - one dwelling unit completely separated on all sides from any other dwelling or structure. single attached (double house) - two dwellings separated by a common wall extending from ground to roof. Single attached (other) - all other dwellings separated by a common wall from ground to roof.	House, apt., flat - every dwelling other than trailers.	The chief difference here lies in the fact that while Canada attempts to distinguish single-unit, double-unit and multi-unit structures by type of dwelling, this is done in the U.S. by means of the "structure" concept mentioned above. Thus the U.S. may have detached and attached structures of one unit only or of more than one unit, which more or less ties in with the Canadian concept of "type of dwelling".

Concept	Canada	United States	Remarks
20. Type of dwelling - Concluded	apt., flat, etc. (duplex) - 2 dwelling units one above the other and adjoining no other structure. apt., flat, etc. (Other) Mobile - any dwelling designed for movement and actually moveable, such as trailer, boat, quarters in a railway car, or house permanently on skids.	Trailer - each occupied house trailer whether mobile or on permanent foundation.	Since trailer-living is not so widespread in Canada as the U.S. it was felt that the broader category of "mobile" would be more suitable for the purposes of this Census. Trailers on permanent foundations were classed as ordinary single-detached dwellings.
21. Condition of dwelling	Three categories, as follows: In good condition. In need of minor repair. In need of major repair.	Two categories, as follows: Not dilapidated: (a) Sound (b) Deteriorating Dilapidated.	Although the terminology to describe condition varies between the two countries, and the U.S. goes into greater detail in defining each category, the concepts are basically similar.
22. Rooms	Both countries advise enumerators to accept the respondent's count of rooms.		The U.S. includes "rooms used for offices by a person living in the housing unit", while Canada excludes "rooms used solely for business purposes". The U.S. gives instructions re partially divided rooms, also, while Canada does not.
23. Source of water supply	Two categories, as follows: (a) Private source on property. (b) Other (municipal mains, etc.).	Three categories, as follows: (a) Public system or private company. (b) Individual well. (c) Other.	Main difference here is in the concept of "private source". In Canada this means a well, spring, or other source located on the property. In the U.S. the "individual well" may be on a neighbouring property serving 5 or fewer houses, while water from springs, creeks, etc. are placed in "other".
24. Fuel	Canada asked specifically for heating fuel only.	The U.S. distinguishes fuels used for heating the housing unit, for cooking and for heating water.	The additional information provided by U.S. should be of considerable value to users.
25. Monthly rent	Rent for month of May, 1961.	Monthly rent for unit.	U.S. includes rent for vacant dwellings. Canada does not.

Concept	Canada	United States	Remarks
26. Value of dwelling or housing unit	Amount expected if sold to a willing buyer.	Amount expected if sold on today's market.	About same concept in both countries with slightly different value intervals. U.S. includes value of vacant dwellings. Canada does not.
AGRICULTURE			
27. Farm	Holdings of one acre or more with agricultural sales of $50 or more for the year.	(a) Places of less than 10 acres with agricultural sales of $250 or more; (b) Places of 10 acres or more with agricultural sales of $50 or more.	In Canada, holdings of less than one acre, and others not satisfying farm criterion, are enumerated on special questionnaire to obtain data on any agricultural operations, (e.g., numbers of livestock). In United States, criteria applied during office processing and farms selected according to definition.
28. Commercial farm	Farms with sales of agricultural products for the year of $1,200 or more.	(a) Farms with sales of agricultural products for the year of $2,500 or more; (b) Farms with sales of agricultural products of $50 to $2,499 provided the farm operator was under 65 years of age, and (i) he did not work off the farm 100 or more days, (ii) the household income from non-farm sources was less than the total value of farm products sold.	While the "farm" definition does not contain too great an element of non-comparability, (since only 5 p.c. of farms in Canada were 10 acres or less in 1956), the difference in the "commercial farm" concept between the two countries could be of greater significance.

Percentage of households by number of persons(1)

Persons	1950 Census of United States	1951 Census of Canada
1	9.3	7.5
2	28.1	21.0
3	22.8	20.3
4	18.4	19.0
5	10.4	12.9
6 +	11.0	19.3
Average	3.5	4.0(2)

(1) A household is defined as all persons occupying a dwelling unit.

(2) Average persons per household in 1956 Census of Canada was 3.9.

Comparisons of family size, Census of Canada, 1956

Item	"Immediate" family concept(1)	"Economic" family concept(2)
1. Number of families	3,711,500	3,642,500
2. Persons in families	14,077,200	14,594,500
3. Persons in families as percentage of population	88.2	91.4
4. Average number of persons per family	3.8	4.0

(1) Corresponding to Canadian definition of a family.

(2) Corresponding to United States definition of a family.

THE TABULATION PROGRAMME FOR THE 1961 CENSUS OF CANADA

Donalda A. MacIntosh, Dominion Bureau of Statistics

Someone has suggested that the Census could be described as being second only to a war in the demands it makes on a government's organizational resources. Perhaps the best measure of the value of this decennial project is the ever-increasing demand by the public for the resulting data. It is in an attempt to satisfy this demand within a reasonably short period after the actual count that each Census sees new procedures and new equipment brought into operation. While it has not been possible or even practical to try to meet every request, the Census of Canada has undertaken in 1961 what is probably its most ambitious tabulation programme. In this regard, the addition of a high-speed computer to the Canadian Census scene has played an important role.

Time will not permit a detailed examination of each set of tabulations. Instead, this paper will attempt to summarize the programme planned for the Census of Population and Housing, pointing out the geographic areas for which the principal tables will be available, and noting, also, the principal areas in which the computer has affected the results.

Provincial data

In planning the tabulation programme, the significance of any detailed data had to be considered in deciding the size of the geographic areas for which the distributions should be provided. Thus, while most of the tables will provide counts for areas below the provincial level, a relatively small number are restricted to Canada and the provinces. Such tables as (a) the distribution of normal families showing age of husband classified by age of wife, and (b) the cross-classification of the experienced labour force by sex, showing detailed occupations by detailed industries, are examples of the kind of tabulations planned at the provincial level only.

Counties and Census divisions

Within the provinces, counties or census divisions are the largest geographic areas. In addition to tabulations for most of the basic population and housing characteristics, cross-classifications at this level will include tables showing such characteristics as sex, marital status, religion, language, and education classified by (a) age, and (b) origin. Basic labour force distributions for these areas will include a count of the experienced labour force classified by marital status and sex.

Municipal subdivisions

The next largest geographic areas, the municipal subdivisions, range from the largest metropolitan areas to the smallest incorporated villages. They include other forms of municipal organization such as the parishes in New Brunswick, municipalities in Quebec, townships in Ontario, municipal districts in Alberta, and so on. The Census will not provide the same amount of information for each of the several thousand such subdivisions across Canada. However, certain very basic characteristics will be made available for each of these areas, regardless of size. These will include population counts distributed by the characteristics of age, sex, marital status, origin, religion, years of schooling, language, and birthplace. Not all of these characteristics will appear in the published reports, but they will be made available for particular communities or other subdivisions on request.

For cities, towns, and incorporated villages of 1,000 population and over, population distributions by broad occupation and industry groups will be available, as well as basic facts concerning household and family composition. For rural types of subdivisions, these same data will be combined for the balance of the county in which they are located.

At the 5,000 population level, certain basic housing characteristics from the 20 p.c. sample will be tabulated, but the full range of distributions for the housing subjects of inquiry will be compiled for rural and urban subdivisions of 10,000 population and over. At this same level, more detailed distributions of population characteristics will be produced, as well as the full occupation and industry distribution of the working force for the cities and towns of 10,000 and over. For Canada's major cities and metropolitan areas of 100,000 and over, these occupational distributions will be extended to include various cross-classifications such as (a) occupations, and (b) industries by age groups and sex, marital status, and earnings of wage-earners. Cross-classifications of general population characteristics such as age groups by marital status and sex, origin groups by religious denomination, and birthplace by period of immigration, will be compiled for individual centres and municipal subdivisions of 30,000 population and over.

In addition to the data already noted for the Census metropolitan areas of 100,000 and over, detailed cross-classifications for all metropolitan areas will be included for most of the population and housing subjects. The component parts of the M.A.'s will be included in the basic counts of such characteristics as age, sex, marital status, birthplace, religion, language, and education and of the total labour force. For these same areas, the 20 p.c. sample of Housing will provide counts of the number of dwellings classified by such characteristics as type of dwelling, tenure, and residential use.

Census tracts

For most purposes, the establishment of Census tracts in Census metropolitan areas and cities of 50,000 and over has fulfilled the needs of a widespread demand for statistics which would isolate areas of change within a community. Not only do they provide a meaningful and useful criterion for the publication of a wide range of census statistics. For municipal authorities, they provide, also, a means of assembling local data which can be viewed against the background of census information compiled on the same basis.

The 1961 Census tract reports will provide again basic population and housing distributions for each Census tract in Canada's major cities. These will include such basic population characteristics as age, sex, marital status, origin, religion, and years of schooling. Household data will show the number of persons and families per household, as well as the number of households with lodgers. The main classifications of family information will relate to the number of children in families and the ages of children.

Housing data will include distributions by types of dwellings, rooms per dwelling, tenure, median rent and year of occupancy. The number of dwellings which possess such household facilities as home freezers, television sets, and passenger automobiles will be shown also. The labour force population will be classified in the Census tract reports by employment status during the week preceding the Census date of June 1, 1961, as well as by such characteristics as class of worker, broad occupation groups, and earnings of wage-earners.

Enumeration areas

To meet the needs of users who urgently require population or housing characteristics for city blocks or other units smaller than those covered in earlier censuses, the 1961 Census is prepared to offer them on the basis of the slightly larger-sized "enumeration area". An Enumeration Area (or E.A.) is the territory covered by one enumerator. In a city, one E.A. usually comprises several city blocks. For 1961, it was planned that the populations of these areas should range from 400 to 800, and in some cases, were even smaller.

The Bureau is not planning to publish Census results by E.A.'s, but does plan to have machine tabulations prepared at cost for basic distributions that may be required to meet the specialized requests for this information. These tables are planned for the intercensal period.

City blocks

The city block is the smallest geographical entity which could be expected to provide a basis for the issuance of census statistics. Because of the tremendous costs involved, however, and the relatively small number of users who would be in a position to make full use of the masses of small figures on this basis, the Canadian Census has never provided block statistics. A cautious step in this direction was taken in 1961, when the enumerators in cities of 50,000 population and over were required to identify the households in each individual block so that counts of total population and households could be assembled on a block basis.

Because of the specialized nature of this material and the relatively small but insistent demand, such counts would have to be made available to users on a cost basis. This, of course, can only be done when the main census results have been compiled and released to the public. Any further steps, such as the provision of population and housing distributions by age, sex, or type of dwelling for city blocks, were not planned, since this would have required block identification on the actual enumeration document, complicating the enumeration processes. The use of expensive tabulating equipment for the compilation of such extremely small sums would not seem to be justified.

Unincorporated villages, etc.

Another type of small area which is recognized by the Census relates to the many thousands of small villages, settlements, and hamlets scattered across Canada which are not municipally organized or incorporated in any way. The populations of such places and their characteristics are included with the particular municipal subdivisions in which they are located. However, a special report is issued at each Census which gives the name, location, and population of each such place which appears to have a population of 50 persons or more. For example, the special report of the 1956 Census contained some 10,000 such place names, their localities, and approximate populations. The word "approximate" is used to qualify these counts, since there are no legal or official boundaries to unincorporated settlements, and identification of a particular community is not always possible.

Changes since 1951

What are some of the additions since 1951 and what has been the contribution of the computer toward extending and facilitating the programme?

Perhaps the most significant addition to the 1961 programme is the inclusion of tabulations for migration, family size, and income data from the 20 p.c. Population (Sample) Questionnaire. Family size and migration inquiries were included in the 1941 Census but were not repeated in 1951. Changes in the patterns of these characteristics during the past decade, however, seemed to indicate the need for their inclusion again in 1961. Total income, on the other hand, is completely new to the Canadian Census.

In each of these fields, a fairly comprehensive tabulation programme has been set up, but decisions as to the tables to be published will be made only after examination of the material. It is intended, however, that a number of tables from each series will be included for publication.

From the migration data, it is expected that information will be available to show the in- and out-migration of the population for various geographic units from which it will be possible to examine the gross movement as well as the direction and pace of migration. Tabulations are planned, also, which will classify the migrant population by type of movement (i.e., intraprovincial and interprovincial migration, movement from contiguous and non-contiguous provinces, movement from rural to rural, rural to urban, etc.). Cross-tabulations of these data with such characteristics as age, sex, marital status, schooling, origin, place of birth, labour force status, occupation, and income will throw some light on the implications of population movements for different parts of the country.

On the basis of tabulations planned for the family size inquiries, the facts about the fertility trends in the past and some of the factors operating in the observed changes and variations in the family formation pattern may be brought to light. Furthermore, the growth potential of local populations as well as the national population may be appraised more realistically than from annual birthrates.

From the income tabulations it is planned to publish three series of tables: (1) income of individuals by selected characteristics; (2) family income by selected family characteristics; and (3) selected household characteristics cross-classified by the income of primary family or the family head. In addition, it is planned to carry out a number of other studies for special release after the main volumes, and to probe into other areas such as the analysis of the population in the older age groups.

Probably the most substantial increase in the volume of data is planned for the employment and economic characteristics from the 100 p.c. count. Data which were available, in some cases, for major industry divisions in 1951 will be available for groups within these divisions in 1961. For example, for certain geographic areas in 1951, data were provided for the number employed in the manufacturing division as a whole. In 1961, the numbers employed in the major manufacturing groups such as the Food and Beverage industries, the Textile and Clothing industries, and the Primary Metal industries, will be provided for these areas. More employment data will be available for married women and for older workers. For the unemployed, such characteristics as their level of education in relation to age, marital status in relation to age, and the period of immigration for the foreign-born unemployed will be compiled. The computer has made possible a considerable number of four-way classifications of labour force data, something not considered in the past. An example of such a table is the classification of detailed industries by detailed occupations by class of worker and sex. The addition of the variable "Class of worker" in this table, incidentally, will provide, for the first time, a distribution of occupations through industry for "paid workers". The computer will provide aggregate and average income for specific geographic areas, as well as for male and female employees in detailed occupations and industries.

Basic counts of family, household, and housing characteristics have been available in the past only down to the level of municipalities of 10,000 and over. In 1961, these counts will be extended to the smallest municipality. In each of these fields, the computer has made possible more detailed cross-classifications and has allowed for an increase in the number of these tables. Classifications of families, for example, will include a considerable volume of data for husband and wife families classified by age, labour force status, and earnings of wife, a tabulation that is new to the Canadian Census and one that should provide interesting information to users interested in working wives. In 1951, household characteristics were tied in with housing characteristics of the household head. In 1961, in addition to similar data, cross-classifications based on household data alone will be run. Cross-classifications of data from the 20 p.c. Housing sample will be extended, also, and, in addition, the computer will make possible a mechanical blow-up of the sample, a task that was carried out manually in 1951.

In spite of the increased volume of tables, it is expected that most of the data required for the reports and volumes will be available earlier than in the past. In the field of general population, for example, it is expected that the cross-classifications by age, origin, and period of immigration will be available from one to four months earlier than in 1951. Similar advances are expected for the other groups of tables based on the 100 p.c. Population count and the 20 p.c. Housing sample.

Unpublished data

It should be pointed out that, although the bulk of data from the tabulations is intended for publication, a considerable amount of material will be available for research, monographs, and other special studies. For example, a set of tables is planned for studies of educational data that will combine population, family, and housing characteristics with schooling and school attendance. Housing tables are planned for the use of such agencies as the Central Mortgage and Housing Corporation, for example, cross-tabulations by gross and contract rent, value of single-detached non-farm homes, condition, number of rooms, etc. Tabulations to meet the needs of the Department of Labour have been incorporated into the labour force programme. For example, a four-way classification will provide them with data (1) by age by years of schooling by marital status and sex, and (2) by age by years of schooling by age and sex. This same table will provide the number of male and female wage and salary earners (1) by age by years of schooling by weeks, and (2) by age by years of schooling and hours.

It should be noted that some of these unpublished data will be provided on a non-priority basis after the tables required for publication purposes have been run. Others will be provided in the intercensal period.

Dates for release of publications

Part of the planning of the tabulation programme included the establishment of priorities to ensure that the tables planned for the reports and volumes would be available when required.

To satisfy the demands of the many users who are interested in summary data and to make such information available as soon as possible, the Census office plans to release an advance series of Population, Housing, and Agriculture reports. The Population reports in the series will include summary data on a number of basic population characteristics for counties and urban centres of 5,000 and over. Summary Housing data will be provided for these same areas. It is expected that most of the Population and Housing reports in this series will be released by mid-1962. The material in the Agriculture reports will be identical to that to appear later in the Agriculture volume but will be available considerably earlier, probably before the end of 1962. (The Agriculture volume (Volume V) is planned for 1963.)

The procedure adopted in 1951 of releasing volume tables as individual reports will be continued in 1961. Volume I, the general population volume will be divided into three parts. It is expected that Part I, the geographic distributions of the Population, and Part II, the counts of the various characteristics, will be completed by the end of 1962. Release of the tables in Part III, the cross-classifications of the Population characteristics, will begin late in 1962. Data on housing and household characteristics will form Part I of Volume II, and families and family characteristics will constitute Part II of this volume. Release of the reports for each part is expected to begin late in 1962 and continue through 1963. Part I of the labour force volume will include occupation data, Part II industry data, and Part III earnings and employment data for wage and salary earners. Issue of each of these series is expected to begin toward the end of 1962.

Since the 20 p.c. Sample Population Questionnaire was not a mark-sense document, a considerable amount of clerical work was required (e.g., preparation of punched cards) before tabulations could begin on the subjects of fertility, migration, and total income. Therefore, it is not expected that release of the reports on these subjects (Volume IV) will begin before 1963.

Volume V, the Agriculture volume, will cover such subjects as the number of farms, areas, tenure, crops, and livestock. As noted previously, release of these reports is planned for the early part of 1963. Reports dealing with the Census of Distribution will constitute Volume VI, release of which is planned to begin in 1963.

The final volume, Volume VII, will include the Administrative Report of the methodology of the 1961 Census as well as a review of the various subjects covered by the Census. Release dates for these reports will probably begin sometime in 1963.

A third series of reports is planned to cover some basic materials not included in the volume series, which will relate to more detailed geographic areas. Included in this special series will be data on specified age groups, origins and religious denominations for census subdivisions, as well as the report presenting the population of unincorporated villages and settlements. This special series will include, also, an Agriculture report giving the number and area of farms for counties and census subdivisions. For the most part, this series cannot be made available before the regular reports in the Volume series.

The remaining group of reports planned to date, the Census tract series, will be issued for 23 cities and Census metropolitan areas, release of which is expected to begin in the latter part of 1962.

In concluding, it should be pointed out that the dates of release of the Census data are subject to some change. It is possible, also, that additional reports of a specialized nature will be planned. A brochure, presenting a complete list of all reports and the tables to be included in each series with their expected date of publication, will be available in the near future.

DISCUSSION

A. Ross Eckler, Deputy Director, Bureau of the Census

Under the terms of a recently negotiated international agreement, I will concentrate my attention upon the papers given by Mr. Marshall and Mr. Lemieux. I should like to say at the outset that I believe that both the authors have done an excellent job in dealing with their subjects and it is difficult to be at all critical about either paper. As might be inferred from Mr. Marshall's statements regarding our close working relationships, the papers are of particular interest to us because they bring to date our information concerning a number of subjects that have been taken up in repeated inter-agency conferences.

For a considerable number of years the Dominion Bureau of Statistics (DBS) and the Bureau of the Census have had an extremely cordial and productive series of relationships concerned with the various operations of common interest to the two organizations. Our joint working committees in a number of subject and administrative areas have been able to keep closely in touch with the experimental work and testing on both sides of the border. We have observed each other's pretests and have welcomed each other in review and evaluation sessions aimed at determining how well a particular operation went off and what changes ought to be made next time. For example, two of our staff members were privileged to participate in the most recent meeting of the Canadian field organization, at which the Canadian staff undertook a very frank and full evaluation of their experiences in taking the 1961 censuses.

I might note that there is one question which has never been settled to our satisfaction. We have never been able to determine whether the Canadian censuses, taken in the years ending in "1" and "6", are pretests for our census taken in the years ending in "0", or whether our decennial census is really a pretest for theirs which is taken 14 months later.

Mr. Marshall's paper makes a real contribution in providing a compact summary of the differences between the census concepts used in Canada and those used by the United States. This represents the results of much careful study and should be a useful reference document for both bureaus, and even more so, for outside organizations.

The close working relationships between our bureaus insure that the differences are not due to ignorance or indifference. Our professional people are fully aware of the differences and are prepared to support the concepts used in each country as being most appropriate in view of national needs. For example, in connection with the choice of census date, the method of enumerating college students, the delineation of standard metropolitan areas, and the attitude toward questions on ethnic origin and religion - there are well-known factors which provide the basis for the differences in approach that have been used.

The Canadian procedures for setting up standard metropolitan areas involve the use of smaller civil divisions than is the case in the United States. As a matter of fact, their present procedure resembles one we used in 1940 to delineate metropolitan districts. As Mr. Marshall notes, the definitions used in Canada lead to relatively small differences between the metropolitan areas and the urbanized areas. We note that the reason for not using counties as building blocks for metropolitan areas is that the counties in Canada are so much larger. Nevertheless, one wonders whether the use of counties would bring in enough rural population to affect the statistics for a metropolitan area. It seems possible that in a great many cases the counties are so sparsely settled that the addition of a whole county with only a small fraction of the total area closely tied to the central city might not significantly affect the metropolitan area population total and its distribution by economic and social characteristics.

The discussion of differences between the United States and Canada in the treatment of the family unit is particularly interesting. I would raise a question as to whether Mr. Marshall is correct in referring to the United States as employing the "economic family" approach. This would seem to imply that the United States concept corresponds closely to the spending unit concept. Perhaps it is closer than the Canadian concept to the spending unit, but it is doubtful whether the agreement is close enough to support Mr. Marshall's terminology. In any case, the differences in practice between the concepts for the two countries prove not to be very great, and with supplemental information available on primary and secondary families, the problems of international comparisons seem not to be serious.

One important difference between the two countries is involved in the timing of the census of agriculture. In both of the last two U. S. censuses of agriculture the traditional "April 1" date was replaced by a fall date, with the advantage that the farmers are reporting at about the end of the crop year. The Canadians, on the other hand, retain the combination of agriculture, population, and housing in the spring of the year. Possibly, they might gain even more than we from shifting the agriculture census to the autumn, for under such a plan, the Canadian census of agriculture could readily be extended to cover production data.

Mr. Lemieux's paper contains a very valuable and clear statement of some of the major problems that affected the taking of the 1961 census in Canada. One cannot read it without being impressed with the highly progressive attitude of the DBS as shown by their willingness to adopt major innovations. This willingness reflects the continuation of an imaginative approach which has characterized the work of the DBS. During the past decade they have taken such steps as the introduction of document sensing for a major census, the decentralization of processing work, a major re-organization of field structure, and the transition from conventional equipment to electronic equipment. In every case the changes introduced have been skillfully adapted to meet their needs, and the effectiveness of the advance planning to insure orderly operations under a new system has been truly remarkable.

An example of this is provided by the dramatic transition to electronic equipment in 1961 described in so matter-of-fact a fashion by Mr. Lemieux. This transition included the full range of applications of such equipment to data processing, including document reading, editing, tabulating, and high-speed printing. It was a very formidable task for any statistical organization to undertake such a transition in advance of a decennial census and it is clear that the extremely competent and careful planning of the DBS personnel has made possible marked improvements in the timeliness of census reports in Canada and presumably will benefit the quality also.

There is one difference between the two countries which I find particularly interesting. In our 1960 census perhaps 80 percent of our population questions were a part of the 25 percent sample schedule, whereas the Canadians put fewer than half of their 1961 questions on a sample basis. Can the difference be explained by variations in the statistical requirements of the two countries, or are sample statistics less acceptable north of the border?

Mr. Lemieux's remarks about the maximum tolerable limit of the population schedule seems to imply greater confidence than we would have that the 1961 Canadian Census was close to the point of overloading. It would be interesting to know whether the statement was based primarily on a priori judgment or whether there has been some experimental work in measurement of public cooperation on schedules of varying length.

One of the most interesting and at the same time troublesome problems that the Canadians faced in 1961 was the question of "ethnic origin". Here the Canadians have a special problem which we are happy not to have shared. On both sides of the border statisticians would agree that in countries like ours the measurement of "ethnic origin" as a part of a census is very rough, at best. Responses are very uncertain for the older stock of the country, who long ago may have lost all identification with the ethnic group of the original male immigrant. Nevertheless, the very strong feelings of the Canadian-French and other groups make it extremely difficult to settle the matter solely on the basis of statistical considerations. The strength of the feelings involved is indicated by the fact that the DBS staff was forced to reprint the schedules at a time when such a task was certainly sure to jeopardize the Census. We can understand the intensity of feeling in Canada, for we encountered a somewhat similar campaign in the United States. The dispute over a proposed question on religion at one stage threatened to affect our census planning substantially and it is quite possible that had we not suspended further consideration of this question the excitement here would have been as great as that which took place in Canada over ethnic origin. This might be hard for them to understand, in view of the fact that they have been asking a question on religion for many decades and doubtless would encounter serious protests if they tried to drop the question.

It is appropriate to pay particular tribute to our Canadian friends for the very important change in field organization introduced as a part of their 1961 Census. Mr. Lemieux described this simply as removing one level of supervision. Essentially, this change meant that the eight regional offices recruited a staff of specialists who were directly in touch with the commissioner, the first line supervisors of the entire body of enumerators. This step is similar to one which we have adopted for our current field operations in the Bureau of the Census, but have never undertaken for a decennial census. Headquarters had to arrange for the appointment of well over a thousand census commissioners, but this seems to have worked out quite well and would work out even better in the future if more time were allowed. We believe that this step that the Canadians have introduced represents a real gain in field organization and shall certainly give it very careful study for future censuses in this country.

I would also like to note with approval the Canadian use of a postal check in order to improve the coverage of the population census. We had hoped to use a similar device in this country as a part of our 1960 census, but because of the budgetary restrictions under which we operated, it was impossible to add this improvement without jeopardizing the continuation of some of the basic features of the census. The number added by the postal check (about 2/10ths of one percent of the population) was undoubtedly less than had been hoped for on the basis of their field test in Galt, Ontario. Indeed, it may have been low enough to raise some questions about the quality of the post-office check. We wonder if the operation included any built-in safeguards to measure the thoroughness of the postal check. Nevertheless, the cost per unit for the persons added was somewhat less than one dollar. Since any means of catching the last percent or two of the population must involve disproportionately high costs, we conclude that this was a profitable outlay on the part of the Canadians. It was, in addition, a good public relations device to help deal with claims of undercounting. We hope to make full use of this experience as we plan for our next census, regardless of whether we follow the Canadian 5-year plan or the traditional United States 10-year plan.

Finally, the pattern of population change in urban areas in Canada seems to be similar to that in the United States. In both countries a disproportionately large percentage of the national growth between 1950 and 1960 is concentrated in metropolitan areas. In these areas the central cities have grown much less rapidly than the peripheral areas. The 1951-61 decline of population in Toronto reminds one of similar developments in a number of our large cities and we are sure that the earlier publication of data regarding declines in many of our central cities must have considerably reduced the number of complaints that otherwise would have come from Toronto officials and others.

VI

TRENDS AS REVEALED BY THE 1960 POPULATION CENSUS OF THE UNITED STATES

Chairman, Dorothy S. Thomas, University of Pennsylvania

Working Paper:

Conrad Taeuber, Howard G. Brunsman and Henry S. Shryock, Jr., Bureau of the Census

Discussion - Talcott Parsons, Harvard University

TRENDS AS REVEALED BY THE 1960 POPULATION CENSUS OF THE UNITED STATES

By Conrad Taeuber, Howard G. Brunsman, and Henry S. Shryock, Jr.
Bureau of the Census

Introduction

The attached set of tables and the map contain some of the more significant and novel figures now available from the 1960 Population Census of the United States. These materials are intended to serve as the basis of the discussion by the panel of social science analyists of the broad implications of recent population trends. The materials are accompanied by only very brief annotations, since it is believed they will tell their own story to statisticians and social scientists.

The three listed authors wish to state that these materials are a staff product of the Bureau of the Census and to acknowledge the contribution made by many staff members of the Population and Geography Divisions in the preparation and compilation of the materials.

Item A. NET MIGRATION, BY COLOR, FOR STATES: 1950 TO 1960

These estimates, made by the vital statistics variant of the residual method, include 2.7 million net immigrants from abroad so that a State is more likely to have a net in-migration than a net out-migration. Nonetheless, 27 States and the District of Columbia incurred a net out-migration over the 1950-1960 decade. In about one-third of the States--and especially in the North, the direction (+ or -) of the net migration was different for whites and nonwhites. The geographic redistribution of whites, Negroes, Indians, and Orientals is evidently continuing at a rapid pace with shifts mostly resembling those in the preceding decade. The South as a whole had a small net in-migration of whites, although this was mainly confined to some of its more atypical States. It would be interesting to determine the extent to which the likely reservoirs of continued out-migration are becoming depleted.

Item B. NET MIGRATION ESTIMATES, 1950 TO 1960, BY COUNTY

Almost half (49 percent) of the 3,000 or so counties in the United States lost population in the 1950's. All of these lost because of net out-migration. In addition, many of the gaining counties increased only because their natural increase more than offset their net out-migration. Of all counties, 29 percent had this type of growth so that, in total, more than three-quarters of all counties had a net out-migration. Obviously, then some counties had rather high rates of net in-migration. In fact, 172 counties (5.5 percent of the grand total) had a net in-migration rate of 30 percent or more.

As shown by the map, the counties with very high rates of net in-migration are mostly outlying counties within metropolitan areas, a few relatively "young" metropolitan areas, and counties in Florida and California. The counties with very high rates of net out-migration, on the other hand, are typically those in depressed areas from which population has been draining away for decades: coal mining areas in West Virginia, Virginia, and Kentucky; the Southern Appalachians and the Ozarks; parts of the Deep South (including northern Florida); the North Woods Cutover; and the low rainfall areas of the Great Plains.

Item C. NET MIGRATION AND NET MIGRATION RATES FOR METROPOLITAN AND NON-METROPOLITAN STATE ECONOMIC AREAS, BY REGIONS: 1950 TO 1960

These two tables are consolidations of the county estimates shown on the map of Item B. The shift of population from non-metropolitan to metropolitan areas is certainly familar by now. Bogue's 1940-1950 estimated net migration rates (+9.2 percent for metropolitan areas and -9.3 percent for nonmetropolitan areas) are not much different from these. Although not shown in the table, it was also found that the net migration rate for nonmetropolitan counties in every region tended to vary directly with the percent of the county's population classified as urban, so that the counties with no urban population had the highest rates of net out-migration. In the past decade, nonmetropolitan areas in the Northeast and the West, however, had but little net migration, although the stream components of the two equilibria may represent quite different situations. The contrast between metropolitan and nonmetropolitan areas, with respect to net migration, is perhaps greatest in the South.

The designated depressed areas in combination had a net out-migration, as we should expect. Those in the North Central States and the West, however, had virtually a zero balance.

Item D. POPULATION INSIDE AND OUTSIDE CENTRAL CITIES OF STANDARD METROPOLITAN STATISTICAL AREAS WITH POPULATION OF AREAS ANNEXED TO CENTRAL CITIES, BY REGIONS: 1960 AND 1950

The new statistics on the population of territory annexed to cities have increased our awareness of the importance of this factor in the growth of incorporated places. The central cities of metropolitan areas gained 4.9 million persons through annexation during the 1950's, an amount equal to 9 percent of their 1950 population. Had it not been for this legal process, many more large cities would have shown population decreases over the decade and the centrifugal movement from central cities to suburbs would have appeared to be even more extensive. Southern cities made particularly effective use of this device to recapture their suburbanites whereas Northern cities were unable to do so to any appreciable extent. The selective contribution of annexations to the population growth of central cities also tended to vary inversely with the over-all size of the metropolitan area. The growth of the entire metropolitan area seemed to bear little relationship to its over-all size but was considerably influenced by geographic location.

Item E. COLOR BY SEX, FOR THE UNITED STATES, BY SIZE OF PLACE, 1960, AND URBAN AND RURAL RESIDENCE, 1950

The geographic redistribution of nonwhites (principally Negroes) has been accompanied by a redistribution in terms of size of place of residence that is just as striking. Less than 50 years ago, a majority of Negroes lived on farms. In 1960 about half lived in the central cities of urbanized areas and about 70 percent lived in urban residences. By several residential indices, the Negro is now more urbanized than the white population. The chief remaining lag is in suburbanization; urban fringes house 23 percent of whites but only 8 percent of nonwhites. This disparity, of course, reflects not only residential segregation but also sharp differences in the quality of housing.

Item F. POPULATION INSIDE AND OUTSIDE CENTRAL CITY OR CITIES OF STANDARD METROPOLITAN STATISTICAL AREAS, BY COLOR AND REGIONS: 1960 AND 1950

The white population of central cities of SMSA's increased nationwide by only 5 percent whereas the corresponding nonwhite population increased by 51 percent. This type of difference was found in the metropolitan areas of every region but was perhaps most striking in the two Northern regions because whites in their central cities actually decreased in numbers. There was also a fairly large percentage increase of nonwhites in the outlying parts of metropolitan areas but the absolute increase involved was quite small.

The influx of Negroes in the large cities of the North and West may lead us to overlook the similar movement in the South. The 1950-1960 rate of growth of the nonwhite population in the central cities of Southern SMSA's was higher than that of Southern whites even though the latter group had a rapid increase of its own.

Item G. POPULATION INSIDE AND OUTSIDE CENTRAL CITY OR CITIES OF STANDARD METROPOLITAN STATISTICAL AREAS, BY COLOR AND SIZE OF AREA: 1960 AND 1950

The excess in the rate of population change of nonwhites over that of whites in the central cities of Standard Metropolitan Statistical Areas was greater in SMSA's with a million or more inhabitants than in smaller SMSA's, partly because the larger SMSA's are not located in the Deep South. The direction of the differences was the same in every size class, however, just as it was in every region. In evaluating the relatively low growth rates of both whites and nonwhites in the outlying parts of SMSA's with fewer than 250,000 inhabitants, we have to make allowance for the heavy losses from annexations there as was shown in Table 2 of Item D.

Item H. POPULATION OF CITIES OF 1,000,000 OR MORE BY AGE, 1960 AND 1950, WITH DECADE CHANGE BY COLOR

As previously noted, a number of our largest cities lost whites by out-migration in the 1950's and had in-migration of nonwhites of a less than compensating volume. These processes had considerable effect on the population structure of big cities by age and color.

In the five cities of a million inhabitants or more, the only age group that everywhere decreased over the decade was 25 to 29--the slim birth cohort of 1930 to 1934. Except for Los Angeles, the losing age group extended beyond this one in both directions, to as low as 15 to 19 and as high as 60 to 64. Among whites alone, even more age groups incurred decennial losses; in fact, only the age groups 10 to 14 years old and 65 years old and over increased in all big-city white populations. The picture was quite different for nonwhites. Only in Detroit, did any age groups (20 to 29) of nonwhites lose population. One result of this pattern of ethnic succession was that the white population is considerably older than the nonwhite in these cities, containing relatively more of the elderly and relatively fewer of children of school age.

Item I. AGE OF THE POPULATION INSIDE AND OUTSIDE STANDARD METROPOLITAN STATISTICAL AREAS: 1960 AND 1950

When we widen the scene from the five largest cities to the total of all SMSA's, we observe decreases at only the age groups 20 to 29 with the largest rates of increase having occurred at 5 to 14 and 65 years and over. The population living outside metropolitan areas had an (algebraically) lower rate of change in every age group than the metropolitan population. The losses in the ages of prime labor-force activity were thus both more widespread and more acute. The residential differentials reflect mainly net migration from nonmetropolitan to metropolitan areas, but underlying both sets of rates of change are the effects of important fluctuations in the size of birth cohorts.

Item J. YEAR OF BIRTH AND AGE OF THE POPULATION INSIDE AND OUTSIDE STANDARD METROPOLITAN STATISTICAL AREAS: 1960 AND 1950

This table makes it easier to compare the size of identical birth cohorts in 1950 and 1960. In metropolitan areas, for cohorts born before 1915, net in-migration was not sufficient to compensate for mortality. It appears that the peak rate of net in-migration occured for the cohort born between 1930 and 1935, i.e., for those 15 to 19 years old in 1950. (A more thorough analysis would take account of survival rates and of differences in coverage and the accuracy of age reporting, at the two censuses.) Disregarding children born during the intercensal decade, we note that, in nonmetropolitan areas, every cohort was reduced in size. The rate of change was also algebraically less for every cohort in nonmetropolitan than in metropolitan areas, although the differences were slight for persons born before 1900. The elderly are least likely to migrate to metropolitan areas, or to migrate at all, for that matter.

An examination of the birth cohorts for the five cities presented in Item H reveals a reduction between 1950 and 1960 of every cohort in the cities of Detroit and Philadelphia. The only increases in the cities of New York and Chicago occur in the cohorts born between 1930 and 1940 (those persons between 10 and 19 years old in 1950). In Los Angeles, increases occurred in all cohorts born in 1915 or later (those under 35 years old in 1950) whereas decreases occurred in all older cohorts.

Item K. COMPARISON OF NUMBER OF PERSONS OF ALL AGES, 45 TO 64, AND 65 AND OVER, APRIL 1, 1950 AND 1960, BY STATES

This item shows the percent of change, by State, for the elderly (65 years and over) and the middle-aged (45 to 64). The number of elderly persons increased from 1950 to 1960 in every State. These increases reflect mainly increases in numbers of births in the nineteenth century with improvements in expectation of life playing a relatively minor role. The presumed out-migration of the elderly even from States like West Virginia and Mississippi was not nearly enough to cancel this increase. On the other hand, the States with mild winter climates were able to attract an extremely disproportionate share of elderly interstate migrants. Thus, Florida and Arizona more than doubled their elderly populations. All States except one (South Dakota) also had more residents 45 to 64 years old at the end of the decade than at the beginning. Almost universal State gains of elderly persons are probably in prospect for the 1970's, therefore.

Item L. PERCENT OF STATE'S POPULATION IN EACH SIZE OF AREA OF RESIDENCE CLASS WHICH IS AGED 65 AND OVER, APRIL 1, 1960, BY STATE

Persons 65 years old and over tend to be most concentrated in villages and least concentrated in the suburbs, although this generalization does not hold for all parts of the country. The heavy concentration of the elderly in villages is most typical of Kansas, Nebraska, Oklahoma, and Missouri, where the proportion of the total population runs up to one-fifth. Florida cannot equal this level in any type-of-residence area although particular towns and cities may be more specialized havens for the aged. California, contrary to popular impression, has a lower proportion of elderly residents than the national average.

Item M. WORKERS BY PLACE OF WORK AND PLACE OF RESIDENCE, DENVER SMSA: 1960

This table illustrates the type of detail on place of work by place of residence now being published in the census tract reports. In fact, those reports also show the central cities separately. The flows of commuters into Denver from the other counties in the SMSA are impressively large except that, in hindsight, there is certainly a question as to whether Boulder County really belongs in the Denver SMSA. Only small proportions of Denver residents work in the outlying counties.

Item N. PLACE OF WORK OF WORKERS FOR SELECTED STANDARD METROPOLITAN STATISTICAL AREAS: 1960

This table gives commuting statistics for a broader geographic coverage but in reduced detail on the streams. These selected figures also come from the census tract reports, which are now being published. This source covers commuting streams within an SMSA or out from an SMSA but not into an SMSA from other counties. The cases where more than one-tenth of the workers commute to jobs outside the SMSA are limited to New England, where SMSA's are defined in terms of towns and hence have much smaller areas. Presumably many of the long-distance commuters from the outlying "rings" of Stamford and Norwalk, Connecticut, work in New York City. It is not at all uncommon for more than half of the workers living in the outlying metropolitan ring to work in the central city.

Item O. HOUSEHOLDS, PRIMARY FAMILIES, AND PRIMARY INDIVIDUALS, BY REGIONS, URBAN-RURAL RESIDENCE, AND COLOR, FOR THE UNITED STATES: 1960 AND 1950

Some people have been confused by the apparent paradox that the average size of household has been decreasing whereas the the average size of family has increased slightly. Here one needs to know the official definitions. Families require a minimum of two related persons. About 16 percent of households, on the other hand, consist of primary individuals, mostly living alone. The increase in family size is essentially attributable to the baby boom. The decline in average size of household reflects the facts that, in our affluent society more old people and more young adults entering the labor force can afford to live alone and that custom increasingly limits the group living under one roof to the nuclear family (husband, wife, and minor children).

The convergence of average size of both households and families continued among the regions and urban-rural groups but not between white and nonwhite groups.

Item P. NUMBER OF CHILDREN EVER BORN PER 1,000 WOMEN 35 TO 44 YEARS OLD, 1960, AND PERCENT CHANGE, 1950 TO 1960, BY COLOR AND URBAN-RURAL RESIDENCE, FOR AVAILABLE STATES

Available State figures on the number of children ever born suggest that completed fertility is increasing everywhere except among white women in the rural South. Urban areas and the nonwhite population show the largest percentage increases, 1950 to 1960, for women 35 to 44 years old. The national replacement quota of roughly 2,030 children per 1,000 women of this age is exceeded almost universally.

In a number of States, the fertility of rural white women exceeds that of urban nonwhite women. Before we can fully assess the effects of the urban environment on Negro fertility, however, we need to know more about the selectivity of rural-urban migration. The very low fertility rate for white women in the District of Columbia reflects the selectivity of both in- and out-migration as well as the relatively low fertility in big cities.

When both urban-rural residence and color are held constant, Virginia, the Carolinas, and Georgia are seen to have relatively moderate cumulative fertility rates, probably below the corresponding national averages. The population "mix," however, produces relatively high over-all rates for these South Atlantic States.

ITEM A

U. S. DEPARTMENT OF COMMERCE
BUREAU OF THE CENSUS
WASHINGTON 25, D. C.

December 4, 1961

NET MIGRATION, BY COLOR, FOR STATES: 1950 to 1960

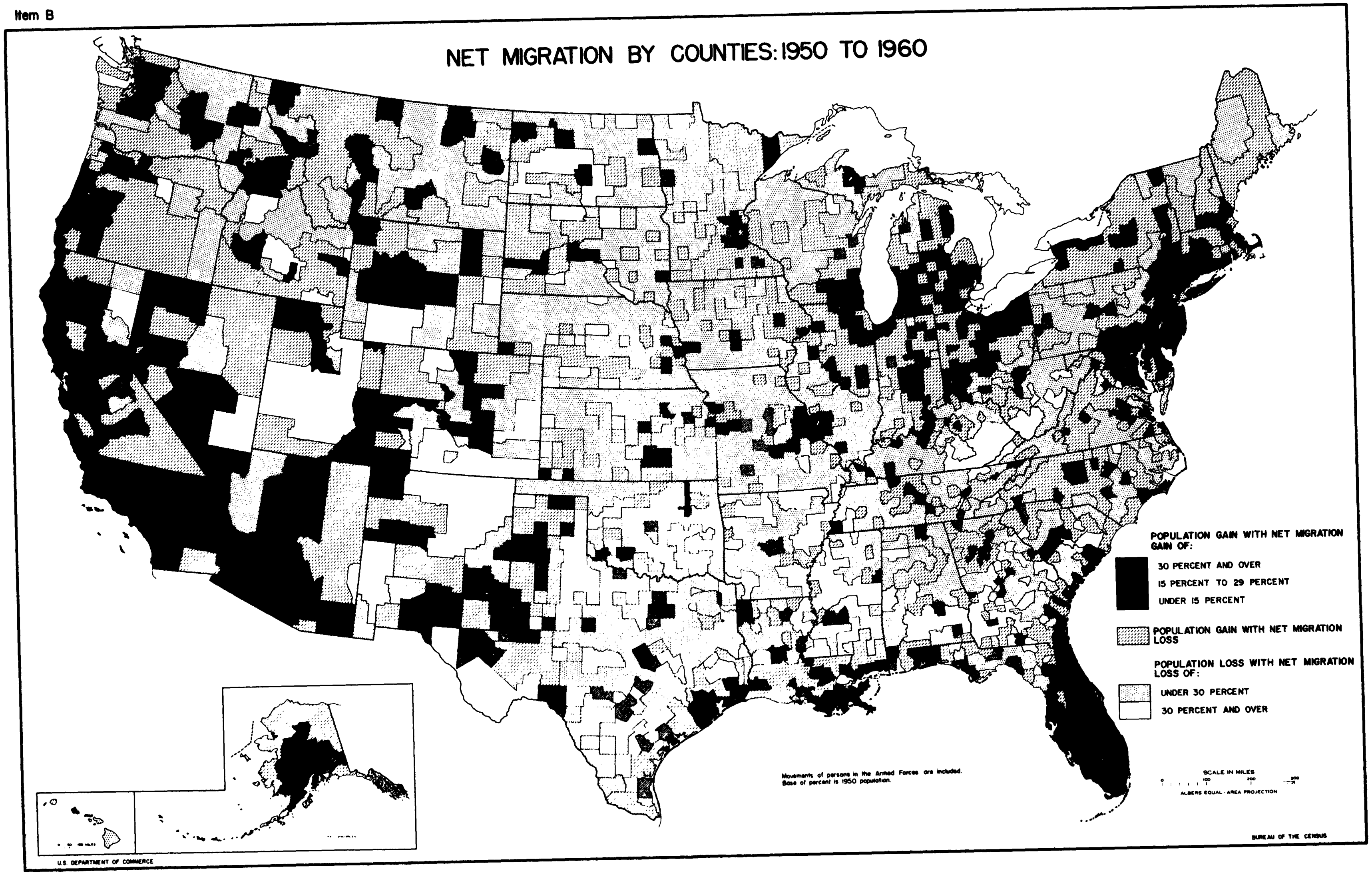
(Net migration comprises both net immigration from abroad and net interregional, interdivisional, and interstate migration according to the area shown. Movements of persons in the Armed Forces are included. Each estimate has been independently rounded from figures computed to the last digit; hence, the sums of parts shown may differ slightly from the totals shown)

Region, Division, and State	Total	White	Nonwhite
United States	+2,660,000	+2,685,000	- 25,000
Regions:			
Northeastern	+ 336,000	- 206,000	+ 541,000
North Central States	- 121,000	- 679,000	+ 558,000
The South	-1,404,000	+ 52,000	-1,457,000
The West	+3,850,000	+3,518,000	+ 332,000
Northeastern States:			
New England	+ 23,000	- 47,000	+ 70,000
Middle Atlantic	+ 312,000	- 159,000	+ 472,000
North Central States:			
East North Central	+ 699,000	+ 178,000	+ 521,000
West North Central	- 820,000	- 857,000	+ 37,000
The South:			
South Atlantic	+ 647,000	+1,189,000	- 542,000
East South Central	-1,464,000	- 845,000	- 620,000
West South Central	- 587,000	- 292,000	- 295,000
The West:			
Mountain	+ 557,000	+ 549,000	+ 8,000
Pacific	+3,293,000	+2,970,000	+ 324,000
New England:			
Maine	- 66,000	- 68,000	+ 2,000
New Hampshire	+ 13,000	+ 11,000	+ 1,000
Vermont	- 38,000	- 38,000	(1)
Massachusetts	- 93,000	- 119,000	+ 25,000
Rhode Island	- 26,000	- 28,000	+ 2,000
Connecticut	+ 234,000	+ 195,000	+ 39,000
Middle Atlantic:			
New York	+ 210,000	- 72,000	+ 282,000
New Jersey	+ 577,000	+ 465,000	+ 112,000
Pennsylvania	- 475,000	- 553,000	+ 77,000

Region, Division, and State	Total	White	Nonwhite
East North Central:			
Ohio	+ 409,000	+ 276,000	+ 133,000
Indiana	+ 63,000	+ 19,000	+ 45,000
Illinois	+ 124,000	- 64,000	+ 189,000
Michigan	+ 156,000	+ 30,000	+ 127,000
Wisconsin	- 53,000	- 82,000	+ 29,000
West North Central:			
Minnesota	- 97,000	- 101,000	+ 4,000
Iowa	- 233,000	- 236,000	+ 3,000
Missouri	- 130,000	- 158,000	+ 28,000
North Dakota	- 105,000	- 103,000	- 2,000
South Dakota	- 94,000	- 90,000	- 5,000
Nebraska	- 117,000	- 121,000	+ 4,000
Kansas	- 44,000	- 49,000	+ 5,000
South Atlantic:			
Delaware	+ 64,000	+ 58,000	+ 6,000
Maryland	+ 320,000	+ 284,000	+ 36,000
Dist. of Columbia	- 158,000	- 213,000	+ 54,000
Virginia	+ 15,000	+ 84,000	- 70,000
West Virginia	- 447,000	- 406,000	- 40,000
North Carolina	- 328,000	- 121,000	- 207,000
South Carolina	- 222,000	- 4,000	- 218,000
Georgia	- 214,000	- 9,000	- 204,000
Florida	+1,617,000	+1,516,000	+ 101,000
East South Central:			
Kentucky	- 390,000	- 374,000	- 15,000
Tennessee	- 273,000	- 216,000	- 57,000
Alabama	- 368,000	- 144,000	- 224,000
Mississippi	- 434,000	- 110,000	- 323,000
West South Central:			
Arkansas	- 433,000	- 283,000	- 150,000
Louisiana	- 50,000	+ 42,000	- 92,000
Oklahoma	- 219,000	- 192,000	- 26,000
Texas	+ 114,000	+ 141,000	- 27,000
Mountain:			
Montana	- 25,000	- 23,000	- 2,000
Idaho	- 40,000	- 41,000	+ 1,000
Wyoming	- 20,000	- 19,000	- 1,000
Colorado	+ 164,000	+ 149,000	+ 15,000
New Mexico	+ 52,000	+ 54,000	- 1,000
Arizona	+ 330,000	+ 340,000	- 10,000
Utah	+ 10,000	+ 9,000	+ 1,000
Nevada	+ 86,000	+ 80,000	+ 6,000
Region, Division, and State			
Pacific:			
Washington	+ 88,000	+ 70,000	+ 18,000
Oregon	+ 16,000	+ 10,000	+ 6,000
California	+3,145,000	+2,791,000	+ 354,000
Alaska	+ 41,000	+ 42,000	- 1,000
Hawaii	+ 3,000	+ 55,000	- 52,000

1/ Less than 500

Item B


NET MIGRATION BY COUNTIES: 1950 TO 1960
POPULATION GAIN WITH NET MIGRATION GAIN OF:
30 PERCENT AND OVER
15 PERCENT TO 29 PERCENT
UNDER 15 PERCENT
POPULATION GAIN WITH NET MIGRATION LOSS
POPULATION LOSS WITH NET MIGRATION LOSS OF:
UNDER 30 PERCENT
30 PERCENT AND OVER
Movements of persons in the Armed Forces are included.
Base of percent is 1950 population.
SCALE IN MILES
ALBERS EQUAL-AREA PROJECTION
BUREAU OF THE CENSUS
U.S. DEPARTMENT OF COMMERCE

U. S. DEPARTMENT OF COMMERCE
BUREAU OF THE CENSUS

Item C

Part 1.--NET MIGRATION AND NET MIGRATION RATES FOR METROPOLITAN AND NONMETROPOLITAN STATE ECONOMIC AREAS, BY REGIONS: 1950 TO 1960

(Net migration comprises both net immigration from abroad and net migration between the areas shown. Movements of persons in the Armed Forces are included)

Region	Net migration, 1950-60[1]	Net migration rate, 1950-60[2]
UNITED STATES	+2,660,000	+1.8
Metropolitan SEA's	+8,149,000	+9.3
Nonmetropolitan SEA's	-5,490,000	-8.6
NORTHEAST	+336,000	+0.9
Metropolitan SEA's	+308,000	+1.0
Nonmetropolitan SEA's	+27,000	+0.4
NORTH CENTRAL	-121,000	-0.3
Metropolitan SEA's	+1,280,000	+5.3
Nonmetropolitan SEA's	-1,402,000	-7.0
SOUTH	-1,404,000	-3.0
Metropolitan SEA's	+2,726,000	+14.9
Nonmetropolitan SEA's	-4,131,000	-14.3
WEST	+3,850,000	+19.1
Metropolitan SEA's	+3,834,000	+28.8
Nonmetropolitan SEA's	+16,000	+0.2

[1]Each estimate has been independently rounded from figures computed to the last digit; hence the sums of parts shown may differ slightly from the totals shown.

[2]Base is 1950 population.

Part 2.--NET MIGRATION AND NET MIGRATION RATES FOR REDEVELOPMENT AREAS, BY REGIONS: 1950 TO 1960

(Data relate to the 212 counties at least half of whose population falls within the 124 Redevelopment Areas designated in accordance with Section 5(a) of the Area Redevelopment Act, as listed in Area Designation Status Report 1, Excludes Puerto Rico. See also notes in Item C, Part 1)

Region	Net migration, 1950-60	Net migration rate, 1950-60
United States	-1,301,000	-7.8
Northeast	-461,000	-6.7
North Central	-41,000	-0.7
South	-804,000	-21.5
West	+5,000	+0.8

Item D

U. S. DEPARTMENT OF COMMERCE
Bureau of the Census
Washington 25, D. C.

Table 1.--POPULATION INSIDE AND OUTSIDE CENTRAL CITY OR CITIES OF STANDARD METROPOLITAN STATISTICAL AREAS WITH POPULATION OF AREAS ANNEXED TO CENTRAL CITIES, BY REGIONS: 1960 AND 1950

Region and component parts of SMSA	1960	1950	Change, 1950 to 1960: Total		Based on 1950 limits of central cities		From annexations		1960 population on basis of 1950 limits of central cities
			Number	Per-cent	Number	Per-cent	Number	Per-cent	
UNITED STATES									
In SMSA's	112,885,178	89,316,903	23,568,275	26.4	23,568,275	26.4	-	-	112,885,178
Central cities	58,004,334	52,371,379	5,632,955	10.8	781,472	1.5	4,851,483	9.3	53,152,851
Outside central cities	54,880,844	36,945,524	17,935,320	48.5	22,786,803	61.7	-4,851,483	-13.1	59,732,327
NORTHEAST									
In SMSA's	35,346,505	31,267,169	4,079,336	13.0	4,079,336	13.0	-	-	35,346,505
Central cities	17,321,731	17,881,490	-559,759	-3.1	-579,874	-3.2	20,115	0.1	17,301,616
Outside central cities	18,024,774	13,385,679	4,639,095	34.7	4,659,210	34.8	-20,115	-0.2	18,044,889
NORTH CENTRAL									
In SMSA's	30,959,961	25,074,674	5,885,287	23.5	5,885,287	23.5	-	-	30,959,961
Central cities	16,510,746	15,836,656	674,090	4.3	-257,583	-1.6	931,673	5.9	15,579,073
Outside central cities	14,449,215	9,238,018	5,211,197	56.4	6,142,870	66.5	-931,673	-10.1	15,380,888
SOUTH									
In SMSA's	26,447,395	19,417,751	7,029,644	36.2	7,029,644	36.2	-	-	26,447,395
Central cities	15,061,777	11,720,837	3,340,940	28.5	615,807	5.3	2,725,133	23.3	12,336,644
Outside central cities	11,385,618	7,696,914	3,688,704	47.9	6,413,837	83.3	-2,725,133	-35.4	14,110,751
WEST									
In SMSA's	20,131,317	13,557,309	6,574,008	48.5	6,574,008	48.5	-	-	20,131,317
Central cities	9,110,080	6,932,396	2,177,684	31.4	1,003,122	14.5	1,174,562	16.9	7,935,518
Outside central cities	11,021,237	6,624,913	4,396,324	66.4	5,570,886	84.1	-1,174,562	-17.7	12,195,799

Item D

U. S. DEPARTMENT OF COMMERCE
Bureau of the Census
Washington 25, D. C.

Table 2.--POPULATION INSIDE AND OUTSIDE CENTRAL CITY OR CITIES OF STANDARD METROPOLITAN STATISTICAL AREAS WITH POPULATION OF AREAS ANNEXED TO CENTRAL CITIES, BY SIZE OF AREA: 1960 AND 1950

Size and component parts of SMSA	1960	1950	Change, 1950 to 1960						1960 population on basis of 1950 limits of central cities
			Total		Based on 1950 limits of central cities		From annexations		
			Number	Per-cent	Number	Per-cent	Number	Per-cent	
ALL SIZES									
In SMSA's	112,885,178	89,316,903	23,568,275	26.4	23,568,275	26.4	-	-	112,885,178
Central cities	58,004,334	52,371,379	5,632,955	10.8	781,472	1.5	4,851,483	9.3	53,152,851
Outside central cities	54,880,844	36,945,524	17,935,320	48.5	22,786,803	61.7	-4,851,483	-13.1	59,732,327
3,000,000 or more									
In SMSA's	31,763,499	25,788,967	5,974,532	23.2	5,974,532	23.2	-	-	31,763,499
Central cities	17,828,227	17,655,217	173,010	1.0	99,318	0.6	73,692	0.4	17,754,535
Outside central cities	13,935,272	8,133,750	5,801,522	71.3	5,875,214	72.2	-73,692	-0.9	14,008,964
1,000,000 to 3,000,000									
In SMSA's	29,818,571	23,858,113	5,960,458	25.0	5,960,458	25.0	-	-	29,818,571
Central cities	12,707,503	12,037,125	670,378	5.6	-270,275	-2.2	940,653	7.8	11,766,850
Outside central cities	17,111,068	11,820,988	5,290,080	44.8	6,230,733	52.7	-940,653	-8.0	18,051,721
500,000 to 1,000,000									
In SMSA's	19,214,817	14,125,628	5,089,189	36.0	5,089,189	36.0	-	-	19,214,817
Central cities	10,126,684	8,340,585	1,786,099	21.4	396,636	4.8	1,389,463	16.7	8,737,221
Outside central cities	9,088,133	5,785,043	3,303,090	57.1	4,692,553	81.1	-1,389,463	-24.0	10,477,596
250,000 to 500,000									
In SMSA's	15,829,067	12,603,137	3,225,930	25.6	3,225,930	25.6	-	-	15,829,067
Central cities	7,750,597	6,671,381	1,079,216	16.2	146,234	2.2	932,982	14.0	6,817,615
Outside central cities	8,078,470	5,931,756	2,146,714	36.2	3,079,696	51.9	-932,982	-15.7	9,011,452
100,000 to 250,000									
In SMSA's	14,497,817	11,525,685	2,972,132	25.8	2,972,132	25.8	-	-	14,497,817
Central cities	8,235,553	6,617,683	1,617,870	24.4	319,286	4.8	1,298,584	19.6	6,936,969
Outside central cities	6,262,264	4,908,002	1,354,262	27.6	2,652,846	54.1	-1,298,584	-26.5	7,560,848
Under 100,000									
In SMSA's	1,761,407	1,415,373	346,034	24.4	346,034	24.4	-	-	1,761,407
Central cities	1,355,770	1,049,388	306,382	29.2	90,273	8.6	216,109	20.1	1,139,661
Outside central cities	405,637	365,985	39,652	10.8	255,761	69.9	-216,109	-59.6	621,746

ITEM E

General Population Characteristics

Table 42.—COLOR BY SEX, FOR THE UNITED STATES, BY SIZE OF PLACE, 1960, AND URBAN AND RURAL RESIDENCE, 1950

[Minus sign (–) denotes decrease. Percent not shown where less than 0.1]

Area and census year	All classes			White			Nonwhite			Percent distribution				
										All classes			White	Non-white
	Total	Male	Female	Total	Male	Female	Total	Male	Female	Total	Male	Fe-male		
UNITED STATES														
1960														
Total	179,323,175	88,331,494	90,991,681	158,831,732	78,367,149	80,464,583	20,491,443	9,964,345	10,527,098	100.0	100.0	100.0	100.0	100.0
Urban	125,268,750	60,733,005	64,535,745	110,428,332	53,631,145	56,797,187	14,840,418	7,101,860	7,738,558	69.9	68.8	70.9	69.5	72.4
Urbanized areas	95,848,487	46,494,210	49,354,277	83,769,935	40,706,094	43,063,841	12,078,552	5,788,116	6,290,436	53.5	52.6	54.2	52.7	58.9
Central cities	57,975,132	27,927,624	30,047,508	47,627,232	22,976,282	24,650,950	10,347,900	4,951,342	5,396,558	32.3	31.6	33.0	30.0	50.5
Urban fringe	37,873,355	18,566,586	19,306,769	36,142,703	17,729,812	18,412,891	1,730,652	836,774	893,878	21.1	21.0	21.2	22.8	8.4
Other urban	29,420,263	14,238,795	15,181,468	26,658,397	12,925,051	13,733,346	2,761,866	1,313,744	1,448,122	16.4	16.1	16.7	16.8	13.5
Places of 10,000 or more	16,172,839	7,838,676	8,334,163	14,561,214	7,070,615	7,490,599	1,611,625	768,061	843,564	9.0	8.9	9.2	9.2	7.9
Places of 2,500 to 10,000	13,247,424	6,400,119	6,847,305	12,097,183	5,854,436	6,242,747	1,150,241	545,683	604,558	7.4	7.2	7.5	7.6	5.6
Rural	54,054,425	27,598,489	26,455,936	48,403,400	24,736,004	23,667,396	5,651,025	2,862,485	2,788,540	30.1	31.2	29.1	30.5	27.6
Places of 1,000 to 2,500	6,496,788	3,149,869	3,346,919	5,995,754	2,909,209	3,086,545	501,034	240,660	260,374	3.6	3.6	3.7	3.8	2.4
Other rural	47,557,637	24,448,620	23,109,017	42,407,646	21,826,795	20,580,851	5,149,991	2,621,825	2,528,166	26.5	27.7	25.4	26.7	25.1
1950														
Total	151,325,798	75,186,606	76,139,192	135,149,629	67,254,991	67,894,638	16,176,169	7,931,615	8,244,554	100.0	100.0	100.0	100.0	100.0
Urban	96,846,817	47,092,839	49,753,978	86,864,031	42,307,596	44,556,435	9,982,786	4,785,243	5,197,543	64.0	62.6	65.3	64.3	61.7
Urbanized areas	[1]69,249,148	33,670,714	35,578,434	61,925,036	30,160,082	31,764,954	7,324,112	3,510,632	3,813,480	45.8	44.8	46.7	45.8	45.3
Central cities	48,377,240	23,432,038	24,945,202	42,041,968	20,402,408	21,639,560	6,335,272	3,029,630	3,305,642	32.0	31.2	32.8	31.1	39.2
Urban fringe	[1]20,871,908	10,238,676	10,633,232	19,883,068	9,757,674	10,125,394	988,840	481,002	507,838	13.8	13.6	14.0	14.7	6.1
Other urban	[1]27,597,669	13,422,125	14,175,544	24,938,995	12,147,514	12,791,481	2,658,674	1,274,611	1,384,063	18.2	17.9	18.6	18.5	16.4
Rural	54,478,981	28,093,767	26,385,214	48,285,598	24,947,395	23,338,203	6,193,383	3,146,372	3,047,011	36.0	37.4	34.7	35.7	38.3
Percent Increase, 1950 to 1960														
Total	18.5	17.5	19.5	17.5	16.5	18.5	26.7	25.6	27.7	...	...	...	...	...
Urban	29.3	29.0	29.7	27.1	26.8	27.5	48.7	48.4	48.9	...	...	...	...	...
Urbanized areas	38.4	38.1	38.7	35.3	35.0	35.6	64.9	64.9	65.0	...	...	...	...	...
Central cities	19.8	19.2	20.5	13.3	12.6	13.9	63.3	63.4	63.3	...	...	...	...	...
Urban fringe	81.5	81.3	81.6	81.8	81.7	81.8	75.0	74.0	76.0	...	...	...	...	...
Other urban	6.6	6.1	7.1	6.9	6.4	7.4	3.9	3.1	4.6	...	...	...	...	...
Rural	-0.8	-1.8	0.3	0.2	-0.8	1.4	-8.8	-9.0	-8.5	...	...	...	...	...
CONTERMINOUS UNITED STATES														
1960														
Total	178,464,236	87,864,510	90,599,726	158,454,956	78,153,040	80,301,916	20,009,280	9,711,470	10,297,810	100.0	100.0	100.0	100.0	100.0
Urban	124,699,022	60,436,481	64,262,541	110,201,999	53,510,814	56,691,185	14,497,023	6,925,667	7,571,356	69.9	68.8	70.9	69.5	72.5
Urbanized areas	95,497,151	46,310,655	49,186,496	83,661,102	40,646,972	43,014,130	11,836,049	5,663,683	6,172,366	53.5	52.7	54.3	52.8	59.2
Central cities	57,680,938	27,777,916	29,903,022	47,546,958	22,935,746	24,611,212	10,133,980	4,842,170	5,291,810	32.3	31.6	33.0	30.0	50.6
Urban fringe	37,816,213	18,532,739	19,283,474	36,114,144	17,711,226	18,402,918	1,702,069	821,513	880,556	21.2	21.1	21.3	22.8	8.5
Other urban	29,201,871	14,125,826	15,076,045	26,540,897	12,863,842	13,677,055	2,660,974	1,261,984	1,398,990	16.4	16.1	16.6	16.7	13.3
Places of 10,000 or more	16,033,777	7,767,482	8,266,295	14,477,915	7,027,558	7,450,357	1,555,862	739,924	815,938	9.0	8.8	9.1	9.1	7.8
Places of 2,500 to 10,000	13,168,094	6,358,344	6,809,750	12,062,982	5,836,284	6,226,698	1,105,112	522,060	583,052	7.4	7.2	7.5	7.6	5.5
Rural	53,765,214	27,428,029	26,337,185	48,252,957	24,642,226	23,610,731	5,512,257	2,785,803	2,726,454	30.1	31.2	29.1	30.5	27.5
Places of 1,000 to 2,500	6,440,164	3,119,578	3,320,586	5,972,660	2,896,883	3,075,777	467,504	222,695	244,809	3.6	3.6	3.7	3.8	2.3
Other rural	47,325,050	24,308,451	23,016,599	42,280,297	21,745,343	20,534,954	5,044,753	2,563,108	2,481,645	26.5	27.7	25.4	26.7	25.2
1950														
Total	150,697,361	74,833,239	75,864,122	134,942,028	67,129,192	67,812,836	15,755,333	7,704,047	8,051,286	100.0	100.0	100.0	100.0	100.0
Urban	96,467,686	46,891,782	49,575,904	86,756,435	42,249,894	44,506,541	9,711,251	4,641,888	5,069,363	64.0	62.7	65.3	64.3	61.6
Urbanized areas	[1]69,249,148	33,670,714	35,578,434	61,925,036	30,160,082	31,764,954	7,324,112	3,510,632	3,813,480	46.0	45.0	46.9	45.9	46.5
Central cities	48,377,240	23,432,038	24,945,202	42,041,968	20,402,408	21,639,560	6,335,272	3,029,630	3,305,642	32.1	31.3	32.9	31.2	40.2
Urban fringe	[1]20,871,908	10,238,676	10,633,232	19,883,068	9,757,674	10,125,394	988,840	481,002	507,838	13.9	13.7	14.0	14.7	6.3
Other urban	[1]27,218,538	13,221,068	13,997,470	24,831,399	12,089,812	12,741,587	2,387,139	1,131,256	1,255,883	18.1	17.7	18.5	18.4	15.2
Rural	54,229,675	27,941,457	26,288,218	48,185,593	24,879,298	23,306,295	6,044,082	3,062,159	2,981,923	36.0	37.3	34.7	35.7	38.4
Percent Increase, 1950 to 1960														
Total	18.4	17.4	19.4	17.4	16.4	18.4	27.0	26.1	27.9	...	...	...	...	...
Urban	29.3	28.9	29.6	27.0	26.7	27.4	49.3	49.2	49.4	...	...	...	...	...
Urbanized areas	37.9	37.5	38.2	35.1	34.8	35.4	61.6	61.3	61.9	...	...	...	...	...
Central cities	19.2	18.5	19.9	13.1	12.4	13.7	60.0	59.8	60.1	...	...	...	...	...
Urban fringe	81.2	81.0	81.4	81.6	81.5	81.8	72.1	70.8	73.4	...	...	...	...	...
Other urban	7.3	6.8	7.7	6.9	6.4	7.3	11.5	11.6	11.4	...	...	...	...	...
Rural	-0.9	-1.8	0.2	0.1	-1.0	1.3	-8.8	-9.0	-8.6	...	...	...	...	...

[1] Figures revised since publication of 1950 reports. The revised 1950 populations are: urbanized areas (69,252,234); urban fringe (20,874,994); other urban, United States (27,594,583); and other urban, conterminous United States (27,215,452). Revised data not available by color and sex. For explanation of changes, see footnotes 3 and 6 on page 7 of table 63.

Item F

U. S. DEPARTMENT OF COMMERCE
Bureau of the Census
Washington 25, D. C.

December 1, 1961

POPULATION INSIDE AND OUTSIDE CENTRAL CITY OR CITIES OF STANDARD METROPOLITAN STATISTICAL AREAS, BY COLOR AND REGIONS: 1960 AND 1950

(Figures in thousands)

Region and component part of SMSA	Total				White				Nonwhite			
	1960	1950	Change, 1950 to 1960		1960	1950	Change, 1950 to 1960		1960	1950	Change, 1950 to 1960	
			Number	Per-cent			Number	Per-cent			Number	Per-cent
UNITED STATES												
In SMSA's	112,885	89,317	23,568	26.4	99,688	80,343	19,344	24.1	13,198	8,973	4,224	47.1
Central cities	58,004	52,371	5,633	10.8	47,655	45,499	2,155	4.7	10,350	6,872	3,478	50.6
Outside central cities	54,881	36,946	17,935	48.5	52,033	34,844	17,189	49.3	2,848	2,102	747	35.5
NORTHEAST												
In SMSA's	35,347	31,267	4,079	13.0	32,383	29,315	3,068	10.5	2,964	1,952	1,011	51.8
Central cities	17,322	17,881	-560	-3.1	14,923	16,305	1,382	-8.5	2,399	1,577	822	52.2
Outside central cities	18,025	13,386	4,639	34.7	17,460	13,010	4,450	34.2	565	376	189	50.3
NORTH CENTRAL												
In SMSA's	30,960	25,075	5,885	23.5	27,714	23,036	4,679	20.3	3,246	2,039	1,207	59.2
Central cities	16,511	15,837	674	4.3	13,668	14,086	-418	-3.0	2,843	1,751	1,092	62.4
Outside central cities	14,449	9,238	5,211	56.4	14,046	8,950	5,096	56.9	403	288	115	39.8
SOUTH												
In SMSA's	26,447	19,418	7,030	36.2	21,192	15,441	5,751	37.2	5,256	3,977	1,279	32.2
Central cities	15,062	11,721	3,341	28.5	11,143	8,865	2,278	25.7	3,919	2,856	1,063	37.2
Outside central cities	11,386	7,697	3,689	47.9	10,049	6,576	3,473	52.8	1,337	1,121	216	19.3
WEST												
In SMSA's	20,131	13,557	6,574	48.5	18,399	12,552	5,847	46.6	1,732	1,005	727	72.3
Central cities	9,110	6,932	2,178	31.4	7,921	6,244	1,678	26.9	1,189	689	500	72.6
Outside central cities	11,021	6,625	4,396	66.4	10,478	6,308	4,169	66.1	544	317	227	71.6

Item G

U. S. DEPARTMENT OF COMMERCE
Bureau of the Census
Washington 25, D. C.

December 1, 1961

Table 1.--POPULATION INSIDE AND OUTSIDE CENTRAL CITY OR CITIES OF STANDARD METROPOLITAN STATISTICAL AREAS, BY COLOR AND SIZE OF AREA: 1960 AND 1950

(Figures in thousands)

Size and component parts of SMSA	Total				White				Nonwhite			
	1960	1950	Change, 1950 to 1960		1960	1950	Change, 1950 to 1960		1960	1950	Change, 1950 to 1960	
			Number	Per-cent			Number	Per-cent			Number	Per-cent
ALL SIZES												
In SMSA's	112,885	89,317	23,568	26.4	99,688	80,343	19,344	24.1	13,198	8,973	4,224	47.1
Central cities	58,004	52,371	5,633	10.8	47,655	45,499	2,155	4.7	10,350	6,872	3,478	50.6
Outside central cities	54,881	36,946	17,935	48.5	52,033	34,844	17,189	49.3	2,848	2,102	747	35.5
500,000 or more												
In SMSA's	80,797	63,773	17,024	26.7	70,823	57,234	13,590	23.7	9,973	6,539	3,434	52.5
Central cities	40,662	38,033	2,629	6.9	32,571	32,785	-214	-0.7	8,092	5,248	2,843	54.2
Outside central cities	40,134	25,740	14,395	55.9	38,253	24,449	13,804	56.5	1,882	1,291	591	45.8
3,000,000 or more												
In SMSA's	31,763	25,789	5,975	23.2	27,713	23,262	4,451	19.1	4,051	2,527	1,524	60.3
Central cities	17,828	17,655	173	1.0	14,395	15,469	-1,074	-6.9	3,433	2,186	1,247	57.1
Outside central cities	13,935	8,134	5,802	71.3	13,318	7,793	5,525	70.9	617	341	277	81.1
1,000,000 to 3,000,000												
In SMSA's	29,819	23,858	5,960	25.0	26,327	21,560	4,767	22.1	3,491	2,298	1,193	51.9
Central cities	12,708	12,037	670	5.6	9,938	10,313	-375	-3.6	2,770	1,724	1,046	60.6
Outside central cities	17,111	11,821	5,290	44.8	16,390	11,247	5,142	45.7	721	574	148	25.7
500,000 to 1,000,000												
In SMSA's	19,215	14,126	5,089	36.0	16,783	12,411	4,372	35.2	2,432	1,715	717	41.8
Central cities	10,127	8,341	1,786	21.4	8,238	7,002	1,236	17.6	1,889	1,338	550	41.1
Outside central cities	9,088	5,785	3,303	57.1	8,545	5,409	3,136	58.0	543	376	167	44.3
250,000 to 500,000												
In SMSA's	15,829	12,603	3,226	25.6	14,381	11,573	2,808	24.3	1,449	1,030	418	40.6
Central cities	7,751	6,671	1,079	16.2	6,710	5,950	760	12.8	1,041	721	319	44.2
Outside central cities	8,078	5,932	2,147	36.2	7,671	5,623	2,048	36.4	408	309	99	32.0
100,000 to 250,000												
In SMSA's	14,498	11,526	2,972	25.8	12,858	10,228	2,630	25.7	1,640	1,297	343	26.4
Central cities	8,236	6,618	1,618	24.4	7,117	5,783	1,334	23.1	1,119	835	284	34.1
Outside central cities	6,262	4,908	1,354	27.6	5,741	4,445	1,296	29.2	521	463	58	12.6
Under 100,000												
In SMSA's	1,761	1,415	346	24.4	1,626	1,309	317	24.2	135	106	29	27.2
Central cities	1,356	1,049	306	29.2	1,258	982	276	28.1	98	68	30	45.1
Outside central cities	406	366	40	10.9	368	327	41	12.6	37	39	-1	-3.8

Item G

U. S. DEPARTMENT OF COMMERCE
Bureau of the Census
Washington 25, D. C.

December 1, 1961

Table 2.--POPULATION INSIDE AND OUTSIDE THE 1950 CITY LIMITS OF CENTRAL CITY OR CITIES OF STANDARD METROPOLITAN STATISTICAL AREAS OF 500,000 OR MORE, BY COLOR AND SIZE OF AREA: 1960 AND 1950

(Figures not adjusted for annexed areas in which the 1960 population was less than 1 percent of the 1960 total population of central city. Color distribution of the population of area annexed to Sacramento estimates. Figures in thousands)

Size and component parts of SMSA	Total				White				Nonwhite			
	1960	1950	Change 1950 to 1960		1960	1950	Change 1950 to 1960		1960	1950	Change 1950 to 1960	
			Number	Percent			Number	Percent			Number	Percent
SMSA's OF 500,000 or more												
In SMSA's	80,797	63,773	17,024	26.7	70,823	57,234	13,590	23.7	9,973	6,539	3,434	52.5
Central cities	38,275	38,033	242	0.6	30,386	32,785	-2,399	-7.3	7,889	5,248	2,641	50.3
Outside central cities	42,522	25,740	16,782	65.2	40,437	24,449	15,988	65.4	2,084	1,291	793	61.5
3,000,000 or more												
In SMSA's	31,763	25,789	5,975	23.2	27,713	23,262	4,451	19.1	4,051	2,527	1,524	60.3
Central cities	17,769	17,655	114	0.6	14,336	15,469	-1,133	-7.3	3,433	2,186	1,247	57.1
Outside central cities	13,994	8,134	5,861	72.1	13,377	7,793	5,584	71.7	618	341	277	81.2
1,000,000 to 3000000												
In SMSA's	29,819	23,858	5,960	25.0	26,327	21,560	4,767	22.1	3,491	2,298	1,193	51.9
Central cities	11,767	12,037	-270	-2.2	9,119	10,313	-1,194	-11.6	2,648	1,724	924	53.6
Outside central cities	18,052	11,821	6,231	52.7	17,209	11,247	5,961	53.0	843	574	270	47.0
500,000 to 1,000,000												
In SMSA's	19,215	14,126	5,089	36.0	16,783	12,411	4,372	35.2	2,432	1,715	717	41.8
Central cities	8,739	8,341	399	4.8	6,931	7,002	-71	-1.0	1,808	1,338	470	35.1
Outside central cities	10,475	5,785	4,690	81.1	9,852	5,409	4,443	82.2	623	376	247	65.6

ITEM H

U. S. DEPARTMENT OF COMMERCE
Bureau of the Census
Washington 25, D. C.

December 1, 1961

POPULATION OF CITIES OF 1,000,000 OR MORE BY AGE, 1960
AND 1950, WITH DECADE CHANGE BY COLOR

City and age	Total		Change, 1950 to 1960			
	1960	1950	Total		White	Nonwhite
			Number	Percent		
NEW YORK						
All ages.......	7,781,984	7,891,957	-109,973	-1.4	-475,779	365,806
Under 5 years.......	686,717	665,889	20,828	3.1	-41,308	62,136
5 to 9 years........	595,847	535,039	60,808	11.4	-982	61,790
10 to 14 years......	575,321	443,599	131,722	29.7	87,187	44,535
15 to 19 years......	486,851	467,065	19,786	4.2	-2,075	21,861
20 to 24 years......	482,522	598,718	-116,196	-19.4	-127,478	11,282
25 to 29 years......	513,629	665,245	-151,616	-22.8	-153,641	2,025
30 to 34 years......	542,769	638,249	-95,480	-15.0	-115,959	20,479
35 to 39 years......	546,966	668,845	-121,879	-18.2	-144,191	22,312
40 to 44 years......	524,381	642,065	-117,684	-18.3	-136,629	18,945
45 to 49 years......	550,310	590,622	-40,312	-6.8	-58,667	18,355
50 to 54 years......	534,526	556,389	-21,863	-3.9	-38,609	16,746
55 to 59 years......	499,493	450,515	48,978	10.9	25,407	23,571
60 to 64 years......	428,825	364,482	64,343	17.7	48,786	15,557
65 to 69 years......	344,063	274,343	69,720	25.4	59,590	10,130
70 to 74 years......	240,101	170,654	69,447	40.7	61,313	8,134
75 years and over...	229,663	160,238	69,425	43.3	61,477	7,948
CHICAGO						
All ages.......	3,550,404	3,620,962	-70,558	-1.9	-398,777	328,219
Under 5 years.......	380,672	327,176	53,496	16.4	-22,643	76,139
5 to 9 years........	312,929	256,150	56,779	22.2	-6,817	63,596
10 to 14 years......	271,083	205,323	65,760	32.0	26,216	39,544
15 to 19 years......	222,540	202,414	20,126	9.9	139	19,987
20 to 24 years......	225,053	276,140	-51,087	-18.5	-62,444	11,357
25 to 29 years......	226,646	321,742	-95,096	-29.6	-103,772	8,676
30 to 34 years......	237,956	311,034	-73,078	-23.5	-91,971	18,893
35 to 39 years......	247,875	303,888	-56,013	-18.4	-73,493	17,480
40 to 44 years......	244,443	273,500	-29,057	-10.6	-41,539	12,482
45 to 49 years......	242,377	250,886	-8,509	-3.4	-17,476	8,967
50 to 54 years......	219,261	235,681	-16,420	-7.0	-26,236	9,816
55 to 59 years......	202,251	211,430	-9,179	-4.3	-23,487	14,308
60 to 64 years......	170,743	171,874	-1,131	-0.7	-10,879	9,748
65 to 69 years......	144,783	124,529	20,254	16.3	14,081	6,173
70 to 74 years......	103,176	74,621	28,555	38.3	23,267	5,288
75 years and over...	98,616	74,574	24,042	32.2	18,277	5,765

ITEM H

POPULATION OF CITIES OF 1,000,000 OR MORE BY AGE, 1960
AND 1950, WITH DECADE CHANGE BY COLOR--cont'd

City and age	Total		Change, 1950 to 1960			
	1960	1950	Total		White	Nonwhite
			Number	Percent		
LOS ANGELES						
All ages	2,479,015	1,970,358	508,657	25.8	303,035	205,622
Under 5 years	249,232	174,120	75,112	43.1	40,784	34,328
5 to 9 years	221,451	132,323	89,128	67.4	60,106	29,022
10 to 14 years	192,416	99,568	92,848	93.3	70,738	22,110
15 to 19 years	147,851	98,314	49,537	50.4	36,310	13,227
20 to 24 years	147,588	138,069	9,519	6.9	-3,635	13,154
25 to 29 years	164,826	174,840	-10,014	-5.7	-21,465	11,451
30 to 34 years	180,882	169,153	11,729	6.9	-1,504	13,233
35 to 39 years	194,128	168,250	25,878	15.4	12,745	13,133
40 to 44 years	176,242	152,400	23,842	15.6	12,484	11,358
45 to 49 years	164,790	140,342	24,448	17.4	14,051	10,397
50 to 54 years	144,059	128,274	15,785	12.3	6,871	8,914
55 to 59 years	130,585	110,089	20,496	18.6	11,418	9,078
60 to 64 years	110,972	95,013	15,959	16.8	10,544	5,415
65 to 69 years	95,477	78,233	17,244	22.0	13,613	3,631
70 to 74 years	73,279	53,040	20,239	38.2	16,901	3,338
75 years and over	85,237	58,330	26,907	46.1	23,074	3,833
PHILADELPHIA						
All ages	2,002,512	2,071,605	-69,093	-3.3	-225,158	156,065
Under 5 years	200,953	184,125	16,828	9.1	-14,255	31,083
5 to 9 years	173,475	152,429	21,046	13.8	-6,750	27,796
10 to 14 years	157,229	129,442	27,787	21.5	7,869	19,918
15 to 19 years	134,608	130,116	4,492	3.5	-5,707	10,199
20 to 24 years	123,804	161,511	-37,707	-23.3	-40,636	2,929
25 to 29 years	122,591	179,643	-57,052	-31.8	-57,189	137
30 to 34 years	131,203	170,054	-38,851	-22.8	-46,277	7,426
35 to 39 years	140,158	166,695	-26,537	-15.9	-35,134	8,597
40 to 44 years	137,783	152,834	-15,051	-9.8	-22,108	7,057
45 to 49 years	135,306	137,886	-2,580	-1.9	-8,395	5,815
50 to 54 years	124,170	132,171	-8,001	-6.1	-12,906	4,905
55 to 59 years	113,442	112,353	1,089	1.0	-8,567	9,656
60 to 64 years	98,987	91,104	7,883	8.7	676	7,207
65 to 69 years	84,500	71,610	12,890	18.0	7,631	5,259
70 to 74 years	60,271	48,042	12,229	25.5	8,409	3,820
75 years and over	64,032	51,590	12,442	24.1	8,181	4,261

ITEM H

POPULATION OF CITIES OF 1,000,000 OR MORE BY AGE, 1960 AND 1950, WITH DECADE CHANGE BY COLOR-- cont'd

City and age	Total		Change, 1950 to 1960			
	1960	1950	Total: Number	Total: Percent	White	Nonwhite
DETROIT						
All ages........	1,670,144	1,849,568	-179,424	-9.7	-362,877	183,453
Under 5 years........	172,367	182,951	-10,584	-5.8	-45,366	34,782
5 to 9 years.........	160,485	140,594	19,891	14.1	-18,483	38,374
10 to 14 years.......	142,296	115,224	27,072	23.5	922	26,150
15 to 19 years.......	111,068	112,574	-1,506	-1.3	-13,871	12,365
20 to 24 years.......	91,077	150,448	-59,371	-39.5	-58,783	-588
25 to 29 years.......	90,748	171,846	-81,098	-47.2	-79,153	-1,945
30 to 34 years.......	105,963	160,770	-54,807	-34.1	-62,094	7,287
35 to 39 years.......	119,950	149,902	-29,952	-20.0	-40,221	10,269
40 to 44 years.......	120,320	134,384	-14,064	-10.5	-25,270	11,206
45 to 49 years.......	114,956	128,927	-13,971	-10.8	-23,235	9,264
50 to 54 years.......	101,855	118,778	-16,923	-14.2	-24,047	7,124
55 to 59 years.......	97,676	101,878	-4,202	-4.1	-13,864	9,662
60 to 64 years.......	83,018	76,200	6,818	8.9	-560	7,378
65 to 69 years.......	70,031	50,198	19,833	39.5	14,577	5,256
70 to 74 years.......	46,634	28,066	18,568	66.2	14,748	3,820
75 years and over....	41,700	26,828	14,872	55.4	11,823	3,049

Item I

December 1, 1961

U. S. DEPARTMENT OF COMMERCE
Bureau of the Census
Washington 25, D. C.

AGE OF THE POPULATION INSIDE AND OUTSIDE STANDARD METROPOLITAN STATISTICAL AREAS: 1960 AND 1950

Age	1960	1950	Change, 1950 to 1960	
			Number	Percent
IN SMSA's				
All ages	112,885,178	89,316,903	23,568,275	26.4
Under 5 years	12,817,940	9,163,578	3,654,362	39.9
5 to 9 years	11,529,454	7,164,155	4,365,299	60.9
10 to 14 years	10,051,320	5,742,150	4,309,170	75.0
15 to 19 years	7,786,334	5,714,514	2,071,820	36.3
20 to 24 years	6,872,190	7,009,490	- 137,300	- 2.0
25 to 29 years	7,145,575	7,785,595	- 640,020	- 8.2
30 to 34 years	7,938,306	7,310,135	628,171	8.6
35 to 39 years	8,315,766	7,058,838	1,256,928	17.8
40 to 44 years	7,599,941	6,403,418	1,196,523	18.7
45 to 49 years	7,010,568	5,702,326	1,308,242	22.9
50 to 54 years	6,141,643	5,214,588	927,055	17.8
55 to 59 years	5,362,755	4,471,869	890,886	19.9
60 to 64 years	4,500,283	3,665,625	834,658	22.8
65 years and over	9,813,103	6,910,622	2,902,481	42.0
OUTSIDE SMSA's				
All ages	66,437,997	62,008,895	4,429,102	7.1
Under 5 years	7,502,961	7,079,563	423,398	6.0
5 to 9 years	7,162,326	6,097,968	1,064,358	17.5
10 to 14 years	6,722,172	5,425,328	1,296,844	23.9
15 to 19 years	5,432,909	4,956,807	476,102	9.6
20 to 24 years	3,928,571	4,539,865	- 611,294	-13.5
25 to 29 years	3,723,549	4,520,356	- 796,807	-17.6
30 to 34 years	4,010,880	4,262,202	- 251,322	- 5.9
35 to 39 years	4,165,343	4,235,640	- 70,297	- 1.7
40 to 44 years	4,000,302	3,837,253	163,049	4.2
45 to 49 years	3,868,917	3,399,452	469,465	13.8
50 to 54 years	3,464,311	3,080,992	383,319	12.4
55 to 59 years	3,067,110	2,780,655	286,455	10.3
60 to 64 years	2,642,169	2,408,738	233,431	9.7
65 years and over	6,746,477	5,384,076	1,362,401	25.3

Item J December 1, 1961

U. S. DEPARTMENT OF COMMERCE
Bureau of the Census
Washington 25, D. C.

YEAR OF BIRTH AND AGE OF THE POPULATION INSIDE AND OUTSIDE STANDARD METROPOLITAN STATISTICAL AREAS: 1960 AND 1950

Year of birth	1960		1950		Change, 1950-1960	
	Age (years)	Number	Age (years)	Number	Number	Percent
IN SMSA's						
(April 1)	All ages....	112,885,178	All ages....	89,316,903	23,568,275	26.4
1955 to 1960	Under 5 years..	12,817,940	--	--	12,817,940	--
1950 to 1955	5 to 9 years...	11,529,454	--	--	11,529,454	--
1945 to 1950	10 to 14 years.	10,051,320	Under 5 years..	9,163,578	887,742	9.7
1940 to 1945	15 to 19 years.	7,786,334	5 to 9 years...	7,164,155	622,179	8.7
1935 to 1940	20 to 24 years.	6,872,190	10 - 14 years..	5,742,150	1,130,040	19.7
1930 to 1935	25 to 29 years.	7,145,575	15 - 19 years..	5,714,514	1,431,061	25.0
1925 to 1930	30 to 34 years.	7,938,306	20 - 24 years..	7,009,490	928,816	13.3
1920 to 1925	35 to 39 years.	8,315,766	25 - 29 years..	7,785,595	530,171	6.8
1915 to 1920	40 to 44 years.	7,599,941	30 - 34 years..	7,310,135	289,806	4.0
1910 to 1915	45 to 49 years.	7,010,568	35 - 39 years..	7,058,838	- 48,270	- 0.7
1905 to 1910	50 to 54 years.	6,141,643	40 - 44 years..	6,403,418	- 261,775	- 4.1
1900 to 1905	55 to 59 years.	5,362,755	45 - 49 years..	5,702,326	- 339,571	- 6.0
1895 to 1900	60 to 64 years.	4,500,283	50 - 54 years..	5,214,588	- 714,305	-13.7
Before 1895	65 and over....	9,813,103	55 and over....	15,048,116	- 5,235,013	-34.8
OUTSIDE SMSA's						
(April 1)	All ages....	66,437,997	All ages....	62,008,895	4,429,102	7.1
1955 to 1960	Under 5 years..	7,502,961	--	--	7,502,961	--
1950 to 1955	5 to 9 years...	7,162,326	--	--	7,162,326	--
1945 to 1950	10 to 14 years.	6,722,172	Under 5 years..	7,079,563	- 357,391	- 5.0
1940 to 1945	15 to 19 years.	5,432,909	5 to 9 years...	6,097,968	- 665,059	-10.9
1935 to 1940	20 to 24 years.	3,928,571	10 - 14 years..	5,425,328	- 1,496,757	-27.6
1930 to 1935	25 to 29 years.	3,723,549	15 - 19 years..	4,956,807	- 1,233,258	-24.9
1925 to 1930	30 to 34 years.	4,010,880	20 - 24 years..	4,539,865	- 528,985	-11.7
1920 to 1925	35 to 39 years.	4,165,343	25 - 29 years..	4,520,356	- 355,013	- 7.9
1915 to 1920	40 to 44 years.	4,000,302	30 - 34 years..	4,262,202	- 261,900	- 6.1
1910 to 1915	45 to 49 years.	3,868,917	35 - 39 years..	4,235,640	- 366,723	- 8.7
1905 to 1910	50 to 54 years.	3,464,311	40 - 44 years..	3,837,253	- 372,942	- 9.7
1900 to 1905	55 to 59 years.	3,067,110	45 - 49 years..	3,399,452	- 332,342	- 9.8
1895 to 1900	60 to 64 years.	2,642,169	50 - 54 years..	3,080,992	- 438,823	-14.2
Before 1895	65 and over....	6,746,477	55 and over....	10,573,469	- 3,826,992	-36.2

TABLE 2.—*Comparison of number of persons of all ages, 45 to 64, and 65 and over, Apr. 1, 1950 and 1960, by State*

State	Total, all ages			45 through 64 years of age					65 years of age and over				
	1950	1960		1950		1960			1950		1960		
		Number	Percent change over 1950	Number	Percent of all ages	Number	Percent of all ages	Percent change over 1950	Number	Percent of all ages	Number	Percent of all ages	Percent change over 1950
Total, 51 States	151, 325, 798	179, 323, 175	+18. 5	30, 724, 245	20. 3	36, 057, 756	20. 1	+17. 4	12, 294, 698	8. 1	16, 559, 580	9. 2	+34. 7
Alabama	3, 061, 743	3, 266, 740	+6. 7	497, 169	16. 2	606, 740	18. 6	+22. 0	198, 648	6. 5	261, 147	8. 0	+31. 5
Alaska	128, 643	226, 167	+75. 8	16, 870	13. 1	28, 103	12. 4	+66. 6	4, 742	3. 7	5, 386	2. 4	+13. 6
Arizona	749, 587	1, 302, 161	+73. 7	130, 440	17. 4	229, 444	17. 6	+75. 9	44, 241	5. 9	90, 225	6. 9	+103. 9
Arkansas	1, 909, 511	1, 786, 272	−6. 5	350, 379	18. 3	372, 902	20. 9	+6. 4	148, 995	7. 8	194, 372	10. 9	+30. 5
California	10, 586, 223	15, 717, 204	+48. 5	2, 249, 830	21. 3	3, 089, 405	19. 7	+37. 3	895, 005	8. 5	1, 376, 204	8. 8	+53. 8
Colorado	1, 325, 089	1, 753, 947	+32. 4	256, 839	19. 4	319, 938	18. 2	+24. 6	115, 592	8. 7	158, 160	9. 0	+36. 8
Connecticut	2, 007, 280	2, 535, 234	+26. 3	445, 093	22. 2	534, 514	21. 1	+20. 1	176, 824	8. 8	242, 615	9. 6	+37. 2
Delaware	318, 085	446, 292	+40. 3	65, 327	20. 5	83, 000	18. 6	+27. 1	26, 320	8. 3	35, 745	8. 0	+35. 8
District of Columbia	802, 178	763, 956	−4. 8	173, 619	21. 6	175, 956	23. 0	+1. 3	56, 687	7. 1	69, 143	9. 1	+22. 0
Florida	2, 771, 305	4, 951, 560	+78. 7	560, 223	20. 2	1, 019, 322	20. 6	+81. 9	237, 474	8. 6	553, 129	11. 2	+132. 9
Georgia	3, 444, 578	3, 943, 116	+14. 5	562, 378	16. 3	712, 776	18. 1	+26. 7	219, 655	6. 4	290, 661	7. 4	+32. 3
Hawaii	499, 794	632, 772	+26. 6	70, 127	14. 0	97, 333	15. 4	+38. 8	20, 419	4. 1	29, 162	4. 6	+42. 8
Idaho	588, 637	667, 191	+13. 3	104, 878	17. 8	122, 517	18. 4	+16. 8	43, 537	7. 4	58, 258	8. 7	+33. 8
Illinois	8, 712, 176	10, 081, 158	+15. 7	1, 978, 606	22. 7	2, 162, 677	21. 5	+9. 3	754, 301	8. 7	974, 923	9. 7	+29. 2
Indiana	3, 934, 224	4, 662, 498	+18. 5	803, 196	20. 4	902, 764	19. 4	+12. 4	361, 026	9. 2	445, 519	9. 6	+23. 4
Iowa	2, 621, 073	2, 757, 537	+5. 2	558, 117	21. 3	561, 764	20. 4	+. 7	272, 998	10. 4	327, 685	11. 9	+20. 0
Kansas	1, 905, 299	2, 178, 611	+14. 3	398, 946	20. 9	431, 588	19. 8	+8. 2	194, 218	10. 2	240, 269	11. 0	+23. 7
Kentucky	2, 944, 806	3, 038, 156	+3. 2	529, 468	18. 0	582, 368	19. 2	+10. 0	235, 243	8. 0	292, 323	9. 6	+24. 3
Louisiana	2, 683, 516	3, 257, 022	+21. 4	465, 753	17. 4	592, 245	18. 2	+27. 2	176, 849	6. 6	241, 591	7. 4	+36. 6
Maine	913, 774	969, 265	+6. 1	184, 078	20. 1	193, 545	20. 0	+5. 1	93, 562	10. 2	106, 544	11. 0	+13. 9
Maryland	2, 343, 001	3, 100, 689	+32. 3	449, 890	19. 2	587, 329	18. 9	+30. 5	163, 514	7. 0	226, 539	7. 3	+38. 5
Massachusetts	4, 690, 514	5, 148, 578	+9. 8	1, 061, 064	22. 6	1, 110, 034	21. 6	+4. 6	468, 436	10. 0	571, 609	11. 1	+22. 0
Michigan	6, 371, 766	7, 823, 194	+22. 8	1, 302, 585	20. 4	1, 504, 930	19. 2	+15. 5	461, 650	7. 2	638, 184	8. 2	+38. 2
Minnesota	2, 982, 483	3, 413, 864	+14. 5	639, 561	21. 4	676, 235	19. 8	+5. 7	269, 130	9. 0	354, 351	10. 4	+31. 7
Mississippi	2, 178, 914	2, 178, 141	(1)	355, 068	16. 3	401, 449	18. 4	+13. 1	152, 964	7. 0	190, 029	8. 7	+24. 2
Missouri	3, 954, 653	4, 319, 813	+9. 2	875, 806	22. 1	935, 709	21. 7	+6. 8	407, 388	10. 3	503, 411	11. 7	+23. 6
Montana	591, 024	674, 767	+14. 2	117, 626	19. 9	125, 224	18. 6	+6. 5	50, 864	8. 6	65, 420	9. 7	+28. 6
Nebraska	1, 325, 510	1, 411, 330	+6. 5	285, 719	21. 6	288, 005	20. 4	+. 8	130, 379	9. 8	164, 156	11. 6	+25. 9
Nevada	160, 083	285, 278	+78. 2	33, 571	21. 0	58, 469	20. 5	+74. 2	10, 986	6. 9	18, 173	6. 4	+65. 4
New Hampshire	533, 242	606, 921	+13. 8	116, 434	21. 8	126, 727	20. 9	+8. 8	57, 793	10. 8	67, 705	11. 2	+17. 2
New Jersey	4, 835, 329	6, 066, 782	+25. 5	1, 102, 801	22. 8	1, 324, 141	21. 8	+20. 1	393, 989	8. 1	560, 414	9. 2	+42. 2
New Mexico	681, 187	951, 023	+39. 6	101, 815	14. 9	143, 568	15. 1	+41. 0	33, 064	4. 9	51, 270	5. 4	+55. 1
New York	14, 830, 192	16, 782, 304	+13. 2	3, 528, 913	23. 8	3, 891, 774	23. 2	+10. 3	1, 258, 457	8. 5	1, 687, 590	10. 1	+34. 1
North Carolina	4, 061, 929	4, 556, 155	+12. 2	624, 686	15. 4	805, 017	17. 7	+28. 9	225, 297	5. 5	312, 167	6. 9	+38. 6
North Dakota	619, 636	632, 446	+2. 1	113, 488	18. 3	119, 941	19. 0	+5. 7	48, 196	7. 8	58, 591	9. 3	+21. 6
Ohio	7, 946, 627	9, 706, 397	+22. 1	1, 680, 420	21. 1	1, 895, 519	19. 5	+12. 8	708, 975	8. 9	897, 124	9. 2	+26. 5
Oklahoma	2, 233, 351	2, 328, 284	+4. 3	430, 882	19. 3	485, 183	20. 8	+12. 6	193, 922	8. 7	248, 831	10. 7	+28. 3
Oregon	1, 521, 341	1, 768, 687	+16. 3	325, 419	21. 4	374, 174	21. 2	+15. 0	133, 021	8. 7	183, 653	10. 4	+38. 1
Pennsylvania	10, 498, 012	11, 319, 366	+7. 8	2, 226, 607	21. 2	2, 452, 213	21. 7	+10. 1	886, 825	8. 4	1, 128, 525	10. 0	+27. 3
Rhode Island	791, 896	859, 488	+8. 5	171, 841	21. 7	184, 016	21. 4	+7. 1	70, 418	8. 9	89, 540	10. 4	+27. 2
South Carolina	2, 117, 027	2, 382, 594	+12. 5	305, 606	14. 4	389, 361	16. 3	+27. 4	115, 005	5. 4	150, 599	6. 3	+30. 9
South Dakota	652, 740	680, 514	+4. 3	131, 341	20. 1	130, 632	19. 2	−. 5	55, 296	8. 5	71, 513	10. 5	+29. 3
Tennessee	3, 291, 718	3, 567, 089	+8. 4	574, 136	17. 4	695, 759	19. 5	+21. 2	234, 884	7. 1	308, 861	8. 7	+31. 5
Texas	7, 711, 194	9, 579, 677	+24. 2	1, 398, 490	18. 1	1, 791, 004	18. 7	+28. 1	513, 420	6. 7	745, 391	7. 8	+45. 2
Utah	688, 862	890, 627	+26. 3	111, 529	16. 2	142, 157	16. 0	+27. 5	42, 418	6. 2	59, 957	6. 7	+41. 3
Vermont	377, 747	389, 881	+3. 2	75, 517	20. 0	78, 450	20. 1	+3. 9	39, 534	10. 5	43, 741	11. 2	+10. 6
Virginia	3, 318, 680	3, 966, 949	+19. 5	560, 404	16. 9	716, 742	18. 1	+27. 9	214, 524	6. 5	288, 970	7. 3	+34. 7
Washington	2, 378, 963	2, 853, 214	+19. 9	483, 875	20. 3	565, 635	19. 8	+16. 9	211, 405	8. 9	279, 045	9. 8	+32. 0
West Virginia	2, 005, 552	1, 860, 421	−7. 2	343, 727	17. 1	372, 307	20. 0	+8. 3	138, 526	6. 9	172, 516	9. 3	+24. 5
Wisconsin	3, 434, 575	3, 951, 777	+15. 1	740, 193	21. 6	805, 702	20. 4	+8. 9	309, 917	9. 0	402, 736	10. 2	+29. 9
Wyoming	290, 529	330, 066	+13. 6	53, 895	18. 6	61, 449	18. 6	+14. 0	18, 165	6. 3	25, 908	7. 8	+42. 6

[1] Decrease of less than ½ of 1 percent.

Item L

TABLE 13.—*Percent of State's population in each size of area of residence class which is aged 65 and over, Apr. 1, 1960, by State*

State	Total	Urban							Rural		
		Total	Urbanized areas			Other urban			Total	Places of 1,000 to 2,500	Other rural
			Total	Central cities	Urban fringe	Total	Places of 10,000 or more	Places of 2,500 to 10,000			
Total, 51 States	9.2	9.2	8.8	9.9	7.2	10.4	9.8	11.0	9.3	12.2	8.9
Alabama	8.0	7.4	7.0	7.3	5.9	8.1	7.6	8.6	8.8	10.0	8.7
Alaska	2.4	2.5	---	---	---	2.5	1.7	4.1	2.3	3.1	2.1
Arizona	6.9	7.3	7.5	7.7	6.3	6.6	7.4	6.1	5.8	6.1	5.8
Arkansas	10.9	10.7	9.8	10.1	7.4	11.1	9.6	12.9	11.0	14.1	10.7
California	8.8	8.9	8.8	10.5	7.3	9.4	8.7	10.3	8.2	10.4	7.9
Colorado	9.0	9.2	8.5	10.8	4.4	11.4	11.1	11.8	8.6	13.1	7.8
Connecticut	9.6	9.9	9.8	10.7	8.8	10.2	10.2	10.1	8.4	13.2	7.9
Delaware	8.0	8.0	7.8	12.4	5.2	10.7	---	10.7	8.0	11.8	7.5
District of Columbia	9.1	9.1	9.1	9.1	---	---	---	---	---	---	---
Florida	11.2	11.9	12.0	13.8	10.4	11.5	10.6	12.5	9.2	12.9	8.7
Georgia	7.4	7.1	6.6	7.5	4.5	8.1	7.9	8.2	7.7	9.2	7.5
Hawaii	4.6	4.4	4.4	4.8	2.1	4.6	4.0	5.5	5.1	6.0	4.9
Idaho	8.7	9.6	---	---	---	9.6	9.6	9.6	7.9	9.8	7.5
Illinois	9.7	9.3	8.8	9.9	6.8	12.0	11.3	12.6	11.1	14.5	10.4
Indiana	9.6	9.3	8.3	9.2	5.7	11.1	10.5	12.0	10.0	12.5	9.6
Iowa	11.9	12.0	10.1	10.7	7.6	14.2	12.0	16.1	11.7	18.7	10.5
Kansas	11.0	9.9	7.8	8.5	7.0	12.1	11.0	14.0	12.8	16.8	12.0
Kentucky	9.6	9.9	9.1	10.4	7.5	11.5	11.3	11.6	9.4	11.8	9.2
Louisiana	7.4	7.1	6.9	7.9	3.9	7.5	7.1	7.9	8.0	9.0	7.8
Maine	11.0	11.1	12.0	12.5	10.1	10.7	10.7	10.6	10.9	12.4	10.6
Maryland	7.3	7.1	6.8	9.0	4.9	9.6	10.7	8.2	7.8	11.2	7.5
Massachusetts	11.1	11.5	11.4	12.4	10.6	12.2	12.3	12.1	9.2	12.0	8.7
Michigan	8.2	7.9	7.4	9.5	5.1	10.5	9.9	11.0	9.0	11.7	8.6
Minnesota	10.4	10.2	9.5	12.6	4.5	12.1	11.5	12.7	10.6	15.1	9.8
Mississippi	8.7	8.1	6.0	6.0	3.9	8.5	8.1	9.3	9.1	10.9	9.0
Missouri	11.7	10.9	10.2	12.2	6.7	13.5	12.1	14.6	13.1	18.2	12.4
Montana	9.7	10.0	8.0	8.3	4.5	11.0	11.3	10.7	9.4	12.9	8.7
Nebraska	11.6	10.8	9.4	9.9	4.1	13.2	12.6	13.8	12.6	19.6	11.2
Nevada	6.4	6.4	6.2	6.8	4.8	7.3	2.9	9.2	6.2	7.0	6.0
New Hampshire	11.2	10.8	11.3	11.5	7.5	10.7	10.9	10.1	11.6	13.2	11.3
New Jersey	9.2	9.2	9.1	10.3	8.8	9.8	9.2	10.4	9.4	12.0	9.1
New Mexico	5.4	5.1	4.7	4.9	4.1	5.4	4.4	7.7	5.9	6.8	5.8
New York	10.1	10.1	9.9	10.8	7.4	12.2	12.9	11.4	9.8	12.6	9.4
North Carolina	6.9	6.8	6.6	6.6	5.4	6.9	6.7	7.2	6.9	8.3	6.8
North Dakota	9.3	8.9	8.9	9.0	.5	8.9	8.2	10.8	9.5	13.3	8.6
Ohio	9.2	9.3	8.9	10.2	7.0	10.9	11.1	10.5	9.0	11.7	8.7
Oklahoma	10.7	9.9	7.3	7.9	4.7	12.9	11.3	15.3	12.0	17.8	11.0
Oregon	10.4	11.3	11.2	13.6	7.6	11.3	11.3	11.3	8.9	9.7	8.9
Pennsylvania	10.0	10.3	9.9	11.0	8.5	11.9	12.0	11.8	9.2	10.9	8.9
Rhode Island	10.4	10.7	11.0	12.8	9.4	9.4	9.1	10.6	8.5	11.6	8.3
South Carolina	6.3	6.4	5.6	7.6	3.7	7.2	7.0	7.3	6.2	7.1	6.1
South Dakota	10.5	10.1	9.0	9.1	5.4	10.5	9.1	12.9	10.8	16.7	9.9
Tennessee	8.7	8.2	7.9	8.7	5.9	8.9	8.2	9.8	9.1	10.6	9.0
Texas	7.8	6.9	6.1	6.4	4.3	9.0	7.9	10.4	10.4	12.1	10.1
Utah	6.7	6.8	7.0	9.2	3.8	6.2	4.9	7.0	6.5	7.9	6.0
Vermont	11.2	11.5	---	---	---	11.5	10.7	12.1	11.0	14.0	10.5
Virginia	7.3	6.6	6.1	7.3	4.4	8.4	8.4	8.3	8.1	9.0	8.0
Washington	9.8	10.2	10.0	12.2	6.0	10.7	10.4	11.1	8.8	12.3	8.3
West Virginia	9.3	10.2	9.2	10.3	6.6	11.2	11.2	11.1	8.7	10.1	8.5
Wisconsin	10.2	9.8	8.7	9.5	7.0	11.8	10.8	13.2	10.8	14.1	10.3
Wyoming	7.8	8.3	---	---	---	8.3	7.9	9.1	7.3	9.1	6.8

Item M

WORKERS BY PLACE OF WORK AND PLACE OF RESIDENCE, DENVER SMSA: 1960

Place of work	Total SMSA	Denver city (county)	Place of residence: Adams County	Arapahoe County	Boulder County	Jefferson County
NUMBER						
All workers	355,501	197,401	41,512	42,778	26,764	47,046
Inside SMSA	330,364	182,744	38,399	40,441	25,299	43,481
Denver city (county)	229,238	165,723	21,317	19,130	2,373	20,695
Adams County	19,634	3,310	13,815	1,252	218	1,039
Arapahoe County	27,078	6,738	1,778	17,153	106	1,303
Boulder County	22,545	404	263	55	21,592	231
Jefferson County	31,869	6,569	1,226	2,851	1,010	20,213
Outside SMSA	6,141	2,639	870	1,030	632	970
Place of work not reported	18,996	12,018	2,243	1,307	833	2,595
PERCENT						
All workers	100.0	100.0	100.0	100.0	100.0	100.0
Inside SMSA	92.9	92.6	92.5	94.5	94.5	92.4
Denver city (county)	64.5	84.0	51.4	44.7	8.9	44.0
Adams County	5.5	1.7	33.3	2.9	0.8	2.2
Arapahoe County	7.6	3.4	4.3	40.1	0.4	2.8
Boulder County	6.3	0.2	0.6	0.1	80.7	0.5
Jefferson County	9.0	3.3	3.0	6.7	3.8	43.0
Outside SMSA	1.7	1.3	2.1	2.4	2.4	2.1
Place of work not reported	5.3	6.1	5.4	3.1	3.1	5.5

PLACE OF WORK OF WORKERS FOR SELECTED STANDARD METROPOLITAN STATISTICAL AREAS: 1960

Standard metropolitan statistical area and component of residence	Number: Total	Number: In SMSA: Central city or cities	Number: In SMSA: Ring	Number: Outside SMSA	Number: Not reported	Percent: Total	Percent: In SMSA: Central city or cities	Percent: In SMSA: Ring	Percent: Outside SMSA	Percent: Not reported
Birmingham, Ala., total	212,557	140,804	58,444	4,482	8,827	100.0	66.2	27.5	2.1	4.2
Central city	121,656	102,905	11,787	2,113	4,851	100.0	84.6	9.7	1.7	4.0
Ring	90,901	37,899	46,657	2,369	3,976	100.0	41.7	51.3	2.6	4.4
Gadsden, Ala., total	28,750	20,138	5,960	1,853	799	100.0	70.0	20.7	6.4	2.8
Central city	18,410	16,193	887	848	482	100.0	87.9	4.8	4.6	2.6
Ring	10,340	3,945	5,073	1,005	317	100.0	38.2	49.1	9.7	3.1
Huntsville, Ala., total	40,923	21,232	18,042	617	1,032	100.0	51.9	44.1	1.5	2.5
Central city	26,694	15,617	10,134	330	613	100.0	58.5	38.0	1.2	2.3
Ring	14,229	5,615	7,908	287	419	100.0	39.5	55.6	2.0	2.9
Mobile, Ala., total	104,144	78,239	20,811	2,464	2,630	100.0	75.1	20.0	2.4	2.5
Central city	71,239	64,404	3,669	1,176	1,990	100.0	90.4	5.2	1.7	2.8
Ring	32,905	13,835	17,142	1,288	640	100.0	42.0	52.1	3.9	1.9
Montgomery, Ala., total	61,586	50,443	8,034	1,855	1,254	100.0	81.9	13.0	3.0	2.0
Central city	50,706	44,952	3,222	1,533	999	100.0	88.7	6.4	3.0	2.0
Ring	10,880	5,491	4,812	322	255	100.0	50.5	44.2	3.0	2.3
Colorado Springs, Colo.,total	55,795	37,248	16,021	1,098	1,428	100.0	66.8	28.7	2.0	2.6
Central city	26,626	19,838	5,535	495	758	100.0	74.5	20.8	1.9	2.8
Ring	29,169	17,410	10,486	603	670	100.0	59.7	35.9	2.1	2.3
Denver, Colo., total	355,501	229,238	101,126	6,141	18,996	100.0	64.5	28.4	1.7	5.3
Central city	197,401	165,723	17,021	2,639	12,018	100.0	84.0	8.6	1.3	6.1
Ring	158,100	63,515	84,105	3,502	6,978	100.0	40.3	53.4	2.2	4.4

Item N

PLACE OF WORK OF WORKERS FOR SELECTED STANDARD METROPOLITAN STATISTICAL AREAS: 1960 (continued)

Standard metropolitan statistical area and component of residence	Place of work									
	Number					Percent				
	Total	In SMSA: Central city or cities	In SMSA: Ring	Outside SMSA	Not reported	Total	In SMSA: Central city or cities	In SMSA: Ring	Outside SMSA	Not reported
Pueblo, Colorado, total	37,159	22,272	13,229	705	953	100.0	59.9	35.6	1.9	2.6
Central city	29,772	19,201	9,259	540	772	100.0	64.5	31.1	1.8	2.6
Ring	7,387	3,071	3,970	165	181	100.0	41.6	53.7	2.2	2.5
Bridgeport, Conn., total	130,092	68,232	38,648	15,647	7,565	100.0	52.4	29.7	12.0	5.8
Central city	61,972	44,126	7,979	4,495	5,372	100.0	71.2	12.9	7.3	8.7
Ring	68,120	24,106	30,669	11,152	2,193	100.0	35.4	45.0	16.4	3.2
Hartford, Conn., total	215,356	99,050	93,471	15,742	7,093	100.0	46.0	43.4	7.3	3.3
Central city	71,781	52,429	12,502	2,722	4,128	100.0	73.0	17.4	3.8	5.8
Ring	143,575	46,621	80,969	13,020	2,965	100.0	32.5	56.4	9.1	2.1
New Haven, Conn., total	124,112	74,738	31,655	11,508	6,211	100.0	60.2	25.5	9.3	5.0
Central city	61,602	46,147	7,239	3,940	4,276	100.0	74.9	11.8	6.4	6.9
Ring	62,510	28,591	24,416	7,568	1,935	100.0	45.7	39.1	12.1	3.1
New Britain, Conn., total	52,123	28,123	10,792	11,869	1,339	100.0	54.0	20.7	22.8	2.6
Central city	33,681	24,321	2,044	6,473	843	100.0	72.2	6.1	19.2	2.5
Ring	18,442	3,802	8,748	5,396	496	100.0	20.6	47.4	29.3	2.7
Norwalk, Conn., total	38,843	19,936	5,260	11,312	2,335	100.0	51.3	13.5	29.1	6.0
Central city	28,332	18,588	931	7,021	1,792	100.0	65.6	3.3	24.8	6.3
Ring	10,511	1,348	4,329	4,291	543	100.0	12.8	41.2	40.8	5.2
Stamford, Conn., total	72,346	31,536	19,365	16,794	4,651	100.0	43.6	26.8	23.2	6.4
Central city	39,426	27,379	3,470	5,688	2,889	100.0	69.4	8.8	14.4	7.3
Ring	32,920	4,157	15,895	11,106	1,762	100.0	12.6	48.3	33.7	5.4

Item N

PLACE OF WORK OF WORKERS FOR SELECTED STANDARD METROPOLITAN STATISTICAL AREAS: 1960 (continued)

Standard metropolitan statistical area and component of residence	Place of work									
	Number					Percent				
	Total	In SMSA: Central city or cities	In SMSA: Ring	Outside SMSA	Not reported	Total	In SMSA: Central city or cities	In SMSA: Ring	Outside SMSA	Not reported
Waterbury, Conn., total	70,878	41,991	19,055	7,216	2,616	100.0	59.2	26.9	10.2	3.7
Central city	42,481	33,103	5,007	2,583	1,788	100.0	77.9	11.8	6.1	4.2
Ring	28,397	8,888	14,048	4,633	828	100.0	31.3	49.5	16.3	2.9
Baton Rouge, La., total	76,576	60,500	10,585	3,031	2,460	100.0	79.0	13.8	4.0	3.2
Central city	54,062	46,490	3,904	1,926	1,742	100.0	86.0	7.2	3.6	3.2
Ring	22,514	14,010	6,681	1,105	718	100.0	62.2	29.7	4.9	3.2
Monroe, La., total	32,734	19,633	10,928	1,444	729	100.0	60.0	33.4	4.4	2.2
Central city	17,301	14,462	1,769	662	408	100.0	83.6	10.2	3.8	2.4
Ring	15,433	5,171	9,159	782	321	100.0	33.5	59.3	5.1	2.1
New Orleans, La., total	298,091	224,964	48,502	6,705	17,920	100.0	75.5	16.3	2.2	6.0
Central city	220,919	192,723	8,679	3,433	16,084	100.0	87.2	3.9	1.6	7.3
Ring	77,172	32,241	39,823	3,272	1,836	100.0	41.8	51.6	4.2	2.4
Shreveport, La., total	98,492	65,306	26,395	2,636	4,155	100.0	66.3	26.8	2.7	4.2
Central city	61,315	51,860	4,910	1,483	3,062	100.0	84.6	8.0	2.4	5.0
Ring	37,177	13,446	21,485	1,153	1,093	100.0	36.2	57.8	3.1	2.9
Portland, Maine, total	45,242	29,844	9,240	3,387	2,771	100.0	66.0	20.4	7.5	6.1
Central city	27,595	22,089	1,663	1,782	2,061	100.0	80.0	6.0	6.5	7.5
Ring	17,647	7,755	7,577	1,605	710	100.0	43.9	42.9	9.1	4.0
Lincoln, Neb., total	63,603	48,135	11,628	1,905	1,935	100.0	75.7	18.3	3.0	3.0
Central city	54,230	45,510	5,387	1,597	1,736	100.0	83.9	9.9	2.9	3.2
Ring	9,373	2,625	6,241	308	199	100.0	28.0	66.6	3.3	2.1

ITEM O

U. S. DEPARTMENT OF COMMERCE
BUREAU OF THE CENSUS
WASHINGTON 25, D. C.

December 6, 1961

HOUSEHOLDS, PRIMARY FAMILIES, AND PRIMARY INDIVIDUALS, BY REGIONS, URBAN-RURAL RESIDENCE, AND COLOR, FOR THE UNITED STATES: 1960 AND 1950

Subject	United States	Northeast	North Central	South	West
Households, 19601/.........	53,021,061	13,521,070	15,377,361	15,503,321	8,619,309
1950...........	42,394,320	11,053,155	12,743,185	12,510,255	6,087,725
Percent increase, 1950-60	25.1	22.3	20.7	23.9	41.6
Average size, 1960.......	3.29	3.21	3.27	3.45	3.15
1950.......	3.42	3.41	3.35	3.64	3.14
Primary families, 19602/...	44,669,793	11,356,380	12,990,239	13,371,743	6,951,431
1950.....	37,758,237	9,896,044	11,339,038	11,355,125	5,168,030
Percent increase, 1950-60	18.3	14.8	14.6	17.8	34.5
Average size, 1960.......	3.65	3.56	3.63	3.77	3.60
1950.......	3.61	3.58	3.53	3.81	3.42
Primary individuals, 19603/	8,351,268	2,164,690	2,387,122	2,131,578	1,667,878
1950..	4,636,083	1,157,111	1,404,147	1,155,130	919,695
Percent increase, 1950-60	80.1	87.1	70.0	84.5	81.4
As percent of hslds, 1960	15.8	16.0	15.5	13.7	19.4
1950	10.9	10.5	11.0	9.2	15.1

Subject	Residence: Urban	Residence: Rural	Color4/: White	Color4/: Nonwhite
Households, 19601/....................	38,315,788	14,705,273	47,765,537	5,045,210
1950......................	28,273,033	14,121,287	38,429,035	3,822,380
Percent increase, 1950-60..........	35.5	4.1	24.3	32.0
Average size, 1960.................	3.18	3.56	3.23	3.84
1950.................	3.27	3.73	3.37	3.90
Primary families, 19602/..............	31,641,435	13,028,358	40,413,378	4,080,673
1950................	24,881,456	12,876,781	34,457,056	3,175,700
Percent increase, 1950-60..........	27.2	1.2	17.3	28.5
Average size, 1960.................	3.57	3.85	3.58	4.34
1950.................	3.45	3.92	3.55	4.18
Primary individuals, 19603/...........	6,674,353	1,676,915	7,352,159	964,537
1950.............	3,391,577	1,244,506	3,971,979	646,680
Percent increase, 1950-60..........	96.8	34.7	85.1	49.2
As percent of hslds, 1960..........	17.4	11.4	15.4	19.1
1950..........	12.0	8.8	10.3	16.9

1/ The number of households is equal to the number of primary families plus the number of primary individuals. 2/ A primary family is a group of two or more related persons including the head of a household. 3/ A primary individual is a household head living apart from relatives. 4/ Conterminous United States.

Source: 1960 Census of Population, Series PC(1)1B, tables 50 and 54; 1950 Census of Population, Vol. II, Part 1, tables 107 and 148; and 1950 Census of Housing, Vol. I, Part 7, table 8.

ITEM P

U. S. DEPARTMENT OF COMMERCE
BUREAU OF THE CENSUS
WASHINGTON 25, D. C.

December 1, 1961

NUMBER OF CHILDREN EVER BORN PER 1,000 WOMEN 35 TO 44 YEARS OLD, 1960, AND PERCENT CHANGE, 1950 TO 1960, BY COLOR AND URBAN-RURAL RESIDENCE, FOR AVAILABLE STATES

(Data based on 25-percent sample, 1960, and a sample averaging 2.4 percent nationally, 1950. Rate or percent change not shown where base is less than 200 women in 1960 or less than 4,000 women in 1950. Percent change not shown where less than 0.1. Minus sign denotes decrease)

Division and State	Urban white		Urban nonwhite		Rural white		Rural nonwhite	
	1960	Percent change, 1950 to 1960	1960	Percent change, 1950 to 1960	1960	Percent change, 1950 to 1960	1960	Percent change, 1950 to 1960
New England:								
Maine	2,530	24.0	...	...	2,999	14.1	...	...
New Hampshire	2,350	15.5	...	...	2,598	16.3	...	...
Vermont	2,456	20.2	...	...	2,962	8.7	...	...
Rhode Island	2,180	21.4	3,033	...	2,363	24.8	...	...
Connecticut	2,120	23.8	2,218	40.2	2,339	13.3	1,740	...
East North Central:								
Wisconsin	2,485	31.7	2,667	...	3,206	15.4	4,624	...
West North Central:								
Minnesota	2,579	37.3	2,477	...	3,367	18.8	5,462	...
North Dakota	2,770	39.4	...	...	3,516	17.4	5,910	...
South Dakota	2,780	29.6	4,942	...	3,223	20.2	4,721	...
Nebraska	2,445	35.2	2,760	...	2,949	14.3	4,115	...
South Atlantic:								
Maryland	2,193	29.5	2,367	27.1	2,549	6.7	3,131	3.7
Dist. of Columbia	1,273	8.2	2,048	44.5	...	...	...	...
Virginia	2,102	24.7	2,385	25.0	2,746	- 2.1	3,602	- 4.7
North Carolina	2,164	12.2	2,560	21.0	2,716	- 9.6	4,230	5.2
South Carolina	2,299	10.6	2,928	35.2	2,764	- 8.7	4,636	10.4
Georgia	2,263	19.4	2,866	34.1	2,967	...	4,899	25.9
Florida	2,124	29.6	2,720	49.6	2,673	- 1.1	3,910	24.8
East South Central:								
Tennessee	2,180	15.5	2,611	36.4	2,899	- 4.7	4,060	19.0
Alabama	2,284	20.5	3,081	29.5	3,107	- 4.1	4,876	21.2
Mississippi	2,318	29.7	3,402	36.8	3,022	8.5	5,251	36.3
West South Central:								
Louisiana	2,417	32.1	3,173	49.7	3,256	8.8	5,136	29.5
Oklahoma	2,258	17.1	2,993	44.2	2,961	- 2.8	4,485	19.8
Mountain:								
Montana	2,675	37.5	3,737	...	3,055	9.0	4,813	...
Idaho	2,819	26.8	3,286	...	3,402	8.9	3,515	...
Wyoming	2,596	18.3	...	...	3,080	20.4	...	...
Colorado	2,429	33.0	2,224	...	3,121	11.3	3,075	...
New Mexico	2,837	21.2	3,167	...	3,640	8.1	4,853	...
Arizona	2,594	10.6	3,252	...	3,017	11.7	5,110	...
Utah	3,188	23.2	2,995	...	3,806	6.7	4,867	...
Nevada	2,108	26.8	2,435	...	2,411	14.8	2,801	...
Pacific:								
Alaska	2,137	(1)	3,294	(1)	2,521	(1)	5,453	(1)
Hawaii	2,363	(1)	2,868	(1)	2,809	(1)	3,679	(1)

1/ Data not available for 1950.

Source: U. S. Census of Population: 1960, Series PC(1), table 51.

COMMENT ON TRENDS SHOWN BY POPULATION CENSUS

Talcott Parsons, Harvard University

The data presented in the excellent memorandum prepared by the staff of the Bureau of the Census for this panel seem in general to confirm the broad impressions many sociologists have of what, in the relevant respects, is going on in American society. They do, however, pin them down more precisely than has been possible before, and on the whole show that the changes are even more massive and rapid than we had thought they were. An overall index of this is of course the very large volume of geographical mobility as such, which is clear evidence that the country is not settling down to a stagnant phase in which change generally is slackening off.

From the point of view of its bearing on the social structure, the most striking development is the rapid increase of the "metropolitanizing" of the population, which has both shifted the rural-urban balance, which has been discussed for so long, substantially further in the urban direction, and has focussed attention on major change in the character of urban communities. Here of course the relative decline of the central cities and the very rapid increase in population of the peripheral parts of the metropolitan areas are the most salient phenomena.

The "industrial-metropolitan complex" as it may be called, has clearly come to a position of dominance in American society, and the end, of course, is not yet. I would like to attempt to conceptualize some of the salient features of this process under three headings, namely inclusion, differentiation and up-grading.

By inclusion I mean the process by which, once a developing new type of structure is relatively well established, increasing proportions of the population come to be included in it. On this point the evidence is striking indeed, On the positive side it is the increasing proportions which are involved in the new conditions of residence and work, while on the negative side the striking thing is the steady draining of the reservoirs of the older type of social organization, not so much rural in general as the types of rural community characterized by a tendency to subsistence agriculture, low educational levels, sometimes prominence of extended kinship groupings and the like - the Appalachian and Ozark highlanders and the rural Southern Negro are the prototypical cases.

Within the new social structure and between it and some of the other components there is then a process of differentiation. By this I do not mean the establishment of just any sort of diversity of social traits - many previous diversities have, as Dr. Taeuber remarked, been lessened in the course of this process. I mean rather the development of specific lines of differentiation between major components of the structure of the system, by virtue of which they develop relations of functional complementarity. This is true of urban and rural communities when the latter progress beyond the subsistence stage so that the countryside produces food and raw materials for urban-industrial populations and the latter produce components of a rural standard of living which would not be possible on a subsistence basis. Internally to the urban community it is the case for differentiation between sub-areas primarily devoted to residence and those primarily devoted to occupational organizations. On this basis it is my thesis that there has been a progressively increasing process of differentiation.

The third aspect of the process I have called upgrading. This is most tangibly illustrated by the case of the rural subsistence reservoirs. By migrating from such communities into urban-industrial ones these populations are both included in the dominant structures and are upgraded with regard to the level of qualifications for social function and expectations of performance. Perhaps the most obvious index is education. Those with the educational level of the traditional rural Southern Negro are obviously disqualified for any but the lowest level jobs in an urban-industrial system. The trend of occupational change, however, has been markedly in the direction, first of reduction of the proportion in the unskilled category, and second of increase of the highest level jobs, notably the professional managerial categories. The whole occupational system has been undergoing a process of upgrading, of which the upgrading of new migrants to the metropolitan communities - and especially of their children - is a major phase.

Agriculture and the extractive industries, which are the economic mainstays of the rural communities as they have been developing, are clearly becoming more differentiated and upgraded. The lower level, subsistence type part of the rural system is the one which has been losing population most rapidly. The high productivity areas of major market agriculture, have not been gaining population because mechanization and applied science have been reducing the need for numbers of manpower, but they have not basically been eroded as communities. The trend has been to performance of specialized functions in the society, and to performing them at increasingly high levels of efficiency.

The development of the metropolitan communities has thus been part of a differentiation process by which functionally higher level rural communities have also come to play a part. There are two other aspects of the rural side of the picture - by Census definition of communities of less than 2500 population. One is the increasing recreational use of rural areas by urban populations; in the broadest sense recreation has become one of the most important "industries" in our society - much of it of course urban, but much not. The other exception is the extreme fringe of the metropolitan areas where many suburbs are classed as rural because the particular political unit still remains under 2500, but in social character they are of course mainly urban.

The new metropolitan community as documented in these figures, may be said to constitute an important new phase in the process of differentiation of urban communities as social structures. The basic quantitative indices of this are the relative decline of the central cities and the figures of growth of the suburban rings.

The preindustrial urban community basically did not differentiate places of residence from places of work. Perhaps the prototypical figure was the artisan whose shop was in the same premises as his residence, but the same was true of the small merchant and the professional practitioner. The cities which developed after the industrial revolution then were organized above all about the differentiation between place of work and place of residence, with the central city being in the first instance the place of work. This of course was the period of urban public transportation systems, the street car, the elevated and the subway.

The new era is in part a function of motor transportation and superhighways and parkways. Not only have average distances travelled in the normal work-home cycle increased, but above all the automobile lends far greater flexibility because of its freedom from specified routes. In any case increasing proportions of the residential community are coming to be more specialized in residential function. This is concomitant with a substantial upgrading of housing standards, most notably the enormous increase of separate one-family dwellings, with yard space. Hence access to the outdoors, with yard and garden and absence of the "asphalt jungle" of the city streets is a major feature which belongs in the category of upgrading.

Meanwhile many of the previous functions of the older central city have been moving out and becoming further differentiated. A far greater proportion of employment opportunity is coming to be found on the outskirts of the city and of course shopping has been greatly decentralized so that the central city department store, though surviving, is of greatly reduced relative significance. Hence transportationwise, in place of the old pattern of movement in and out of the central city along the radii, there has developed a much more complex criss-crossing pattern between residential, industrial and other subcommunities over the whole range of the metropolitan circle.

This means in turn that the central cities have acquired far more specialized significance, centering on high concentrations of special function like banking and finance, legal services, government of course and a few others. These are cases where personal access of the most important people to each other is a great asset. New York city is of course the premier example. Such a concentration of course requires a substantial support base of service-facilities which makes them centers above all of clerical employment.*

* With respect to this whole picture of the developing metropolitan community I am much indebted to Dr. Winston White of Harvard University.

An interesting point of a special feature of this differentiation process is brought out in the data presented (Item O), namely that 16% of households "consist of primary individuals, mostly living alone" and not of families. Clearly this documents a new stage in the process of the "isolation of the nuclear family". The trend has been very strongly toward the extrusion from households which are primarily composed of families of persons who are not family members. The general decline in domestic service, particularly living-in, is well known. A second category which has declined greatly is that of "roomers", but in addition to this, relatives not members of the nuclear family, especially parents or siblings of one of the married couple, have greatly declined. The increase of households consisting of "primary individuals" is of course the other side of this coin. One of the very interesting features of it is the problem of how far these persons are isolated in the sense of "being abandoned" and how far they have gained a new independence, which under previous circumstances was not possible. The desirability of the change from the point of view of nuclear families seems to me on the whole to be clear, though not without important exceptions. I suspect that the independence is also valued by the "isolates" in a larger proportion of cases than is generally recognized. Partly this is because new modes of communication and transportation make it possible to combine independence of residence with the desired amount of contact to a considerably higher degree than previously.*

* The basic data on this problem area down to 1957 are excellently brought together and analyzed in Glick, American Families. This figure, however, represents a sharp accentuation on the trend which Glick emphasizes.

I would like finally to comment on one further major point. Along with the decline of relative magnitude of the central cities - taking account also of the processes of annexation - a very striking finding, made very clear in these data, is the major influx of Negroes - and partly, especially in New York of Puerto Ricans - into these central areas. The racial problem of course is in many respects a special one, and it is important not to underestimate the difficulties which the inclusion and upgrading of the nonwhite population in American society entails.

Nevertheless, because of the high salience of the particular case, it is easy to overlook the extent to which the problem of the urban nonwhite masses is primarily a class, rather than a "caste" problem, in a special sense. This is the sense in which it is not a static matter of a frozen status of an underpriveleged group, but one which fits into a dynamic pattern of "ecological succession". The reference point of course lies in the "reservoirs" the striking drainage of which is so clearly documented in this material.

The big city slum has long been the first port of call for the upwardly mobile immigrant elements into the urban world, at the turn of the century above all the peasant masses from Europe. Various indigenous elements have of course also been involved along the line. In the perspective of the general situation it seems clear that the rapid increase in the non-white lower class element in the cities is the latest phase in a process of long standing. The most important inference to be drawn from this is the high probability that it will prove to be temporary. The dispersion of the Negro into the new suburbia is as yet minimal, but there are clear signs that it is beginning. There is every reason to assume that the forces making for its acceleration are very powerful and that as early as the 1970 Census a major change will be visible. Two forces in particular operate in that direction, namely educational upgrading, and the erosion of the bottom layers of the occupational hierarchy, probably now accelerating because of automation.

The intriguing possibility is that this is not only the latest but will prove to be the last of the long series of "peasant" infusions into the American urban community. Indications point strongly in that direction. On the one side is the demonstrated absorbtive power of the metropolitan-industrial complex as I have called it; this clearly points toward the increasing predominance of "middle-class" patterns of life. On the other side is the exhaustion of the older reservoirs from which such population elements can come which lie within American jurisdiction, and the improbability that mass lower class immigration from outside will again be permitted. Considering the structuring of the "social problem" at least since the Civil War period, the gradual disappearance of the older type of urban lower class will certainly be one of the major social transformations of our time - incidentally one which was not foreseen by the social prognosticators of the last generation and is barely being considered even now.

From the point of view of the general social scientist the development of the U.S. Census is indeed a great boon. It is one of the developments which has contributed most to placing the task of empirical generalization about our society on a solid empirical base. I am indeed grateful to have had access to these early compilations of the 1960 results and to have been given the opportunity to suggest a few points about their significance in the light of sociological analysis of the trends of development of our society.

VII

DEVELOPMENTS IN SCIENTIFIC AND TECHNICAL PERSONNEL STATISTICS

Chairman, Philip M. Hauser, University of Chicago

THREE YEARS AFTER THE HAUSER COMMITTEE REPORT ON SCIENTIFIC AND TECHNICAL PERSONNEL DATA

Thomas J. Mills, National Science Foundation

Introduction

Before launching into this topic, I would like to make clear that I have no intention of presenting a carefully balanced objective presentation of this subject. The topic of scientific and technical personnel is such a broad one, and involves so many different programs and approaches that only with the greatest difficulty can one individual become thoroughly conversant with all aspects of it. Such observations as are contained in this paper are prepared from the vantage point of one who has been intimately concerned in Government with collection and analysis programs in this area for several years. Necessarily many important non-Government programs will not be discussed, and time limitations will not permit a full treatment of even all those within Government. However, I do not want to let the opportunity pass without saluting the fellow-laborers in this vineyard; particularly some of those in universities, in the professional societies and in research organizations who are taking an increasing interest in studies of scientific manpower. We are delighted to share the work with them, and very much welcome the contributions they are making.

The "Hauser Committee" Report of 1958

The so-called Hauser Committee Report was the product of a special Advisory Panel appointed jointly in 1957 by the National Science Foundation and the then existing President's Committee on Scientists and Engineers. Both the Foundation and the President's Committee had found their programs handicapped by the lack of authoritative information on the supply and requirements for scientific manpower. The Bureau of the Budget, properly concerned with increasing Government survey activities in this field, had also requested the Foundation to "develop a program for collection of needed supply, demand, employment, and compensation data with respect to scientists and engineers and for such other professional groups as it considers appropriate" at about this same time.

These developments led to the appointment of the Advisory Panel in June 1957. The Panel was requested to review and evaluate available information, and to recommend such inquiries or other measures as it believed would provide additional information needed for policy formulation in the field of scientific and technical personnel.

The membership of the Advisory Panel reads like a "Who's Who" of those concerned with this topic. Membership was purposely selected to be broadly representative of non-Government organizations in order to secure as favorable a reception of its findings by non-Government groups as possible. Members were Philip M. Hauser, University of Chicago and our respected chairman for this session; Philip H. Coombs, then of the Ford Foundation and now Assistant Secretary of State for Educational and Cultural Affairs; Henry David of the National Manpower Council and now President of the New School for Social Research; Colman R. Griffith of the American Council on Education and now at the University of Illinois; M. H. Trytten of the National Research Council; Ralph J. Watkins of Brookings; and Dael Wolfle of the American Association for the Advancement of Science. Secretariat services were provided by Surveys and Research Corporation, and five Government agencies, including the National Science Foundation, the President's Committee, and the Departments of Labor, Commerce, and Health, Education, and Welfare, supplied consultant services.

The Committee completed its deliberations and reported to its sponsoring agencies in the Spring of 1958. Based on this report, the Foundation released its recommendations as "A Program for National Information on Scientific and Technical Personnel" which has come to be known as the "Hauser Committee Report."

The Report reviewed briefly the growing needs for data on scientific manpower, evaluated the existing sources and types of data, and recommended a program of additional research and surveys covering about 15 topics. A "most urgent" priority rating was assigned to three projects in the categories of definition and classification of scientific personnel and jobs, periodic establishment surveys of the resources of scientific and technical personnel, and periodic studies of demand. Other recommendations in the Report ranged from sample population surveys to research on community attitudes toward scientific personnel. Some of its recommendations were broad in scope; others dealt with rather narrow specialized surveys.

It is not our purpose at this session to catalog in detail the many ways in which these recommendations have been implemented over the past three years. Detailed treatment is being given to the selected topics of periodic establishment reports, studies of demand, and the extremely interesting topic of Census-related studies by the other participants. This paper will touch on some of the ways in which the programs have been financed and coordinated. Some attention will also be given to an overall appraisal of progress and the importance of the Hauser Committee Report as a landmark in the identification of data needs for Government decision making.

Program Development

Soon after the release of the Report, the National Science Foundation was designated by the Bureau of the Budget to act as the "focal agency" for developing the recommended program. After the Federal Government English is unwrapped from this expression, we can interpret it as meaning to act in a leadership role with the Government agencies in planning, determining priorities, and publication of data in this field. This designation has simplified the task of developing a coherent program embracing many topics and involving the programs of a great many other agencies, Government and non-Government.

It is well known that the Government's statistics program operates on a decentralized basis involving many different agencies in its collection aspects. Data relating to scientific and technical manpower are no exception to this rule. For example, the U. S. Office of Education is the agency generally collecting statistics on education, the Department of Labor usually collects data on employment, and the Bureau of the Census is most frequently concerned with surveys of population. The usual pattern is that the responsible individual agency seeks appropriations from Congress for that part of the program for which it considers itself responsible. The advantage of this method of financing is the close association between responsibility for performance and accountability for funds. The disadvantages include the possibility that specialized needs may be lost sight of or subordinated in larger general programs and a lack of flexibility in finding financing for new programs on short notice.

The National Science Foundation has been able to utilize the resources available to it for manpower data and studies in ways which strengthen the existing institutional arrangements. For the most part its resources are used to support, through transfer of funds, the logical extension of other agencies' programs into the desired fields. Such support for continuing series will normally be provided for an initial period of perhaps two or three years. Upon completion of this period, support will usually be withdrawn with the understanding that the program extension has become of sufficient importance to the other performing agency as to warrant direct appropriation for the purpose. Support is likewise frequently provided for single-time studies. Under such arrangements, the performing agency benefits through the additional resources provided to it as contrasted with the alternative prospect of finding a competitor in its field. The general program benefits through the ability to move resources into priority areas as they develop on relatively short notice in contrast with the ponderous machinery required by regular budgetary cycling. This principle is working well with respect to several surveys and studies being performed in the Bureau of Labor Statistics and Office of Education.

Transfer of funds appropriated to the Foundation has also been instrumental in utilizing the resources of some non-Government, especially non-profit, agencies to perform some of the research and studies recommended by the Hauser Committee Report. Such agencies, particularly the professional societies, universities, and research organizations, frequently have highly competent research staff who can be interested in a wider program when a relationship can be found to their own specialized preoccupations. In this way, advantage can be taken of the services of well-qualified staff, many of whom would not be available to Government on a more conventional basis. The Foundation has been able to support studies on science teacher qualifications and teaching loads at the American Association for the Advancement of Science and on traits and characteristics associated with successful careers in science at Columbia and Harvard Universities to cite a few examples. Currently, an experimental program at the American Institute of Physics is being supported to determine the practicability of developing in a professional society setting a research and information center dealing with manpower problems in specific disciplines.

Let it be clearly recognized that any Government program of studies and research can only be successful if the Congress--especially the Appropriations Committees--is willing to authorize the necessary appropriations to make it so. The programs of the Foundation--as well as those of other Government agencies--in this area have been singularly fortunate over the past few years in this regard. Growing interest in the Congress in the twin topics of science and scientific manpower has been generally effective in producing relatively sympathetic appropriation actions. There is no question but that this willingness to provide the additional funds necessary is primarily responsible for the very considerable progress which has been made.

Program coordination has been directed toward attaining a maximum Government return from each new study, and at the same time imposing as minimal a burden as possible on respondents. For example, the establishment surveys of scientific personnel (recommended by the Hauser Committee) have been designed to meet the specific data needs of not only the Foundation, but of the Atomic Energy Commission and the National Institutes of Health as well. Special tabulations of the National Register of Scientific and Technical Personnel are prepared to meet the needs of other agencies, including the Civil Service Commission, the Public Health Service, and the Department of Labor. A genuine effort to lighten the reporting burden on respondents includes adoption of standard definitions, uniform reporting dates, and generous use of advisory groups.

The Present Status of Hauser Committee Recommendations

Progress in carrying out the recommendations of the Hauser Committee Report has been extensive. Without being exhaustive in this brief review or encroaching on the papers of my colleagues on this program, it can be said that some actions have developed with respect to each recommendation. Many agencies and programs have had a part in this effort, some of which has perhaps been performed without any particular reference to the Report or, in some cases, possibly without knowledge of it. However, most of the recent developments can be attributed to the general acceptance of the recommendations of the Report as constituting the elements of a desirable program. Particularly noteworthy in this respect are some of the studies of the Office of Education in the areas of enrollments in science and technical training at all levels of the formal educational system; sources and extent of support for graduate training in the sciences; and a new study of science offerings in the non-public secondary schools. Establishment reports of the Bureau of Labor Statistics, the Office of Education, and the Civil Service Commission now provide an annual series on employment of scientists and engineers. The 1960 Census related study being developed with the Bureau of the Census and the National Opinion Research Center holds great promise of providing bench mark data on the pool of science manpower. Better information on the transition between the baccalaureate and enrollment in graduate study or entry into the labor force is available from recent studies by the Bureau of Social Science Research and the Women's Bureau. Staff studies by the Foundation have investigated such topics as the extent of retention in our educational system of our most able students; the contribution of immigration to the American science manpower pool; and the estimated demands for scientific personnel implicit in some of the rapidly expanding Government supported programs.

Lest our picture of progress appear too complacent, I would like next to touch upon some areas urgently requiring more attention. These include both the topics for which data are needed and the additional occupational groups which would be included.

Among the priority subject matter topics for which research is desired is the development of measures which will permit assessment of quality. For example, our science manpower resources are frequently classified into groups by educational attainment; i.e., doctorate, baccalaureate, or no degree. Under this system any possessor of a bachelor's degree is considered equivalent to any other as are the doctorate holders. Yet it is well known that ability-wise, the lowest graduate of some of our colleges will be better trained on any reasonable comparative basis than the highest in some other colleges. Scientific advances are usually attributed to a mere handful of individuals. In assessing manpower resources in a given field, how does one measure this qualitative factor? Similarly, how might one assess the quality of science instruction being offered at our institutions of higher education?

Another topic for exploratory research involves the utilization of scientific and technical personnel. Frequently, apparent shortage situations are assessed as "poor utilization." What kind of measures might be proposed as providing indexes of utilization practice? Such measures would be particularly useful in any assessment of feasibility of scientific programs from the standpoint of manpower resources.

The decisions now being made on almost a daily basis by the science program managers in the aggregate affect significantly the economic and social fabric of the Nation. An annual research and development expenditure of $14 billions now represents a little less than 3 percent of GNP. Scientists and engineers now constitute about 2 percent of the civilian labor force. Implementation of the Hauser Committee Report recommendations are now making possible a more rational basis for decision making in this area.

The mobility of scientific and technical personnel is a suggested third topic which requires study. The speed with which technological advance is pervading the economy means that university training is not able to anticipate or even keep pace with specialized manpower requirements. Yet manpower requirements are met, if slowly in some cases, through occupational mobility supplemented by short-term training as witnessed by the way manpower requirements were met in nuclear reactor engineering in an earlier period and in oceanography at present. What is the extent of such mobility, and between what areas of science can it be anticipated are the kinds of questions for which answers are sought by the analyst evaluating a proposed program calling for substantial demands upon our science manpower resources.

Our knowledge of manpower resources in the social science and technician occupations is another area of relative weakness. We now have relatively little information on the numbers and characteristics of the work force in the social sciences, an area which promises to require rather large numbers, and which will be in sharp competition with the other professions for well qualified and trained youth. Technicians are one of our fastest growing occupations, and yet, we know little on how many we have or how many are being trained. In both these cases, work has been handicapped by a lack of agreement of what should be included in such occupations.

Conclusion

In conclusion, it may be fairly said that the Hauser Committee Report constitutes an important landmark in the development of information on the Nation's scientific and technical personnel resources. Remarkable progress in developing the recommended programs has been achieved in the past three years. While much, of course, remains to be done before the comprehensive program recommended by the Report is fully implemented, it is now possible to estimate the general dimensions of many of the troublesome problems previously little known. It is particularly fortunate that this ability does now exist and the Report, coming when it did, provided a certain amount of lead time in developing data which make it possible.

Finally, the future should hold a mechanism for comprehensive review and evaluation of our situation in this area before very long if we are to avoid "hardening of the arteries", which is likely to affect program no less than people. It is suggested that such review ought to take place within the next two years, but in any event not later than 1965.

PERIODIC ESTABLISHMENT SURVEYS OF EMPLOYMENT IN SCIENCE AND ENGINEERING

Robert W. Cain, National Science Foundation

The starting point for the quantitative appraisal of the supply of scientific and technical manpower is the number of persons employed as scientists, engineers, and technicians and as teachers of these personnel. In fact, this was recommended in the Hauser Committee Report1/ to be one of the "most urgent"projects--that is, the basic counts of persons who are classified as scientific and technical personnel. The term "establishment surveys" as used here refers to surveys of employers and is not confined to "establishment" in the industrial classification sense.

GENERAL FACTORS There are of course varying methods used in the collection of scientific manpower information. Each of them has advantages in given situations, and the most appropriate method applied depends upon several factors such as timing, precision required, and relative expense. In general, the respondent chosen must possess the information or be in a position to secure it and must be willing to supply it.

Collection of data directly from individuals is especially appropriate where detailed information regarding professional or personal characteristics is desired. This includes such information as age, educational attainment, or employment and professional history. Foremost among such programs of collection are the population surveys of the Bureau of the Census and those made in connection with the National Register of Scientific and Technical Personnel of the National Science Foundation.

Employers and employing units or establishments are usually the appropriate respondents for data collections dealing with employment of personnel. There are several reasons for this. The employer represents a repository of information on the individuals in his employ. Also, he judges on a relatively disinterested basis for all employees their field or discipline, their work activity, and their actual employment status.

Two other general factors are considered in the collection of scientific manpower information --the completeness of coverage of respondents and the collection method used. In each case it must be determined whether a universe or partial-coverage survey is to be employed. Complete-coverage surveys are generally more costly and more time consuming. However, in some cases where a high degree of precision or a large volume of detailed characteristics is to be counted, the universe approach is used. In most cases a partial-coverage method is utilized in employer-reporting surveys, particularly when a carefully selected sample base can be employed. Personal interviews, mail questionnaires, or telephone and telegraphic surveys are all applied in collection of information. By and large, however, most surveys of employment of scientific and technical personnel use mail questionnaires, frequently preceeded by detailed pretesting and supplemental interviewing of large employers or those with complex situations.

This paper refers primarily to the National Science Foundation's program in the area of establishment surveys of scientific and technical personnel. There are many organizations, both Government and private, which have contributed much to our knowledge of the Nation's resources of scientists and engineers. Some of the surveys made by these organizations have been supported financially by the Foundation, others have been carried out in coordination with NSF studies; others have been less directly connected with the Foundation or other Government agencies but have contributed valuable data and experience. Though the primary purpose of Foundation-sponsored establishment surveys of manpower is to provide basic counts of employed individuals, they have provided a basis for other types of manpower studies. One such study is the intensive analysis made by the Bureau of Labor Statistics of employment in mathematical work in industry and Government. On the basis of returns to the regular surveys of employment of scientific and technical personnel, the Bureau surveyed employers of mathematicians and persons engaged in mathematical work. These employer surveys of course are essential to such research work as the recently completed study by the BLS of The Long-Range Demand for Scientific and Technical Personnel.2/

Basically, establishment surveys have been made in the several broad sectors of the economy which employ scientists, engineers, technicians, and teachers in these fields. These sectors are: Private industrial concerns, colleges and universities, the Federal Government, State and local governments, and nonprofit organizations. Such surveys have been carried out on varying bases. The largest employers--industry, educational institutions, and the Federal Government--have been surveyed more frequently than other sectors. Furthermore, in the last decade scientific manpower data collection has been in a developmental or even experimental phase. Each survey has been more comprehensive and more accurate. As concepts and definitions have been clarified, as the need for additional data has arisen, survey questionnaires and techniques have been developed to meet these situations. As many "old hands" know, many of our well-established surveys in the general employment area, or in other data-collection programs, have finally been accepted only after many years of trial and error.

1/ A Program for National Information on Scientific and Technical Personnel, NSF 58-28, National Science Foundation, 1958.

2/ The Long-Range Demand for Scientific and Technical Personnel: A Methodological Study, NSF 61-65, National Science Foundation, 1961.

INDUSTRIAL SURVEYS The first studies of industrial employers of scientific and technical personnel which can be termed comprehensive started in the 1950's. One of the pioneering projects was a nationwide survey of some 2,000 industrial companies made by the Bureau of Labor Statistics for the Department of Defense in 1952.3/ This survey collected information on research and development scientists and engineers in companies having large defense contracts or who were otherwise large employers of such personnel. As part of the first series of studies conducted shortly after it was organized, the Foundation sponsored a more extensive Bureau of Labor Statistics project in the first comprehensive survey of research and development expenditures and manpower in industry.4/ Also collected at that time were 1953-1954 data on scientists and engineers in all activities and on technicians in research and development. A similar survey collected information on research and development expenditures and manpower for the year 1957.

In 1959, the Foundation supported the first survey oriented directly to employment of scientific and technical personnel, followed by a similar study in 1960; both were carried out by the Bureau of Labor Statistics.5/ A sample of about 11,000 companies selected from the lists of the Bureau of Old Age and Survivors Insurance (as were the 1953-54 and 1957 surveys) was surveyed, and employers were asked to report their scientists and engineers by field and by work activity or function.

Up to this time in these industrial surveys data were collected on a companywide basis. Thus, all employment for each reporter was classified in the one industry believed to represent the major activity of the company. This procedure has obvious defects in dealing with multi-industry companies, and it was decided after consultation with many officials in industry and Government to request reports from a sample of establishments as the best method of improving the industry classification of data. Therefore, the sample for the 1961 survey was drawn from lists of establishments reporting in State Unemployment Compensation programs, and was stratified by industry and size of company. In addition, the sample was supplemented from lists of industrial laboratories and other establishments known to employ large numbers of scientific and technical personnel. All together, nearly 15,000 units were included in the questionnaire mailing list. The response to this survey was excellent; replies were received from about 90 percent of the units in the sample. About 10 percent of these establishments were included in consolidated reports of some 200 companies. A 1962 survey is now in its final planning stages and will be conducted along the lines of the 1961 survey. Questionnaires probably will be mailed in early February 1962.

Another development in connection with these industrial surveys is that beginning with 1962 they will be supported through regular Department of Labor appropriations. This is a welcome development in that it enables the Department to carry out these studies in conjunction with its established programs and at the same time enables the Foundation to use its funds to support other areas of data collection.

COLLEGE AND UNIVERSITY SURVEYS In past years, several types of studies have covered the area of employment by colleges and universities. For example, the Office of Education has regularly collected information on the numbers of personnel employed in these institutions. However, these surveys have not provided information on the fields or occupations with which the Foundation and others are concerned. In 1955, the Foundation as a part of its first series of studies conducted a survey covering academic year 1953-54.6/ This survey provided the first comprehensive data on employment of scientists and engineers in these institutions together with information on the types of activities in which they engage. Since this 1953-54 survey, the Foundation in cooperation with the Office of Education carried out a study of research and development expenditures and professional manpower in educational institutions covering academic year 1957-58 (the manpower information related to April 1958).7/ At the present time, the Foundation and the Office of Education are cooperating in a similar study covering academic year 1960-61. This latter survey includes a detailed breakdown of academic disciplines.

FEDERAL GOVERNMENT SURVEYS In this area, the National Science Foundation has relied on the regular collection program of the Civil Service Commission. Periodically, the Commission has collected information on the occupations of employees of the Federal Government. Comprehensive surveys were made in 1931, 1938, 1947, 1954, and 1957. In 1954, the NSF in cooperation with the Commission collected additional information on the activities of Federal employees related to science--such as research and development, scientific data collection, scientific information, and testing and standardization.8/ In 1958, the

3/ Scientific Research and Development in American Industry, A Study of Manpower and Costs, Bulletin 1148, Bureau of Labor Statistics, 1953.
4/ Science and Engineering in American Industry, Final Report on a 1953-54 Survey, NSF 56-16, National Science Foundation, 1956.
5/ Scientific and Technical Personnel in Industry 1960, NSF 61-75, National Science Foundation, 1961.

6/ Scientific Research and Development in Colleges and Universities, Expenditures and Manpower; 1953-54, NSF 59-10, National Science Foundation, 1959.
7/ Scientists and Engineers Employed at Colleges and Universities, 1958, NSF 61-38, National Science Foundation, 1961
8/ Scientific Manpower in the Federal Government, 1954, NSF 57-32, National Science Foundation, 1957.

the Commission began a regular annual collection, in the fall of each year, and the NSF has included a supplemental form covering those Federal personnel engaged in or administering research and development. Surveys were carried out in 1958, 1959, 1960, and 1961.9/ The Commission is now developing plans for a special "work history" sample of Federal employees. The general outline of this study calls for a sample selected from each agency on the basis of social security numbers. The Foundation hopes the sample will be large enough to permit detailed tabulations for selected scientific occupations.

STATE GOVERNMENT SURVEYS In 1954, the Foundation supported a pilot study of scientific activities in six State governments. The study was conducted by the University of North Carolina and covered California, Connecticut, New Mexico, New York, North Carolina, and Wisconsin. However, until recently there had been no comprehensive study of the employment of scientific and technical personnel in State governments. In 1959, the Foundation supported the Bureau of Labor Statistics work in the first full-coverage survey.10/ This survey covered some 3,300 separate agencies in all 50 States. At the present time, the BLS is completing plans for a second such survey to be conducted in early 1962.

NONPROFIT ORGANIZATION SURVEYS The area of nonprofit organizations was also covered in the first series of studies of science activities for the period 1953-54. The emphasis in this area as in others at this time was on research and development activities. Research institutes and commercial laboratories were surveyed by Syracuse University. Battelle Memorial Institute surveyed trade associations, professional and technical societies, and other cooperative groups. Similar coverage was made of philanthropic organizations in a study made by Russel Sage Foundation; however, no manpower information was collected from these latter organizations. In late 1958, the Bureau of Labor Statistics under sponsorship of the Foundation conducted an integrated survey of nonprofit organizations, collecting employment information on scientific and technical personnel for January 1957 and January 1958.11/ The survey covered philanthropic foundations, voluntary health agencies, research institutes, professional and technical societies, science museums, zoological and botanical gardens, and arboretums. The Foundation is now starting another series of surveys in the nonprofit area. A survey of private foundations is now underway, and other organizations will be surveyed in the near future.

9/ Scientists and Engineers in the Federal Government, October 1958, NSF 61-43, National Science Foundation, 1961.
10/ Employment of Scientific and Technical Personnel in State Government Agencies, Report on a 1958 Survey, NSF 61-17, National Science Foundation, 1961
11/ Scientific Research and Development of Nonprofit Organizations, Expenditures and Manpower, 1957, NSF 61-37, National Science Foundation, 1961.

LOCAL GOVERNMENT SURVEYS Data on employment of scientific and technical personnel by local governments are almost completely lacking. The only effort to date consisted of a pilot project carried out by the Bureau of Labor Statistics in six States. This study was primarily an attempt to determine the availability of information on scientific and technical personnel in local government organizations. It is hoped that at least a limited survey in this area will be launched next year--perhaps covering major local governments.

SELF-EMPLOYED SECTOR SURVEYS Employment of scientific and technical personnel by private consulting firms has been covered in the industrial surveys. However, aside from the limited information available from the Census Bureau and from the National Register almost nothing is known about the volume of employment in this sector. In one sense, it should not be included under the general approach of employer-reporting, or establishment, surveys. This employment--that is, the number of scientists and engineers who are not connected in some manner to an employer--is believed to be relatively small.

PROBLEMS CONNECTED WITH EMPLOYER-REPORTING SURVEYS
In any data collection program continuing problems exist to harass the collector. The following is a review of some major problems--not given in terms of how the problems have been solved or not completely solved, but simply as an exposition of them.

Definitions. What is an engineer, a scientist, a technician? What is an employer of such personnel? How can research, development, administration be best defined?

Coverage of occupations. It has been mentioned that the Foundation generally includes the following in its surveys: Scientists, engineers, technicians, and teachers of these personnel. Are we concerned with social scientists? Yes, but this has been a more recent movement and presents some special problems. What is a social scientist in industry, in Government, in a university?

Sampling of employers. Since universe surveys probably will not be used in most situations, how should the samples be chosen? In some areas a universe listing of employers does not exist. How is the list kept current with the rapidly expanding number of newly emerging employers of scientists and engineers? What are the problems connected with changing industrial classifications?

Periodic reporting. How often should each major sector be surveyed? In some areas, employment can shift fairly rapidly, particularly in defense-supported industries. In the college and university area, what is the best time in the academic year for reporting?

Coordination with other surveys. Several agencies of the Federal Government and other nongovernmental organizations have special requirements for data. Can all or most of these requirements be accommodated in one survey? Is it better to expose a respondent to one long, detailed questionnaire or to several relatively specialized ones?

Geographic information. How can the need for information on geographic location be met? To date such questions have not been included in employer surveys directly. Some provide this information--such as the colleges and university sector--but when carried out on a sample basis, some accuracy is lost.

Relationship to surveys of individuals. Surveys of individuals have been regarded as much more appropriate when dealing with personal or professional characteristics. How can information from the Census Bureau surveys or the National Register or other organizations be correlated with employer surveys?

"Crash" studies. No matter how well a manpower data collection program is planned, there probably always will be special needs. Such a special case faces the Foundation and other Federal agencies at the present time. How many scientists and engineers are engaged in work which is directly related to the planned space programs? How many additional personnel will be so engaged with certain expenditures of funds over the next 10 years? How many of the Nation's scientists and engineers are involved in health-related research?

FUTURE PLANS FOR EMPLOYER SURVEYS Very simply, the NSF's program of studies is to provide, within resources provided by Congress and with resources of other organizations, a comprehensive and current analysis of scientific and technical personnel in all sectors of the economy. The Foundation seeks to identify needs for information, build on existing programs of data collection, and select priority areas for support of new information collection. The Foundation generally does not carry out surveys itself but rather supports, wherever necessary and appropriate, existing data collection programs of other organizations. In recent years, the several Federal agencies concerned with scientific manpower information have joined to cosponsor certain projects. Our goal then is to have available current information on the Nation's resources of scientific and technical personnel on a continuing basis. In some areas this will require annual (or even more frequent) surveys; in others less frequent data collection will be required.

STUDIES OF DEMAND FOR SCIENTIFIC AND TECHNICAL PERSONNEL

Harold Goldstein
Bureau of Labor Statistics, U.S. Department of Labor

It is ironic that greatest current interest attaches to long-range projection of demand for precisely that small part of the working population for which projections of demand are most difficult to make--scientific and technical personnel. Demand for most types of workers can be related to the occupational requirements of an existing technology in meeting an economic demand, and to projections of the volume of that demand, by means of traditional methods of economic analysis. If we are aware of the latest scientific discoveries and technological innovations, we can make judgments as to their possible effects on existing occupational patterns. When the new technology is introduced in the first plant in an industry we have an actual model of the occupational patterns toward which the whole industry will be moving as the other plants are modernized over a period of years. It is a good deal more difficult to get clues as to the technological factors affecting demand for the very workers who are at the forefront of technological change--the scientists who discover new principles and the engineers who develop new applications.

Nevertheless, projections of scientific and technical personnel requirements are needed as much as, or even more than, projections of requirements for other occupations: they are critical for economic growth and national defense; expansion of educational facilities has to be planned in terms of the output of engineers and scientists required 8 to 10 years hence; and this planning has to take place immediately in the face of an expansion in the college age population in the mid-1960's. To illustrate the urgencies here, the number of 18-year olds--the age of entrance to college--will increase by 1 million, or over one-third, in a single year, 1964 to 1965.

Because the need is so pressing, we must cast about for means. On closer examination, we find that a major part of the demand for scientific and technical personnel can be related to some measure of economic activity that can be projected independently. For that large number of scientists and engineers engaged in teaching, requirements can be based on projections of population and of college enrollments, allowing for all of the variables in the ratio of faculty to students. Requirements for engineers and scientists in work directly related to the industrial process--in production supervision, design, analysis, etc.--are affected by the levels of activity in the various industries. Only one-third of industry's scientists and engineers engaged in research and development are somewhat remotely related to current production activity. Here the volume of research which will be supported (by government, business, universities, etc.) is the determining factor and we may hope that even this is not completely unpredictable.

In an attempt to develop a methodology for making these projections, the National Science Foundation supported a research study by the Bureau of Labor Statistics which has just been published by the Foundation.1/ The purpose of the study was to analyze the problem of making projections, break it down into its component parts, and lay out a systematic methodology which would be capable of selective improvement as better or more up-to-date data become available, and which would point to areas where further research is needed. It should be understood that the purpose of a methodology for long-range economic projections is not to serve as a mechanism for producing estimates untouched by human hands, so to speak, but rather to marshall the relevant data so that better judgments can be made as to the future demand.

The approach used was what has been called a "segmental method" in the report of an advisory Committee to the National Science Foundation on research on scientific manpower.2/ In the past, demand for scientists and engineers has been projected on the basis of trends in the relationship of total employment in these occupations to the total population, or the total labor force, or the gross national product. These projections of broad trends have been subject to a substantial margin of error because the quantities dealt with were so large and general that one had little basis for exercising judgment with respect to future trends in the relationships. The theory behind segmental studies of demand is that when one carefully examines the trends in employment of scientific and technical personnel within each economic segment--each sector of industry, colleges and universities, Federal, State and local government agencies, nonprofit organizations, etc.--one can draw upon a knowledge of the factors behind these trends to make the projections more realistic. Such segment-by-segment projections can be reviewed from time to time in the light of current statistics on employment and economic activity in each segment; reasons for differences from the projected trend can be examined, and revised projections can be made with much greater assurance that we are close to hard reality than would be possible with an overall projection of demand. In this procedure there is danger of ignoring an industry which is virtually unknown today but which may be a major employer of scientific and technical personnel in 10 or 20 years. However, actual experience shows that it is rare than an unknown industry becomes a major employer over so short a period.

In the examination of trends in scientific and technical employment in each sector of the economy, we try to relate this employment to an independent measure which is causally related to demand in the occupation, which can be projected independently, and for which we have data for the past that make it possible to examine the relationship and project it into the future. In most cases, the best measure is total employment in the sector. In the case of colleges and universities, the best independent measure is total enrollments.

It will be obvious that by projecting future demand by a relationship of past employment of scientists and engineers to some other variable we are explicitly assuming that the number of scientists and engineers employed has been equivalent to the demand, and that there was no unmet need. In the absence of statistics on job vacancies in these fields, we cannot say that this is entirely a valid assumption. It might be argued, for example, that if shortages existed past employment reflects supply rather than demand. On the other hand, in engineering at least, there has always been a substantial influx of persons without an engineering degree, many of them men without any substantial amount of college-level education at all. This suggests that supply is flexible and that over the last decade levels of employment have reflected something above the total supply and closer to the "true" level of demand.

Using total employment in each industry as the independent measure against which to relate scientists and engineers is a method that has been followed not only in the United States, but also in France and Great Britain.3/

In order to develop projections by industry for use in this study, the Bureau of Labor Statistics used an economic model which it had developed for its general program of occupational manpower requirements projections. This model was based on a projection of the population by the Bureau of the Census, a projection of the total labor force, and a projection of the gross national product which would be generated by this growing labor force. Conditions of high levels of business activity--of "full employment"--were assumed. The growth in economic activity over the decade of the 1960's, as measured by the GNP, was computed at about 50 percent and the growth of the labor force as a whole was estimated at about 20 percent. Projections of employment in each industry were developed on the basis of the demand that would be generated for each industry's product under these economic conditions. It was found that employment in such important technical-personnel-employing industries as electrical equipment, metal products, machinery, instruments, and chemicals is expected to increase rather more than the average for all industries.

In each sector past trends in the ratio of scientific and technical personnel to total employment were examined. We had data on employment of scientific and technical personnel by industry covering only a brief span of time. The first survey sponsored by the National Science Foundation was for 1954, and subsequent surveys provided data for the years 1957, 1958 and 1959. Some rough comparisons can be made with data from the Censuses of Population for 1940 and 1950, but the most reliable and consistent data are available only for these few recent years--a period of rapid growth in scientific research and development. This is a major limitation of the present study which should be corrected as time goes on. Surveys of scientific and technical employment in industry are being continued on an annual basis by the Bureau of Labor Statistics and we should be able to make much better judgments on trends in this employment as we develop more experience.

The trend shown by the four surveys in the ratio of scientific and technical personnel employment to total employment was projected for each industry separately to 1970. It was possible to make a careful evaluation of the accuracy of these projections in only two industries--chemicals and electrical equipment. For these industries much published and unpublished data on employment and related factors were studied and a large number of interviews were held with leading research and other officials in major companies. In these interviews the trends in employment of each scientific and technical occupation were were discussed, the factors affecting their future employment were explored, and a judgment was arrived at as to future changes. When these were summarized it was found that the results supported the original projections in the case of the chemical industry, but in the electrical equipment industry a somewhat lower level of employment was indicated that had been originally projected. This experience suggests that similar interviews should be carried on--although perhaps not as extensively--in every major industry employing scientific and technical personnel. Since the projections were heavily influenced by the trends in the 1954-59 period, it is particularly important to determine whether this period is representative of the longer-range trend.

The conclusions of this first model are, therefore, to be viewed as tentative and subject to revision. They may, however, be of some interest. As I remarked earlier, the labor force is expected to grow by about 20 percent over the decade of the 1960's. Our other studies have suggested a growth for all professional and technical occupations, as a group, of about double this rate, or 40 percent. The present study showed scientific and technical occupations growing at double the rate for all the professions, that is, at about 80 percent. Demand for engineers was projected to increase from about 780,000 in 1959 to 1,480,000 in 1970, or about 90 percent. Demand for scientists is projected to grow a bit more slowly, from 310,000 in 1959 to 550,000 in 1970. Among the sciences, demand for mathematicians and physicists is expected to grow the fastest in this period, each of them more than doubling. Least growth is expected for geologists.

To translate these projections of growing demand into terms on which policy decisions can be made, we need to look at the implications for the educational institutions. To begin with, in order to project the number of trained persons the colleges and universities would have to produce to meet this growth in demand, it is necessary to allow for the replacement of scientists and engineers who die or retire. Deaths and retirements were estimated from data on the age distribution of engineers and scientists by application of age-specific death and retirement rates developed from actuarial tables of working life for all males.4/ It was estimated that in the 1960's 17,000 engineers and 4,000 scientists may be expected to die or retire annually. When these figures are added to the annual average net growth in demand for engineers (64,000) and scientists (21,000) it may be estimated that 80,000 persons will have to enter engineering annually and 25,000 will have to enter scientific employment annually to meet the projected demand.

What are the prospects for this many entrants? In engineering, one must allow for a substantial influx of persons who do not have engineering degrees. The number of such entrants has amounted to 23 percent of the total entrants in recent years. If this pattern should persist in the 1960's, some 18,000 persons would come into the profession in this way annually, leaving 62,000 entrants to be provided by the engineering schools. It does not seem likely that the output of engineering schools will come anywhere near this average level over the decade. In recent years the number of freshmen enrolled in engineering has declined, while total freshmen enrollments in all higher education increased. On the basis of these freshmen enrollments, the graduates in engineering (with bachelors' degrees) are likely to be a few thousand below the 1960 level of 38,000 in each year through 1965. Only by a heroic increase in graduations for the remainder of the decade could an annual average of 62,000 graduates be attained. While an increase in output of graduates should be sought, it will be necessary for industry, government, and other employers of engineers to improve the utilization of their present engineering staffs, rather than count on an expanded supply alone.

With respect to scientists, on the other hand, it was calculated that if the total college enrollments expand as projected by the Office of Education (a projection which assumes that a slightly increasing proportion of the rapidly growing population of college age will attend college), and if the proportion of students selecting scientific careers remains the same as in 1960, the output of students will be adequate to meet the demand. However, there will be differences by field of science, and most certainly by level of training. There would be no surplus of science graduates to meet the shortage of engineering graduates. Needless to say, there is little ground for complacency in this projection; an increase in enrollments of the magnitude projected would require very great expansion of facilities and recruitment of additional competent teachers--and this should start at once.

These projections suggest that a substantial increase in the output of engineering graduates from the Nation's higher education system is needed, without cutting into the output of science graduates. The opportunity to attain this increase arises after the middle of the decade, when we will get a one-third increase in our 18-year-old population. However, the factors behind the recent decline in engineering enrollments, despite a rise in total freshman enrollments as well as other aspects of student motivation, are worth looking into before a substantials expansion in engineering education is projected.

Since the purpose of the study was to develop methodology and since it was found that in many areas much more data and research are needed, the above conclusions should be viewed as illustrative projections only. The National Science Foundation is supporting a continuation of this study in which more precise methods and more up-to-date data can be used to revise various components of the projection. On the demand side, a revised economic model will be developed and projections of employment for each industry will be made which reflect recent economic and technological developments. In examing trends in the ratio of scientific and technical personnel in each industry, data for 1960 and 1961 (if available) will be examined, and to the greatest extent possible separate examination will be made of the trends in employment in each occupation individually within each industry. These trends will be reviewed in interviews with leading executives in some of the major industries. Attempts will also be made to examine projections of research and development expenditures.

On the supply side, recent trends in college enrollments and graduations will be examined, the results of a recent follow-up survey of college graduates made for the National Science Foundation by the Bureau of Social Science Research5/ will be examined to see what patterns are developing in the employment choices of graduates in each field, and other sources such as 1960 Census of Population will be used if available.

These improvements should make the projections somewhat more reliable. We view this activity as a continuing one, since projections need to be reviewed from time to time and lessons drawn from experience in sharpening out judgments and techniques. In this perspective the present study should be viewed as a first step.

1/ "The Long-Range Demand for Scientific and Technical Personnel--A Methodological Study," National Science Foundation, NSF 61-65, 1961.

2/ "A Program for National Information on Scientific and Technical Personnel," National Science Foundation, NSF 58-28, August 1958.

3/ "The Long-term Demand for Scientific Manpower," Advisory Council on Scientific Policy, October 1961, London, H.M. Stationery Office, Cmnd. 1490. See also "Forecasting Manpower Needs for the Age of Science," Organization for European Economic Cooperation, Office of Scientific and Technical Personnel, Paris, September 1960.

4/ "Tables of Working Life: Length of Working Life for Men," Bureau of Labor Statistics, U.S. Department of Labor, Bulletin No. 1001, August 1950.

5/ "The 1958 College Graduate--Two Years Later," Bureau of Social Science Research, Inc., Washington, D. C., 1960 (draft of a report prepared for the National Science Foundation).

A PROGRAM OF CENSUS-RELATED STUDIES OF SCIENTIFIC AND TECHNICAL PERSONNEL

Seymour Warkov, National Opinion Research Center

The Advisory Panel to the National Science Foundation and the President's Committee on Scientists and Engineers recommended that a survey be undertaken of ". . . a large sample of persons recorded in the 1960 Census enumeration as college graduates or as persons currently or last employed in scientific and technical positions, whether college graduates or not, to determine relationships between training and subsequent occupation."[1] Seeking to implement this recommendation, the National Science Foundation requested that the National Opinion Research Center, affiliated with the University of Chicago, prepare a planning statement[2] on a series of post-enumeration studies of scientific and professional workers and college graduates.

The planning statement provided the basis for implementing this recommendation of the Advisory Panel. These studies will employ as their basic reference point the 1960 Census of Population which classified one-fourth of the population by a number of key characteristics such as education and occupation. Using the Decennial Census as the frame, samples ranging in size from one thousand to five thousand persons per occupational title are being drawn from over forty scientific and professional occupations classified as "Professional, Technical and Kindred" workers under the three-digit occupational code. Once these persons are identified, information collected in the 1960 Census will be supplemented at comparatively low cost by mail questionnaire later this winter.

The questionnaire, now undergoing extensive field testing, covers four topics: first, job duties; second, career patterns and job mobility; third, training history; and fourth, social and professional characteristics.

<u>DUTIES</u>: Certain important questions about professional and other workers and their job activities cannot be adequately answered by the present Census occupational classification system. For instance, the word "research" is not included in the scheme, and "researchers" are scattered among many different categories. Although elementary and secondary teachers are separated, college professors and instructors are not grouped by level or type of school, and "mathematics professors and instructors" cover a range from junior college to advanced graduate level teaching. The Census also cannot reveal information about multiple jobs (the teacher who also consults, the engineer who also teaches) and "joint appointments."

Perhaps even more important, however, is that for many jobs, the exact meaning of the job title is unknown. Although great effort has gone into detailed studies of what factory workers and farmers do on their jobs (time and motion studies), little is known about the actual job requirements and duties of many professional jobs. Thus, John Smith's job as a "research chemist" may or may not involve teaching classes, training of assistants, consulting to industry and government, writing proposals for research, contract negotiations, project administration and purchasing, hiring and firing, reading up on the literature, writing reports, and so on. Similarly, what engineers, accountants, economists, and technicians "really do" is not actually known.

It may be assumed that each occupational label actually covers a range of these clusters of activities. Under the umbrella of "engineering" may be found people who are mostly administrators or salesmen and people who are mostly technicians and draftsmen. Similarly, college professort range from the classroom teacher to the man who essentially directs a research organization; and psychologists include people who counsel and give therapy to troubled individuals (or groups), experimental physiologists and "human engineers" who design and improve equipment.

What these professionals do in terms of their particular cluster of duties is only one aspect of their jobs; a full picture entails additional information about the organization of their work. Historically, the professional was a "solo practitioner" who provided services to a particular client or groups of clients. He either was self-employed, or like the teacher, worked alone even though employed by an organization. Today, professional work increasingly is done in groups. The research team in the laboratory, group practice, program directors, section leaders: all these groups and/or titles reflect a growing organization of professional work. In order to understand these jobs, then, we must know whether the professional has a superior, collaborating professionals and subordinates.

Therefore, we plan to consider the details of job activities, and the organizational context of work to provide information on these questions. Each subject in the study will be asked to indicate what duties were part of his job and

---

[1] "A Program for National Information on Scientific and Technical Personnel," <u>National Science Foundation</u>, NSF 58-28, p. 6.

[2] I am indebted for many of the ideas presented at this session to James A. Davis, Senior Study Director, National Opinion Research Center, who prepared the planning statement.

which ones are key elements. A number of analyses may be made from such data.

First, by re-weighing the stratified sample of respondents, it should be possible to make national estimates for 1960 of the proportion of professional or highly educated workers in terms of categories not covered by the Census: e.g., the proportion involved in research, the proportion who teach, the proportion who are involved in administration, etc.

Second, these materials can enhance the utility of data derived from the 1960 Census. By factor analysis or analogous statistical techniques it may be possible to describe dimensions of activities or clusters of activities which will aid in interpreting the Census materials. Thus, it may be possible to find sub-types within a Census category differentiated in type of work, or by inter-correlating Census categories over activities (i.e., comparing pairs of jobs in terms of similarity or difference in activities), locating occupations which, although classified separately, have similar patterms of work.

In order to assess interpersonal environments of professional work, questions are directed at three levels of organizational hierarchy. Above the respondent, we wish to know whether he has a boss, and the key elements in the work activity of this superior. At the same level as the professional, we ask whether there are coworkers and what they do. At the administrative level below the respondent, questions are asked about his subordinates, if any, and their key activities. These data will permit a description of the division of labor among professionals and scientists within the context of the employing organization.

TRAINING: High levels of advanced training are the hallmark, in fact the definition, of most professional and technical jobs. Compared with the general population, the people in these jobs have such high levels of education that to consider training as a problem seems, at first glance, a false issue. However, many of the fields in question are developing at such a rapid rate that professional training as little as five or six years old may be outmoded, unless supplemented by more recent ideas and knowledge. Also, if it is correct that professional jobs are actually complexes of specific duties rather than applications of single skills, many professionals may be well trained for parts of their jobs, but undertrained for others.

As a first step in identifying relationships between training and subsequent employment, we plan to collect information on formal training in detail beyond that of years completed as reported in the Census. Data can be gathered on the specific degrees and fields of specialization, both graduate and undergraduate, specific institution, and sources of financial support for undergraduate or professional training. Second, we include questions concerning "brush-up" training by asking about correspondence courses, in-service training programs, apprentice training and so on. When all these data are matched against specific occupations and clusters of job duties, they can give a fairlyprecise description of the training of American scientists and technicians. Furthermore, statistical analysis of job histories and background characteristics of the respondent can provide insight into the processes accounting for differences in levels of training and duties performed on the job.

CAREER PATTERNS AND MOVEMENT: Much more needs to be known about processes over time by which trained people are allocated to various jobs and institutions, the career paths which characterize a profession and the flow of professionals between various types of employers. Procedures to secure such information include the following: As a start, schedules will be collected early in 1962. Thus, a number of the respondents will no longer be working at the jobs which they reported to the Decennial Census field worker. It will, of course, be necessary to ask these changers about their April 1960 jobs to make the analysis in terms of the occupations enumerated in 1960. This situation will provide a natural design for studying job changes, as it will be possible to contrast changers and non-changers, and to contrast the old and new jobs of changers.

At the same time, the two year interval will not give enough of a time spread to develop findings on long-run career patterns. Therefore, each respondent will be asked selected items of information on his first full time job held after reaching age 25. It should be stressed that these materials will not be complete job histories, since it is not uncommon for civil engineers, for example, to have had so many jobs in so many places that a complete job history would require a lengthy questionnaire. Rather, it is planned to get brief over-all characterizations of the mobility history since finishing school sufficient to separate respondents into broad categories: those who have always had the same job with the same employer; those who have always worked for the same employer, but in different types of jobs; those who have done essentially similar work, but for a large number of employers, etc.

Some possible uses of these two sets of data are as follows: First, the recent mobility data will provide an assessment of the flow between various employers, geographical regions, and functions. Are more professionals moving from academic to industrial jobs than vice versa? Are more people moving into administration than out of it? Are certain regions of the country attracting a high share of the mobile professionals? These and similar questions may be answered from the current mobility data.

The retrospective job history data also will make possible the establishment of typical and variant career histories for specific occupations and for respondents with specific types of training. They should give clues on continuity in professional careers, in the routes to particular destinations in the job market (how does one get to be a project director, consultant, professor of medicine, government project director?), and by comparison of the early jobs of older and younger respondents, perhaps some insight into historical shifts in the career patterns of professionals. Taken as a whole, these findings should tell much about the interaction of supply and demand, and perhaps locate high and low turnover situations which can be exploited to improve the recruitment of professionals.

Finally, the questionnaire will permit us to supplement the current fund of information in the following three areas: (1) Professional characteristics: This includes memberships in professional associations, information which will help assess the validity of membership lists in professional societies as inter-censal indices of professional size. (2) Fertility: This includes the number and ages of the respondent's children as a possible factor in social mobility and as a measure of the reproductive rates of a crucial segment of the society. (3) Occupational attitudes: Although this research is not an attitude survey, a few questions are asked about the relative importance of, and satisfaction with, selected aspects of occupations. These data should permit us to identify the cluster of values which characterize specific occupations and provide additional clues on continuity of employment and turnover among persons employed in April 1960 as engineers, natural scientists, physical scientists, etc.

Although a somewhat truncated description of the scope of these studies has been presented today, clearly the data to be gathered in this post-censal survey of scientific manpower will be relevant for any formulation of policy which seeks to relate training to subsequent employment in these fields. Furthermore, we expect that the data will aid substantially in advancing a comparative analysis of the professions in American society.

VIII

RECENT FINDINGS IN HEALTH STATISTICS

Chairman, Paul M. Densen, New York City Department of Health

THE INCREASING MORTALITY FROM CHRONIC RESPIRATORY DISEASES

Harold F. Dorn, National Heart Institute, National Institutes of Health

A declining death rate is generally assumed to be an almost inevitable concomitant of increased medical knowledge and improved public health and medical practices. If pressed, thoughtful persons may admit that man is not immortal and that the death rate must either cease to decrease or else decrease at an almost imperceptible rate at some future time; but probably few persons believe that this is, in fact, imminent.

Although reliable mortality statistics are not available for the entire period, the death rate among people of western European origin probably has been declining for at least two centuries. Apart from the acute communicable diseases primarily of early childhood and adolescence, diarrhea, enteritis, and dysentery, the group of diseases for which one of the largest relative declines in mortality has occurred are those of the respiratory system. This group includes such familiar disorders as tuberculosis, influenza, pneumonia, bronchitis, emphysema, cancer of the lung, and acute upper respiratory infections such as the common cold.

At the beginning of this century, 26 percent of the deaths in the United States were attributed to these diseases of the respiratory system. Three of them, influenza and pneumonia, and tuberculosis were the two leading causes of death. By 1940, the percentage of deaths attributed to this group of diseases had dropped to 11. In 1959, diseases of the respiratory system accounted for only 8 percent, or one out of every 12 deaths in the United States. The numbers of deaths reported as caused by specific diseases of the respiratory system during 1959 are shown in Table 1.

Table 1. Number of deaths attributed to diseases of the respiratory system, United States, 1959.

Diagnoses		Number of deaths
001-008	Tuberculosis	10,627
162,163	Cancer of lung	34,302
160,161,164	Other respiratory tract cancer	3,111
241	Asthma	4,984
470-475	Acute upper respiratory infections	971
480-483	Influenza	2,845
490-493	Pneumonia	52,194
500-502	Bronchitis	3,840
526	Bronchiectasis	2,263
527.1	Emphysema	7,728
525	Non-occupational chronic interstitial pneumonia	3,570
510-524, 527.0, 527.2	Others	6,220
Total		132,655

Source: Publications of the National Office of Vital Statistics.

Asthma has been included since some physicians may use this term for disorders diagnosed by others as emphysema, bronchitis, etc.

Pneumonia still is the leading cause of death among the diseases of the respiratory system, but tuberculosis which formerly ranked second has been displaced by cancer of the lung. Indeed tuberculosis now causes fewer deaths than a group of respiratory disorders which, for want of a better term, have been referred to as chronic nonspecific respiratory diseases. Included in this group are conditions diagnosed as bronchitis, bronchiectasis, chronic pulmonary fibrosis, chronic interstitial pneumonia, and emphysema. These diseases, which are assigned numbers 500-02, 525, 526, and 527 in the International Statistical Classification of the World Health Organization, caused 18,825 deaths during 1959. There are good reasons to include bronchial asthma in this group which would raise the number of deaths to 23,809. However throughout the remainder of this discussion I shall group together the conditions coded to numbers 500-02, 525, 526, 527 of the International Statistical Classification and shall refer to these rather loosely as chronic respiratory diseases.

In the past these diseases have not received much attention in medical literature in the United States. There seems to have been a feeling that these terms should be used as a diagnosis of last resort and applied to conditions for which a more precise and definite term is not appropriate. Whatever the reason, these disorders have never been recorded as important causes of death in the United States until the past decade, in contrast to Great Britain where bronchitis has been an important cause of morbidity and mortality for many years.

The rather slight medical importance attached to bronchitis and emphysema has been reflected in rules for classifying and tabulating death certificates on which these diseases were entered as primary or contributory causes of death. Prior to 1949, the principal or underlying cause of death shown in national mortality statistics was determined by an arbitrary hierarchy of diseases shown in the Manual of Joint Causes of Death. Bronchitis, emphysema, and similar diseases had a low priority in this hierarchy, so that when these were entered on a death certificate together with a more "important" disease the latter was selected as the underlying cause irrespective of the attending physician's statement.

Some information concerning the effect of these procedures on the number of deaths that were classified as caused by these diseases in official publications of mortality statistics can be obtained from special tabulations of contributory causes of death for 1925 and 1940. In 1925, the number of deaths for which chronic respiratory diseases were classified as a contributory cause of death was 74 percent greater than the number for which these were classified as the primary or underlying cause.

In 1940, the number of death certificates with chronic respiratory diseases classified as contributory was 2.5 times the number for which these were considered primary. The principal diseases that took precedence over chronic respiratory diseases were cardiovascular diseases and other respiratory disease exclusive of tuberculosis. Corresponding tabulations for years subsequent to 1940 have not been published but there is no reason to believe that there has been an increase in the proportion of all death certificates with a diagnosis of one or more of these chronic respiratory diseases for which these diseases have been selected as the underlying cause of death. On the contrary, it is possible that this proportion may have decreased and that the number of deaths for which chronic respiratory diseases are considered to be a contributory cause of death may now be three or more times the number for which these are considered to be the primary cause.

An additional factor affecting the comparability of mortality statistics for this group of respiratory diseases over the past three decades is changes in their classification in the International List of Causes of Death and its recent revisions, The International Statistical Classification. The code numbers assigned to these diseases in recent years are as follows:

Diagnosis	7th Revision (1958-)	6th Revision (1949-1957)	5th Revision (1939-1948)	4th Revision (1930-1938)
Bronchitis	500-502	500-502	106	106
Chronic interstitial pneumonia	525	525	114	114
Bronchiectasis	526	526	106	106
Emphysema	527.1	527.1	113	113

In the Fourth and Fifth revisions of the International List, number 114 also included silicosis, pneumoconiosis, gangrene of lung, and abscess of lung which are assigned to Nos. 521, 523, and 524 of the Sixth and Seventh revisions. The number of deaths assigned to emphysema, No. 527.1, in the Sixth and Seventh revisions was not published by age, sex, and color for 1949 to 1957 inclusive so that No. 527 as a whole has been used unless otherwise specified. In addition to emphysema, No. 527 includes deaths assigned to atelectasis, hernia of lung, stenosis and ulcer of the bronchus, acute pulmonary edema, and mediastinitis. Nearly 85 percent of the total number of deaths assigned to this number are attributed to emphysema.

The number of deaths assigned to the following code numbers have been used to compute the death rates shown in this paper: Sixth and Seventh revisions, 500-502, 525, 526, 527; Fourth and Fifth revisions, 106, 113, 114. The death rates for 1930-32 and 1939-41 were multiplied by .774 and those for 1948 were multiplied by .810 to correct for the inclusion of diseases that were not assigned to Nos. 500-02, 525, 526, and 527 of the Sixth and Seventh revisions.

A major change in the method of coding and classifying causes of death was incorporated in the Sixth revision. Not only was the entire classification changed but the procedure for selecting the underlying cause of death also was changed. The Joint Cause Manual was discontinued and the selection of the underlying cause of death was based on the statement of the physician. No attempt has been made to adjust the mortality rates shown here for lack of comparability brought about by the changes in coding and classifying causes of death that followed the introduction of the Sixth revision. Instead, rates are shown for 1948, the last year during which the Fifth revision was used, and 1949-51, the first three years of use of the Sixth revision. Most of the difference in the magnitude of these two rates may be attributed to the change in procedures following the introduction of the Sixth revision.

Even though the Joint Cause Manual was abandoned at the end of 1948 as the principal basis for selecting the underlying cause when two or more diseases are entered on a death certificate, a number of joint cause coding rules have subsequently been adopted by the National Office of Vital Statistics for use when the sequence of diseases given by the certifying physician is considered to be implausible. The only basis for appraising the effect of these rules upon the number of deaths assigned to the respiratory diseases discussed in this paper is a special tabulation showing the number of deaths assigned to each code number from a 10 percent sample of deaths occurring in 1958 and which were coded by both the Sixth and Seventh revisions. There was no change in the classification numbers for these diseases in the two revisions so that the difference in the number of deaths assigned to each cause may be attributed to changes in rules for selecting the underlying cause of death.

The following numbers are the ratios of the number of deaths assigned to each cause following the rules for the Seventh revision to the corresponding number assigned following the rules of the Sixth revision. The ratios which are from unpublished data provided by the National Office of Vital Statistics are:

500-02	Bronchitis	106.7
525	Other chronic interstitial pneumonia	99.4
526	Bronchiectasis	114.7
527	Other diseases of lung and pleural cavity	85.1

For the group as a whole the ratio was 0.95.

The increase in the number of deaths assigned to bronchiectasis was due primarily to the fact that deaths attributed to both bronchitis and bronchiectasis, which had been assigned to bronchitis by the rules of the Sixth revision, were assigned to bronchiectasis by the rules of the Seventh revision. The net increase in the number of deaths assigned to bronchitis resulted from a change in coding rules whereby a combined diagnosis of asthma and bronchitis, which formerly had been assigned to asthma, was assigned to bronchitis. It is apparent that substantial changes in the number of deaths assigned to specific causes of death may be brought about by changes in coding rules even though the basic classification of disease remains unchanged.

Since it is believed that a clear distinction in usage of the terms, bronchitis, bronchiectasis, emphysema, pulmonary fibrosis, chronic interstitial pneumonia, etc., is not made by practicing physicians, and since changes in the coding rules and the system of classification used in computing mortality statistics have shifted the assignment of a significant number of deaths from one diagnosis to another at various times in the past, it is probable that the most valid presentation of the trend in the death rate is given by considering this group of diseases as a unit. Although death rates will be shown for certain specific diseases the main emphasis will be upon the trend in the death rate for the entire group of diseases coded to Nos. 500-02, 525, 526, 527 of the International Statistical Classification.

The trend in the age-adjusted death rates from the main diseases of the respiratory system is shown in Figures 1 and 2. The sharp

Fig. 1. United States White Males

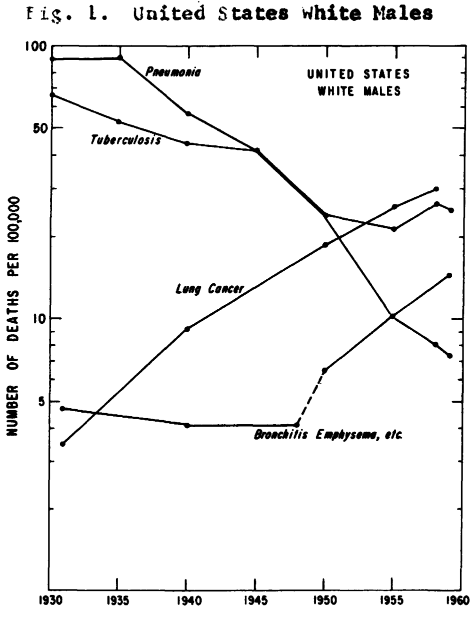
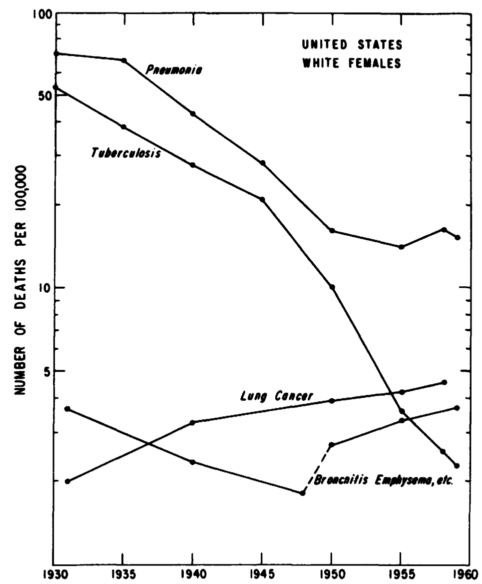


Fig. 2. United States White Females



decline in the death rate from pneumonia and tuberculosis is well-known. The controversy concerning the relationship of cigarette smoking to lung cancer has attracted widespread attention to the increase in the death rate from this disease. The recent increase in the death rate from chronic respiratory diseases including bronchitis and emphysema is not so well-known.

To what extent does the reported increase in the death rate from this group of diseases since 1950 reflect a real increase in incidence? This question cannot be answered accurately from existing information. It is obvious from the sharp increase in the death rate when the Sixth revision was introduced that procedures for the classification and coding of death certificates have greatly affected the reported trend in the death rate. The increase in the mortality rate for this group of diseases that has been recorded since 1949 almost certainly began before that date, but its beginning has been forever concealed by past methods of classifying official mortality statistics. It is plausible to suppose that the death rate from chronic respiratory disease as determined from death certificates has been increasing for as long as two or three decades but that the rate of increase may have accelerated during the past decade.

Granted that the number of death certificates with a diagnosis of bronchitis, emphysema and similar disorders has increased in recent years, is this simply a reflection of improved methods of diagnosis and a change in nomenclature and habits of diagnosis on the part of practicing physicians combined with a decline in the death rate from tuberculosis and pneumonia, two major respiratory diseases whose presence may have masked bronchitis, emphysema, and allied conditions in the past? These factors undoubtedly account for part of the reported increase but any estimation of the relative magnitude of their effect can represent only an expression of personal opinion since objective data on this point are not available.

Raymond Pearl suggested that the respiratory system is more susceptible to disease than any other organ system. It is in continuous contact with the external environment and is not as well adapted to resist the action of harmful agents as is the skin which is similarly exposed. This diathesis may be expressed by the development of several specific diseases. If the lethal form of any one of these diseases is controlled, it will soon be replaced, in whole or in part, by an increase in the fatal form of some other respiratory disease.

The data in Figures 1 and 2 may be interpreted either as supporting or not supporting this hypothesis. There is no indication that the death rates from lung cancer and chronic respiratory diseases among white females, in the near future, will even reach the level of the death rates from pneumonia and tuberculosis recorded as recently as 1950. The contrary is true for white males. If the present rate of increase continues, the combined death rate from lung cancer and chronic respiratory diseases by about 1970 will reach the level of the combined death rate from tuberculosis and pneumonia in 1940.

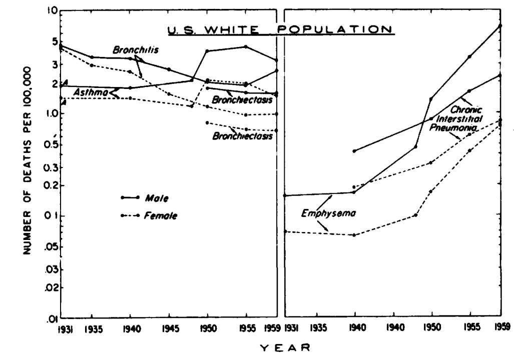
It seems fully as plausible, however, in the present state of knowledge to interpret the sex difference in the trend of mortality from these diseases as being not so much a manifestation of a constitutional susceptibility to disease as a response to specific exogenous agents.

The number of deaths from bronchitis, emphysema, and allied diseases currently reported in official vital statistics is a gross understatement of the number for which these diseases have contributed to the fatal outcome. The number of deaths for which these diseases are coded as the primary or underlying cause may be no more than one-third of the number for which they are listed as one of the causes of death. A tabulation of multiple causes of death is needed to provide more precise information on this.

The recent trend in the death rate from chronic respiratory diseases resembles that for cancer of the lung. The rate of increase is much greater for males than for females. The increase in the death rate among females is no greater than that which might reasonably be attributed to more accurate diagnosis and improved case finding. The amount and rate of increase for males in comparison to those for females are so great that a real increase in the incidence of these diseases is plausible even though the proportion of the increase that may be real cannot be stated with precision.

The trend of the death rate since 1930 for some of the specific diseases included in this group of chronic respiratory diseases is shown in Figure 3. A consistent increase has occurred

Fig. 3. U. S. White Population



only for emphysema and chronic interstitial pneumonia. The death rate for emphysema shown in Figure 3 is based on deaths assigned to No. 527.1 of the International Statistical Classification which is defined as emphysema without mention of bronchitis. If both emphysema and bronchitis are mentioned on a death certificate, this death is assigned to No. 502.0, bronchitis with emphysema, and is tabulated with other deaths assigned to chronic bronchitis. The number of death certificates with both emphysema and bronchitis is about 15 percent of the total number showing emphysema with or without mention of bronchitis.

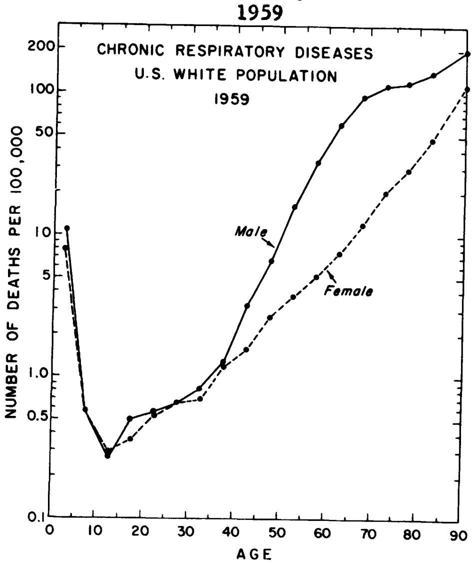
The term, chronic interstitial pneumonia, includes diseases coded to No. 525 of the International Statistical Classification of which the principal ones are fibrosis of the lungs, chronic pneumonia, chronic inflammation of the lungs, cirrhosis of the lungs, and interstitial pneumonia. All of these are presumed to be nonoccupational.

It is evident from Figure 3 that emphysema and chronic interstitial pneumonia are rapidly becoming the most frequently reported diagnoses among this group of diseases. In spite of the fact that current coding practices generally give preference in the selection of the underlying cause of death to other diseases the combined number of deaths assigned to these two diseases now is greater than the number assigned to respiratory tuberculosis.

Until nomenclature and usage become more uniform, the most valid representation of the trend in mortality from chronic respiratory diseases will be given by combining the diseases now coded to Nos. 500-502, 525-527 of the International Statistical Classification. However information concerning the number of deaths assigned to each of the diseases coded to these members as well as the number of deaths for which these diseases are coded as a contributory cause of death should be published in order to provide data for judging the effect of coding practices and variations in diagnostic terminology upon the trend in the death rate.

Mortality from this group of diseases is relatively low until about 40 years of age (Figure 4). The death rate for children less than 10

Fig. 4. Chronic Respiratory Diseases
U. S. White Population
1959



years of age is considerably greater than that during adolescence but this is due to acute bronchitis which is not the same as the disorders diagnosed as chronic bronchitis, emphysema, etc., of late adult life.

After age 40 the death rate increases rapidly especially for white males for whom the rate at ages 65-69 is nearly 100 times that at ages 35-39. Prior to age 40 there is very little difference in the death rate for males and females, but after this age, the rate increases much more rapidly for males. The largest relative difference in the death rate for the sexes occurs between 60 and 70 years of age when the rate for males is 8 times that of females.

__Summary__

1. The death rate from diseases of the respiratory system has declined sharply since 1900. In 1900 twenty-six percent of all deaths were attributed to these diseases; in 1959 only 8 percent were attributed to these diseases.

2. During recent years, the death rate for certain respiratory diseases that formerly were unimportant as causes of death has been increasing rapidly. The first of these diseases for which an increase was observed was cancer of the lung which now causes more deaths than any other respiratory disease except pneumonia.

3. During the past decade, the death rate has increased significantly for a group of chronic respiratory diseases including bronchitis, bronchiectasis, pulmonary fibrosis, chronic interstitial pneumonia, and emphysema.

4. Due to lack of uniformity in the use of these diagnoses and to changes in the classification and coding rules used in the preparation of mortality statistics, the length of time that the death rate for these chronic respiratory diseases has been increasing as well as the level it has reached cannot be accurately ascertained.

5. The number of deaths attributed to these diseases in 1959 was nearly twice the number attributed to tuberculosis. Even so, this number possibly represents only about one-third of the total number of deaths for which this group of diseases contributes to the fatal outcome.

6. Most of the increase in the death rate is due to an increase in the number of deaths attributed to emphysema and chronic interstitial pneumonia.

7. The death rate for this group of chronic respiratory diseases is low prior to age 40, but it increases nearly 100-fold between the age groups 35-39 and 65-69 for white males. The death rate for males greatly exceeds that for females after age 40.

__Legends to Figures__

__Figure 1.__ Age-adjusted death rate per 100,000 white male population from pneumonia (490-493), tuberculosis (001-019), lung cancer (162,163), and chronic respiratory diseases (500-502, 525-527) United States, 1930-32 to 1959.

__Figure 2.__ Age-adjusted death rate per 100,000 white female population from pneumonia (490-493), tuberculosis (001-019), lung cancer (162, 163), and chronic respiratory diseases (500-502, 525-527), United States, 1930-32 to 1959.

__Figure 3.__ Age-adjusted death rate per 100,000 white population by sex from asthma (241), bronchitis (500-502), chronic interstitial pneumonia (525), bronchiectasis (526), and emphysema (527.1), United States, 1930-31 to 1959.

__Figure 4.__ Number of deaths per 100,000 white population by age and sex from chronic respiratory diseases (500-502, 525-527), United States, 1959.

SOME STATISTICAL PROBLEMS INVOLVED IN A ONE-DAY CENSUS OF HOSPITAL PATIENTS

J. D. Colman and Nathan Morrison, Associated Hospital Service of New York

This is a report on work in progress, and presents some of the statistical problems that have arisen and the methods that are being used in studying them.

A brief description of the New York hospital census will be helpful as background. On May 10, 1961, all the short-term general care hospitals in the lower 17 counties of New York State, including New York City, Nassau, Suffolk, and Westchester, participated in the preparation of single-page case forms for each patient in the hospital at 11:59 p.m., and double-page forms for each patient who was discharged alive or who died on that day.

The single-page form (which was also the first page of the double-page form) contains items on the patient's demographic characteristics (age, sex, color, marital status, place of residence), admission diagnosis, date of admission, surgical procedures, if any, type of hospital accommodation, and method of payment. The second page of the double-page form contains items on the tests and treatments received by the patient, drugs administered, nursing care, and some summary figures on the hospital charges.

Each of the 230 hospitals in the 17 county area submitted forms for their cases as they were completed over a period of time (which began on May 11, 1961) without having to meet any deadline. We were more interested in accuracy and completeness than in speed. These forms have been carefully checked and edited and, when necessary, returned to the hospitals for correction or for additional information.

All types of hospitals, voluntary, proprietary and governmental, (including veterans' administration and public health service hospitals) in the area cooperated in this census. Hospitals which had both long-term and short-term divisions were asked to report only on patients in the short-term divisions.

A total of 51,000 cases are included in the study, 46,000 involving patients who were in the hospital at 11:59 p.m. on May 10th and 5,000 persons who were discharged or who died on that day. As of December 28, 1961, about 50,000 of the case forms have been received from the 230 hospitals, and edited, coded, and keypunched. Complete data have been received from 220 hospitals. About a thousand more cases are expected to come in during the next three weeks. Most of these are cases sent back for correction. The cutoff has been fixed as January 19, 1962.

On the basis of pilot analysis of the records of a few selected hospitals, machine tabulation has already been started.

The study has been a community enterprise from the beginning and could only have been carried out with the whole-hearted cooperation of all of the hospitals. The forms, procedures, tabulation plans, and statistical codes were developed in consultation with representatives of many institutions, both governmental and voluntary, experienced in working with such data. The study attempted to utilize and profit from the experience of the Professional Activities Survey, the Cleveland area study of 1956-57, the Michigan study (undertaken in 1958, and recently completed) and all past studies in this country and in Canada of a similar nature.

For all of the 51,000 hospital patients in the study, the single or first page form yields information on about 16 basic variables. For the 5,000 persons who were discharged or who died on May 10, 1961, the second page of the form yields additional information, some of which will need to be analyzed carefully on a qualitative rather than a quantitative basis. The data on the drugs administered offer many challenges, for example. However, it will be possible here to use the experience gained in two pilot studies in Michigan and in New Jersey.

In studying the data for the 51,000 cases, the approach will be to analyze the basic variables separately and in various combinations and to search systematically for meaningful interpretations.

Some of the basic variables are dichotomous (e.g., sex: male or female, and race: white or non-white.) Other variables have many more than two possible values (e.g., admission diagnosis: several hundred possible codes, and age: from the newborn to persons aged 100 years.)

It is planned to use a variety of statistical techniques in analysing the data ranging from frequency distributions, centering constants, to any appropriate multivariate analysis technique such as point biserial correlation coefficients for determining the degree of relation between a dichotomous variable and a continuous variable, multiple and partial correlation, factor analysis, discriminant functions, and canonical correlation.

For the 5,000 persons who were discharged or who died on May 10, 1961, the date of admission - (the actual hour of admission and discharge is shown on the form) - is available and thus the length of stay distribution will be tabulated for these complete stays.

For the 46,000 persons who were still in the hospitals at 11:59 p.m. on May 10, the length of stay distribution will represent, of course, incomplete stays of various kinds.

A review of the published work and of some unpublished materials on hospital stay distributions reveals a number of promising approaches that can be taken in analyzing the two types of distributions arising out of the May 10th census.

The useful and interesting analogy between actuarial life tables and hospital stay tables has been mentioned in several forums and is explored in a 1953 article in the American Journal of Public Health by Robins and Sachs in which admissions, discharges, and patients remaining in hospitals are treated similarly to the births, deaths, of a living population in a life table.

Reference should also be made, in this connection, to a 1957 article in the American Journal of Public Health by Mortimer Spiegelman entitled "The Versatility of the Life Table."

An interesting approach is presented in a paper (not yet published) by Clifford A. Bachrach entitled "Estimation of Length of Hospital Stay from Discharge Data." One sentence may be quoted from this paper: "Given a distribution of length of stay of discharged patients, the stationary population model permits us not only to predict the mean stay-to-date, but also permits us to estimate all other characteristics of the distribution of length of stay-to-date of patients remaining in the hospital." It will be possible, of course, to test this approach by using the length of stay distribution of the 5,000 discharged cases to predict the length of stay distribution of the 46,000 persons remaining in the hospital on May 10th, and comparing the prediction with the actual experience.

It will also be possible to apply the approach suggested in two 1959 papers by Balintfy and Flagle of Johns Hopkins. Balintfy's paper is entitled "A Stochastic Model for the Analysis and Prediction of Admissions and Discharges in Hospitals." His abstract, at the beginning of the paper, states:

> "Frequency distributions indicate that the length of stay of any particular patient in a hospital may be regarded as the joint effect of a large number of causes acting in ordered sequence during the time of recovery and defining the probability of discharge as a stochastic function of the length of treatment. Making use of this relation, statistics of expected figures of daily discharges and admittances have been developed as conditional values at given occupancy. The effect of chance and seasonal fluctuation, and the possibilities of smoothing is demonstrated by the calculation of the method of forecasting expected occupancy levels.
>
> "The paper reveals the basic statistical laws governing the admission, discharge, and length of stay distributions in hospitals. It proves that both admissions and discharges are compound Poisson processes following the negative binomial distribution and that length of stay follows the logarithmic normal distribution. This information is used to predict the daily census and indicate the direction of further research."

At the conclusion of his paper, Balintfy suggests the possibility of applying Markov processes in describing the variation in a hospital census over time. He also notes that the results concerning the forms of the distribution of arrivals and length of stay suggest the consideration of multichannel queuing models for analysis of hospital systems.

Some of these methods and models described by Balintfy and Flagle can be applied to the analysis of the May 10th hospital census data. The 51,000 cases form a large enough group to permit study of various sub-groups.

Another approach for examining the flow of patients through the hospital has been suggested by R. W. Revans of the University of Manchester in a 1959 article. Revans would study the major stages of a patient's experience in the hospital. It would be possible to test variations of this approach on a sample of the cases in the May 10th Census.

In addition to the explorations outlined above, the specific purpose of the study is to describe with reasonable precision the major elements of hospital service used by the population in the Greater New York area in 1961.

REFERENCES

1. Clifford A. Bachrach - "Estimation of Length of Hospital Stay from Discharge Data" - to be published (Journal of American Statistical Association.)

2. Joseph L. Balintfy - "A Stochastic Model for the Analysis and Prediction of Admissions and Discharges in Hospitals" - Proceedings of the Sixth Annual International Meeting, The Institute of Management Sciences, 1959.

3. Anthony F. Bartholomay - "On the Linear Birth

and Death Processes of Biology as Markov Chains" - Bulletin of Mathematical Biophysics, Vol. 20, 1958.

4. Charles D. Flagle - "The Problem of Organization for Hospital Inpatient Care" - Proceedings of the Sixth Annual International Meeting, The Institute of Management Sciences, 1959.

5. P. H. Leslie - "On the Use of Matrices in Certain Population Mathematics" - Biometrika, Vol. 33, Part III, November, 1945.

6. A. G. McKendrick - "Applications of Mathematics to Medical Problems" - Proceedings of the Edinburgh Mathematical Society, Vol. 44, 1925-26.

7. R. W. Revans - "The Hospital as an Organism: A Study in Communications and Morale" - Proceedings of the Sixth Annual International Meeting, The Institute of Management Sciences, 1959.

8. Morton Robins and Rose Sachs - "An Application of Biometry to Hospital Statistics: Hospital Stay Tables" - American Journal of Public Health, February, 1953.

9. Mortimer Spiegelman - "The Versatility of the Life Table" - American Journal of Public Health, March, 1957.

A METHOD FOR ANALYZING LONGITUDINAL OBSERVATIONS ON INDIVIDUALS IN THE FRAMINGHAM HEART STUDY

Harold A. Kahn, National Heart Institute, National Institutes of Health

The Framingham Study was established in Framingham, Massachusetts in 1948 to measure the development of cardiovascular disease in a sample of the town population aged 30-59 and to determine factors associated with disease onset. Data are obtained primarily from a scheduled biennial examination conducted in a clinic maintained especially for this study. Although the examinations are intended to be at 2 year intervals, they are often longer or shorter by several months and in some cases are very much longer whenever one or more complete examination cycles is missed. Additional description of the Framingham Study together with some of the substantive findings to date can be found in an article by Kannel, W.B., et al., in ANNALS OF INTERNAL MEDICINE 55:33 July 1961. For the present let's concentrate on the repeat examination feature of the study.

Consider where we are today. Approximately 4400 individuals have had 5 or more examinations, about 3400 of these have had 6 or more and about 900 have had 7. However, analyses up-to-date have been based almost entirely upon associating the characteristics found on the first examination with subsequent outcome regarding disease. What we are trying to do now is to associate with subsequent outcome, all of the measurements made on a variable over the entire study period. As a specific example suppose that we consider the measurement of total serum cholesterol on successive examinations. Two hypothetical individuals might have readings as follows:

"A" 220-240-260-280-300 (died of CHD)

"B" 280-260-240-220-200-180-170 (still living)

Our interest is in finding the association between serum cholesterol level and development of coronary heart disease (CHD). Obviously we get somewhat different views of this association if we look at all of the measurements made or if we look at only the first examination results. It seems reasonable to presume that using all rather than part of the data available will lead to a better description of the relationship. The question is how to do this. For convenience we will refer to cholesterol measurements throughout the rest of the paper although obviously the procedures discussed are applicable to other variables and with appropriate modification to attributes as well.

One approach would be to consider in a single unit all the cholesterol measurements made on an individual. He could then be classified according to various functions of these data. However, since all the measurements are used, the observation time after establishment of the individuals classification category is very small. In effect, we would have "used up" the time under study to establish a classification and would have little time left over in which to see what is the subsequent incidence of disease associated with various classifications. It would, of course, be possible to wait for enough time to elapse following the series of cholesterol measurements so as to observe incidence rates associated with various categories. Such a waiting period implies not using later measurements to up-date the individuals classification, since to do so would bring us right back to practically all classification and almost no subsequent observation time.

There are two possible sub-divisions of the method suggested above. (1) restrict the analysis to just those who have not experienced the event being studied over the entire period during which cholesterol measurements were made or (2) also include any individuals who are no longer under observation at the end of the measurement period because the event being studied did occur at sometime during the measurement period. In case (1),we would be limiting our study to that subset which "survived" the measurement period and could not properly apply our findings to the total population. Case (2) has a different defect which can be illustrated by considering two individuals with cholesterol measurements as follows: "C" - 200-200-200 (Died of CHD); "D" - 200-200-200-220-240-260-280. If we classify the trend of measurements for each individual as a unit we shall categorize individual "C" as one showing a level trend and "D" as one showing an upward trend. We would then associate the unfavorable outcome observed for "C" with level trend experience and await subsequent developments in order to associate an outcome with the upward trend observed for "D". The method of treating the individual as a unit is not satisfactory in that individual "D" did have level trend experience for a period which was identical to that for "C". However <u>no event occurred</u> to "D" immediately following <u>this level trend experience</u> and because his total experience is considered as a unit the method makes no provision for recording what is in fact the truth: that two individuals showed level trend experience for a similar length of time and that only one of these had an event. A method which counts cholesterol classification categories which are followed by an event, but doesn't count an identically classified category which is not followed by an event, cannot be recommended. The method lacks a clear way of determining how many years of observation are associated with a particular classification. We could make a reasonable approach toward answering this question if we considered the total classification period in segments and assigned various segments to those cholesterol categories with which they would be identified if an event occurred during the segment, but this would lead us away from treating the classification measures on an individual as a unit. We turn now to description of an alternative approach which is not restricted in this way and which uses all of the classification data and all of the observations regarding occurrence of events.

This approach depends upon breaking up the entire observation for each individual into periods beginning with each examination date and extending up to, but not including the next examination date. Each of these observation periods is classified according to cholesterol category as of the start of the period. However, there is no reason to exclude data on cholesterol classification available from prior examinations. We do, in fact, use various functions of all of the cholesterol measurements made up to and including the examination which defines the beginning of the observation period. Thus, we can categorize an observation period according to cholesterol measurements such as the value at the beginning of the period, the average, the maximum, the standard deviation or the slope using all examination data up to and including the beginning of the period.

We have defined an observation period as the interval from one examination date to the next one (with appropriate modification when there is no "next" examination because of death, or loss from observation). Since the longer the period the greater the chance that a new case of disease may occur in it, we relate the occurrence or non-occurrence of the event being measured to the length of the observation period within which it had some chance to happen.

At this point it may help to introduce the following notation:

Let L_{ij} = length of observation period following jth exam on ith individual

X_{ij} = serum cholesterol measurement as of the jth exam for ith individual

Y_{ij} = 1 or 0 according to whether the event being measured (say the occurrence of myocardial infarction) occurred or not to the ith individual during the observation period following jth exam.

We will compute the statistic $\frac{\Sigma\Sigma Y_{ij}}{\Sigma\Sigma L_{ij}}$ for which the corresponding X_{ij} are all in a specified cholesterol classification category. For example the cholesterol category might be one in which the average value of all measurements made prior to and including the beginning of the interval is greater than 260 mg%. The numerator would then be the sum of the cases occurring in intervals classified as above. The denominator would be the total length in years of the observation periods so classified. The statistic $\frac{\Sigma\Sigma Y_{ij}}{\Sigma\Sigma L_{ij}}$ is an incidence rate per person year which we can associate with the particular cholesterol classification for which it was computed.

The entire procedure is essentially one of pooling together data from different individuals, as well as from different time periods for the same individual for which the common element is cholesterol status according to some specified definition. For instance an incidence rate associated with a cholesterol average of less than 200 mg% might be made up of data from 2 consecutive time periods for individual A plus two non-consecutive intervals from individual B, etc. Obviously such a wholesale hashing together of available data would not be very helpful without appropriate restraints. One of these is to combine data in age-sex specific groupings so that there is a reasonable approximation to homogeneity of combined experience. Ten year groupings are probably adequate in this regard. Another restriction is to omit observation periods for which the risk of developing an event is zero. Thus all observation periods subsequent to the occurrence of an event are excluded from the calculation of an incidence rate of that event.

We have defined the length of the observation interval according to the following rules: If there is a subsequent examination date, the interval ends on that date. If there is no subsequent examination we have considered the interval indeterminate and have not used it for events such as angina pectoris or hypertension which are not likely to be discovered except by examination. For events such as myocardial infarction or death which have a high probability of detection in the study mechanism, the observation interval for individuals still living in Framingham, ends on the anniversary date following an event (if there is one) or the cut-off date of the tabulations whichever is earlier. For those who have moved out of town the observation period ends on the estimated moving date.

In order to compute age-specific incidence rates it is necessary to assign observation periods and events to age classes such as 45-54, 55-64, etc. If the entire observation period is within a single age class, there is no problem. If it should overlap two, the length of observation period is assigned on a pro-rata basis to each except that the older age class assignment is cancelled if the event being measured took place in the younger age class. This is in keeping with the rule to avoid counting observation periods for which the risk of an event is zero.

Events such as myocardial infarction or death can be easily allocated to correct age classes. Events such as angina pectoris or hypertension are arbitrarily assigned to the mid-point of the interval in which they were discovered.

The incidence rates for different cholesterol categories that result from applying the above definitions have an unusual property in that the same individual may be contributing his experience to the incidence rates for both high and low cholesterol. Of course such incidence rates are not independent and significance tests of the difference in rates between groups are complicated thereby. Fortunately, the co-variance between rates of the size we are concerned with here (say .01 per person year) is sufficiently small in comparison to the variance that we may safely neglect it. Derivation of the variances and covariances of the incidence rates for these various categories is presented in the following appendix.

The most important feature of the incidence rates described is that they are able to use all the cholesterol information available in relating cholesterol category to disease outcome. Given a long term prospective study with sequential measurements such as the Framingham Heart Study it should be possible in time to compute incidence rates associated with cholesterol categories such as the following:[1]

- the most recent value
- the mean
- the variance
- the maximum
- the slope, or
- the length of time during which all measurements have exceeded a specified value

APPENDIX

All of the following should be understood as referring to rates for a particular classification category for cholesterol and to a particular age-sex group.

As before: Y_{ij} = 1 or 0 depending on whether the event being measured happened or not to the ith individual during the observation period following the jth exam

L_{ij} = length of the observation period following the jth exam for ith individual (in years)

----

[1] Computer programming of the Framingham data began in July 1961.

$$p = \frac{\Sigma\Sigma Y_{ij}}{\Sigma\Sigma L_{ij}}$$ is computed for ij in which the classification variable (cholesterol) is in a specified category and for which the L_{ij} are in a particular age-sex group.

Among all the possible samples which could have been chosen from the Framingham Adult population we are going to restrict our interest to that subset with the same amount of observation time in each classification category as the actual sample we have. This seems like a reasonable restriction which both simplifies and improves our results since variance associated with variation in observation period is somewhat extraneous to our interest.

We shall also postulate that incidence rates are constant for all observation periods within any specific age, sex and cholesterol bracket. Actually this is not true, but it is probably sufficiently close for our purposes.

Designating P as the true incidence rate, we have,

$$E(Y_{ij}) = PL_{ij}$$

where the P and L_{ij} are of such size that their product lies between 0 and 1.

$$V\left[p\right] = \frac{1}{\left[\Sigma\Sigma L_{ij}\right]^2} \Sigma\Sigma\, E\left[Y_{ij} - PL_{ij}\right]^2$$

$$= \frac{1}{\left[\Sigma\Sigma L_{ij}\right]^2} \Sigma\Sigma\left[E\, Y_{ij}^{\,2} - P^2 L_{ij}^{\,2}\right]$$

$$= \frac{1}{\left[\Sigma\Sigma L_{ij}\right]^2} \Sigma\Sigma\left[P\, L_{ij} - P^2 L_{ij}^{\,2}\right]$$

$$= \frac{P}{\left[\Sigma\Sigma L_{ij}\right]^2} \left[\Sigma\Sigma L_{ij} - P\,\Sigma\Sigma L_{ij}^{\,2}\right]$$

Of course, we do not know P and would substitute the sample observation p in place of it. A necessary qualification on this formula is that it is appropriate only if $\Sigma\Sigma L_{ij} > P\,\Sigma\Sigma L_{ij}^{\,2}$. In the Framingham study with most observation periods 2 years long and annual rates of about .01, $\Sigma\Sigma L_{ij}$

will generally be about 25 times as big as $P\,\Sigma\Sigma L_{ij}^{\,2}$. The previously stated restriction that $0 < PL_{ij} < 1$ is sufficient to assure that the variance of p will always be positive.

Although the preceding form is the one used in computing, it is helpful to get an idea of what this variance amounts to under the assumption that all L_{ij} are of equal length say L. This means that $\Sigma\Sigma L_{ij} = mL$ where m is the number of observation periods contributing data to the incidence rate under consideration. Then:

$$V\left[P\right] = \frac{P}{m^2 L^2}\left[mL - P\,mL^2\right]$$

$$= \frac{P}{mL}\left[1 - P\,L\right]$$

which is noticeably different from the variance of a binomial variable in that the denominator is person years rather than persons and the numerator is not of the form P(1-P) but P(1-PL).

The coefficient of variation is insensitive to the units in which the observation is measured, i.e., changing from person years to person months will divide the rate by 12 and the variance by 144.

To find the covariance between p_1 and p_2 where p_1 is the sample incidence rate for an age-sex cholesterol category and p_2 is a similar rate but for a different cholesterol category, we recall that:

$$p_1 = \frac{\Sigma\Sigma Y_{ij}}{\Sigma\Sigma L_{ij}} \quad \text{for cholesterol category 1 and}$$

$$p_2 = \frac{\Sigma\Sigma Y_{ij}}{\Sigma\Sigma L_{ij}} \quad \text{for cholesterol category 2.}$$

To help distinguish cholesterol categories we shall introduce "1" and "2" into the notation preceding the subscript for individuals. The two rates will in general be made up from data for different sets of individuals and it will also help to change notation for individuals in group "2" from i to k. Since the covariance between rates depends upon the same individual contributing to both rates and isn't affected by the number of observation segments into which that individuals experience is divided, we shall simplify our notation by dropping reference to subscript j for observation periods. However it should be understood that all references to an individual are for the sum of his observation periods appropriate to the cholesterol and age classification being considered. Thus we have:

$$p_1 = \frac{\sum_{i=1}^{n_1} Y_{1i}}{\sum_{i=1}^{n_1} L_{1i}} \quad \text{and} \quad p_2 = \frac{\sum_{k=1}^{n_2} Y_{2k}}{\sum_{k=1}^{n_2} L_{2k}}$$

The covariance of p_1 and p_2 is:

$$CV\left(p_1 p_2\right) = E\left[p_1 - E(p_1)\right]\left[p_2 - E(p_2)\right]$$

$$= E\,p_1 p_2 - E\,p_1\,E\,p_2$$

$$\text{where } p_1 p_2 = \frac{\left(\sum_{i=1}^{n_1} Y_{1i}\right)\left(\sum_{k=1}^{n_2} Y_{2k}\right)}{\left(\sum_{i=1}^{n_1} L_{1i}\right)\left(\sum_{k=1}^{n_2} L_{2k}\right)}$$

In order to see how large the covariance might be under extreme conditions we will assume that $n_1 = n_2 = n$ which would only be the case if all individuals contributed at least some of their experience to each of the rates. We can then break up the numerator into two terms.

$$p_1 p_2 = \frac{\sum_{i=k}^{n} (Y_{1i} Y_{2k}) + \sum\sum_{i \neq k}^{n} (Y_{1i} Y_{2k})}{(\sum^{n} L_{1i})\,(\sum^{n} L_{2k})}$$

The first term in the numerator is the product of $Y_1 Y_2$ values for the <u>same individuals.</u> Since the only possibilities here are 0.0, 0.1 and 1.0 we can substitute zero for this term. 1.1 is not possible, nor is it even defined, because we do not include observation periods after an event.

$$E\,p_1 p_2 = \frac{\sum\sum_{i \neq k} E\,(Y_{1i} Y_{2k})}{(\Sigma L_{1i})\,(\Sigma L_{2k})}$$

Since $i \neq k$ the Y_1 and Y_2 values are for different persons and therefore independent. Thus we get:

$$= \frac{\sum\sum_{i \neq k} P_1 L_{1i} P_2 L_{2k}}{(\Sigma L_{1i})\,(\Sigma L_{2k})}$$

$$= \frac{P_1 P_2}{(\Sigma L_{1i})\,(\Sigma L_{2k})}\left[(\Sigma L_{1i})(\Sigma L_{2k}) - \sum_{i=k} L_{1i} L_{2k}\right]$$

$$= P_1P_2 - P_1P_2 \frac{\sum_{i=k} L_{1i} L_{2k}}{(\Sigma L_{1i})(\Sigma L_{2k})}$$

$$CV\left(p_1p_2\right) = P_1P_2 - P_1P_2 \frac{\sum_{i=k} L_{1i}L_{2k}}{(\Sigma L_{1i})(\Sigma L_{2k})} - P_1P_2$$

$$= - \frac{P_1P_2 \sum_{i=k}^{n} L_{1i}L_{2k}}{(\Sigma L_{1i})(\Sigma L_{2k})}$$

For purposes of comparison with the variance of a specific rate, we will make additional simplifications to that we made previously and assume that all L_{1i} and L_{1k} are equal and that each is made up of $\frac{m}{n}$ periods of length L where m is the number of observation periods and n the number of persons observed. Then the covariance of p_1p_2 =

$$- \frac{P_1P_2 \, n \left(\frac{m}{n}L\right)\left(\frac{m}{n}L\right)}{(mL)^2} = - \frac{P_1P_2}{n}$$

which is identical with the covariance of binomial proportions for the special case assumed here.

We now compare the absolute size of the covariance between two equal incidence rates with the variance of one of them.

$$\frac{\text{Covariance } (P_1P_2)}{\text{Variance } P_1} = \frac{P^2 \, (mL)}{nP \, (1-PL)}$$

$$= \frac{P}{(1-PL)} \; \frac{mL}{n}$$

For P of .01, L of 2 years and each person contributing about 10 years experience, the ratio of covariance to variance is about 1 to 10. Recalling that we have artificially magnified the covariance by presuming that all persons have equal experience in both classifications, it is reasonable to ignore the covariance at least for the Framingham Study. For these data it is not sufficiently large in comparison to the variance to require its use when considering the standard errors of differences between rates.

IX

PANEL DISCUSSION: STATISTICS AND SURVEYS AS LEGAL EVIDENCE

Chairman, Eli S. Marks, Case Institute of Technology

Panel:

W. Edwards Deming, Consultant in Statistical Surveys
Arnold J. King, National Analysts, Inc.
James A. Bayton, National Analysts, Inc.

SESSION ON STATISTICAL AND SURVEY EVIDENCE IN THE COURTS ANNUAL MEETING OF THE AMERICAN STATISTICAL ASSOCIATION

Eli S. Marks, Case Institute of Technology

This session is focused on the use of surveys and statistics as legal evidence. We are particularly interested in survey evidence and the problem of how one gets the courts to examine whether survey results are statistically correct rather than whether they are mathematically correct.

National Analysts has had some experience with this problem in the du Pont case where they did a survey regarding which Dr. Benjamin Tepping testified (before the United States District Court in Chicago). More recently, some of their staff appeared in a hearing before a Federal Trade Commission examiner with respect to a survey presented as evidence in a case involving the Borden Milk Company. In both these cases (as in others in which have involved survey evidence) the question of the admissibility of survey evidence has come up -- the critical point being, of course, that the idea of statistical data as distinct from individual testimony is a relatively novel concept in the courts. In both of the cases in which National Analysts was involved, the court and the examiner accepted their survey results as statistical results, over the objections of the government attorneys who persisted in trying to treat the survey data as individual reports subject to attack in terms of their individual validity.

The idea that data which are inaccurate when taken one at a time, may have quite adequate accuracy when considered as statistical aggregates is, of course, a fairly sophisticated concept and one which should (in my opinion) be expounded by the statistical profession as a whole rather than by individual statisticians. In organizing this session, I suggested that the discussion be slanted toward what can be done by the statistical profession as a whole to educate legal thinking on this matter.

Another phase of the discussion is the question of confidentiality and the assessment of the reliability of survey evidence. Dr. Eckler has outlined the recent Supreme Court decision which permits the government to subpoena the manufacturer's copy of reports submitted to the Census Bureau. The confidentiality of the Census Bureau's copy is, of course, legally protected. The decision in this case is at best a nuisance and at worst a negation of the primary purpose of the legal provision for confidentiality. That is, the purpose of confidentiality is to insure complete and truthful reporting to the Census Bureau by assuring respondents that their reports can never be used against them. Of course, the respondents can still protect themselves by not keeping copies of reports sent to the Census Bureau. This would, of course, mean difficulty in checking on incomplete or ambiguous entries. More important is the probability that respondents with anything to conceal will in the future supply statistical data only to the extent that compliance is legally required and legally enforceable!

The problem of confidentiality also came up in the cases in which National Analysts was involved. The opposition (government) attorneys have wanted access to individual reports -- their argument being that, if the court grants the admissibility of survey evidence (over their objection), they and the court should have an opportunity to assess the accuracy and validity of the evidence submitted. Adhering to the position that survey data constitutes "hearsay evidence" the government attorney in one case demanded the right to cross-examine the individual respondents in open court. National Analysts, of course, opposed this on the basis of the confidentiality of the information -- i.e., indicated that the data were secured under a promise of confidentiality and could (probably) not have been secured without this promise -- but I think we need, in the future, to go farther.

The principle of confidentiality must, of course, be upheld in its own right and as a matter of ethics and practical business. Beyond that, however, we need to get across to the courts and to the legal profession, that error in individual statistical reports does not demonstrate satisfactorily error in a statistic based on these reports. To do this we cannot "sit on our expert testimonies" and assert that the court must accept our statement that the data are accurate, without ever attempting to make an independent assessment of this fact. Making such an independent assessment is, of course, not a simple matter, but I do not agree with one opinion that has been expressed that the accuracy of survey data should be a matter of expert testimony with each side being allowed to present the opinions of their experts and may the best experts win! Certainly, we want to insist that the assessment of the accuracy of a survey is a matter of examination by experts, but we all know that the only way to begin to resolve a controversy with respect to the validity of a survey result (as distinct from its variance) is to repeat the survey on the same or a different sample -- i.e., demonstrate the reproducibility of the survey results.

ON STATISTICAL SURVEYS AS LEGAL EVIDENCE (ABSTRACT)

W. Edwards Deming, Consultant in Statistical Surveys

The principles that govern successful presentation of a statistical survey as legal evidence are much the same as for the presentation of a scientific paper, namely, say what you know, say it so that your audience can understand it, and don't say what you don't know, as every word is subject to challenge.

It is important that the statistician stay within his field of competence. If he steps out of his field and into someone else's, he will make statements that he can not substantiate, and which will almost certainly conflict with the testimony of other witnesses who are qualified to answer the question that the statistician is not.

There are well recognized statistical standards for the statistician to follow. It is important that he observe the limitations as well as the power of statistical theory. Part of the statistician's job on an engagement is to provide statistical controls that will detect departures from instructions and other operational blunders, so that he may be in position to testify whether the departures detected do or do not preclude meaningful use of the theory of probability for generalizations to the frame that the sample came from.

The theory of probability enables the statistician to generalize to the frame that the sample came from, provided the statistical controls show that the sampling procedures as actually carried out did not depart too far from the specifications. Generalizations made by theory, under proper conditions, can not reasonably be challenged, as they are mathematical consequences, with standard objective interpretations the world over.

Generalization to other material or conditions not included in the frame, or to future dates, requires substantive knowledge, not statistical knowledge. It is important for the statistician's testimony to differentiate clearly between the assertions that he can make on the basis of probability, and those that he or others make on the basis of judgment. Anyone has a right to say what he believes, and to sign his own name, but he has not the right to use the theory of probabilities where it does not apply.

An important part of the statician's engagement is to separate, in the planning stages, the responsibilities of the statistician and of the expert in the subject-matter. This is important from the standpoint of good science; and also to avoid conflicting responsibilities and conflicting testimony.

The full paper will contain examples of testimony and of statistician's reports.

STATISTICS AND SURVEYS AS LEGAL EVIDENCE

Arnold J. King, National Analysts, Inc.

Although sample surveys have been received as legal evidence as early as 1940[/1] it was not until June 1959 that the findings of a survey were given serious consideration in an important legal decision. The break-through occurred in a Federal Trade Commission hearing involving nine ice cream companies when Examiner John Lewis [/2] gave considerable weight to a sample survey in rendering his decision. Then, shortly after this, on October 2, 1959 in the U. S. District Court, Judge Walter J. La Buy[/3] gave material weight to the findings of a sample survey in an important decision regarding the divestiture of du Pont holdings of General Motors stock by du Pont. More recently there was another Federal Trade Commission hearing involving The Borden Company[/4] in which a sample survey was received as evidence.

National Analysts conducted the three surveys referred to above. During these hearings I have found myself pressured more and more by the Federal counsel toward the unprotected disclosure of the names of the respondents. This was done on the grounds that the names were needed to check the reliability of the survey data. But, the pressure to disclose the names is not applied until the survey is about to be presented as evidence. This turns out to be a legal trap which leaves only two clear-cut courses of action that can be taken - both leading to undesirable results. One course of action that can be taken is to disclose the names for unrestricted use. This leaves open the possibility of using the "Public Witness" procedure of subpoenaing the respondent against his will and forcing him to testify and be cross examined about the information which he voluntarily gave to the interviewers. The American people have willingly given information about themselves and their businesses to statistics-gathering agencies because the information has by-and-large been held in confidence. A violation of the principle of confidentiality and anonymity would undermine much of the statistics in this country, whether it is being collected by government agencies, universities or businesses. All a person has to do is to turn to other countries where these principles have been violated to see how difficult, if not impossible, it is to obtain accurate statistics of the kinds obtained in this country. More on this subject later.

The other course of action is to withdraw the survey from evidence. If this is done the question arises as to whether the research company can collect the money due it for conducting the survey. The research company might not only be subjected to this loss but might also be held liable for all other costs associated with this action on their part. If the research company does not take the loss then the respondent to the case must shoulder it. The financial loss to one or the other party is not an important matter in itself. It is an important matter if the courts are deprived of the kinds of information that is basic to the resolution of complex litigated matters that can only be obtained accurately by using scientific survey methods to obtain it.

Whenever we have conducted a sample survey to be used as legal evidence we have agreed to cooperate in the checking for reliability in the following way:-

(a) To make available to the counsel the names of each and every staff person who had anything to do with any phase of the survey, including all of the interviewers who participated in the survey, so that they could be questioned on any aspect of the survey. If necessary, they could be placed on the stand and cross examined.

(b) Every letter, memorandum, report or writing of any kind accumulated in the files in connection with the design and the carrying out of the survey, copies of all the interview sheets, containing substantially the verbatim replies of the interview (with the names of the survey respondents removed) would be

1/ United States v. Aluminum Company of America, 35 F. Supp. 820, 823-24-S.D. N.Y. 1940

2/ F.T.C. Hearings - Docket Nos. 6172 through 6179 and 6425, involving nine ice cream companies

3/ U.S. District Court for the Northern District of Illinois, Eastern Division, Civil Action No. 49C-1071, U.S. vs. du Pont, General Motors Christiana Realty Company, and Delaware Realty Corporation.

4/ F.T.C. Hearings - Docket No. 7129, involving The Borden Company.

made available to counsel with the understanding it would be held in camera.

In the dairy manufacturers and the du Pont cases, the surveys were offered and received in evidence under the above agreement and considerable weight was placed on them in the court's decision, without any pursuit of individual survey respondents. The respondents' names were not made public; counsel for both parties agreed that, before offering any of the survey papers in evidence the names would be removed therefrom. However, in the Borden case the government counsel insisted that they be given the names of all of the respondents for their unrestricted use. We realized that the government counsel wanted to satisfy themselves that the interviews were actually conducted and in exactly the manner described in our instructions. With this in mind we did agree to provide the government counsel with a limited number of respondents' names for restricted use. How this was to be done was spelled out in a written memorandum. (See Pages 6 through 11).

Sample surveys provide the courts the opportunity to utilize the scientific principles and procedures developed by the statisticians, the economists, and those working in the fields dealing with human behavior in order to obtain some of the kinds of information that are basic to sound legal decisions. It is only possible through the use of these principles and procedures that some of the kinds of information needed can be obtained accurately enough that the legal decision could confidentially be based upon them. Data collected for scientific investigation also has to be assessed for validity and reliability before it is accepted as scientific proof and methods have been developed for doing this. These methods must be applied in assessing the reliability of sample surveys submitted as court evidence. Since the surveys are based upon samples, the reliability of the sample estimates must be assessed from the viewpoint of sampling variation and biases and any statement of accuracy must be expressed as probabilities derived from sample frequency distributions. The mere fact that a respondent report is shown to be in error does not in itself prove that the sample estimates are useless. Errors of this type may, for example, in the aggregate be held within tolerance limits through a quality control system. If these scientific concepts are used to check the reliability it would necessitate a considerably different approach to the checking procedures than are sometimes used by the legal profession. Dr. Bayton will address his remarks to this point. It is up to the scientist to point out to the courts what these concepts are and how they can be applied. Who has more at stake in this matter and who is better equipped to do this than the American Statistical Association? The legal profession and the sampling organizations are looking to them for guidance.

I am presenting to you as a case in point the procedure followed in the Borden case. It is possible that we may have gone too far in yielding to counsel demands. It must be emphasized that there was and is a perfectly legitimate aspect to these demands. The court and the parties to a legal case should examine critically the question of whether survey evidence is what it purports to be and the question of limitations and inaccuracies of the data presented. It seems to me, however, that the criteria and procedures for such an examination should be primarily statistical rather than legal. It is high time that the ASA and other responsible professional organizations (e.g., the American Marketing Association) explore these implications thoroughly and then take a definite position and provide leadership in securing legal acceptance of this position. The individual research organization cannot by itself sustain a sound position. This requires the weight of the profession and, to date, the statistical profession as such has not discharged its responsibility for setting and supporting appropriate standards of professional behavior as they relate to the disclosure of information from survey respondents. The following quotations are taken verbatim from the memorandum sent to the Federal Trade Commission in the Borden case:

(1) There will be (a) a series of ten interviews conducted by you alone; (b) a series of twenty interviews conducted jointly by one of you and one of Borden's counsel; and (c) any and all additional joint interviews which you may desire (subject to our option, if the aggregate number of joint interviews should exceed 100, to content before the Hearing Examiner, on the basis of the circumstances then obtaining, that you are not entitled to pursue the matter further; but without any present agreement on your part that the number shall be anything less than the entire, all of the respondents).

(2) Selection of respondents to be interviewed will be entirely in your discretion. Neither National Analysts nor Borden nor anyone on their behalf will contact any such respondent in any way in advance of the interview. By way of further assurance as to that, you may if you wish inform us, in advance of each interviewing trip, only as to the location where you are to be met, and wait until the interviewing trip is underway before designating (by segment number and interview number) any respondent who is to be approached for interview.

(3) You (along with your representative, in the case of joint interviews) will be accompanied to the respondent's door by a male interviewer from National Analysts. In no instance will this be the interviewer who saw that particular respondent during the National Analysts' study.

(4) After ascertaining that the person at the door answers to the same name as shown for that address in National Analysts' study records, the National Analysts interviewer will open the conversation as follows:

"Good ___________. I am __________ from National Analysts, a market research firm in Philadelphia. You will recall that some months ago one of our people called on you in connection with a study we were making on evaporated milk. Today this (these) gentlemen would like to ask you some questions about that. This is Mr. ____________ who is a lawyer with the Federal Trade Commission; (and this is Mr. _______ __________, who is one of our client's lawyers.")

(5) After the respondent's willingness to be interviewed has been indicated expressly or by fair implication, the National Analysts' interviewer will excuse himself and go back to the car, and the interview will proceed.

In the event that any respondent should state an unwillingness to be interviewed, nothing further will be asked; and that respondent will, of course, not be counted as among the 10 or 20 or 100, as the case may be, mentioned in paragraph 1 above.

(6) In each interview, whether conducted by you alone or jointly with one of us, you will be free to conduct your questioning entirely as you see fit, without restriction or interruption on our part. After you have completed your questioning in a joint interview, our representative will be entitled to ask such additional questions as he may wish.

(7) The 10 interviews to be conducted by you alone will be deemed of a purely exploratory nature, and no effort will be made to adduce in any form any evidence in this case relating to those particular interviews.

(8) As to each joint interview, you and our representative will endeavor to stipulate as to all matters which either of us may wish to have covered. Such stipulations may be offered in evidence, in whole or in part, as either party may see fit in usual course, subject to the other party's right to object in the same manner as if the interviewed respondent were personally present on the stand.

(9) You will in no event, by any means, whether formal or informal, and regardless of whether the procedures set out in this letter are pursued to a conclusion or abandoned, further contact or pursue any respondent whose name is made known to you in accordance with the foregoing procedures. This will apply, for example, in respect to any respondent who may be unwilling to be interviewed, and it will apply without regard to anything that may happen or fail to happen during or after any interview.

(10) If at any point during the conduct of the above-mentioned series of interviews you should for any reason (whether because of what you

> might deem to be an excessive number of refusals to be interviewed, or otherwise) conclude that such procedures were inadequate from your standpoint, you would be entirely free to abandon those procedures and to proceed before the Hearing Examiner in whatever way you might think fit, on the basis of the circumstances then obtaining.

The government counsel in the Borden case refused to agree to the above proposal, still insisting, upon unrestricted use of names. So it was necessary to ask for a ruling on it by the examiner and I submitted the following affidavit in support of our position.

"Throughout my association with National Analysts, which has been continuous since 1948, it has carefully followed a policy of not knowingly or willingly disclosing to persons outside its organization the identity of persons interviewed in the conduct of studies and surveys. This policy conforms with the non-disclosure policies and practices of all other reputable survey research organizations and is supported and policed by our industry societies and associations.

This fundamental policy in not disclosing the identity of respondents extends throughout the Company's operations. The psychologists, who constitute a substantial part of the professional staffs of our own and other survey research organizations, must operate within the following principle, set out in the ethical standards of psychologists, adopted by the Council of Representatives of The American Psychological Association (Principle 4.32-1):

> The identity of research subjects must not be revealed without explicit permission. If data are published without permission for identification, the psychologist is responsible for adequately disguising their source.

This principle is recognized and practiced by our other staff members in devising and directing scientific sample surveys. Our interviewers are indoctrinated and trained to understand the importance of obtaining and maintaining the confidence of respondents.

Accordingly, our interviewers explicitly or implicitly convey to respondents the assurance that their response will be kept anonymous. While the extent to which specific and explicit assurances are made may vary from survey to survey, depending upon the nature of the information being sought, this principle of confidentiality and anonymity is basic to all of our work. By the use of the code systems and the simple expedient of cutting or tearing off name and address information from response sheets, not even the National Analysts' client, or in this case its legal counsel, can identify respondents.

The following passage from National Analysts' "Manual for Interviewers" is representative of the emphasis given the non-disclosure policy.

> The willingness of the average citizen to give information must at all times be protected. Researchers, including interviewers, must always be considerate and honest in their dealings with the public, or else the usefulness of the survey as a tool in economic and social research will soon be at an end. No person should take part in a survey without a sense of responsibility to the public from which the survey sample is drawn. Any betrayal of confidence or unscrupulous use of "sample individuals," remarks, or of the final results, on the part of any person connected with a survey organization, is a breach of responsibilities. That survey organization cannot long endure. In the final analysis, the continued use of marketing research as a useful social tool depends upon the spontaneous cooperation of our respondents, and as researchers we must never lose sight of that fact.

As to scientific work for legal purposes, it seems clear to me that a rule that organizations who conduct such work may be subjected to court-ordered unrestricted disclosure of names of respondents, with the ensuing unprotected intrusion upon them, would sound the death knell for scientific legal survey work. If there were such a rule, no reputable scientific research firm could in good conscience undertake a legal survey of the evaporated milk type without first disclosing to its respondents the possible future involuntary involvement demanded here. Such a disclosure would greatly reduce the response rate and thus introduce a selectivity bias which would render these scientific research surveys unreliable.

Of immediate concern to me as chief executive of National Analysts is the grave impact which an unrestricted disclosure of the identity of respondents in the Borden survey would have on our organization. To begin with, respondents pursued for involuntary reinterviews and further unwanted questioning after a promise of anonymity, expressed or implied, would be rightfully indignant and even incensed that National Analysts, in revealing their identity, subjected them to inconvenience, annoyance and discomfort. Such respondents are part of the public whose continued confidence National Analysts must have. The promise of non-disclosure, by policy and practice, to the National Analysts' interviewers who took part in this survey would have been broken. It follows that the disillusionment of our staff of interviewers with National Analysts, occasioned by any such disclosure, would greatly impair their effectiveness in future surveys. They could hardly be expected to continue to strive conscientiously to create the confidential interview relationships necessary for unbiased responses. Certainly, the over-all quality of their work, in which we have invested so much, would be reduced and I have great fear that many would refuse to remain with us.

Concern of several of our professional men is already evident. Two of our key staff members have expressed their unreserved disapproval of the unrestricted disclosure of names of respondents in the Borden survey, and have advised me that they would consider resigning if National Analysts discloses the names of respondents to persons outside our organization except under the circumstances and conditions expressed below.

We have made every effort to work out a program which would enable us to keep faith with our survey respondents and our employees and yet permit government counsel to go into the field and check into our conduct of this survey.

The essential ingredients of this program, from National Analysts' standpoint, are the voluntary participation by survey respondents in reinterviews or in testimony or both, and the agreement that any persons refusing to be reinterviewed (or, having been reinterviewed, refusing to testify) would not be contacted again.

It is the uncoerced consent to reinterview and the uncoerced consent to testify which I feel relieves National Analysts of its burden of protecting the anonymity of respondents."

In this case the examiner agreed to our position and ruled against disclosure of respondents' names without unrestricted use. Examiner Abner E. Lipscomb stated [4] "We believe there is considerable virtue in the insistence of respondent's counsel upon keeping confidential the names of the persons interviewed." This does set a precedent but does not necessarily prevent the next examiner or judge ruling to the contrary.

In summary, I have placed before the Association an issue upon which they should take a stand. In taking a stand the Association must decide between (1) should the respondent's name in sample surveys, used in legal cases, be kept strictly confidential? (2) Should the names be disclosed without restricted use? (3) Should names be disclosed with the consent of the respondent and with a restriction placed upon their use? If so, what are the principles and procedures to be followed? In this decision consideration should be given to (a) the use of scientific principles and procedures for checking the reliability and validity of the sample surveys' findings, (b) to the principle of confidentiality and anonymity and to the consequences of any violation of these principles, and (c) to the framework of legal principles and procedures within which the legal decision must be made. This is an important issue and is one in which the American Statistical Association and its members have a great deal at stake in the stand they take. If a position is taken, I am sure it will, in the future, weigh heavily upon the Court's decision as to how far the research organization should go in the disclosure of respondents' names. This is a pressing issue and I trust that the Association will, in the near future, state its position.

STATISTICS AND SURVEYS AS LEGAL EVIDENCE

James A. Bayton, National Analysts, Inc.

The issue as to the disclosure of the names of respondents could be resolved if courts could be persuaded that these respondents do not stand in the conventional legal position of witnesses. Persuasion of the courts to this point would seem to depend upon their acceptance of a particular basic approach to the matter of survey data as legal evidence. The key factor in this approach is the establishment of the fact that the survey data are presented solely as being the result of research activity. The critical point to be established is that since this is the result of a research activity the challenges must be made in terms of research principles. If there is any fundamental matter that is uppermost in a researcher's mind it is that his results must be able to withstand challenge. In fact, one could say that all of the care that goes into the designing of a research project (whether for strictly "academic" or for applied purposes) has the purpose of anticipating challenge to the results which will be produced. The researcher designs his project in terms of research principles; he expects the challenges to be made in terms of research principles. A challenge which takes the form of demanding the names of the respondents comprising the sample is invoking a research principle. What research principle or criterion is involved in such a challenge?

The two essential research principles are validity and reliability. Any specific challenge must be striking at either the validity of the research data or its reliability.

Validity refers to whether the techniques used "tapped" dimensions or parameters that are pertinent to the issue at hand. Obviously, if the issue at hand was some matter related to price, a questionnaire directed only to the aesthetics of brand labels would be of dubious validity. Challenges to validity of survey data can be met by giving the challenger the questions asked and the answers given to the questions (with names of respondents deleted). Inspection of the questions and their answers could lead to quite legitimate claims as to unsatisfactory validity of the survey data.

When, however, the challenge is a demand for the names of the respondents, the attack has to be upon the reliability of the survey data, rather than upon its validity. The heart of this issue is: "You asked a valid question, but I don't believe the percentage you give as the answer is correct - I believe that something is producing a wrong answer to the problem, in terms of that percentage you show." The search now must turn to locating the "something" that might be causing an error. Once again, we must insist that the challenge to reliability is legitimate.

Reliability of survey data (assuming initial validity) is a function, of course, of two factors - sampling error and bias. Expert investigation of the probability methods used in the sample design and the methods of calculation of the sampling errors would reveal weakness in reliability with respect to the criterion of sampling error. Knowing the actual names of the respondents would make absolutely no contribution to an attack on the reliability of the data from the point of view of sampling error.

The challenge represented in the demand for the names of the respondents must, therefore, be directed toward the possibility of exposing some form of bias in the results. What are the potential sources of bias in sample survey research?

The sample might be biased - a matter of expert investigation of the definition of the universe sampled, method of drawing the sample, characteristics of the sample of respondents actually interviewed (income, age, and/or other relevant characteristics). Names of the respondents are not needed for this challenge.

The questions, per se, might be biased - "Don't you feel that the price of Brand X is much too high?" The question is valid because it refers to the pertinent issue - the price of Brand X; it is obviously biased. Inspection of the questions as written reveal the status of this problem.

The question, per se, is not biased but the interviewer inserts bias by inflection of voice, adding comments, etc. - the interviewers are available to be called as witnesses.

Bias might enter in the analysis (especially coding) of the interview material - this can be inspected in the interview material made available (with names deleted) and the codes used. If one wishes, he could recode the data and experts could testify as to the merits of any differences in results.

Bias could enter into the interpretation of the data. For example, does 25 percent represent "many people" saying something or does it represent "only a minority." The interpretative phrase used might depend upon the interest of the speaker (or writer) in the issue being studied.

With respect to the names of the respondents, the argument has to be that the respondents were a sample and were of no importance as specific individuals. The particular people who fell into the sample were, at best, fortuitous. Assuming that at some point in the sample design a system of randomization was used, the specific respondents would all have been different, if one had started at a different place in a table of random numbers. The researcher would argue that, with all other factors held constant, the data would be the same (within sampling error) regardless of the shift to a different set of respondents. The respondents in survey research have no more than the status of guinea pigs in biological laboratory research. The respondents are used solely as a sample to permit generalization back to some universe. Their particularity as Henry Smith, Mary Harrison, etc. has no status for the research purpose. (There is, of course, the question of honesty - were the so-called respondents actually interviewed? An independent research organization could make a check on this matter.)

If after expert investigation the challenger still is not convinced that the survey data are reliable, he has available to him the ultimate challenge - replication. Hold all factors constant - definition of the universe, sample design, questionnaire, coding procedure, etc. - and see whether the same results (within sampling error) are obtained.

Even further, the challenger could initially conduct his own survey, in anticipation of a survey to be presented by the opposition. If this survey produces different results experts could be called to investigate and to testify as to the sources of the difference.

In summary, Arnold King has asked that the Association take a stand on this issue. He poses three possibilities - strict anonymity of respondents; disclosure with the consent of the respondent but restriction upon the use of the information; and disclosure without restriction as to use.

The speaker advocates that the proper rule should be complete anonymity of respondents. It is believed that the respondent's awareness of anonymity makes a positive contribution to the reliability of the data. It is further argued that anyone who questions the results of a survey has open to him the avenues of challenge cited above. Finally, if the rule is to be otherwise, with disclosure demanded, we can anticipate that a valuable source of information in legal matters will be erased.

Without wishing to insult the intelligence of the courts, if survey research data are to be admitted as evidence, the courts will have to become knowledgeable with respect to the rules and principles of research or rely upon expert advice and testimony as to these rules and principles. With respect to the names of the respondents, the heart of this matter is that they do not have value or status as particular individuals. This is especially true when the implication is that <u>only</u> those particular persons who comprised the sample could possibly have had the attitudes revealed by the survey. From the point of view of sampling theory, a start from a different place in a table of random numbers would have produced a different set of particular individuals but the results should be essentially the same, in spite of this change to a new group of particular respondents

X

THE PRESENT STATUS AND FUTURE PROSPECTS OF SOCIAL STATISTICS

Memorial Session in Honor of Professor Samuel A. Stouffer

Chairman, Conrad Taeuber, Bureau of the Census

Samuel. A. Stouffer and the Progress of Social Statistics - Frederick F. Stephan, Princeton University

Social Statistics, Social Science and Social Engineering - Philip M. Hauser, University of Chicago

Institution and Imagination in Statistical Progress - Nathan Keyfitz, University of Toronto

SAMUEL A. STOUFFER AND THE PROGRESS OF SOCIAL STATISTICS

By: Frederick F. Stephan, Princeton University

Our purpose in meeting here this evening is two-fold. First, we wish to take a broad look at "social statistics" and consider the directions in which this realm of statistical activity may develop advantageously in the future. Second, but not secondarily, we seek to recall the many contributions to social statistics made during his career by Samuel A. Stouffer, who, were he alive today, would be up here fulfilling our first purpose more skillfully than I.

Perhaps no one can match Sam Stouffer as a pioneer in social statistics and social research. Few men in his generation have had his unique combination of far-sightedness and practicality, ingenuity joined to persistent and concentrated effort in pursuit of elusive truth, high standards of research craftsmanship, and restless intellectual curiosity.

These remarkable qualities of his, as well as his great contributions to the progress of social statistics and social research, are vivid memories for many of you who are here tonight and who knew him as an inspiring teacher, helpful adviser, stimulating colleague, and warm friend. For others who were not fortunate enough to have known him this well a brief review of his professional career may serve to show the variety of his interests and achievements. It will also remind us of his connection with many of the trends in social statistics which we plan to discuss together.

Samuel A. Stouffer's Professional Relations to Social Statistics and Social Research

Sam Stouffer became interested in a career as a social scientist after he had taken his bachelor's degree and, indeed, a M.A. in English. He was already well started on a career in journalism, following the footsteps of his father who was publisher of a newspaper in a small city in Iowa. Quite by accident, while on vacation, he read one of the classic books on sociology and, with a newsman's knack for spotting what is significant, he decided to become a sociologist. He entered the University of Chicago as a graduate student. It was there that he became interested in statistics and in related aspects of scientific method. Perhaps the greatest influence in this direction came from L. L. Thurstone who taught him the elements of statistical method and a good deal of psychophysics. W. F. Ogburn, who had just come to Chicago from Columbia to introduce courses in statistics in the Department of Sociology, also had a hand in arousing Stouffer's interest in statistics and shaping his outlook.

Stouffer took his Ph.D. in 1930 with a dissertation on "An Experimental Comparison of Statistical and Case Study Methods of Attitude Research."(1) This was before the day of modern attitude surveys; it was, in fact, a validation of a newly developed Thurstone scale on attitudes toward Prohibition. Characteristically Stouffer sought to do genuinely scientific work on a controversial subject. His report was a thorough presentation and scrutiny of his experimental data, restrained in its claims and conclusions, and carefully tested in many ingenious ways to make sure the mutual confirmation of the two approaches to the determination of attitudes was not spurious. It stands as a masterpiece in attitude research and a clear indication of the kind of research Stouffer was to do during his subsequent career.

Stouffer taught statistics briefly at the University of Chicago, in place of Ogburn who was busy with a gigantic study of social trends. He became convinced that the social research of the future would need men with better training in mathematics and mathematical statistics. Here again he pioneered. He applied for and was awarded a Social Science Research Council fellowship to study for a year with Karl Pearson and R. A. Fisher. After he returned from London to teach statistics in the Sociology Department at the University of Wisconsin he prepared a number of papers on the application of tests of statistical significance to specific studies in sociological research including one with Clark Tibbitts on a problem in criminal statistics involving the analysis of experience in predicting the risk of violation of parole. Thus he immediately applied what he had learned of mathematical statistics to current problems in sociology. He also encouraged his students to obtain a good grounding in mathematics and mathematical statistics. This was well in advance of the great interest that sociologists now show in such preparation.

Stouffer also had a deep concern with the improvement of the data of social statistics collected by state and federal agencies and he sought to extend their usefulness by supplementary projects for data collection. In the early 30's, he worked with the Registrar of Vital Statistics in Wisconsin on a study of trends in birth rates in that state based on data obtained by a well executed mail survey coupled with the birth registration records.(2) He also collected data by mail on marriage and divorce to supplement the incomplete reports then obtained by the Bureau of the Census(3).

It was in this period that Stouffer contributed two judicious discussions of current sampling practice.(4) He continued to have an active interest in sampling and in the modern methods which were just then gaining attention and being developed.(5)

It is not surprising that Stouffer was appointed a member of the Committee on Government Statistics and Information Services which was formed by the American Statistical Association and the Social Science Research Council at the beginning of the Roosevelt Administration to advise in the improvement and development of the statistical work of the Federal Government. He also served for 5 months as a member of the technical staff of the Committee and subsequently became a member of the staff of the newly formed Central Statistical Board.(6) Among the subjects to which he gave special attention were population, marriage and divorce, and vital statistics.

Stouffer became Editor of the Journal of the American Statistical Association toward the end of 1934. He almost became full-time Secretary of the Association as well. Had he done so the improvement of social statistics collected by government agencies might have remained his principal interest. However, he decided to return to teaching and academic research. He accepted a professorship at the University of Chicago but in his return to academic life he did not escape frequent calls to serve elsewhere as an adviser and researcher. He organized a series of studies for the SSRC on research that could be conducted to determine the social consequences of the great depression of the 30's. He was chairman of a Committee on Prediction of Personal Adjustment which studied logical and technical problems of prediction in such fields as vocational selection, school success, marriage and rehabilitation of criminals.(7)

He also made a study of the Population Census for the National Resources Committee,(8) prepared estimates of the effect on population growth of the reduction in the marriage rate during the depression of the 30's, participated in the International Population Congress in Paris in 1937, and served as a principal staff member in Gunnar Myrdal's comprehensive study of The American Negro. In the midst of all this, he devoted himself as fully as he could to an abstract problem in migration, namely how to explain the distribution of migrants from a common place of origin as they spread over the pattern of places to which they move. What excited him most was the theory he developed in terms of the pattern of the final location of migrants from all sources (intervening opportunities) rather than geographic distances.(9) He returned to this problem with a refinement of his theory only a few years ago.

Had it not been for the Second World War, Stouffer might well have become absorbed in the development of mathematical theories in sociology, pioneering further and inspiring others to join him in exploration. Instead he became director of the professional staff of the Research Branch, Information and Education Division of the War Department, in 1941 and spent 5 years directing highly original and significant studies of the attitudes of soldiers and officers on a great many matters pertinent to programs and policies in the Army and Air Force. Immediately after the War he and a group of his colleagues prepared the well known and monumental reports on The American Soldier and Measurement and Prediction. His work in this great research enterprise coupled with his previous achievements led Princeton University to confer on him in 1948 an honorary degree of Doctor of Science.

In 1946, Stouffer went to Harvard as Professor of Sociology and Director of the Laboratory of Social Relations. Here as elsewhere he led a busy and fruitful life, working with colleagues and students on many studies including social mobility and education, role conflicts, motivation, attitudes toward communism and civil liberties, and finally, with the Population Council, on motivation for the control of fertility in economically underdeveloped areas. He served as President of the American Association for Public Opinion Research and of the American Sociological Association. He also was chairman of the Committee on Measurement of Opinion, Attitudes and Consumer Wants appointed jointly by the National Research Council and Social Science Research Council.

It may appear that Stouffer's early interest in the improvement of the data of social statistics was displaced by his later emphasis on theory and his involvement in studies of controversial subjects and important public problems. This is not quite so; again he was pioneering. He pleaded for a shuttling back and forth between applied research and pure research as the best way social scientists could make progress since thereby they might more readily bring ideas into effective relations with data. He held out controlled experimentation as the ideal research method but also pleaded for rigorous statistical analysis, with more than two or three appropriate variables held constant or controlled, as a good approximation to the ideal when experimentation is not feasible. This calls for greater detail in tabulation, especially more cross tabulation, and places more exacting requirements on the collection of social statistics.

Sam Stouffer was humble in his view of his work. He stated in the opening paragraphs of The American Soldier that "Science...is cumulative, in the sense that a scientific achievement is most successful when it stimulates others to make the concepts and techniques it has used look crude and become obsolete as rapidly as possible." This will be very difficult for us to do but it is worth our greatest effort; he has given us a great challenge as well as a great example. We will remember him not only this evening but again and again as we continue the quest for better social statistics and work our way ahead in the many directions in which he pioneered.

Past Progress and Current Accomplishments

Sam Stouffer's unique contributions to social statistics were part of a larger development that started a long time ago with the first sporadic assessments of human and economic resources and the first quantitative determinations of risks and probabilities in human affairs. These two grand traditions followed different courses and at times were in conflict but repeatedly they were joined in ways that served individual and national needs and advanced the development of social science. Stouffer was particularly successful in blending data and abstract analysis.

The statistical tradition concerned with systematic collection of data on population, birth, deaths, marriages, migration, education, and many related aspects of family life and individual behavior has made remarkable progress in recent years considering the great difficulties that impede it and the substantial costs of this kind of data collection. About one third of the thousand pages of the Statistical Abstract of the United States are devoted to social statistics and these pages present only a very condensed summary in small type of the information which is published by the numerous collecting agencies of the Federal and State Governments and private organizations. We are all greatly indebted to the people who serve in such agencies to produce these data. They work day after day and year after year performing tasks that are often routine, uninspiring, and tiresome. From the top administrators to the clerks, interviewers, and machine operators their work is made more difficult by public indifference and inertia, by the natural human errors made by other people, and often by the lack of adequate facilities and appropriations. While some of them are lax in their performance, most of them are conscientious, highly competent, and motivated by the highest objectives of public service.

We have no time to review the progress made in each specific field of social statistics but any of us can find for himself what has been accomplished in fields of special interest to him by comparing the data available now with data available say 30 years ago. This will produce some shocks as well as some pleasant surprises; the progress has not been uniform and some fields have made little advance or even fallen back.

We can more readily take a general view of the progress of the last three decades. There are a number of prominent landmarks to guide us. At the time that Sam Stouffer was getting started in statistics our Association had a Committee on Social Statistics which arranged three joint sessions with the American Sociological Society and published the papers in 1930 in a volume in Statistics in Social Studies, edited by Stuart A. Rice, the chairman of the Committee.(10) A few years later Rice completed a comprehensive survey of social statistics for the President's Committee on Social Trends. Unfortunately his reports are not readily accessible though they were reproduced and circulated at the time. The various reports of the Committee on Government Statistics and Information Services provide many more landmarks, as do the pertinent parts of the Mills and Long report on the Statistical Agencies of the Federal Government prepared for the Hoover Commission and published in 1949.(11) These reports are dominated by problems of economic statistics but they attempt to deal comprehensively with the objectives, organization, and performance of statistical data collection. There are many other landmarks for specific fields within social statistics, two examples of which may be cited: 1) the mammoth collection of papers on demography edited by Hauser and Duncan in 1959 (12) and 2) the reports on criminal statistics by Beattie in 1959 and 1960.(13) Finally there is that masterly appraisal of the present situation and future prospects of social statistics presented to us by Bowman, Gall, and Rubin at Stanford last year.(14) I purposely avoid repeating what they have said so well and commend it to you for reading and rereading.

If I may be so bold as to attempt to summarize what we may expect to learn from a general survey of social statistics guided by these landmarks I will do so in the following terms:

1. Maintain past achievements and move ahead.

By dint of repeated efforts we have developed a great body of data in the various fields of social statistics, particularly in the field of population. Yet there are serious gaps and deficiencies still to be eliminated and there are many unrealized opportunities for improvement.

2. Get things done by cooperation within the existing organization of statistical activity.

Our decentralized systems of statistics in the Federal and State Governments enable us to collect many statistics as by-products of administrative procedures which otherwise would be too expensive or difficult to obtain. Yet this involves serious problems of coordination and comparability as well as great variations in the quality and performance of statistical work. Consequently it is very important to have the leadership and coordinating functions of such agencies as the Office of Statistical Standards in the Budget Bureau and similar offices in the State capitols plus a general willingness everywhere to work out problems cooperatively in the absence of central control of statistical operations. We can not readily change this system and we should do all we can to make it work well inspite of its weaknesses and inconsistencies.

3. Determine needs more specifically.

We need greatly a clearer formulation of the needs to be served by various sets of statistical data and more precise specification of the accuracy, detail, and other characteristics of the data that are essential to their effective use. Resources should not be utilized for the collection of unimportant data when they are needed for the collection of other data of critical importance. Unfortunately, the demand for data is expressed too much in terms of continuing what has always been done or in terms of pressure groups rather than clear demonstration of value and need.

While the widespread use of social statistics for many kinds of individual decisions and actions in addition to their general use in the legislative, judicial and executive functions of national, state and local governments, makes it difficult to assess the need for particular kinds of statistical information, the inevitable choices made in the regular operations of the data collecting agencies deserve better guidance and justify a major effort to determine what is needed enough to warrant the cost of obtaining it.(15)

4. Continue technical progress and extend its applications further.

Great progress has been made in the technical aspects of data collection through standardization of classification schemes, improvement of the techniques of collecting data at the source, advances in data processing equipment, and the application of modern mathematical statistics in sampling, estimation, quality control, and analysis. The benefits of this progress are still to be fully realized in many realms of social statistics and there are still some technical problems to be solved in every field. Further advances may be expected but there are difficulties connected with maintaining comparability with similar data collected in the past, shortages of adequately qualified personnel, and practical details of operation to be worked out.

5. Gain public understanding and support.

Social statistics have an increasingly important part to play in a free-enterprise, democratic society such as ours as well as elsewhere around the world. While there are differences in opinion about the appropriate role of the Federal, State and local governments in some programs such as education and public welfare, it is clear that they are of concern to individual citizens and that citizens need dependable statistical information about them. It is true that statistics play a central part in other countries where state planning in socialistic and totalitarian governments require it for regulating their economies and the actions of their citizens. In a free society individual citizens need ample statistical information to regulate themselves. Their associations, companies, and representative governments need statistical information on many subjects for which legislation, policy determinations, and administrative decisions must be provided to keep the free society free and healthy and to make it prosper. We should not resent the common jokes about statistics but we should concern ourselves about prevalent misunderstanding and unwarranted attacks as well as ignorance of the important role statistics play in a free democratic society.

6. Cultivate the interplay of data and ideas that advances science and enhances the value of applied statistics.

Finally, as Stouffer recognized so well, not only good statistical data but valid analysis and theorizing are necessary for the advancement of our understanding of human life in all its ramifications and, indeed, for the effective application to statistical information to important, even controversial, public problems. This means that our collection of data will be guided by the development of our ideas, our social sciences, just as in turn our ideas will be tested and developed by their fruitful association with data. I argued two years ago for greater articulation of theoretical concepts and statistical definitions.(16) Now I would only plead that social statisticians work more closely with other social scientists and that we again arrange joint sessions with them as the Committee on Social Statistics did 32 years ago.

A Program for the Future

As social statisticians we have many common interests and similar problems. Although we tend to be absorbed in our particular subdivisions and specialties we have profited from each other's experience in the past and borrowed in a neighborly way many useful ideas and procedures. We can and should do this increasingly in the future. In my opinion there are several rather definite ways in which we can do this within the Social Statistics Section and in the other opportunities we have to work together.

1) Annual review. We should continue the practice of making a comprehensive survey of recent developments in social statistics. This review should not be patterned too closely after Ray Bowman's paper last year or this paper of mine but should be adapted freely to the possibilities of broadening and strengthening our comprehension of our common purposes and how well they are being accomplished. Each year there may be a special emphasis that is appropriate for its time.

2) Planning and specifications. We should study together by individual research projects and special committee inquiries the complex problems of determining what is needed, what is feasible, and what should be chosen among the almost infinite variety of data that conceivably can be collected. The problems of deciding when data should be taken, how often, from what sources, in what detail, in what relation to other data, and with what degree of accuracy are familiar to all who have had some responsibility for initiating or for revising a system of data collection. These problems get more tightly interconnected as we elaborate the system and tune it up to meet more exacting requirements. Hence we can well justify devoting more time to preliminary planning and systematic analysis of data collecting operations than we could when everything was easier and simpler. Moreover our accumulating experience should crystallize and definite principles should emerge as we work on this problem of design and decision. The discussions at meetings of our Section should contribute to this part of the program at least in examining well chosen cases if not in effecting the final synthesis of experience.

3) Development of new concepts and reformulation of old ones. We should explore actively the possibility of enriching the vocabulary of basic concepts we use in classification and measurement. Examples are common in every field. The evolution of the concept of unemployment is familiar to most of us. Just now it is being examined again; it may be improved or clarified to meet better the current need for statistical information about this important public problem. In the 1960 Census of Population the concepts of dwelling unit and urban population were revised. We need better concepts of social class or socio-economic status, occupational mobility, social cohesion, education, health, and many other matters of importance both to social scientists and to the public. Our Section can well discuss this phase of the program and stimulate work on these problems elsewhere.

4) Development of new and improved means for measuring, classifying, and combining variables. Most of our measurements have been made by simple counting such as size of family, years of schooling, or weeks of unemployment. The Committee on Social Statistics was interested in the possibility of constructing an index of social welfare. Stouffer was greatly interested in Thurstone's and Guttman's systems of scaling and in Lazarsfeld's latent structure analysis. These are only beginnings in the creation of measuring procedures that are needed for the analysis of social statistics and their more effective application to practical problems. The long-range program of the Section may well allot a good fraction of our time and attention to progress in measurement.

5) Appraisal and testing. As we become more specific about the needs to be met by our data, we become more concerned about their quality. We can no longer afford to be blissfully ignorant of their shortcomings but must test and validate to make sure they meet reasonable specifications and serve effectively the purposes for which they are collected. Too often in the past we have assumed that numbers, and the words that give them meaning, when printed in black and white or reported by cooperative respondents are correct as given.

We have only recently come to a clear realization that virtually all data are subject to some degree of inaccuracy and most collections of data are to some degree incomplete. Instead of assuming that a set of statistics is correct until proven guilty of error we should consider every set on probation and possibly inaccurate, incomplete, and imperfectly defined until sufficient evidence is produced for assuming otherwise.

Many agencies still shrink from public discussion of the deficiencies of their data. Others, surprisingly, are not aware of these deficiencies. The progress made by the Census Bureau in its Post-Enumeration Survey and other tests of its data is an example of what can be done. (17)

The trend in this direction is evident in the papers on validation and response errors presented at recent meetings of the International Statistical Institute. (18) At times an independent examination of important statistical results is needed to give the public assurance, if they are trustworthy, or reveal their deficiencies more clearly, if they are not. The Section may well discuss this increasingly important aspect of all important statistical reports.

Regular appraisal practices are the mark of professional maturity. Those agencies are most advanced which risk embarrassment in determining the accuracy of their statistical data and reporting their determinations frankly and fully to the users. They are to be commended in leading the way that others should follow. It might be appropriate for our Section or our Association to give public recognition from time to time to accomplishment in assessing data and informing the public of the result.

Conclusion

Five major parts of a long range program have been suggested. Clearly they are all interrelated; certainly others should be added. In the spirit of Sam Stouffer's own endeavors and the grasp he had of what still needed doing, as well as in continuation of the efforts of all the many others who contributed to the progress of social statistics in the past, we should strengthen our present efforts and, where necessary, start new work so that our statistical work can better fulfill the expectations we have for it and better serve our free society.

FOOTNOTES

(1) Abstracts of Theses, Humanistic Series, University of Chicago Press, 1932, Vol.VIII, pp. 263-270.

(2) "Trends in the Fertility of Catholics and non-Catholics," American Journal of Sociology, Vol. 41, 1935, pp. 143-166.

(3) Samuel A. Stouffer and Lyle M. Spencer, "Recent Increases in Marriage and Divorce," American Journal of Sociology, Vol. 44, 1938-39, pp. 551-554.

(4) "Sociology and Sampling," in L. L. Bernard, Ed., Fields and Methods of Sociology, New York, Ray Long and R. R. Smith, 1934.

"Statistical Induction in Rural Social Research," Social Forces, Vol. 13, 1935, pp. 505-515.

(5) Frederick F. Stephan, "History of the Uses of Modern Sampling Procedures," Journal of the American Statistical Association, Vol. 43, 1948, pp. 12-39.

(6) Government Statistics, New York, Social Science Research Council, Bulletin 26, 1937, 174 pp.

(7) The Prediction of Personal Adjustment, Social Science Research Council, Bulletin 48, 1941, 455 pp.

(8) "Problems of the Bureau of the Census in Their Relation to Social Science," Section 7 of Research - A National Resource, Report of the Science Committee to the National Resources Committee, Vol. 1, Washington, U. S. Government Printing Office, 1938, pp. 195-232.

(9) "Intervening Opportunities: A Theory Relating Mobility and Distance," American Sociological Review, Vol. 5, 1940, pp. 845-867.

(10) Philadelphia, University of Pennsylvania Press, 1930, 222 pp.

(11) Frederick C. Mills and Clarence D. Long, The Statistical Agencies of the Federal Government, New York, National Bureau of Economic Research, 1949, 201 pp.

(12) Philip M. Hauser and Otis Dudley Duncan, Editors, The Study of Population, An Inventory and Appraisal, Chicago, University of Chicago Press, 1959.

(13) Ronald H. Beattie, "Sources of Statistics on Crime and Correction," Journal of the American Statistical Association, Vol. 54, 1959, pp. 582-592. Ronald H. Beattie, "Criminal Statistics in the United States - 1960," Journal of Criminal Law, Criminology and Police Science, Vol. 51, 1960, p. 49.

(14) R. T. Bowman, Alexander Gall, and Israel Rubin, "Social Statistics: Present Conditions, Future Needs and Prospects," in Proceedings of the Social Statistics Section, 1960, Washington, American Statistical Association, 1961, pp. 74-81.

(15) An example of such effort is the Appraisal of Census Programs, Report of the Intensive Review Committee to the Secretary of Commerce, Washington, U.S. Government Printing Office, 1954, 119 pp.

(16) "Relations of Some Social Science Concepts to Statistical Data," in Proceedings of the Social Statistics Section, 1959, Washington, American Statistical Association, 1960, pp. 170-171.

(17) The Post-Enumeration Survey: 1950, Technical Paper No. 4, Washington, U.S. Government Printing Office, 1960, 93 pp.

(18) P. Depoid, "Rapport sur le degre de précision des statistiques démographiques," Bulletin de l'Institut International de Statistique, tome XXXV, 3ème livraison, 1957, pp. 119-230. Henri Bunle, "Sur les erreurs entachant les statistiques et des recensements de la population et de l'état civil," loc. cit., pp. 283-288.

See also the many papers on non-sampling error, validation, and error assessment in the volumes of the Bulletin which contain papers presented at the Stockholm, Brussels, and Tokyo meetings.

SOCIAL STATISTICS, SOCIAL SCIENCE AND SOCIAL ENGINEERING

Philip M. Hauser, University of Chicago

The present status and future prospects of social statistics in relation to social science and social engineering is the subject of this essay. This subject has been selected for two reasons. First, it touches upon the roles of social statistics that engaged Sam Stouffer's mind, heart and much of his prodigious energy; and second, it represents an effort to minimize overlap with what I anticipated would be comprehensive and excellent papers by Fred Stephan and Nat Keyfitz focused more directly upon social statistics per se and on Sam Stouffer's significant contributions to their development.

The framework for my discourse was well set forth over fifty years ago by Richmond Mayo-Smith, Professor of Political Economy and Social Science at Columbia College.1/ After some preliminary discussion he observed:

"We get from statistics indications of relations which maintain themselves with a persistence and constancy that give us an impressive sense of the reign of law in the social actions of men. It is this reign of law which we desire to see revealed." (p. 15)

Mayo-Smith also observed that "We are surrounded by sociological or social problems which urgently demand solution. We cannot wait for the completed science; we must seek to understand the conditions affecting the particular problem before us. This may be called practical sociology. Everywhere in this domain we find statistics a useful instrument of investigation. . . ." (p. 16)

The first of these quotations serves admirably as a concise summary of the role of social statistics in social science; and the second, their role in social engineering. As a preliminary to consideration of social statistics in relation to social science and social engineering it is well to define the term as used in this paper. "Social statistics" is considered here as embracing not only the figures which flow from tabulations of census, survey, registration, administrative records, and the like but, also, the statistical methodology which is employed in social research and the substantive findings of social research stated in quantified form.

SOCIAL STATISTICS AND SOCIAL SCIENCE

Social science like any science has as its objective the achievement of knowledge independent of the "caprice" and "willfullness" of man, in contrast with knowledge based on the methods of "tenacity," "authority" and "intuition" as discussed by Morris R. Cohen and Ernest Nagel.2/ Such knowledge, which can more explicitly be stated as consisting of predictability and explanation, can be achieved only by means of the method of science, the method of "reflective inquiry"3/ or research. Statistical methods wherever applicable provide more efficient and effective prediction than alternative social science methods. The fact that a quantified projection or prediction is to be preferred to a non-quantified one needs no elaboration. It may also be argued that "explanation" based on statistical research is also to be preferred to alternative forms of explanation in social science. Social science explanation based on social statistics produces "probabilistic" explanation of a type not as rigorous as the "deterministic" explanation possible in branches of the natural sciences. But it constitutes a more rigorous and objective explanation than other types of explanation possible in the social sciences, which in the sense discussed by Nagel, include the "functional," "teleological" and "genetic" forms of explanation.4/

The Experimental Method. The most efficient and rigorous method for obtaining knowledge that science has yet devised is that represented by the controlled experiment--a method that necessarily involves measurement.5/ That is, to detect the impact of an experimental variable on a dependent variable "before" and "after" measurements are necessarily involved; as is also the measurement of the experimental variable itself. Even though the use of the controlled experiment has relatively limited application in social science, its use is by no means inconsequential. On the contrary the experimental method employed in social researches has resulted in important gains in knowledge--as, for example, in small group research.6/

The use of the experimental method even though limited, then, is the first of the ways in which we may note social statistics are involved in social science. In the employment of the experimental method the social variables utilized in the experiment are quantified and change is measured. Moreover, in the conduct of the experiment not only are the various scales of measurement and descriptive statistics employed but, also, the techniques developed by statisticians for the design of experiment and statistical inference. Although the use of the experiment in social science is limited, the statistical methods employed in conjunction with the experiment may well be viewed as a point where statistics on its most sophisticated level contributes to social science. And the knowledge which flows from experimental studies in the form of predictability and explanatory propositions may appropriately be considered a part, even though as yet a limited part, of the fund of "social statistics" as well as, simultaneously, of social science.

Statistical Method. Most social science research necessarily utilizes methods other than the controlled experiment. These methods best achieve the objectives of science, however, to the extent that they approximate the experimental method. Thus, the method of the controlled experiment has great indirect utility in social science in serving as a yardstick by which all other methods may be evaluated and their strengths or weaknesses assessed. Stouffer in his classical article, "Some Observations on Study Design,"7/ has provided an excellent demonstration of the manner in which the experiment can serve as a framework for assaying other methods.

The research methods which come closest to the controlled experiment in the study of social phenomena are the methods of statistics. Mathematical statistics has produced a wide spectrum of specific techniques for the study of social phenomena which in varying degree approximate or simulate the experimental method. It may be argued that the research methods involved are mathematical statistical methods not "social statistics." But generic statistical methods utilized in social research may appropriately in this context be considered as included in the methodological component of "social statistics," and the product of research based on these methods is appropriately included as part of the substantive body of "social statistics." This is, after all, usage analagous to that of the generic term "statistics" itself, which possesses both substantive and methodological connotations.

Social science research is becoming increasingly quantitative either by means of statistical methods or direct mathematical approaches.[8/] It is possible, therefore, to say that social science is becoming science, in contrast to its 19th and early 20th century philosophical state, largely by reason of the impact of the statistical method. The application of statistical method to social phenomena and the body of propositions and theory emanating therefrom are playing a major role in the advancement, or perhaps more accurately the achievement, of social science.

It must be acknowledged that there are still some scholars who feel that social science cannot be pursued in the imagery of natural science, and that the experimental method and statistical techniques approximating it are not applicable to social phenomena.[9/] But it is probably true that they are a dwindling minority and that the view I am presenting is now the modal position in a number of the social science disciplines, e.g., psychology, social psychology, economics, sociology, demography; and is gaining ground in others, e.g., anthropology and political science.

The specific ways in which social statistics has contributed to the emergence of social science may be simply set forth to serve as a frame for pointing to future developments. The social statistician has done much to measure the phenomena in which the social scientist is interested; to collect, process and tabulate social data; to describe and analyze the data ("explain"); and to project discerned patterns or sequences (trends) into the future ("predict"). Much has been achieved over the years, as Stephan and Keyfitz at this meeting, and others from time to time have shown in respect to each of these areas of activity.

<u>Statistical Bottlenecks</u>. Although much has been achieved, much, also, remains to be done, as I am sure the other papers in this session will agree. For my part I should like to point to a few of the obstacles to the development of social science which urgently require the concerted efforts of social statisticians.

<u>Measurement</u>. The first of these bottlenecks constitutes, in my judgment, the greatest contemporary obstacle to the development of social science. It lies in our inability successfully to measure many of the phenomena we wish to treat as dependent or independent variables. Our conceptual frameworks are in a number of areas ahead of our ability to obtain adequate metrics. As a result we are often able to explain only pitifully small proportions of the total variances of our dependent variables, not necessarily because the hypothesized relationships are not there but, rather, because we work with such confounded variables that we cannot detect the relationships. There is a great need to concentrate on obtaining better measurement of our variables--to get measurements that more clearly match our concepts, as Stephan noted two years ago in a paper before this section.[10/]

A good example in point is to be found in our classification of urban and rural population. Urbanization has had profound effects on man's way of life, and many efforts have been made to trace its impact by means of the census type of measurement of "urban" as an independent variable. But "urban" as measured in a census includes persons of third or subsequent urban descent together with newly arrived in-migrants from generations of living in the Appalachian highlands or the Mississippi delta. It includes persons who have been relatively segregated in ethnic enclaves within the urban place as well as persons of similar origin who have become completely assimilated. It includes residents of the central city and of the suburb; of persons in, as well as outside metropolitan areas and so on. There can be no doubt that the classification "urban" as we generally use it in social research is a badly confounded variable that obscures much of the order and regularity we seek, and that impairs our ability to improve our predictability and explanation.

There is at the present time considerable variation among the social sciences in the attention and energy devoted to measurement. The psychometrician who devotes several years to the development of a single scale certainly stands in sharp and meritorious contrast to the social survey operator who develops a 30-page questionnaire in several weeks--or even a few days. There can be little doubt that the sample social survey, which has been one of the most effective research instrumentalities developed in social statistics, is also producing a mass of spurious information by reason of poor measurement of the type described.

As another aspect of the problem of measurement much more work is needed to achieve higher order measurement scales in working with social phenomena. Success in this effort would have the salutary dividend of increasing the power of the methods of analysis and inference which may be employed in social research; and may result, therefore, in higher orders of predictability and fuller explanation. In general, it is probably true that a large proportion of social scientists remain unaware of the great differences that obtain in the power of admissible statistical operations when, for example, a ratio-scale can be utilized in lieu of a nominal scale.

<u>Error</u>. Many of our statistics still contain relatively large elements of error which can be eliminated. Sampling error has, of course,

become the easiest component of error to manage. But the various other types of error, as enumerated by Deming some years ago[11] can certainly be better controlled than they yet are, as a result of the quality control methods which have evolved, of the post-enumerative survey, and of other developments.[12] Data that do not meet as high standards of reliability, validity and precision as may be required for our research objectives stand as obstacles to the advance of social science. Perhaps the time has come when all social statistics, and especially those based on a canvass of the population, should routinely carry in addition to a report of sampling error some indication of the magnitude of other elements of error. Certainly we now possess enough knowledge to do much more of this than ever was possible before.

Longitudinal Data. Another bottleneck in the development of social science which the social statistician can help to break is to be found in the disparity between social science projections and the actual course of events. The demographer has learned the hard way to label his efforts to foresee the future as "projections" rather than "predictions." What can be done to improve social science predictability? At least part of the answer, I believe, lies in the expansion and elaboration of longitudinal, as opposed to cross sectional, statistics.

One of the greatest weaknesses of social science, in general, lies in the drawing of diachronic conclusions from synchronic and, therefore, inadequate data. Longitudinal conclusions are always dangerous when based on cross-section data, and especially so when the information is also inadequate in various other ways. Yet most of the data which the social scientist uses for projection or prediction are of this character. They consist of cross-section information drawn from a relatively small number of points in time, containing gross and confounded categories with volatile and often conflicting components. A good illustration is given by the many attempts to predict fertility behavior from current statistics. Just as the demographer has learned the limitations of cross-section birth-rate data and the value of the cohort approach, so the social scientist, in general, with the assistance of the social statistician, must greatly increase his fund of longitudinal data.[13] In this manner dynamic patterns and sequences could be discerned that would undoubtedly provide a better basis for projection and prediction.

Methods of Analysis. Less of a bottleneck, but, nevertheless constituting some obstacle to the advance of social science is the uneven development of statistical methods for the analysis of social, as compared with natural, science data. Mathematics and mathematical statistics, both, were in large measure produced by scholars concerned with physical and biological science and with engineering problems. Only relatively recently, as social science has achieved the ability to pose new problems for the mathematician or mathematical statistician, has concerted effort been devoted to the evolving of techniques specifically oriented to social science types of research problems.[14] Such effort, it may confidently be expected, will increase and, especially, as more social scientists achieve high mathematical and statistical competence. The social scientist himself, cum statistician, is both more apt to be able to define the problem and to be motivated to find the solution than the non-social science oriented methodologist.[15]

New Uses of Computer. Another item that merits discussion is the new opportunity for a major inflection point in the advancement of social science by means of social statistics afforded by the advent of the computer. Electronic processing of statistical data has already produced great dividends both in the handling of mass data[16] and in the statistical analysis of social data.[17] But only a beginning has been made. For thus far, the computer has been used largely routinely as a glorified punch-card tabulator or for the conduct of time-honored forms of analysis.

The computer has yet to be fully exploited for social statistical analysis of a type not feasible before its appearance. For example, in addition to n-dimensional cross tabulations of data in which the unit of analysis is the person, household, age-group, geographic area, etc., tabulations are now possible of patterns and sequences of characteristics--of profiles of the units, instead of the units themselves.[18] Such tabulations would permit utterly new types both of micro- and macro-analysis that could prove exciting and probably greatly improve the power of social science to predict and explain. To gain the most from this possibility the social sciences must cultivate the training of scholars who combine statistical with computer competence.

Use of Theory. The increasing role of social statistics in social science methodologically and substantively, has helped push into the background the type of speculative and philosophical activity usually, but mistakenly in my judgment, called "theory" in the social sciences. Theory, in the imagery of science, is neither "speculation"--the undisciplined drawing of global generalizations without the benefit of empirical research; nor "philosophizing"--the disciplined drawing of generalizations without the benefit of empirical research. Theory in science may be considered as tentative generalization in the form of predictability or explanation based on empirical research and pointing to further research.[19]

At the present time it is correct, I think, to say that a wide gap still exists between empirical research and "theory" in social science to the detriment of both. There can be no doubt but that the advancement of social science is being badly retarded by this gap. No one saw this more clearly than Stouffer who was equally impatient with "talky-talk" sociology, on the one hand, and blind data collection on the other. He felt that progress in social science could be achieved primarily through the interplay of theory and empirical research--either one alone was sterile. He left some brilliant examples of the way in which such interplay could be effected in his work on migration and norms of behavior.[20]

The social statistician, having begun the task of using methods of empirical research and building a fund of empirical data, needs increasingly to turn to the construction of theory. The social scientist who specializes in "theory" by the same token, needs increasingly to turn to the use of empirical research. Social Science may be expected to achieve important progress as the interplay between theory construction and empirical research is increased and, it may be hoped, as these two functions are increasingly combined in the person of the well-trained social scientist.

Training. The speed at which many of the specific obstacles to the advancement of social science can be reduced will, in large measure, depend on the extent to which the social scientist obtains improved mathematical and statistical training. Although notable improvement has taken place over the years, the average social science student still tends to grumble about even the limited mathematical and statistical course work now generally required. In many cases it is undoubtedly true that the student just lacks quantitative aptitude--and that is why he is in social science. But many social science students lack not mathematical aptitude as much as incentive and motivation to obtain the desired competence. Social science faculties could remedy this situation relatively quickly, without undue hardship on students, if more of them would prescribe at least some pathways in social science training in which high mathematical and statistical competence are prerequisites. A number of social science departments have already done this--but the number is still relatively small. Much more needs to be done to encourage social science students to achieve higher mathematical and statistical training.

This is not to say that statistical and mathematical methods are the only methods by which social science may advance. But there can be little doubt that more and better social scientists cum statisticians would greatly improve social science output. Certainly there is no evidence, yet, that we have too many well-trained social statisticians.

SOCIAL STATISTICS AND SOCIAL ENGINEERING

The increased volume and quality of social statistics over the years are more directly traced to social engineering, than to social science, needs. For the demand for more statistics in virtually every realm, including the social, has steadily grown as our society has become more complex and more interdependent.[21/] As Mayo-Smith stated in 1910 "we cannot wait for the completed science" to find solutions for urgent social problems. Social statistics, in the sense of quantitative information about problems, are collected primarily to provide a sounder basis than is otherwise available for policy formulation and administrative action. In general, there are few realms of social action, in government or out, which are not dependent on social statistics for intelligence about the problems with which they deal, or which would not be better off if they had adequate statistics. I am sure Martin Gainsburgh's Presidential Address on "The Statistics We Live By" will furnish eloquent testimony of the social engineering utility of statistics on a number of fronts.[22/] And it should be added that statistical developments in respect of decision-making, in general, have greatly increased the utility of social, as well as other types of statistics, for action programs.

Social engineering needs have in large measure generated social statistics. But it is also true, of course, that social science has had an important role in determining the specific form that social statistics have taken. For the social scientist, although he has not had a completely free hand, has typically been the designer and operator of projects to collect and disseminate social statistics. The fact that social statistics often are the direct product of efforts to solve social problems is, on the one hand, the reason for their proliferation; and, on the other hand, the reason for their frequent inadequacies from the standpoint of social science. Social statistics have in large measure taken the form they now possess by reason of the interplay between social engineering and social science needs; and it is likely that a similar interplay will continue to influence the course of their development.

The great bulk of social statistics are government collected--federal, state and local--in response to needs not only of government but, also, of business, labor, education, civic enterprises, health and welfare organizations and a host of other types of organizations and groups.[23/] Additional social statistics, generally not as widely available, are generated by universities and research organizations; and by various non-governmental activities, largely as a by-product of administration.

Let us first briefly consider government statistics and then turn to the non-governmental.

Government Statistics. Government statistics have greatly expanded since the 1930's when the "bloodless revolution" of the New Deal placed the federal government into a number of new fields of activity. Preparation for, and participation in, World War II, and the requirements of the Cold War have had a similar effect. In consequence, the federal government itself has become the most important and largest single consumer of its own social statistical product. The fact that important policy decisions and administrative action of the federal government are increasingly dependent on the factual picture provided by social statistics has led to great scrutiny of, and much improvement in, many of the specific statistical activities of the government. The improvements effected in the Monthly Report on the Labor Force contribute a good case in point.

The status of social statistics produced by the government was excellently reviewed at a meeting of this Section just a year ago by Bowman, Gall and Rubin.[24/] In that paper, as well as in Stephan's paper at this session, some shortcomings of the federal statistical system were noted and indications given for improvement. Stephan makes a good case for furthering the improvement of statistics within the established decentralized statistical and administrative agency framework.[25] But it is also appropriate to observe that the federal statistical system was never really planned, that it has relatively few comprehensive reviews, and even fewer comprehensive overhaulings. It may well be that it would be in the interest of social engineering and, also, social science, for the statistical system as a whole to be not only reviewed but, also, overhauled from time to time, although not necessarily every 17 years as prescribed for another matter by Tom Jefferson. An occasional hard look and effective reshuffling could conceivably improve the utility of social statistics for social engineering purposes--as well as increase the statistical output per unit of cost.[26]

Non-Governmental Statistics. Social statistics produced by non-governmental organizations have also proliferated over the years and especially in the post-war period. Advances in the sampling of human populations and in social survey methods, in general, have made it easier for non-governmental organizations to do their own fact finding, or to utilize non-profit or commercial research agencies for the purpose. In consequence, social statistics has become part of the mass communication diet of the American people in an unprecedented way, directly through the public opinion polls, and indirectly, through the increased dependence of action agencies on social statistics for, and explanation of, action.

The quality of non-governmental social statistics varies widely. Some of the statistics, especially those collected by university affiliated organizations such as the Survey Research Center at the University of Michigan and the National Opinion Research Center at the University of Chicago, meet the highest standards. Other batches of social statistics, particularly those collected by commercial agencies under budget pressure, leave much to be desired. Especially distressing is the failure of a number of such organizations fully to publish the details of the designs, procedures and forms that they use. As the nation becomes more and more dependent on social statistics for effective and efficient action, the pressures for improved social statistics from all sources may be expected to increase. Fortunately, most of the important producers of such data have become increasingly professionally minded and strive to maintain and improve the quality of data as may be observed in the activities of the American Association of Public Opinion Research.[27]

Non-governmental as well as government agencies produce much in the way of social statistics as by-products of conducting their own activites. Some of these data are of high quality and are widely circulated as in the case of the Statistical Bulletin of the Metropolitan Life Insurance Company. Most of such material, however, is compiled and used for specific organization purposes and never enters the public domain. With increasing interdependence and government interventionism it may be anticipated that many of the social statistics now considered private will become public. The statistics relating to employment practices, for example, are in this process now under the pressure of increasing state and federal "fair employment" activities.

In general, the future of non-governmental as well as governmental social statistics may be considered a bright one in the sense that they may be expected both to be improved and to become increasingly important for social engineering purposes. It is hardly necessary to mention that one of the developments that justifies this optimistic outlook is to be found in the growing professionalization of statistical activities, as the growth and programs of the American Statistical Association testify.

CONCLUDING OBSERVATIONS

Social statistics, methodologically and substantively, has played a prominent role both in social science and social engineering. Social statistics, in turn, have proliferated largely as a result of social engineering, especially government, needs. But they have achieved a relatively high level of quality and utility largely as the result of the impact of social science.

Social statistics may be expected to achieve an even more important place than they now possess in social science in the years ahead. The natural sciences can make greater use of the experimental method and a direct mathematical approach than can the social sciences. In consequence, although statistical methods are useful in the natural sciences, they assume a much greater importance in the social sciences in that they represent the major means by which social phenomena can be quantified and the experimental method approximated or simulated. This is likely to remain the situation for many years to come, although direct mathematical approaches are increasing in the social sciences.

Similarly, social statistics is likely to achieve even greater prominence and utilization in social engineering. Our society continues to grow more complex and interdependent, and policy and program, therefore, require ever more intelligence for sound direction. The most efficient form of such intelligence yet devised is afforded by statistics, including social statistics. Moreover, developments in statistics, as well as electronics, have materially increased the efficiency and effectiveness of the decision making process itself.

It is fitting to close this paper at this Memorial Session in Honor of Professor Samuel A. Stouffer with acknowledgement of our indebtedness to him for the remarkable character of his contribution to social statistics both in social science and in social engineering. It is difficult to traverse the field of social statistics over the past thirty years without crossing Sam's tracks--directly or indirectly. No one has done more over this period either to advance the cause, or the science and art, of social statistics. I have not elaborated on his specific contributions here because I have already done so elsewhere,[28] and because I anticipated the other papers would do so.

Despite his premature death, Samuel A.

Stouffer is still in a position to exert great influence on the future of social statistics. Today, about thirty years after the completion of his own statistical training which he achieved relatively late in his career, he may yet be viewed as a prototype of the well-trained social scientist. It is a sad commentary that the vast majority of social scientists, including even the majority of recent Ph.D.'s, have not matched his training in methods of research. In this thought lies an additional well-deserved tribute to Sam; and a basis for sombre reflection for the rest of us about how much remains to be done to produce competent students really equipped to advance both social science and social engineering by means of greatly improved social statistics.

FOOTNOTES

1/ Richmond Mayo-Smith, Statistics and Sociology, New York: The Columbia University Press, 1910, Ch. 1.

2/ Morris R. Cohen and Ernest Nagel, An Introduction to Logic and the Scientific Method, New York: Harcourt, Brace & Co., 1934, Ch. X.

3/ Ibid., pp. 195ff.

4/ Ernest Nagel, The Structure of Science, New York: Harcourt, Brace & World, Inc., 1961, Ch 2, 14.

5/ Cohen and Nagel, op.cit., XIII; E. Bright Wilson, Jr., An Introduction to Scientific Research, New York: McGraw Hill, 1952, Ch. 4.

6/ Harold H. Kelley and John W. Thibaut, "Experimental Studies of Group Problem Solving and Process," in Gardner Lindzey, Handbook of Social Psychology, Cambridge, Mass.: Addison-Wesley Publishing Co., 1954, Vol. II, Ch. 21. See bibliography of Ch. 21 for further references.

7/ Samuel A. Stouffer, "Some Observations on Study Design," American Journal of Sociology, Vol. LV (1950), pp. 355-361.

8/ Paul F. Lazarsfeld and Morris Rosenberg (eds.), The Language of Social Research, Glencoe, Ill.: The Free Press, 1955; Paul F. Lazarsfeld (ed.), Mathematical Thinking in the Social Sciences, Glencoe, Ill.: The Free Press, 1954.

9/ Herbert Blumer, "Sociological Analysis and the 'Variable,'" American Sociological Review, Vol. XXI (1956), pp. 683-690.

10/ Frederick F. Stephan, "Relations of Some Social Science Concepts to Statistical Data," Proceedings of the Social Statistics Section, 1959, Washington: American Statistical Association, 1960, pp. 170-171.

11/ W. Edwards Deming, "On Errors in Surveys," American Sociological Review, Vol. IX (1944), pp. 359-369.

12/ Philip M. Hauser and Otis Dudley Duncan, "The Data and the Methods," in Philip M. Hauser and Otis Dudley Duncan (eds.), The Study of Population: An Inventory and Appraisal, Chicago: University of Chicago Press, 1959, Ch. 3. See bibliography, especially references to work of Morris Hansen, William Hurwitz, Leon Pritzker and A. Ross Eckler.

13/ E.g., P.K. Whelpton, Cohort Fertility, Princeton: Princeton University Press, 1954.

14/ E.g., Robert R. Bush and Frederick Mosteller, Stochastic Models for Learning, New York: John Wiley & Sons, 1955; H.A. Simon, Administrative Behavior, New York: Macmillan, 1947; Leo A. Goodman, "The Use and Validity of a Prediction Instrument. II. The Validation of Prediction," American Journal of Sociology, Vol. LVIII (1953), p. 510; "Measures of Association for Cross Classification" (with William Kruskal), Journal of the American Statistical Association, Vol. LIV (1959), p. 123; "Snowball Sampling," Annals of Mathematical Statistics, Vol. XXXII (1961), p. 148.

15/ The career of Professor Leo A. Goodman of the Department of Sociology and Statistics at the University of Chicago is a good case in point.

16/ The use of UNIVAC by the U.S. Bureau of the Census for tabulation of the 1960 Censuses of Population and Housing is a case in point. See Richard A. Hornseth, "Programming the Population Census," Proceedings of the Social Statistics Section, 1959, Washington: American Statistical Association, 1960, pp. 200ff.

17/ E.g., James S. Coleman, "The Use of Computers in the Study of Social Structure: Interaction in a 3-Person Group," Proceedings of the Social Statistics Section, 1959, Washington: American Statistical Association, 1960, p. 42.

18/ This possibility was proposed by my colleague, Professor O.D. Duncan, in discussions about next steps in the development of tabulations of social data. The idea merits intensive exploration.

19/ Philip M. Hauser and Otis Dudley Duncan, op.cit., p. 15ff.

20/ Samuel A. Stouffer, "An Analysis of Conflicting Social Norms," American Sociological Review, Vol. XIV (1949), 707-717.

21/ Philip M. Hauser and William R. Leonard (eds.), Government Statistics for Business Use, New York: John Wiley & Sons, 1956, Ch. 1.

22/ Martin Gainsburgh, Presidential Address, American Statistical Association, 1961. Will appear in March 1962 issue of Journal of American Statistical Association.

23/ Bureau of the Census, Statistical Abstract of the United States, 1961, Washington: Government Printing Office, 1961 (issued annually). See especially "Bibliography of Sources," pp. 957-994.

24/ R.T. Bowman, Alexander Gall, and Israel Rubin, "Social Statistics: Present Conditions, Future Needs and Prospects," in Proceedings of the Social Statistics Section, 1960, Washington: American Statistical Association, 1961, pp. 74-81.

25/ Frederick F. Stephan, "Samuel A. Stouffer and the Progress of Social Statistics," to appear in Proceedings of the Social Statistics Section, 1961 of the American Statistical Association.

26/ Philip M. Hauser, "Non-random Observations on Statistics from Washington," The Federal Statistics Users' Conference, Washington: Fourth Annual Meeting, September 1960, pp. 65-74.

27/ See Public Opinion Quarterly, the journal of the American Association of Public Opinion Research.

28/ "Samuel Andrew Stouffer, 1900-1960," The American Journal of Sociology, LXVI (1961), pp. 364-365.

INSTITUTION AND IMAGINATION IN STATISTICAL PROGRESS

Nathan Keyfitz, University of Toronto

My contribution to this session in honour of Professor Stouffer is based on a very brief acquaintance with him, and leans on his published work and on the comprehensive papers of Professors Stephan and Hauser which precede this one. In the course of working out these remarks I also talked with a number of people, including Mr. William Stouffer, brother of the man we are honouring. He told me something of the life of the Stouffer family in Sac City, Iowa, about the turn of the century when Sam was born. The family owned a weekly newspaper, the Sac City Sun, on which Sam worked after taking a master's degree in English at Harvard, and it seems likely that the contact with men and affairs, both at home and later as a reporter, helped to kindle in the young man the deep curiosity about how society works that was to characterize him throughout his life. After three years of reporting, Stouffer decided that the way to more solid knowledge of the operation of society lay through social science, and he enrolled at the University of Chicago, whose department of Sociology had been given its shape during the 1920's by another ex-reporter, Robert E. Park.

Stouffer's mature work included powerful insights into the means of finding out about society, some of which he exploited himself, some of which were exploited by students and colleagues. As an example of his style of thinking I shall refer to experimental design. He points out that a proper experiment can be designed to find the effect of a film on the attitudes of viewers, of a course on students, of training on soldiers, or of other "treatment" on those who are subjected to it.[1] The proper procedure includes observing the group before and after treatment and observing also, for purposes of control, an untreated group. This gives four cells to be entered with observation of attitudes: before and after, for each of the experimental and the control groups. In addition it ought to be decided at random, say by the toss of a coin, which of the two groups is to be treated and which used for control, and all this ought to be repeated in a number of trials or replications.

Stouffer shows what happens to this means of drawing conclusions when it gets into the hands of people who are careless or ignorant. They leave out the random allocation and the replications, but this still provides the four cells on which some kind of judgment might be made. But then they say that since it is the effect of the experiment that is sought, the control group is unnecessary; this brings them down to two cells: before and after in the experimental group. But after all it is only in the situation after treatment that the result shows, so why bother with before, which, if needed, can easily be constructed by recollection. Armies have always judged the worth of their training by the fine qualities of their men after training; Stouffer persuaded the United States Army to carry out a proper experiment. Suitable objective tests of physical stamina were devised, and their use in the experiment showed that "men who had been in the Army six months to a year and had been subjected to the old-fashioned conditioning system made little better scores on tests of strength or of stamina than did new recruits."[2] A new program of training was tried which proved to be clearly superior in a controlled experiment. On the strength of these results the Army scrapped its traditional procedures.

Such tests always meet resistence, for fear of the doubt they will throw on what has been the basis of day-to-day work; people who can bravely stand up to the thorough investigation of the illusions of others will fiercely resist any tampering with their own, and, because their cooperation is usually required, this has made progress slow. Professor Stephan has told us how agencies resist discussion of the accuracy of their data. Stouffer combined a realization of the need for doing the experiment with the ability to persuade those concerned to see its merits--I am told that his argument went up the army command in a chain rising through General Osborn to General George C. Marshall, then Chief of Staff. Senior officers who initially had supported a regulation unconditionally forbidding surveys of men in the forces later sat up nights reading Stouffer's reports. His persuasiveness worked horizontally as well as vertically and infected many colleagues, and the style of thought of which he was one of the first proponents has spread widely since the 1930's. Census and survey methods began to be tested; radiologists began to make duplicate readings of X-ray films; professors began to arrange duplicate gradings of essays. When someone asks today whether a long schedule will be less well filled out by respondents than a short one, the answer is not as obvious as it was before Stouffer and Hansen and others began to change the outlook of all of us. It may be that people get tired with the length of the questionnaire and their attention flags; it is also possible that as the successive questions are asked their interest increases. Only the type of experiment Stouffer was discussing can tell. As a result of many such experiments we now know that there may be a large difference in a census result arising out of seemingly small differences of procedure; that even with the identical procedure a random component is revealed in successive repetitions of the survey. The two preceding papers have cited many other examples of Stouffer's imaginative approach to problems.

To make the points which are to follow, it is convenient to compare Samuel Stouffer with John Graunt, who worked in London in the 17th century.

[1] Samuel A. Stouffer, "Some Observations on Study Design," AJS, 1950, 55, 355-61.

[2] Samuel A. Stouffer, "A Study of Attitudes," The Scientific American, CLXXX (1949), No. 5, 11-15.

He too was a social scientist, but, living in the days before foundation support and fellowships, he could start on his life's real work only after he had made good as a haberdasher and was able to retire. Graunt's Observations on the Bills of Mortality was published in 1662; we can now celebrate the 300th anniversary of this remarkable book. It should be celebrated for a number of reasons; for one, because like Mayo-Smith, who has just been cited by Professor Hauser, Graunt expressed a belief in the regularity of measurable events in the lives of men; he was able to trace some of these regularities even in his very crude data—the fact, for instance, that there are more births of boys than of girls but that the death rates of boys are higher, so that the sexes are about equal in the older population.

Graunt's life-long struggle to derive useful information from incomplete materials made him a backer of more deliberate methods of securing and publishing data. He spoke long and persistently of the need for a complete census. His voice had to be joined by many others before it was heard; a complete and regular census did come in England, but not until 139 years after the Observations were published. Starting about the end of the 18th century censuses were institutionalized in all advanced countries. This meant that considerable numbers of people came to depend on the census for making decisions on many matters. One indicator of whether or not an activity is an institution is secured by trying to think how we would go about our business without it. If doing without it is inconceivable, then it passes this test; it is institutionalized. It is hard for us to realize that prior to 1700 the population of Europe was not known with even approximate accuracy, and today we are not sure whether there were more people in 1700 than in 1300. Where once differences from century to century were of little concern, it is now essential to know how population is changing from year to year. Periodic censuses are part of our lives. Insofar as we can more easily imagine ourselves being ignorant of family budget patterns than of census information, we can say that censuses are more institutionalized. Censuses and vital statistics were the answer to John Graunt's problems; there is no equally simple answer to Stouffer's.

The way in which censuses become institutions has immediate application to underdeveloped countries. Towards the end of his life Stouffer became interested in world population, and the data available for its study engaged his attention. He had visited Puerto Rico, and he and his wife were planning a trip to India. The Population Council had secured his help in its work of extending knowledge in this field. He would have had to cope not only with the absence or inadequacy of censuses in many countries but also with the fragility of those which are set up in underdeveloped areas. Having few ties in the needs of the business community, government, labour, and the other elements of the community whose relation to the U.S. Census Professor Hauser has mentioned, such censuses are not easily brought into existence nor, once brought in, sustained. An excellent census was carried out in the Argentine in 1914, as its dusty reports testify. But it stands like a monument in a desert; it had no ties with the lives of men, was subject to no claims and pressures from other institutions, nor could it exert any, and when the man who had envisioned it left the scene he and his work were virtually forgotten; the next census was not taken until 33 years later. It is not such isolated efforts that are needed, but a continuing institution of census-taking, as independent of personalities as it can be made, as intertwined with other institutions as can be arranged.

When the Dominion Bureau of Statistics in recent years proposed to improve its statistics of crime, it began not by publishing volumes of data but by getting in touch with the police, finding out their needs, even making them aware of needs which they had not been conscious of before. Police chiefs and statisticians entered into fruitful collaboration on the design of the survey; former police personnel were taken on to the staff of DBS. We can look for a permanent institutional structure out of this process.

Professor Stephan has referred to the centralized and decentralized, Professor Hauser to the governmental and private, organizations in which statistical data are collected. Both have declared, as I have above, the dependence of these on other organizations outside the field of statistics. I would add only the effect on them of a different kind of institution—the association of statisticians with one another. If the individual statistician working for a soap company was oriented solely to his superiors, then he would be a maker or seller of soap and not really a statistician. It is the fact that he internalizes the norms of fellow statisticians in other organizations—including competing soap producers—and insists on applying these in his work that makes him a professional. His boss may be able to tell him to what problems he must direct himself, but cannot instruct him on how to tackle them. The role of the American Statistical Association, with which Stouffer was associated in a number of capacities, in developing norms and orienting statisticians to them is evident enough that it needs no underlining here.

New institutions often come into existence where there is conflict and struggle. Not only the work of Graunt, but even more that of the statisticians who followed him in the various states of Europe during the 17th and 18th centuries, constituted strong pressure for statistics of population. The mercantilists and their rulers had to know the number of their people, as well as their wealth, in order to promote measures that would advance their states in the military and economic competition of the day. To the demands of scholars like Graunt were added the requirements of those who thought that the more population the better for the state, and for whom censuses would record progress in making their state stronger than the neighbouring ones. Professor Stouffer's work with the War Department began in earnest within days of Pearl Harbour. The Canadian statistical system arose in part out of the dissatisfaction with the data available for guiding the economy during World War I. The strongly pressed claims of the Canadian provinces and of the American states for representation in the national legislature were resolved by using census figures and constituted the initial raison d'être of the census in both countries. The relation of the modern competitive economy to the demand for marketing and other statistics is familiar.

Sometimes the conflict out of which statistical work has been institutionalized is not military, political or economic, but intellectual. An example is the controversy between Malthus on the one side and the perfectionists and mercantilists on the other—the latter agreeing on little save the desirability of greater population. Discussion led to an avid searching for relevant data, and hence pressure for better censuses and vital statistics. Science thrives on the desire of people to prove themselves right, on condition that there is a scientific public able to compel the contestants to conduct their argument according to the rules of evidence.

In tracing the life of a scholar and man of action it is natural to seek to find what presently existing institutional forms can be attributed to his work. Of Dr. R.H. Coats, President of this Association in 1938, who died a few months before Professor Stouffer, one can say with some confidence that the centralized Canadian system is in a certain quite definite sense his achievement. The Ottawa situation was reasonably fluid when he arrived on the scene about the time Stouffer was born; Coats directed its flow along lines that he thought were good; he faced opposition, but was able to overcome it. If he had not lived and worked in Ottawa, or had taken some field other than statistics for his endeavours, Canada would probably not have her present arrangement for collecting and publishing data.

I have directed questions to those who knew and worked with Stouffer at different points in his career, but no correspondingly clear-cut assertion concerning his influence emerges from what I was told. He was advisor to the Bureau of the Census prior to 1940 and exerted much influence towards professionalization of the staff, giving strong moral support to the work that Morris Hanson, W. Edwards Deming, and others were beginning to do. He became Director of the Laboratory of Social Relations at Harvard and initiated many projects there. Prior to that he inspired people at the University of Chicago. One of his first students at the University of Wisconsin, before he had yet secured his Ph.D., was Harold Dorn, who followed him to London to study under Pearson. In what sense is the present work of the Bureau of the Census, of the statistics and sociology departments at Harvard and Chicago, of the Institutes of Health, the result of Stouffer's effort? At this point in tracing the history of our profession I was brought to a halt. All I learned was that much that now goes on has some relation to the personality of Stouffer; according to universal testimony he inspired colleagues, students, and administrators; he left them with questions and ideas which were combined with many others before they became embodied in continuing survey and research institutions.

I suggest that this arose partly from his own personality and partly from the situation in which he worked, which is that in which we all work. With the exception of a few items such as a five year census, what seems to be needed today is not so much the immediate establishment of more series that are definitely namable and specifiable. Issues have become infinitely more subtle since John Graunt called for the institution of censuses. To argue this point, which will be my last, I shall make use of the initial work of a continuing committee on social statistics of which Dr. Taueber is chairman. It was founded by Professor Stephan when he was chairman of the Social Statistics Section of ASA, and will I think help overcome the handicap each year's program committee faces as it starts afresh in its one-year term; by pursuing subjects from year to year and making long-range plans it will give continuity to this part of the ASA program, as well as furthering the cause of the Section in other useful ways.

Its initial deliberations help to outline the present frontier of social statistics. On that frontier is knowledge of how to conduct surveys of the ability of people to predict their own behaviour. Dr. Taueber points out in effect the formal resemblance despite complete difference of content among surveys of intentions to purchase automobiles and houses; the number of children that couples intend to have; the party a person intends to vote for. This resemblance ought to make it possible to transfer experience from one field to another. It happens to tie in also with a part of Stouffer's work that has just been mentioned. He was chairman of an important SSRC Committee on the Prediction of Personal Adjustment, which concerned itself with prediction by neutral observers of success of couples in marriage, of students at school, of criminals on parole. Under Stouffer the committee reached some points at which its insights have not since been surpassed.

Again Professor Stouffer was vitally concerned at every point in his career with sharpening the tools of measurement. He measured the values of college students and the attitudes of soldiers. His work in the army came to a stop at one point before the apparently impenetrable problem of measuring fear. During this meeting of the ASA I have heard of the need for measuring the quality of medical care, the effectiveness of treatment for alcoholism and drug addiction; the degree of mental illness and quality of treatment; the factors which bear on the sentencing behaviour of judges.

One of the main features of American and European society at the moment is rapid upward mobility. The entire occupational spectrum is moving upward; in North America it has passed the point at which over half the labour force is in white collar occupations. Through what sequence of occupational changes do people cross the manual-whitecollar line? From what social groups do schoolteaching, engineering, medicine, and other professions recruit? What is the degree in which the educational system is being upgraded, especially at its upper end? To what extent is upward mobility reflected in changes of place of residence, from the countryside to the cities, and from the centre of cities to the periphery? Is the movement of people to the suburbs accompanied to any important extent by the movement of the industrial establishments in which they work, a process which would reduce the amount of commuting? What is the strength of the move back to the centres of cities, and who are the participants?

Attention is also being directed to statistical measurement of features of particular groups. The incomes, health, welfare services received by older people come to the fore with the increasing proportion of them in the population. At younger ages the decline virtually to zero in death rates deprives vital statistics of some of their interest and has begun to bring into existence on a periodic basis statistics of health and sickness, but these are not yet by any means in finally satisfactory form. Professors Stephan and Hauser have offered important indications of where the present statistical frontiers lie.

The preceding argument has distinguished between imaginative thinking in statistics and the establishment of institutions turning out data. The former is the basis of the latter, but often the links are indirect. It is institutions that get the work of the world done in a reliable way; being in some degree independent of individual talent they enable the majority of mankind to pursue useful careers. Imaginative thinking, on the other hand, as often dissolves old institutions as it founds new ones. It seems that what is needed at the present stage in the history of statistics is not so much a greater quantity of institutionally provided data as the better defining of statistical variables, better knowledge of survey methods, in some instances the replacing of existing series by quite different new ones. Unlike John Graunt, Stouffer could not simply call for the establishing of censuses, or any other equally clear-cut answer to our problems. The inspiration he gave was both more subtle and more fundamental. This brings me to the same somewhat depressing conclusion as the other papers on this program: that what our age needs most is the element hardest to come by—men with the training, curiosity, honesty, persistence, imagination and brilliance of Samuel A. Stouffer.

XI

CONTRIBUTED PAPERS IN SOCIAL STATISTICS

Chairman, Joseph Steinberg, Bureau of the Census

FACTORS AFFECTING FAMILY INCOME: RESULTS OF A CONDITIONAL PROBABILITY ANALYSIS

James Morgan, University of Michigan
Martin David, U. S. Department of Treasury (On leave from University of Wisconsin)
Harvey Brazer, U. S. Department of Treasury (On leave from University of Michigan)

We present here an analytical model used to explain family income in the United States, a brief description of the statistical methods, and some selected results. The analysis of family income is itself part of a larger study of income distribution and redistribution, and of patterns of intergenerational change, to be published by McGraw-Hill Book Company. The study was supported by a grant from the Ford Foundation and smaller supplementary grants from the Office of Education and the Federal Office of Vocational Rehabilitation. Wilbur Cohen was a fourth principal researcher on the total project.

In the United States, some of the forces affecting family income are decisions and acts of the individuals themselves. Income is no longer a predetermined, exogenous factor to the individual family. Hence it is important to know how past history, outside forces, and individual decisions jointly determine family incomes, and their distribution.

Family income is made up of components, each affected by its own set of forces. We think of this process as a series of steps. At each step something is being determined as the result of outside forces and the acts and decisions of the individuals. An individual's earned income is the product of three things, each of which can be analyzed separately: whether he works, how long he works, and at what hourly rate. We thus have a whole set of things to explain, namely the items in the boxes of Chart 1.

As you can see, we ran ten multivariate analyses, relying on two-way tables to examine the other less important components such as transfer income from outside the dwelling, income taxes, and the income added by secondary units living with relatives.

You can think of the results of the analysis as allowing the derivation of an expected value of the income of a given family as follows: using various factors, estimate the probability that the head worked. Then use a table of random numbers and decide whether that particular head worked. If he worked, find expected values for hourly wage rates and hours and multiply one by the other to estimate the head's earnings. A similar estimate of expected value can be made for capital income, and income (money saved) from growing food or fixing up the house. The sum of these plus some small earnings of minors provide an estimate of the factor income of the unit excluding the wife's earnings. That income and other explanatory factors are then used to predict whether the wife works, her wage rate and hours. The resulting total "gross factor income" then helps explain transfer income, income tax, and whether the unit lives with relatives, either becoming dependent upon them, or providing housing for them. The estimated family income then includes income added by secondary units.

The decisionsinvolved at the various stages are clearly not independent of one another. Many of them are alternatives, and hence jointly determined, such as whether the head of a unit should work longer hours or let his wife go to work. We cannot be so circular as to use the head's income to explain why his wife works, and use his wife's income to explain how long the head works. We treat these joint decisions by making one conditional on the other. Since one can always convert a joint probability into a set of conditional probability statements, the overall results will be the same whichever we treat as being determined first, and thence influencing the next step.[1]

The particular sequence we have chosen makes some empirical sense, too, we feel. It assumes that transfer incomes are largely determined by other income of the unit, not the reverse, and that the decision to live with relatives is usually made in the light of what the unit's income is, though a few people might decide to live with relatives in order to loaf.

At each stage, then, the thing to be explained is seen as the result of various exogenous forces and of previous things already determined. In some cases the order is inevitable. We can analyze the decision of how long to work only for those who had decided to work (though the possibility of part-time work may have influenced the decision to work).

Space does not permit us to present more than a small sample of the results, a full set of which will appear in a book to be published by McGraw-Hill in 1962. They come from a national probability sample interviewed in the spring of 1960.

The analysis uses smaller units than the family: spending units, or where possible adult units -- adults or adult couples with their children, sometimes called nuclear families because the analysis of the decision to combine to form larger units with higher incomes was part of the scheme. A glossary of terms is provided.

The method of analysis at each stage was essentially multiple correlation of regression with dummy (one or zero) variables, though the IBM program is actually an iterative one, not a matrix inversion.[2] It is an extension of multiple

[1] That is, the overall predictions will be the same. The estimated effects of different factors will depend on the order, to the extent that the factors in the second analysis are correlated with the outcome at the previous stage which is used in the second analysis as one of the predictors.

[2] See Daniel Suits, "Use of Dummy Variables in Regression Equations," Journal of the American Statistical Association, 52 (December, 1957), 548-551.

correlation to the situation where the explanatory factors are membership in subclasses like age groups, rather than numerical variables. Like any regression, it minimizes the summed squared errors of predictions made by the derived formula. We use dummy variables even where it might be possible to form a numerical scale of the explanatory factor. This has the advantage that we need make no restrictive assumptions about the linearity of the effect.

The most important restriction of this form of analysis is the assumption that each explanatory factor affects the dependent variable in an additive manner, regardless of the value of the other explanatory factors. This assumption of additivity is, of course, only an approximation to reality. Where interaction effects seemed likely to be important, we built them in from the beginning by using joint classification on two dimensions at once, or checked later by rerunning the whole analysis for part of the population, e.g., white, nonfarmer males not yet retired.

The analysis used various number of predictors, depending on which of the available measures seemed appropriate. Where the same factor, such as race or education, is used at each stage, its total effect in the determination of family income can be inferred from its effect at each stage.

We shall discuss here only one of these analyses, that explaining the hours worked during 1959 by the 86 per cent of the heads of spending units who worked at all. Clearly, purposeful decisions, like taking on a second job or working overtime, and involuntary results of outside forces, like illness or unemployment, affect the hours a man works during a year. Indeed, our theoretical model groups the various characteristics used to explain hours worked into constraining factors (like local unemployment and the worker's physical stamina and capacity for long hours of work) and motivational factors (like the pull of wages and the need for money and the desire to get ahead).

Each of these has, as proxy measures for it, a number of measured characteristics of the individual, his family, or the situation he faces. Sometimes the same measure is a proxy for more than one theoretical construct. When we realize also that some are measured more accurately than others, it becomes clear that the interpretations of the results requires some judgment beyond the usual statistics.

However, the results are presented in terms of the importance and the significance of sets of dummy variables, and for each set, both the unadjusted and adjusted (multivariate) coefficients are presented.[3]

[3]For those not familiar with dummy variables they are merely variables which take on only one of two values, one or zero, one if the individual belongs to a particular class on a particular characteristic. In the case of "race," there is a dummy variable "nonwhite" which is equal to one if the individual is nonwhite, otherwise it is zero.

Table 1 shows the sets of subclasses in order of their importance. We are clearly interested in the importance of each set, such as "hourly wage rate." The measure we use is analogous to the beta coefficient of multiple correlation.[4]

The significance test is somewhat less legitimate, treating the adjusted coefficient as a set of means and comparing the variance estimate derived from them with that derived from the varaince unexplained by the whole regression. The measure is probably somewhat nonconservative. The different ranking of race as to importance and significance is, of course, meaningful. A significant difference applicable only to a small proportion of the population aids prediction less than an equally significant difference affecting large groups.

We have dealt with the problems of an additive analysis where interactions exist, by building them into the characteristics from the beginning; for instance, education and age form a joint classification, as do sex, marital status, and children in the "adult unit composition," and a personality measure and an attitude in "need-achievement and attitude toward hard work."

You will notice that we use both "factual" and attitudinal variables, as well as local conditions and family background measures.

We have selected only a few of the specific classifications in Table 1 to duscuss here because of the shortness of space.

Age and education interact in their effect on earnings, and the joint classification, used elsewhere in explaining wage rates was also used in the analysis of hours worked. (See Table 2.) It turns out that not only do those with more education earn more per hour, but particularly when they are older, they manage to work more hours. Most of these differences are the result of differential unemployment, not different motivation. The differences between the third and fourth columns represent the multivariate adjustments for the effects of the other predictors used, and are relatively small and unsystematic here. For instance, the top row shows that high school dropouts under 25 worked 267 hours less than the average of 2092, but that after adjustment for other things, their low age and education account for only 164 hours of the difference.

In regression, one cannot have dummy variables for each class of a characteristic without overdetermining the system, but the coefficients can always be converted into a set, one for each subclass, with a weighted mean of zero for each set, and this makes the constant term in the predicting equation equal to the grand mean. Our iterative program produces its output in this form in the first place.

[4]The beta coefficient is generally considered not so good as rerunning the analysis without that factor to observe the loss in predictive power, but better than a coefficient of separate determination.

Table 3 deals with an interesting economic problem of the supply of labor in response to wage rewards, and illustrates the advantage of multivariate adjustments for other factors (like age and education), for only after adjustment does the negative relation between wage rates and hours worked become apparent. Why is the adjustment so dramatic? Because there are a number of factors causing spurious positive relations between wages and hours: professionals work longer hours at higher wages, because they are professionals, not because of the wages. The disabled work shorter hours at lower wages, yet few would attribute the shorter hours to the lower wages. Similarly, the young and uneducated have more unemployment and lower wages, but we should not attribute the unemployment to the low wages.

Why the negative relationship? Is leisure a superior good with a high income elasticity so that increases in real income from a higher wage rate are partly used to purchase more leisure? Perhaps so. Or perhaps there are minimum standards, that people who earn less per hour than others in their subclass, feel compelled to work longer to achieve an acceptable standard.

Table 4 shows a small but probably significant tendency for those with plans for the future that will require funds, to work longer hours now. The reduction in the effect is the result of the fact that the middle aged, more highly educated people, have more plans. The "no plans group" is composed of some persons who do not plan and others who have no living parents or children in school for whom to plan. Plans to send children to college seem to have more impact than plans to help parents, and having both types of plans seem to have an effect roughly equal to the sum of the two independent effects.

Finally, because of intrinsic interest, and the possibility that with better measures the factor might prove still more important, we look at Table 5. It is a joint classification according to a personality measure, and an attitude, both of which appear to have some effect on hours worked. The personality measure is an index of the achievement motive, behind which there is a great deal of theory and laboratory experimental work, and some analysis of cultures and subcultures.[5] The possibility is suggested that achievement motivation may affect economic progress through the behavior of the masses as well as through the initiative of the small entrepreneurial class.

Theoretically the achievement motive, the propensity to derive satisfaction from overcoming obstacles by one's own effort in situations where one's own performance can be compared with some standard of excellence, is developed by early childhood factors, and changes only slowly if at all under the impact of subsequent experience. It seems to affect how many hours people work. In other parts of our study it also affected wage rate, education completed by the head, and by his children, and education planned for the children who had not finished with school.

The subjective probability or belief that hard work leads to success in this world, rather than luck or help from friends, is a measure of a more volatile attitude which is assumed to be subject to change according to one's experience. Atkinson's theoretical model has this attitude interacting with the achievement motive, their product being the resultant motivation to act.[6] In our case, to act means to work long hours. Perhaps because of other constraints on hours worked, our data would indicate something less than a multiplicative relationship.

Summary

We have attempted to provide something of the outline of our analysis, and the flavor of its results. The multivariate dummy variable procedure with selected interaction effects built in is seen to be useful in distilling the effects of different factors. (Professor Orcutt and his colleagues make extensive use of this method in preparing data for their simulation model of the economic system, reported earlier at these meetings.)

The analysis does not deal with the fact that explanatory factors may be at different levels in the causal chain, even when we break decisions or results down as we have. For instance, if achievement motivation is indeed fixed in early childhood it may help determine how much education the individual completes but education could not affect achievement motivation. Putting them both in a simultaneous multivariate analysis allows education, through which the achievement motive operates, to take credit for something which is ultimately the result of prior motivation.

Ultimately what is needed in the analysis of data is more flexible multivariate analysis procedures which take account both of interaction effects, and of the logical sequences which are possible when one explanatory factor can affect a second explanatory factor, but cannot be affected by it.

[5]See John W. Atkinson (ed.), Motives in Fantasy, Action and Society (Princeton: D. VanNostrand, 1958); and David McClelland, The Achieving Society (Princeton: D. VanNostrand, 1961).

[6]John Atkinson, "Motivational Determinants of Risk-Taking Behavior," Psychological Review, 64 (November, 1957), 339-372.

Chart 1

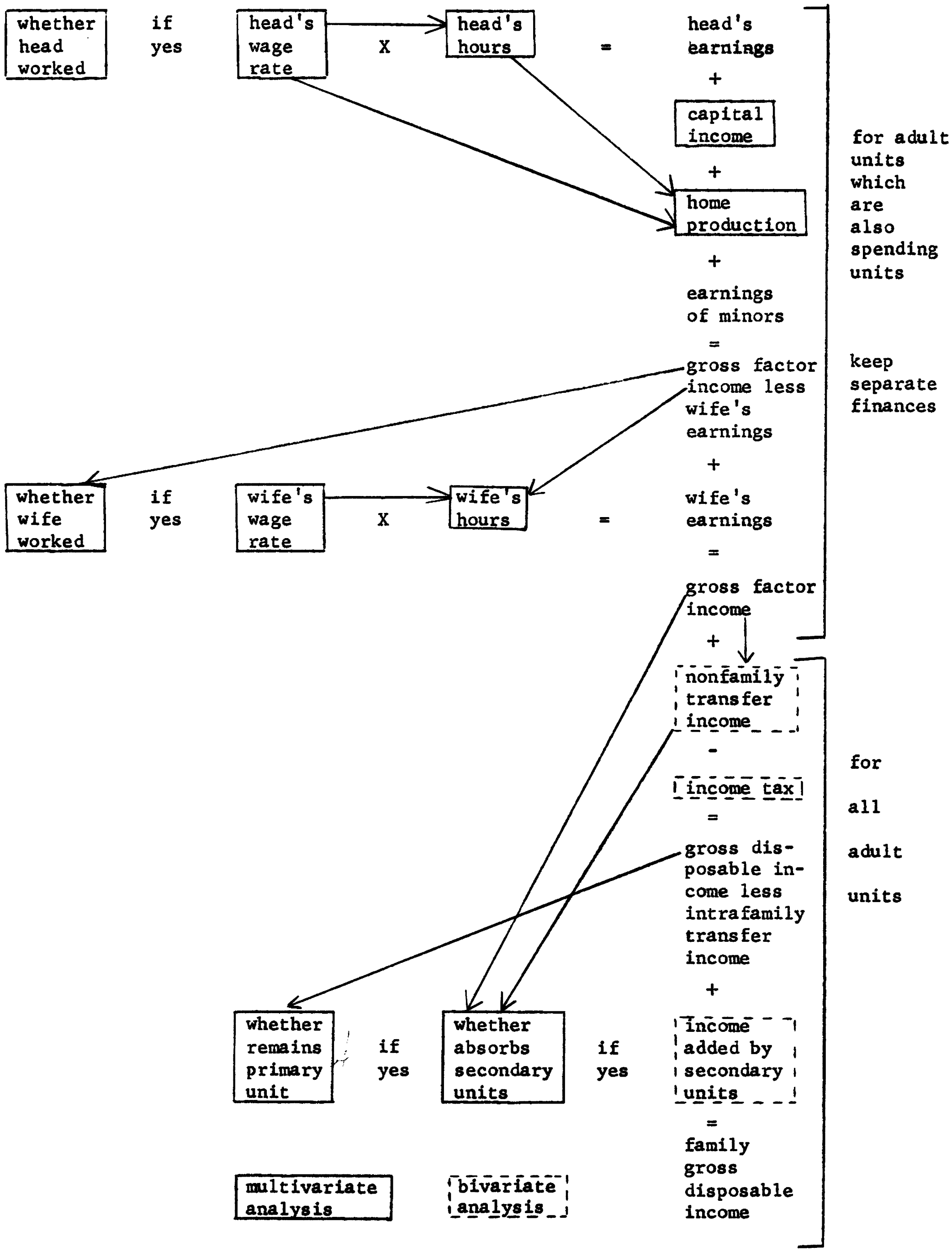

whether head worked
if yes
head's wage rate
X
head's hours
=
head's earnings
+
capital income
+
home production
+
earnings of minors
=
gross factor income less wife's earnings
+
whether wife worked
if yes
wife's wage rate
X
wife's hours
=
wife's earnings
=
gross factor income
+
nonfamily transfer income
-
income tax
=
gross disposable income less intrafamily transfer income
+
whether remains primary unit
if yes
whether absorbs secondary units
if yes
income added by secondary units
=
family gross disposable income
multivariate analysis
bivariate analysis
for adult units which are also spending units
keep separate finances
for all adult units

Table 1

Characteristics used to explain hours worked by spending unit heads
(heads who worked during 1959)

Characteristics of spending unit heads	Indexes of: Relative importance (Beta coefficient)	Significance (F-ratio)
adult unit composition	.291	42.4**
education and age	.258	11.7**
occupation	.254	56.5**
hourly earnings	.234	24.0**
physical condition	.115	15.4**
plans to help parents or children	.084	8.2**
need-achievement score and attitude toward hard work	.066	2.5*
religious preference and church attendance	.063	2.3*
race	.044	6.8**
extent of unemployment in states	.032	0.9
difference in education of heads and wives	.025	0.4
immigration of heads and fathers	.018	0.6

** significant at probability level of .01
* significant at probability level of .05

Table 2

Hours worked: deviations for education and age
(for spending unit heads who worked during 1959)

Education and age	Number of cases	Per cent of spending unit heads	Unadjusted deviations*	Adjusted deviations*
0 - 11 grades				
under 25	76	2.5	-267	-164
25 - 34	251	7.8	72	46
35 - 44	297	9.5	69	90
45 - 54	332	10.4	59	106
55 - 64	269	8.5	-167	-53
65 and older	104	3.8	-776	-699
12 grades				
under 25	72	2.5	-101	-34
25 - 34	97	3.4	215	133
35 - 44	129	4.5	150	160
45 - 54	78	2.6	178	130
55 - 64	32	1.1	-220	-90
65 and older	5	.2	83	-205
12 grades and some college or nonacademic training				
under 25	84	3.0	-404	-376
25 - 34	142	5.0	134	50
35 - 44	139	5.0	323	224
45 - 54	99	3.5	185	177
55 and older	77	2.8	-234	-159
college degree				
under 35	98	3.4	-126	-312
35 - 44	82	2.9	280	106
45 - 54	53	1.9	90	45
55 and older	53	1.8	111	95

* deviations from grand mean of 2092 hours

Table 3

Hours worked: deviations for hourly earnings
(for spending unit heads who worked in 1959)

Hourly earnings	Number of cases	Per cent of spending unit heads	Unadjusted deviations*	Adjusted deviations*
none**	23	0.7	-426	-82
$0.01 - 0.74; negative amount	347	8.3	195	369
$0.75 - 0.99	175	5.0	-17	228
$1.00 - 1.49	397	12.5	-120	84
$1.50 - 1.99	337	12.0	74	114
$2.00 - 2.99	714	26.6	5	-52
$3.00 - 3.99	331	12.1	-34	-219
$4.00 - 4.99	130	4.8	-70	-269
$5.00 and over	115	4.1	2	-301

* deviations from grand mean of 2092 hours

** self-employed who just covered expenses

Table 4

Hours worked: deviations for plans to help parents or children
(for spending unit heads who worked during 1959)

Plans to help parents or children	Number of cases	Per cent of spending unit heads	Unadjusted deviations*	Adjusted deviations*
no plans to send children to college or help parents	1255	41.2	-153	-60
plan to help parents	466	15.6	-13	-7
plan to send children to college	520	17.8	168	68
plan to send children to college and help parents	328	11.5	277	117

* deviations from grand mean of 2092 hours

Table 5

Hours worked: deviations for attitude toward hard work and need-achievement score (for spending unit heads who worked during 1959)

Need-achievement score and attitude toward hard work	Number of cases	Per cent of spending unit heads	Unadjusted deviations*	Adjusted deviations*
hard work is more important than luck; need-achievement score is in:				
high range	631	22.3	28	28
middle range	921	31.4	39	25
low range	440	14.2	10	-29
hard work is less important than luck; need-achievement score is in:				
high range	120	3.9	0	18
middle range	228	7.3	-151	-72
low range	139	4.0	-231	-181
need-achievement score not ascertained	90	3.0	22	73

* deviations from grand mean of 2092 hours

Glossary of Terms

Adult Unit A person 18 or older, his spouse if he is married, and any children under 18 who live with him and for whom he is responsible.

Adult Unit Composition Marital status, sex, number of children under 18 for spending unit heads.

Dependent Adult Unit Any adult unit which does not contain the head of a spending unit. An adult unit which does not contain the major earner in the spending unit.

Family All occupants of a dwelling unit who are related to each other by blood, marriage, or adoption.

Gross Disposable Income Disposable money income, nonmoney transfers, money saved by home production, 6 per cent on investment in owner-occupied home. An alternative definition: gross factor income, total nonfamily transfers, net intrafamily transfers, less income tax. Total money and nonmoney income available to the unit after taxes.

Home Production Labor which the unit expended in growing some of its own food, performing repairs and improvements in the home. The value of home production is the respondent's estimate of the money that the unit saved by doing this work itself.

Hourly Earnings For heads and wives who are not self-employed businessmen or farmers, the ratio of wage and salary income to hours worked. For self-employed businessmen and farmers, wages and salaries were imputed by deducting a reasonable return on the capital investment in the enterprise from the business or farm income (6 per cent of capital in business enterprises, 6 per cent of livestock and equipment, 5 per cent of property in farms). Earning rates were calculated on the basis of these imputed wages and salaries. If the business or farm did not earn enough to cover the imputed return on capital, earning rates may be negative.

Need-Achievement Score A measure of the extent to which spending unit heads differentiate in favor of high status, high reward occupations, thought to be a measure of need-achievement.

Physical Condition Spending unit head's report of the presence or absence of an illness physical condition, or nervous condition which limits his work.

MEASUREMENT OF NONSAMPLING ERRORS IN A SURVEY OF HOMEOWNERS' EXPENDITURES FOR ALTERATIONS AND REPAIRS

John Neter, University of Minnesota and Joseph Waksberg, Bureau of the Census

I. Background and Purpose of Study:

A survey designed to provide quarterly data about expenditures for maintenance, repairs, alterations and additions to residential properties was assigned high priority among the construction statistics programs when the responsibility for these programs was given to the Bureau of the Census.

After examining alternative methods for collecting such statistics, it was decided to conduct personal interviews with owner occupants of 1 to 4 household properties in their own homes. These constitute the bulk of residential property owners. Information from renters was obtained in the same way. A high proportion of responses in these interviews were based on respondent's memories of expenditures.

Personal interviewing and reliance on respondent's recall are commonly used in consumer expenditure surveys though relatively little is known about the accuracy of the results obtained in this way. Previously conducted studies, designed by market research firms and university analysts were limited in scope and based on samples so small that definitive conclusions were difficult to reach. However, these studies did indicate the kinds of problems which could be anticipated.

Consequently, a program was designed to study response errors in the survey of residential alterations and repairs. The immediate purpose was to provide a rational basis for establishing permanent survey methods. However, the results of the research have bearing on the design of other consumer expenditure studies which are faced with similar problems. In fact some of the subjects studied, such as the impact of memory recall over variable periods of time, the effect of using different household members as respondents, and the effect of conditioning of respondents, are significant in many kinds of surveys and the results of the research may be of more general applicability in survey methodology.

In the current survey of residential alterations and repairs, data are also being collected by mail from owners of large properties and nonresident owners of small properties. However, the experimental studies are restricted to the interview survey in which only owner occupants of 1 to 4 family homes and renters are included. Although a study of renters expenditures was included in the original research plan it soon became apparent that the expenditures reported by renters are such a small part of the total that they have virtually no impact on the total estimate. Consequently, all of the analyses made are of owner expenditures.

II. Subjects Studied:

The research program was designed initially to provide information about the following subjects:

1) Effect of length of recall period on recall of expenditures
2) Effect of length of reference period on recall of expenditures
3) Effect of choice of household member as respondent on recall of expenditures
4) Extent of shifting of expenditures in time by respondents
5) Effect of repeated interviews in household
6) Effect of using detailed, probing questionnaire as compared with the use of a less intensive one

As a second step, information will be obtained about the items below. Information on these subjects is not yet available.

7) Effect of using alternative types of questionnaires
8) Feasibility of using mail questionnaires
9) Use of diary techniques
10) Use of double-sampling procedure

III. Experimental Design:

Most of the work that has been done, is in connection with the first five items listed above. These items were studied by measuring expenditures for identical periods of time by different random samples of households with a different enumeration technique used for each sample. An alternative research method, involving record checks, was considered but records of sales or contracts on a broad enough scale to provide useful results do not seem to exist.

More direct methods of measuring the erroneous allocation in time of expenditures and the effect of obtaining replies from more than one respondent in a household have been tried recently, but the results are not yet available. Nor have procedures been developed for studying the subjects listed for future study (Items 7 to 10 above). Therefore, the remainder of this report will describe only the initial research program results.

The effect of using detailed probing questionnaires as compared with the use of a less intensive form (listed as Item 6 above) was studied by adding a relatively simple questionnaire to the Bureau's Current Population Survey for one quarter and comparing the results with the reports for the same quarter obtained in the regular survey of residential alterations and repairs. Since a variety of enumeration methods were used for the regular survey, there were reasons for differences in addition to those caused by questionnaire design. Consequently, comparisons have been made between the CPS supplement results and data from a subsample of the regular survey for which the same length of recall and choice of respondents were used as in the CPS supplement. Comparisons were also made between the CPS supplement results and those from the full Survey of Residential Alterations and Repairs sample.

Assigning the wrong time period to expenditures, effect of length of reference period, recall problems and conditioning of respondents (items, 1, 2, 4 and 5 listed above) were studied by using the following enumeration procedures:

1) Unbounded type 1 -- In this type of interview, the enumerator has no information from a previous interview about the household expenditures and asks only about expenditures for the preceding month.

2) Unbounded type 6 -- Enumerator has no information from a previous interview about the household's expenditures and asks about expenditures during the preceding 6 months. The respondent allocates the expenditures he reports to the specific months in which he thinks the expenditures were made.

3) Bounded 1 month -- This describes the second and third interviews in the same household. The interviewer asks about the preceding month only, but he brings with him a record of the previous interviews which he reads to the respondent in order to eliminate any possibility of duplication of expenditures, and to fix the time period of reference more clearly in the respondent's mind.

4) Bounded 3 months -- This interview is similar to the bounded 1 month, but is performed after a lapse of 3 months from the previous interview in the household and covers expenditures during the 3 months. The respondent also allocates expenditures to the specific months in which they were made.

In order to study conditioning and problems associated with the respondent's allocation of expenditures by months, it was necessary to keep households in the sample for more than one interview. This feature also appeared useful in that it took advantage of whatever correlation over time existed for identical households and segments, minimizing the sampling error of tne differences. Consequently, a rotation scheme was established in which, to the extent possible, all of the different enumeration techniques were used in the same households over the course of time. It was not possible to conduct unbounded 1 and 6 month interviews in the same households but all the other procedures were used for each household.

The rotation scheme operated in the following way. Each household remained in the sample for four interviews. Half of the first interviews were unbounded type 1 and the other half unbounded type 6. In the second and third interviews the bounded 1 month interview was used. The fourth and last interview was of the bounded 3 month type.

The rotation system used to implement this randomization is shown on the attached chart. Each panel is an independent random sample of 50 segments containing an expected 300 households. Two new panels were introduced each month. Once the rotation system became stabilized, during any month the enumeration comprised about 2400 households distributed as follows:

Unbounded type 1 -- 300 households
Unbounded type 6 -- 300 households
Bounded 1 month -- 1200 households
Bounded 3 months -- 600 households

For the study of the effect of alternate household members as respondents, the sample households were also randomized and a specific respondent designated for interview in each random group. This randomization was done independently of the type of interview pattern described above. The randomization was in four groups and these were designated for interview as:

1) Head specified as respondent
2) Wife specified as respondent
3) Joint interview with head and wife specified
4) Respondent unspecified

The designated respondent remained the same for all 4 interviews. During any month 40 percent of the households were designated for nonspecified respondent and each of the other 3 types of designation was used in 20 percent of the households.

The identical questionnaire was used for all of the interviews. It was designed to provide intensive interviewing and enumerators were instructed to probe carefully for reports of expenditures and their timing. Before the first interview a letter was sent describing the nature of the survey and indicating the scope of expenditures covered by it. During the second and subsequent interviews the enumerators were given a shuttle record containing the information reported previously and were instructed to read it before asking the questions about expenditures. The questionnaire provided for the reporting of each maintenance, addition or repair job separately, to improve reporting, permit classification of expenditures by type, and make possible the analysis of response accuracy by type or size of job.

IV. Method of Analysis:

Comparisons of expenditures were made for identical time periods using different enumeration techniques and involving different lengths of recall periods. Tabulations of this type have been made for estimates of the number of jobs reported and the value of expenditures for these jobs. However, the value of expenditures is subject to such large variances that, with the sample sizes used, very few inferences can be made. A clearer picture of the factors affecting response accuracy is obtained by comparing estimates of the number of jobs, by size of job. These estimates are subject to much smaller sampling variability. Of particular interest are the estimates of the number of jobs over $100 and over $500, which account for about 75 percent and 35 percent, respectively, of all expenditures.

In order to minimize the effect of sampling variability, the longest time period, for which data were available, was used for each comparison. This period varied for different kinds of analyses because the rotation plan did not provide data for some enumeration methods in the early or late months of the survey. For example, data for 3 month recall bounded interviews were not available until the survey had been conducted for 4 months. As a result, different comparisons use different time periods, though each comparison shows the effect of different enumeration methods within the same time period.

In the analysis of the results, it has been assumed that the bounded 1 month interview provides the greatest accuracy. Evidence of bias for some of the procedures therefore follows from comparisons made with this type of interview. This assumption is not made in the more detailed studies whose results will be available later.

The problems of assigning expenditures to the wrong time period and the effect of differing lengths of reference periods are studied by comparing: (a) the unbounded with the bounded 1 month interviews, (b) the unbounded type 1 with type 6 interviews, (c) the bounded 3 month and unbounded type 6 interviews. Two aspects of time-reference problems are thus studied: displacement from the time before the one designated as the desired recall period and displacement within the desired recall period (usually from the earlier to the current months). For most of the analysis, however, only data from the most recent 3 of the 6 month recall period have been analyzed.

Problems of recall failure are examined by comparing the bounded 3 month with bounded 1 month interviews. Data for the unbounded type 6 interview also provide some evidence about recall.

Conditioning is studied by: (a) comparing the first and second bounded 1 month interviews, and (b) comparing the trends throughout the four interviews for each type of respondent.

Information about choice of respondent is based on comparisons among the sample groups designated for each type of respondent. Enumerators were not always successful in interviewing the specified respondent. Their rate of success varied from about 75 percent for the husband-wife designates to 90 percent for the wife designates. It would have been possible to restrict the analysis to the cases in which the specified respondents were enumerated, but this would have led to uncertainty about the extent to which the groups were samples of the same universe.

An additional source of information on conditioning and effectiveness of different household members as respondents is data on nonresponse rates.

V. Preliminary Results:

Preliminary summaries of results are shown in the attached tables. More detailed data will be published at a later date. They show that major differences exist among procedures. In many cases the reasons for the differences are clear, but some differences can be explained by more than one hypothesis. These cases are being studied by work in progress. Until this work is completed, all conclusions are tentative.

Inference cannot be drawn on any one subject by examining a single table because there are relationships between length of reference period, time reference problems, recall failure and even, to some extent, conditioning. However, by combining the results of the various tables, the following conclusions seem to be indicated:

1) Reporting Expenditures in Wrong Month- There is strong evidence that it has a major impact on estimates. In general, it results in respondents tending to report expenditures in a more recent month than actually made. This phenomenon has been observed before, and is frequently referred to as "telescoping". It seems to have a greater effect on reporting of large jobs than small ones. For example, the allocation of expenditures by month within a three month bounded interview period indicates that the overstatement of the report for the most recent month is of the order of 40 percent for jobs of $100 or more and 5 percent for jobs under $20. (See Table 2). However, if some effect of conditioning, as described in paragraph 4 below, is assumed the estimated over-

statement for jobs of under $20 can be more like 20 percent than 5 percent.

2) Recall Failure - Large jobs are apparently remembered over a long period of time and reported well. Smaller jobs are susceptible to serious losses. The three month recall resulted in losses of nearly 40 percent of jobs under $20, but no measurable decreases in reporting of jobs of $100 or more. (See Table 2).

3) Effect of Length of Reference Period - Some of the small expenditures seem to be missed when there is a long reference period. This is distinct from the recall failure discussed in the previous paragraph which relates to the effect of differing lengths of recall periods. For the same recall period (one month), the six months interviews showed a smaller number of jobs than the one month interviews. (See Tables 1 and 2). However, the magnitude of this effect is not very clear since the data contain the combined effect of length of reference period, length of recall period and telescoping. In any case, no appreciable effect appears to exist for the larger and more important jobs.

4) Conditioning - There is a little evidence of a loss of small jobs after repeated interviewing, but the larger jobs seem to be as well reported. (See Table 3). This is further supported by evidence of an increase in households reporting no expenditures between the 1st and 2nd of the bounded 1 month interviews. Repeated interviewing, however, had no effect on nonresponse rates. (See Table 5).

5) Choice of Respondents - This seems to have no important effect on the estimates. (See Tables 4 and 5).

6) Use of Simpler Questionnaires - This resulted in an important downward bias. (See Table 6). However, the conditions under which the interviews were performed were different, and the training of the CPS enumerators on this survey was not as extensive as for SORAR, so that further work is necessary before final conclusions can be drawn.

VI. Implications for Survey Design:

If the conclusions described above are borne out by later studies, they imply that a survey to measure maintenance, repair, additions and alterations on residential properties should have the following properties:

1) Long-term (at least 3 month) recall period for large expenditures.
2) A shorter-term recall period for small expenditures.
3) Coincidence of recall and reporting periods to avoid respondents having to allocate timing of expenditures within the interview period.
4) Bounded interviews for data about larger expenditures.
5) Limited number of consecutive monthly contacts with the same respondents.
6) A detailed and probing questionnaire in order to stimulate maximum recall.

VII. Applicability to Other Expenditure Surveys:

Most expenditure surveys which are based on respondents' memories are subject to similar types of problems. The need for a probing type questionnaire has been reported in other studies. Telescoping and problems associated with the length of recall period have also been described in the past. In fact, these have been noted in quite dissimilar surveys, such as ones covering illness records.

However, it is possible that the magnitude of the effect of these items varies greatly from survey to survey, depending on conditions peculiar to a particular survey. For example, we suspect that one reason for the major impact of telescoping is the nature of the measurement process in this survey. An expenditure for a job is reported in the month the job ended; an expenditure for the purchase of materials is reported in the month the purchase was made. There is frequently some ambiguity as to the exact date on which a job is finished, and this could be responsible for much of the telescoping. This ambiguity does not extend to the date of purchase of materials. Consequently, additional analyses are being made of the data on materials purchases to see whether telescoping occurs to the same extent on these expenditures. This will shed some light on whether the telescoping occuring here is something inherent in a respondent's memory or whether it is a reflection of a survey which attempts to assign a time period to an action extending over a period of time.

There is another illustration of the way specific surveys may be affected differently. Woolsey, in a small-scale survey on illnesses, reported that the effect of telescoping was quite different, depending on whether the questionnaire was constructed to inquire about the most recent month first and then go backwards in time, or to inquire first about the earliest month and proceed in the reverse order.

It is probable that importance of the effects discussed in this paper will depend on the subjects covered by the survey, the kind of questionnaire used, the interviewers' training and experience, the respondents' motivation and other aspects of the survey.

Table 1. Analysis of One-Month Recall: Number of Jobs by Size of Job, for Estimates Based on One-Month Recall With Different Enumeration Methods, Feb. 1960-Mar. 1961

(in millions of jobs)

Size of Job	Enumeration Method			Percent of Bounded 1 Month	
	Unbounded Type 6 1/	Unbounded Type 1	Bounded 1 Month	Unbounded Type 6 1/	Unbounded Type 1
All jobs	213.9	300.6	215.1	99.4	139.7
$ 1-9	89.4	145.1	112.2	79.7	129.3
10-19	40.9	47.2	35.5	115.2	133.0
20-49	40.7	52.7	34.6	117.6	152.3
50-99	17.6	25.6	13.5	130.4	189.6
100-499	19.5	25.3	15.7	124.2	161.1
500+	5.8	4.6	3.5	165.7	131.4
		SUMMARY			
Under 20	130.4	192.3	147.8	88.2	130.1
20-99	58.3	78.3	48.2	121.0	162.4
100+	25.3	30.0	19.2	131.8	156.2

1/ Obtained by using only the most recent month of the 6 for which data were collected.

Note: Most of the differences are statistically significant. The c.v. of the difference in estimates of total jobs between either of the unbounded and the bounded 1 month is of the order of 3-4 percent. The c.v. on jobs over $100 is about 10 percent.

Table 2.- Analysis of Three Month Recall: Number of Jobs by Size of Job for Estimates Based on One, Two and Three Months Recall, June 196 -Jan. 1961

(in millions of jobs)

Enumeration Method and Recall Period	Size of Job			
	All Jobs	Jobs Under $20	Jobs of $20 to $99	Jobs $100 and over
Bounded 1 month	116.3	79.1	25.2	12.0
Bounded 3 months				
Average	84.1	49.4	22.5	12.1
Expenditures reported as having been made:				
1 month before interview	133.1	82.7	34.1	16.4
2 months " "	70.7	37.7	20.9	12.0
3 months " "	48.4	27.9	12.6	8.0
Unbounded type 6				
Average	87.8	45.8	26.7	15.3
Expenditures reported as having been made:				
1 month before interview	122.5	69.4	35.0	18.0
2 months " "	75.0	34.6	26.9	14.5
3 months " "	65.0	33.3	18.2	13.5
	Percent of bounded 1 month estimates			
Bounded 3 months				
Average	72%	63%	89%	101%
Expenditures reported as having been made:				
1 month before interview	114	105	135	137
2 months " "	61	48	83	100
3 months " "	42	35	50	68
Unbounded type 6				
Average	76	58	106	128
Expenditures reported as having been made:				
1 month before interview	105	88	139	150
2 months " "	65	44	107	121
3 months " "	56	42	72	113

Note: c.v. of difference between total jobs for bounded 1 month and average bounded 3 months or unbounded type 6, is about 3 percent. For jobs over $100, it is about 10 percent. Most of the percentages shown above are significantly different from 100 percent.

Table 3. Analysis of Conditioning: Comparison of Number and Sizes of Jobs Reported in 1st and 2nd Bounded One Month Interviews, March 1960–March 1961.

(in millions of jobs)

Size of Job	1st Bounded Interview	2nd Bounded Interview	2nd as % of 1st
All jobs	215	195	91
$1-9	114	105	92
10-19	36	32	89
20-49	34	30	89
50-99	13	12	90
100-499	15	14	93
500 and over	3	3	100

Note: c.v. of difference for total jobs is about 3 percent; c.v. for jobs $1-9 is 4-5percent. For other classes, the c.v. is higher and the individual differences are within 2 standard errors.

Table 4. Analysis of Choice of Respondent: Mean Number of Jobs per Household Reported by Each Type of Respondent, by Size of Job, Feb. 1960-Mar. 1961 (Estimates Shown for Three Different Enumeration Methods)

Size of Job and Respondent Designated	Bounded 1 Month	Unbounded Type 6 1/	Unbounded Type 1
Total Jobs			
Head	7.3	6.0	8.9
Wife	6.0	6.9	9.7
Head & Wife	7.0	7.1	9.7
Unspecified	6.6	6.7	9.2
Jobs under $20			
Head	5.0	3.7	5.6
Wife	4.1	4.6	6.6
Head & Wife	4.6	4.3	6.2
Unspecified	4.6	3.8	5.8
Jobs $20-99			
Head	1.6	1.7	2.6
Wife	1.5	1.8	2.1
Head & Wife	1.6	2.0	2.3
Unspecified	1.4	1.8	2.5
Jobs $100 and over			
Head	.7	.6	.7
Wife	.4	.5	1.0
Head & Wife	.8	.8	1.2
Unspecified	.6	1.1	.9

Note: For total jobs, c.v. of difference between head, wife, and head and wife is about 5 percent for bounded 1 month and 10 percent for the other two groups. It is somewhat less for comparisons with unspecified respondents. Most differences shown, therefore, are not statistically significant.

1/ Only the most recent month of the 6 for which data were collected has been used in this table, to provide a uniform recall period.

Table 5. Analysis of Conditioning: Noninterview Rates by Designated Respondent and Order of Interview, July - Dec. 1960

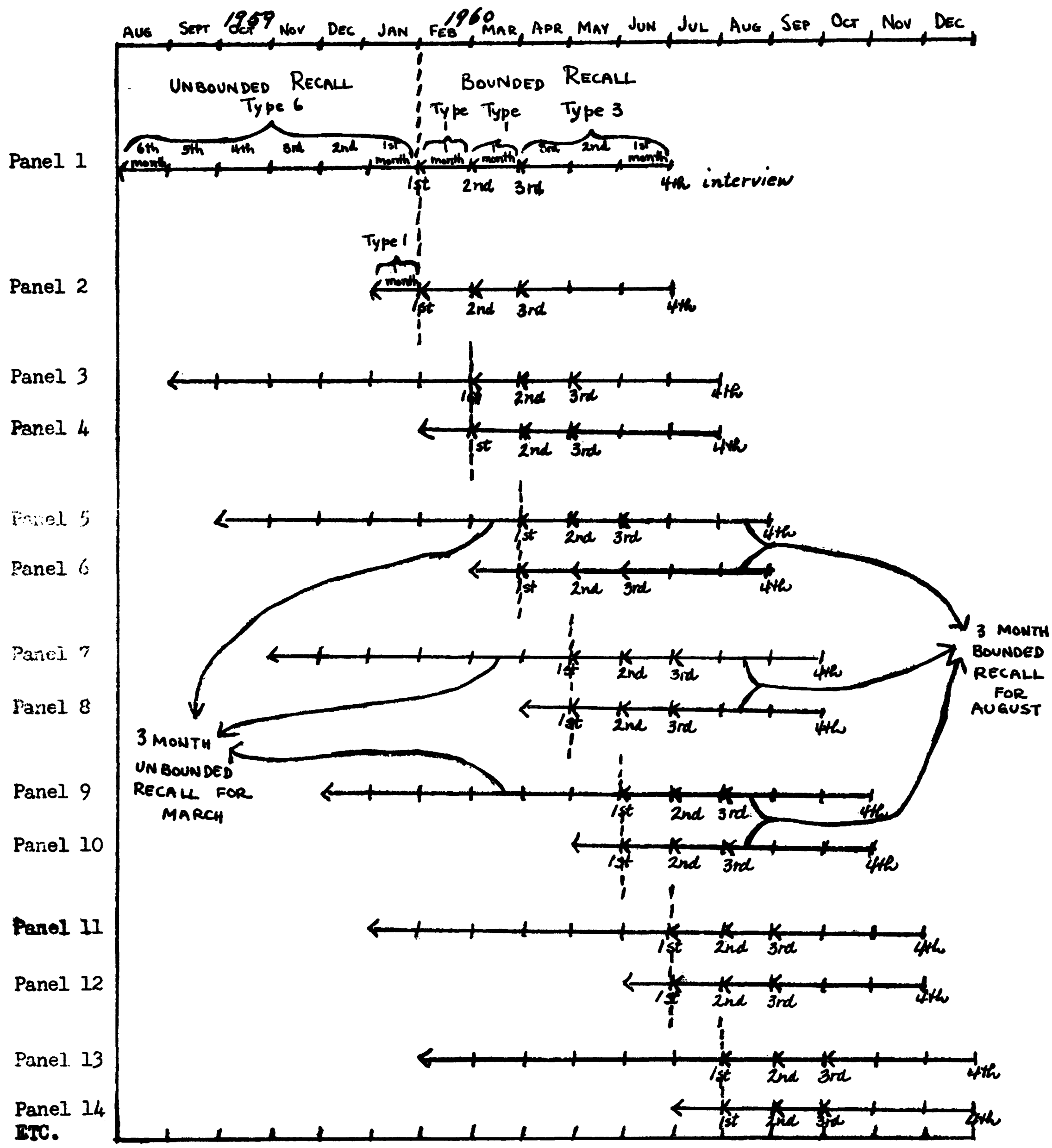
Respondent Designated	Percent Noninterview			
	1st Interview	2nd Interview	3rd Interview	4th Interview
Head	8.9	8.6	7.9	6.7
Wife	7.3	8.6	7.9	6.3
Head & Wife	7.9	7.8	7.1	8.1
Unspecified	6.5	6.4	6.3	5.9

Table 6. Analysis of Simple Questionnaire(as CPS Supplement 1/)for Collecting Expenditure Data: Number of Jobs Estimated in CPS Supplement as Percent Reported in the Regular Survey Oct-Dec. 1960

Size of Job	% of Bounded 1 Month Estimate	% of Bounded 3 Month Estimate	% of Unbounded Type 6 Estimate
Total	43	63	66
Under $20	29	52	61
20 - 99	58	71	70
100 and over	100	85	74

1/ CPS supplement included 2 random subgroups - telephone enumeration was permitted in one group and not in the other. The analysis here is restricted to the group not permitting telephoning, to eliminate this as a source of difference. The number of jobs reported by the group permitting telephoning was 86 percent of those reported by the group for which telephoning was not permitted.

Chart Showing Rotation Scheme Employed for
Survey of Residential Alterations and Repairs



SOME EFFECTS OF INTERVIEWER-RESPONDENT INTERACTION ON RESPONSES IN A SURVEY SITUATION

Daniel O. Price and Ruth Searles, University of North Carolina

Two years ago the Institute for Research in Social Science at the University of North Carolina received a grant from the Rockefeller Foundation to study "the changing position of the Negro" in the U. S. The major portion of the grant was to be invested in a study of Negro political participation in the South. Before beginning field work on the political participation study, we were faced with the question, "How much difference would it make in the validity of our data if members of the field staff who interview Negroes were themselves Negroes or whites?" Behind this was the larger question, "To what extent do the social and personal characteristics of interviewers affect the information they elicit from respondents in a survey situation?"

A number of studies[1] have demonstrated that interviewees tend to respond to questions with answers which the interviewer might be expected to desire. Reporting on an NORC study in 1942 in which equivalent samples of Negro respondents were interviewed by whites and by Negroes, Hyman observed, "on almost all the opinion and attitude questions, the white interviewers obtained significantly higher proportions of what might be called by some 'proper' or 'acceptable' answers."[2]

There are different ways to conceptualize the effects of the interviewer upon the information he elicits in response to survey questions. One way is to assume that there is a single complete and accurate answer to each question for each respondent and that the interviewer, by his attitude, demeanor, or apparent social background, may "bias" the respondent's report of that answer. Another is to view the interaction between interviewer and respondent as the result of psychological forces playing upon the perceptual field of each. A third is to view the interview situation less as a person-to-person event than as a sample of more general role phenomena. Using this last approach, we assume that the respondent's report reflects what he considers appropriate to his role vis-a-vis the interviewer and that there may be a variety of "socially true" answers, depending on how the respondent views his role-partner.

Because of our interest in the changing position of Negroes, we hoped to use the "social role" aspect of the interview to draw inferences about changing norms governing role behavior. In his report of the NORC study, Hyman noted that Negroes interviewed by whites gave "proper" or "acceptable" answers. As we thought about the "proper" or "acceptable" answers to the questions we wanted to ask, we recognized that there were at least three different versions of proper role behavior for Negroes in the South, each associated with a mode of social organization and the time period in which that mode has predominated.

There is the propriety associated with an agricultural economy and a relatively simple pattern of social organization, i.e., with concentration of employment in the primary industries, with a simple division of labor and minimum development of occupational specialties . In this setting, it is "proper" for a Negro to be relatively unskilled, uneducated, and unambitious; demonstrating acceptance of his subordinate position as his "place;" showing that his concern for status is limited to that future beyond his worldly life. Status in this social setting may be advanced by dependence on a benevolent white patron and by simple piety.

There is the proper role behavior associated with a more industrialized economy and an urban pattern of organization, i.e., with a shift of the labor force into secondary and tertiary industry, greater development of occupational specialties, and a concomitant increase in the importance of training and skill. In this setting, it is proper for Negroes to be educated, to be ambitious for security, to be provident in saving time and money for investment in the future. Status in this setting is advanced by education and respectability.

Finally, there is a proper pattern of behavior associated with a more advanced urbanism with its still greater emphasis on and opportunity to acquire occupational skills and formal education, its higher standard of living and its reduced power differentials between social classes. In this setting it is acceptable for Negroes to be concerned not only with security but with civil rights, and to advance their status not only by education but also by political, legal, and organizational means.

The three modes of social organization we have described exist simultaneously in the South today. Each has its version of proper behavior for Negroes. We hoped to find these versions reflected in the answers given by Negroes in different social settings. Changes in the South, as elsewhere, have tended to give Negroes increasing power vis-a-vis whites. We hoped to find these changes reflected in whatever differences might appear between the answers given by Negroes to white interviewers and to Negro interviewers in different settings.

Four sites were selected for field work: two rural counties and two cities in North Carolina, each of which had a high proportion of Negroes in the population.[3] Negro respondents in rural and urban settings were interviewed by Negro and by white interviewers. Each interviewer conducted interviews in both rural and urban sites.

Table 1 shows the number of interviews conducted according to the respondent's residence and interviewer's race.

Our principal hypotheses were that:

1. Differences in social organization will be reflected in differences in response as between rural and urban settings.

2. Differences in the social position of Negroes vis-a-vis whites will be reflected in the magnitude and direction of difference between answers given to Negro interviewers and those given to white interviewers in rural and urban settings.

3. Differences between answers given to Negro interviewers and answers given to white interviewers will reflect the accepted means by which Negroes can gain status in each social setting. Specifically,

   a. With regard to church membership and participation, there will be no difference in response attributable to the interviewer's race in either rural or urban settings.

   b. With regard to educational aspirations, there will be a difference attributable to race in the rural setting but not in the urban setting.

   c. With regard to civil protest and political participation, there will be a marked difference in response attributable to race in the rural setting and a real but less marked difference in the urban setting.

Data relevant to these hypotheses are presented in tables 2 through 8.

Tables 2 and 3 which present data about church attendance and participation show a difference between rural and urban settings--though not in the direction we had anticipated. Thinking of the city as a more secularized environment, we had expected less church participation there than in the rural setting, forgetting the importance of the church as a voluntary association in the city and the barriers to church attendance in the country which might be presented by the counties' poor roads and the relative isolation of many of our respondents. The prediction concerning difference in response according to interviewer's race holds--there is no significant difference between the answers given to white interviewers and those given to Negro interviewers in each setting.

Responses to questions about educational aspirations are indicated in tables 4, 5, and 6. Response to questions about the amount of education the respondent would want his son or daughter to have reveal important differences between rural and urban environments, with urban interviewees more likely than rural ones to want their children to complete college. Comparing responses obtained by white and by Negro interviewers, we find that the over-all differences in urban settings are not great nor do they show a consistent pattern. In the rural setting, on the other hand, Negro respondents are consistently more likely to advocate a college degree for their children when interviewed by Negroes than when interviewed by whites.

Responses to questions concerning varying aspects of social change, civil protest, and political information are given in tables 7 through 11. We were interested to see the strength of endorsement for social change, for the sit-ins, and school integration in both rural and urban settings, whether the interviewer was Negro or white. On top of this over-all endorsement of change, there was the anticipated difference between rural and urban environments, city respondents even more eager for it than country ones.

Differences according to the interviewer's race were found, as predicted. Rural respondents, especially, were more reluctant in reporting to whites than to Negroes that they advocated frequent changes in the way our country is run, that they approved of the sit-ins, that they thought a Negro mother should send her first-grade daughter to a previously white school, and that they correctly recalled the names of the candidates in the Democratic gubernatorial primaries.

Differences in response attributable to the interviewer's race were significant at the .01 level when a question about sit-ins was asked directly: "How do you feel about these Negro college students sitting at lunch counters where only whites were served before and waiting to be served?" Interestingly, however, when a question about sit-ins was put more impersonally--"About how many Negroes in North Carolina would you say approve of the sit-ins?"--the differences attributable to interviewer's race were not statistically significant.

Table 11 suggests the sensitivity of southerners to political participation by Negroes. Our field work was done in the summer of 1960, about 6-8 weeks after the hotly contested Democratic gubernatorial primaries in which a segregationist vs. "Southern moderate" contest was paramount. Differences of 20 percentage points or more between the answers given to white and to Negro interviewers indicate that our respondents certainly did not assume consensus between whites and Negroes on the propriety of Negroes being informed about political affairs.

Up to this point, we have interpreted our data in light of known differences in rural and urban social organization and hypothesized relations between these organizational differences and the norms governing role behavior for Negroes.

While this view is appropriate and secures considerable support from our data, we would be foolish to ignore other or supplementary interpretations of the effects of interviewer-respondent interaction on the data obtained in survey interviews. For example, what about the bias introduced by the interviewer's attitudes or social background? We have noted a reluctance by Negroes to admit as much support of integration to white interviewers as to Negro interviewers. Does that mean that Negro interviewers secured more valid data? Probably it does, yet we should not ignore the possibility that Negro respondents were reluctant to admit any hesitation about integration to our enthusiastic, well-educated, middle class Negro interviewers; nor the possibility that our interviewers gave cues as to their feelings or failed to record accurately.

In an attempt to isolate the manifold effects on interview data made by characteristics of the respondent and the interviewer, a second approach was taken in analysis of data. Less formal and more empirical than the analyses discussed thus far, this approach utilized an extensive analysis of variance. Five variables most nearly meeting the requirements of quantitative variables were chosen.

As shown in table 12, these dependent variables were (1) educational aspiration for son, (2) educational aspiration for daughter, (3) the amount a job should pay in order for a high school senior to quit school and go to work, (4) attitude toward sit-ins, and (5) attitude toward school integration. We recognize that none of these variables meets the criterion of equal intervals. It is unlikely that any quantitative variable in social science research really meets this criterion, for even with a variable as "quantitative" as income, a difference of $1,000 does not have the same social significance at the $3,000 level that it does at the $25,000 level. However, exploratory studies can aid the development of theory by generating new questions, hunches, and hypotheses even if all the necessary assumptions for statistical accuracy are not met. This is not to argue for a flippant attitude toward assumptions but simply to note that exploratory studies can be fruitful even though the results may be inaccurate statistically. The final standard for judging results must be meaningfulness and relatedness to a broader context.

Educational aspirations for sons and daughters were measured in terms of number of years of education reported as desirable. These two variables behave very similarly, though not identically, and are clearly not independent of each other. In trying to measure the amount of earnings that would justify a high school senior quitting school, the top interval is, of necessity, open-ended since some respondents said he should not leave school under any conditions.

Attitude toward the sit-ins was indicated by responses to the question, "How do you feel about these Negro college students sitting down at lunch counters where only whites were served before and waiting to be served?" The scale ranges along 6 points from "strongly approve" to "strongly disapprove" with 3 being "neutral." The variable which we have called "attitude toward school integration" developed from answers to the following question: "A Negro mother has a daughter entering the first grade next year. Should she apply for her daughter to go to a nearby school which has always had white students or should she send her to a Negro school where her other children went?" Responses were classified on a 5-point scale from "strongly favoring the Negro school" to "strongly favoring the white school."

An analysis of variance program on the UNIVAC 1105 was used to estimate the nature and statistical significance of the relationship between each of the dependent variables and the independent variables. The 5 columns in table 12 stand for the dependent variables as identified at the top of the table, and "x" indicates a relationship that is statistically significant at the .05 level.

Looking at table 12, we see that educational aspiration for a son is related to 5 independent variables in a statistically significant way. All but one of these, "education of interviewer," are characteristics of the respondent or of his household. Apparently, education of interviewer had an influence here with the surprising effect that respondents reported higher aspirations to interviewers with less education. The only independent variable related to educational aspirations for son which is not also related to educational aspirations for daughter is the number of veterans in the household. The latter variable was included to give some indication of contact with other styles of life. Apparently it only affects the educational aspirations of sons, and then in a negative fashion. Further analysis of the data is being undertaken in an effort to explain the direction of association of these variables.

The report of educational aspiration for a daughter seems to be more sensitive to the interviewer's characteristics than the report of aspirations for a son. It is significantly related to the race of the interviewer, education of the interviewer, and the interviewer's educational aspirations for a young woman. How much of this association is due to the respondent reaction to the interviewer's characteristics and how much of it is due to the interviewer's bias in reporting is unknown. The fact that a significant association exists between the interviewer's educational aspirations for a young woman and the respondent's educational aspirations for a daughter suggests that it might be due more to the interviewer's than to the respondent's reaction.

Dependent variable number 3, "the amount of pay which would justify quitting high school" is very interesting in that 8 out of the 10 variables to which it is related are characteristics of the interviewer. The fact that this variable is associated with the supervisor's rating of the interviewer on two characteristics would indicate that the interviewer might have been providing obvious cues to the respondent in this question. Further analysis of this variable and its relationships to the interviewer's characteristics within various categories of respondent's characteristics are being undertaken in an effort to explain some of the findings. Data on additional characteristics of the interviewer are available, and their relations to the variables discussed here are being examined.

Dependent variable 4, "attitude toward sit-ins," is surprising in that it shows an association with only three independent variables. Since this attitude question was more controversial than those previously discussed, one might have expected greater variations in relation to interviewer characteristics here than in the others. It is possible, of course, that the poor quality of the attitude measurement may be obscuring other associations.

On attitudes toward sending a daughter to a previously white school, 7 of the 8 variables with which this variable shows a significant association are characteristics of the respondent. This is the only one of these 5 dependent variables that gave any evidence of statistical interaction between interviewer and respondent characteristics. It is significantly related to the rural-urban residence of the respondent by race of interviewer. Several other variables examined by other methods have shown significant association with race of interviewer in rural areas but not in urban areas.

The last page of the distributed material shows the directions of the significant associations presented in table 12. The direction of several of these associations is unexpected, and analysis of these and other available data is continuing in an effort to provide additional information on the effects of interviewer characteristics on data gathered in differing situations.

FOOTNOTES

1. Cf., e.g., Herbert H. Hyman, et al, Interviewing in Social Research (Chicago: University of Chicago Press, 1954), 138-170; Robert L. Kahn and Charles F. Cannell, The Dynamics of Interviewing (New York: John Wiley & Sons, Inc., 1958), 58-64, 143-148, 179-202; Gerhard E. Lenski and John C. Leggett, "Caste, Class, and Deference in the Research Interview," American Journal of Sociology LXV (1960), 463-467.

2. Hyman, op. cit., 159.

3. One purpose of this study was to estimate the amount of variability in response which might be attributed to certain gross social and attitudinal characteristics of the interviewer, but it also served as a pre-test of interview items for a number of subsequent studies. In the course of interviewer training, we discussed various sources of bias in reporting and interview techniques to minimize them. We talked about the possible influence on response of the interviewer's race. The pre-test purpose of the study was emphasized.

Since we did not intend to generalize our findings to any particular population, ours was a purposive sample. Residential areas were selected in rural and urban sites, and an interview was conducted with one adult member of the household in all dwelling units in the selected areas. The urban sample was taken in Raleigh and Greensboro, N. C., and the rural sample in Caswell and Granville counties. White respondents were interviewed by white interviewers in each of the four sites, but these data are not reported here. In the cities, the upper-class Negro residential areas were over-sampled in order to obtain sufficient representation of upper-class Negroes. Residential segregation was less marked in the counties. Areas of concentration of Negro population were selected as beginning points for interviewing in rural areas, with the interviewers working out from these points to obtain additional Negro and white respondents. Interviewers were assigned to areas by supervisors, the only restriction being that Negro interviewers did not interview white respondents.

There were 14 Negro and 12 white interviewers. During the first 2 weeks of the field work, half of the Negro and half of the white interviewers were assigned to an urban area and the other half to a rural area. During the second 2 weeks, those who had been working in urban areas were switched to rural areas, and vice versa.

That the rural-urban difference was associated with a difference in social organization is indicated by 1950 Census figures on the industries in which adult males were employed. The two rural counties had an average of 65 percent in primary industry (mostly farming) and 18 percent in tertiary industry. This is in contrast to the cities, in which only one percent were employed in primary industry and 68 percent in tertiary industry.

TABLE 1. Distribution of interviews with Negroes by interviewee's residence and age and by interviewer's race.

Interviewer's Race and Respondent's Residence	Under 30	Age of Respondent 30-39	40-49	50 and Over
Negro Interviewer				
Rural	53	57	44	85
Urban	79	77	81	120
White Interviewer				
Rural	22	23	27	42
Urban	38	41	23	64

TABLE 2. Responses of Negro interviewees to question, "Do you attend church?" by age, residence, and interviewer's race.

	Percent Indicating Attendance Once a Week or Oftener, or Every time Church Meets			
Interviewer's Race and Respondent's Residence	Under 30	Respondent's Age 30-39	40-49	50 and Over
Negro Interviewer				
Rural	23	25	34	35
Urban	42	39	54	59
White Interviewer				
Rural	14	35	30	38
Urban	51	58	65	56

Age differences significant at the .05 level
Residence differences significant at the .001 level
Interviewer's race not significant.

TABLE 3. Responses of Negro interviewees to question, "Do you belong to any committees or organizations in the church, besides just being a member of the church?" by age, residence, and interviewer's race.

	Percent Indicating Committee or Organization Membership in the Church			
Interviewer's Race and Respondent's Residence	Under 30	Respondent's Age 30-39	40-49	50 and Over
Negro Interviewer				
Rural	32	38	52	43
Urban	33	47	41	56
White Interviewer				
Rural	14	23	30	43
Urban	36	46	48	59

Age differences significant at the .05 level.
Residence differences significant at the .05 level.
Interviewer's race not significant.

TABLE 4. Responses of Negro interviewees to question, "About how much education would you want your son to have today in order to get a good start in life?" by age, residence, and interviewer's race.

Interviewer's Race and Respondent's Residence	Percent Indicating Son Should Finish College or More			
	Respondent's Age			
	Under 30	30-39	40-49	50 and Over
Negro Interviewer				
Rural	68	67	65	62
Urban	82	86	76	79
White Interviewer				
Rural	45	50	48	48
Urban	76	95	79	71

Age differences not significant.
Residence differences significant at .001 level.
Interviewer's race significant at .05 level.

TABLE 5. Responses of Negro interviewees to question, "About how much education would you want your daughter to have today in order to get a good start in life?" by age, residence, and interviewer's race.

Interviewer's Race and Respondent's Residence	Percent Indicating Daughter Should Finish College or More			
	Respondent's Age			
	Under 30	30-39	40-47	50 and Over
Negro Interviewer				
Rural	68	70	68	69
Urban	84	88	79	81
White Interviewer				
Rural	41	54	48	48
Urban	76	88	83	66

Age differences not significant.
Residence differences significant at .001 level.
Interviewer's race significant at .05 level.

TABLE 6. Responses of Negro interviewees to question, "If you had a son in his last year of high school, how good a job should he be offered for you to think he should quit school and go to work?" by age, residence, and interviewer's race.

Interviewer's Race and Respondent's Residence	Percent Indicating Son Should Not Quit School Under Any Circumstances			
	Respondent's Age			
	Under 30	30-39	40-49	50 and Over
Negro Interviewer				
Rural	66	70	80	68
Urban	60	81	71	72
White Interviewer				
Rural	55	62	57	73
Urban	44	77	65	77

Age differences significant at .05 level.
Residence and interviewer's race not significant.

TABLE 7. Responses of Negro interviewees to question, "Which of these statements do you think is better: 'It is not a good idea to make changes in the way our country is run,' or 'We must frequently make changes in the way our country is run,'" by age, residence, and interviewer's race

Interviewer's Race and Respondent's Residence	Percent Endorsing "We must frequently make changes in the way our country is run"			
	Respondent's Age			
	Under 30	30-39	40-49	50 and Over
Negro Interviewer				
Rural	89	86	91	81
Urban	90	99	92	90
White Interviewer				
Rural	81	64	69	81
Urban	79	93	87	66

Age differences not significant.
Residence differences significant at .05 level.
Interviewer's race significant at .05 level.

TABLE 8. Responses of Negro interviewees to question, "How do you fell about these Negro college students sitting down at lunch counters where only whites were served before and waiting to be served?" by age, residence, and interviewer's race.

	Percent Approving of Sit-ins			
	Respondent's Age			
Interviewer's Race and Respondent's Residence	Under 30	30-39	40-49	50 and Over
Negro Interviewer				
Rural	67	67	67	69
Urban	92	95	87	84
White Interviewer				
Rural	40	52	56	43
Urban	86	90	77	58

Age difference not significant.
Residence differences significant at .001 level.
Interviewer's race significant at .01 level.

TABLE 9. Percentage distribution of responses of Negro interviewees to question, "About how many Negroes in North Carolina would you say approve of the sit-ins?" by age, residence, and interviewer's race.

	Respondent's Age and Residence							
Interviewer's Race and Respondent's Answer	Under 30 Rural	Under 30 Urban	30-39 Rural	30-39 Urban	40-49 Rural	40-49 Urban	50 and Over Rural	50 and Over Urban
Negro Interviewer	100	100	100	100	100	100	100	100
Most	40	47	29	58	41	47	30	46
A great many	26	35	38	34	30	37	38	42
Some, few, or none	34	18	34	8	30	16	32	12
White Interviewer	100	100	100	100	100	100	100	100
Most	40	61	28	53	40	44	26	36
A great many	20	24	17	33	40	30	19	40
Some, few, or none	40	15	56	13	20	26	55	25

Age differences not significant.
Residence differences significant at .01 level
Interviewer's race not significant.

TABLE 10. Responses of Negro interviewees to question, "A Negro mother has a daughter entering the first grade next year. Should she apply to have her daughter go to a nearby school which has always had white students, or should she send her to a Negro school where her other children went?" by age, residence, and interviewer's race.

	Percent Indicating Mother Should Send Daughter to Previously White School			
	Respondent's Age			
Interviewer's Race and Respondent's Residence	Under 30	30-39	40-49	50 and Over
Negro Interviewer				
Rural	29	39	29	41
Urban	61	64	70	59
White Interviewer				
Rural	19	18	22	23
Urban	56	63	35	53

Age differences not significant.
Residence differences significant at .001 level.
Interviewer's race significant at .01 level.

TABLE 11. Responses of Negro interviewees to question, "Do you remember the names of the men who ran for Governor of North Carolina in the recent primary election?" by age, residence, and interviewer's race.

	Percent Giving Two or More Correct Names and No Incorrect Names			
	Respondent's Age			
Interviewer's Race and Respondent's Residence	Under 30	30-39	40-49	50 and Over
Negro Interviewer				
Rural	45	45	48	43
Urban	79	75	68	59
White Interviewer				
Rural	18	17	35	30
Urban	53	72	52	34

Age differences not significant.
Residence differences significant at the .001 level.
Interviewer's race significant at the .01 level.

TABLE 12. Relationships between independent and dependent variables in the study of interviewer-respondent interaction. (An x indicates statistical significance at the .05 level.)

Dependent variables:
1. Educational aspirations for son
2. Educational aspirations for daughter
3. Wages sufficient to justify quitting
4. Attitude toward sit-ins (high school)
5. Attitude toward school integration

Independent variables	1	2	3	4	5
Respondent Characteristics					
1. Urban-rural residence				x	x
2. Education	x	x			
3. Education of head of household					x
4. Registered voter or not	x	x	x	x	x
5. Home ownership					
6. Sex					x
7. Age				x	x
8. Number of children in household					
9. Time in South for head of household					
10. Status of highest status job in household			x		
11. Sex of person with highest status job					
12. Number of veterans in household	x				
13. Skin color of respondent					x
14. Sources of income other than employment (welfare, pensions, etc.)	x	x			x
15. Number of children and highest occupation					
Interviewer Characteristics					
16. Urban-rural background			x		
17. Race		x	x		
18. Sex					
19. Education	x	x			
20. Family income			x		
21. Interviewer's educational aspirations for a young man		x	x		
22. Interviewer's educational aspirations for a young woman			x		
23. Interviewer's attitude toward sit-ins			x		
24. Supervisor's evaluation of interviewer's interviewing technique			x		
25. Supervisor's evaluation of interviewer's use of status symbols			x		
Interviewer-Respondent Combinations					
26. Rural-urban background of interviewer by rural-urban residence of respondent					
27. Education of interviewer by education of respondent					
28. Sex of interviewer by sex of respondent					
29. Rural-urban residence of respondent by race of interviewer					x

Direction of Significant Associations Shown in Table 12

1. High educational aspirations for son are associated with
   - High educational level of respondent
   - Being a registered voter
   - Few or no veterans in the household
   - All income in household being from wages (no welfare payments, pensions, etc.)
   - Low educational level of interviewer

2. High educational aspirations for daughter are associated with
   - High educational level of respondent
   - Being a registered voter
   - All income in household being from wages (no welfare payments, pensions, etc.)
   - Interviewer being a Negro
   - Low educational level of interviewer
   - Interviewer having high educational aspirations for a young woman

3. A statement that a job should pay very high wages before a person should quit as a high school senior and go to work is associated with
   - Being a registered voter
   - Someone in the household having a high status job
   - The interviewer having a rural background
   - The interviewer being a Negro
   - The interviewer having a low family income
   - The interviewer having high educational aspirations for a young man
   - The interviewer having low educational aspirations for a young woman
   - The interviewer being in favor of sit-ins
   - The interviewer's supervisor evaluating his interviewing technique as poor
   - The interviewer's supervisor saying that interviewer minimized status symbols.

4. Being in favor of sit-ins is associated with
   - Urban residence
   - Being a registered voter
   - Being young

5. Being in favor of sending daughter to a previously white school is associated with
   - Urban residence
   - High education of head of household
   - Being a registered voter
   - The respondent being male
   - Young respondents
   - Light skin color of respondent
   - Receiving other income than wages (welfare payments, pensions, etc.)
   - White interviewers in rural areas and Negro interviewers in urban areas

RECONCILIATION PROBLEMS AMONG THE SOCIAL ACCOUNTS[1]

Richard Kosobud, Survey Research Center, The University of Michigan

Social accounting data are prepared by a number of different producers or sources. Seemingly like categories of data have different transactions or transactor content. Consumers of such data are confronted with reconciliation problems. For example, the sum of all the items of government expenditure presently in the Bureau of the Census type accounts does not equal total government expenditure for a like period as reported in the financial records of governments. Nor would the sum of all government expenditures equal the total as estimated by the National Income Division of the U.S. Department of Commerce. As transactors and transactions contained in one social account system are treated differently in another, we may anticipate different cyclical or trend patterns in the series due solely to the preparation of the facts. If the intention is to prepare analyses of the data, the results are partly dependent upon data preparation practices. This is an uncomfortable fact.

This study presents a method of distinguishing classes of reconciliation problems, and a description of the data preparation practices which lead to such problems.

A narrow definition of a reconciliation problem is adopted. Reconciliation is required whenever two or more producers of social accounting data follow different rules for the treatment of single transactions between parties or actors. Such rules cover the inclusion or non-exclusion of certain types of transactions in the accounts. They also include the assignment of individual actors to larger groups, and the assignment of a transaction to a time period. Apparently comparable categories have revealed differences of five to ten per cent or more.[2] Also, cyclical variation in this percentage appears discernible.

It is tempting to resolve this pervasive, and frequently irritating problem by a decision to utilize only one source. But the situation with respect to the different social accounts is that each major source of social accounting data yields unique detail, or covers a unique period of time, or yields unique groupings of actors. Each provides information not currently available in the others. It could be argued that reconciliation problems are a consequence of the varying goals of the producers of social accounting data. The main point of the discussion is that difficult and sometimes intractable reconciliation problems could be eased if social accounting data were presented so that consumers of data could make the adjustments.

Common Features of Social Account Systems

It will be desirable to marshall some of the distinguishing features of social accounts which are common to all the systems to be examined before we set out the data preparation practices which separate them. These features explain in good part the rapid growth of social accounting information, and the increasing use of such data to describe the roles of actors in an interdependent economy. It will be recalled that if each party to a transaction reckoned in money units maintains a double-entry system, the transaction may be made to yield six entries--two for the social accounts. Also useful information about the circumstances of the exchange may be obtained. The minimum two entries on the books of each party yield information about the functional area in which the transaction took place (e.g., in the case of local government units, transactions in the areas of education, recreation, etc. could be distinguished); the type of account (operating, income, or wealth); and the object of the transaction (consumption goods and services, labor service, raw material, equipment, etc.). For the economy as a whole, we obtain an entry classifiable as an "incoming" (receipt, revenue) for one transactor, and an "outgoing" entry (expenditure) for the other. These transactions may be prepared so as

a) to reveal the identity of parties to transactions,
b) to reveal the circumstances of the transaction such as the function, account, and object,
c) to yield a balanced system; i.e., a complete presentation of flow items and consequent changes in stock items,
d) to yield consistent sub-totals and totals (as happens in the case of data prepared by one source, but not among the various sources of social accounting data as we shall see).

Although individual unit accounts, budgets, financial records, profit and loss statements and balance sheets remain the conceptual primary source of data (even if actors such as households do not keep them), it may be more efficient to obtain data on transactions from tax returns, social security account transactions, industry sales records, etc. For our purposes, it is easier and sufficiently accurate to think of the individual account and transaction undergoing processing.

Data Preparation Decisions

Producers of social accounting data must make a number of important decisions about the treatment of transactors and transactions. Various kinds of agents or units in the economy can be distinguished such as consumers, businesses, government units, and others. A decision must be made about grouping individual transactors in one category or another, or to exclude them from the accounts. We single out one of these decisions to illustrate the consequences of two major sources deciding differently in one instance. The Bureau of the Census statisticians sharply distinguish government units from other units while the income and product account statisticians group government enterprise transactions with private enterprise.

Another group of decisions concern the type of transactions. Should all transactions be included, making the system a record of actual transactions, or should some transactions be discarded? For example, in the preparation of input-output tables, a decision must be made as to the inclusion of transactions in existing assets and financial claims. Input-output statisticians do not include them and concentrate on newly produced commodities. U.S. Bureau of the Census statisticians do include them in their reports on government revenues and expenditures. The Federal cash budget as well as the accounts of businesses and households include such items. The treatment of capital account transactions is one of the most delicate decisions, and problems confronting the social accounting statistician. Newly produced capital commodities enter into input-output tables and income and product data. Other capital account transactions are presented in flow of funds accounts. Reconciliation is possible if such distinctions are maintained.

Input-output statisticians make efforts to retain information on parties to transactions by industry, whereas in the other systems such direct information is not maintained. Social accounting statisticians must decide whether or not to group transactions by calendar year or by fiscal year. A similar problem arises when decisions are made to group transactions on a cash or accrual basis over time. In the case of the national income and product accounts, the record of actual transactions is altered by the preparation of imputations of the value of commodities not purchased. Examples are the estimates of food grown by farm households for their own use and the value of goods and services provided by the military establishment for personnel. These imputations create a discrepancy between the income and product accounts and other systems. It will be useful to present these differences in a more explicit manner.

The Transactions Matrix Approach to the Problem

We propose a sharper attempt to cut through the thicket of problems of reconciling social accounting data. The tool for this effort is a formal scheme, a transactions matrix with associated grouping matrices. We seek gains in treating systematically the issues described previously. The cost lies in the fact that it is a more roundabout method.

A transactions matrix, (t_{ij}), is introduced which depicts individual actor or unit exchanges reckoned in money. By convention, columns of entries assigned to an actor will be understood to reveal outgoing or expenditure transactions while rows of entries reveal incomings or revenue (or receipts) of actors. Sometimes this arrangement is referred to in terms of "purchases from" (columns) and "sales to" (rows). However, transfer transactions, subsidies and intergovernmental transactions are of interest. We shall adopt the view, therefore, that places a value upon completeness of description of transactions above all else. We want (t_{ij}) to contain what actually happens, not what we think ought to happen.

Each exchange produces an outgoing entry in a column which when located identifies the party for whom it is an incoming. There is no need to be inhibited at this stage by limitation of data processing equipment. The dimensions of the underlying transactions matrix can be thought of as being equal to the number of households, business, non-profit units, federal, state, and local government units of all kinds, public authorities, corporations, trust funds, agencies, commissions, etc. Transactors may be further identified by size, regional location, or other characteristics. Such identification would permit consolidation at a later stage into various groups which are thought to be homogeneous. The dimensions of the matrix are now further expanded so as to permit transactions to be classified as to function, account, and object. Revenue rows may be expanded so as to distinguish type of revenue or income (tax source, labor or property income, etc.). These define some of the possibilities.

Clearly, present data processing equipment does not permit us to deal with a transactions matrix of dimension even close to the one we have sketched. Simplification is required.

Processing operations in which transactors and transactions are grouped or eliminated find their analogue in grouping matrices A for rows and B for columns with which we can operate upon matrix T. These operations may be represented succinctly:

1) $\sum_j \sum_k a_{ij} \; t_{jk} \; b_{kl} = ({}^{r}il),$

where a, t, and b indicate elements in matrices. If we think of the indices as running $j = 1, \ldots, J$ and $k = 1, \ldots, K$, then $J > I$ and $K > L$. Grouping matrices are defined to contain as elements either 0 or 1. They may facilitate, conceptually, our systematic treatment of the processes of preparation of varied social accounting data. The grouping matrix with which we pre-multiply T will either consolidate row entries or send them to zero. The post-multiplication matrix will operate upon columns of T in like manner. The fact that we reduce T to more manageable size is a striking advantage but not without its costs. Not only will consistency problems arise among sources, but also aggregation problems may be revealed by this approach.

Operations with grouping matrices may be classified logically and associated with current practices. If A and B are identity matrices, we reproduce the entries of T in R. If A and B are row and column vectors, respectively, each element of which is 1, we reduce all transactions to a scalar. These are the extremes.

It is apparent that we may group actors with respect to expenditures by strings of 1's in rows of A, and we may group actors with respect to receipts by strings of 1's in columns of B. Rows and columns may be rearranged in the transactions matrix by rearranging the rows or columns of the grouping matrices in their identity form. Transaction entries will be washed out wherever zeros are found in the grouping matrices--an important operation. In instances in which the grouping matrices are vectors, the existence of zeros annihilates information. This important example is worth a simple illustration:

$$(1\ 1\ 1) \begin{pmatrix} 0 & 4 & 2 \\ 2 & 0 & 1 \\ 1 & 3 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = (6).$$

Processing operations may be classified so as to correspond to operations executed by grouping matrices:

a) Transactors may be grouped in varying degrees (reduction in the dimension of T).
b) Government transactors may be grouped with private actors (if all government actors were initially placed side by side, the operation may be thought of as an interchange of rows or columns which places a government actor among private businesses, for example. The operation is carried out conceptually by interchanging the rows or columns of a grouping matrix).
c) The extent of direct information as to the identity of both parties may be reduced--and extinguished in the cases in which R is a vector, or scalar.
d) Transactions may be grouped (loss of detail as to function, character, or object represented by a further reduction in the dimensions of T).
e) Transactions may be sent to zero (by $a_{ij} = b_{kl} = 0$).
f) The underlying transactions matrix may be torn apart to make room for alterations such as imputed expenditures or revenues (no grouping matrix analogue).
g) Transactions may be netted to a scalar for one actor; i.e.,

$$(\sum_j a_{ij} - \sum_i a_{ij} = r).$$

h) The underlying transactions matrix may be altered in its time dimension so as to convert it from a fiscal to calendar basis, or from a cash to an accrual basis.

Differences in the preparation of data by sources may be portrayed by constructing grouping matrices which correspond to their practices. Grouping matrices may then be compared, and any difference in elements along any row or column, or any interchange of rows or columns will result in a reconciliation problem. Differences in dimensions need not lead to such problems, although it will rarely be possible to reverse the degree of consolidation of series without, for example, inspection of worksheets. Whenever the grouping matrices are vectors, direct information as to the identity of the parties is lost.

This survey of logical possibilities is more than an exercise. The intention is to show that social accounting data as currently prepared may be conceptually derived by grouping operations from a common, underlying transactions matrix.

Derivation of Some Contemporary Social Accounts

Consider the derivation of budget information such as general fund transactions for local government units, or the federal administrative budget. The first step in our program is to partition T so that all transactions among actors of no interest to us at the moment are sent to zero. We may think of this activity as preparing another matrix whose elements are matrices, only one of which is of interest to us. Expenditure transactions of the government unit are reduced to a vector. In the government matrix element the grouping matrices appear as follows

$$A = (11\ldots1)$$

$$\text{and } B = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ 0 & & & 0 \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ 0 & \ldots & \ldots & 0 \end{pmatrix}$$

We obtain R as a vector of budget expenditures by function, department, or character. If the dimensions of T are N by M, then A is 1 x N, and B is M by P. A vector of revenue transactions may be obtained by interchanging the roles of A and B. These data differ from those presented in financial reports (such as Treasury Statements) in that special fund, trust fund, and government sponsored enterprise transactions are not customarily included in administrative budgets. By expanding the row or column dimensions of the grouping matrices we may obtain the "cash budget" which includes these remaining transactions.

The landscape becomes much more obscure when we consider the important income and product accounts. Again, our end products are expenditure and revenue vectors in which direct information about both parties to a transaction is lost. All transactions on capital account other than transactions involving newly produced capital goods are washed out. If other accounting systems do not distinguish transactions in the same way, reconciliation may be impossible. Imputations may be viewed as an alteration of the underlying transactions matrix, and an alteration of its property of being a record of actual transactions. Another practice that leads to problems of reconciliation is the conversion of transactions to a calendar year from a fiscal year basis and to an accrual from a cash basis.

A major reconciliation problem arises in the large number of government transactions that are sent to zero in the process of obtaining aggregates such as national income. Some of these transactions are presented separately. However, all transactions involving land, old assets, and financial items are sent to zero. Another major problem is that transactions of government enterprises are grouped with those of business, a step which may be viewed as an interchange of rows and columns of T brought about by interchanging corresponding rows and columns of A and B in identity matrix form. Entries other than those classified as value added are then sent to zero.

Similar conceptual operations yield input-output tables. The grouping matrices for the reduction of T to existing input-output tables are not equivalent to those utilized for the derivation of income and product information. An important distinction in operations is that the grouping matrices for input-output tables are not vectors. Transaction flows among actors are preserved, as is well known. This difference does not necessarily lead to irreconciliable data as it represents in concept different degrees of consolidation. Also transactions in land and old assets, and transactions in financial assets are customarily sent to zero in the preparation of input-output tables. Some tables have also been prepared with government enterprise transactions grouped as intermediate product (conceptually all intermediate-type transactions could be placed in a north-west partition of (t_{ij}) by appropriate interchange of rows and columns of identity grouping matrices).

Reconciliation problems in the data prepared by these two sources[3] do arise from imputation practices that we have discussed. Problems also arise from the practice of netting items in the income and product accounts, a practice not followed in input-output tables. In the latter, government enterprise transactions are not consolidated with private business transactions.

Flow-of-funds or money flow data are reported to be designed as a record of actual transactions with emphasis upon financial transactions. In this instance, the grouping matrices operate upon (t_{ij}) so as to send no transactions to zero. Inspection of these accounts reveals also that while the detail about currently produced goods and services is less than that of other sources, some information as to parties of transaction, especially financial, is preserved. That is, each grouping matrix is greater in at least one dimension than its associate in the case of the income and product accounts. We note that money-flows data are on a cash basis, and contain no imputations. The degree of consolidation of transactions and transactors except for the financial sector is greater than in the case of other sources. We recall also that money-flows data are available over a shorter period of time. A good deal of netting of transactions is carried out.

The gap between money-flows data and income and product data seems larger than the gap between the latter and most input-output tables as presently constructed. What we have shown is that the difference between these accounts arises from current practice, not from any basic conceptual or a priori conflict. There would seem to be little logical difficulty to overcome in bringing the various accounts into a consistent framework, without impairing prevailing objectives.

Less well developed social accounting data on government revenues and expenditures are available from Bureau of the Census publications. These, however, possess desirable characteristics. They are with one minor exception consistent with cash budget items and therefore make up a balanced system. Operations such as sending transactions to zero, imputing transactions, and grouping actors with those of different motivations are avoided in the preparation of these facts. Data are given for fiscal periods in terms of cash transactions. We lack an isomorphism between these series and cash budget series only to the extent that some items which are net in the latter are gross in the former. This reconciliation problem can easily be repaired by seeking more detailed information available separately in financial records. Census data possess further useful characteristics in that they reveal function in some detail although not account or object. This detail is available at intervals from 1902 to date.[4]

It may be objected that this argument is incomplete, and unfair, in that no mention has been made of the objectives of these varied practices which cause reconciliation problems. Ought not there be a discussion of the goal of estimating an aggregate welfare measure and the fine adjustments required in that case? It must be noted that our aims are not always well served by data prepared for these particular welfare judgements. To the extent that accounts describe actual transactions, the consumer of data rather than the producer could decide upon adjustments. A multiplicity of aims could be served.

FOOTNOTES

[1]Mr. Britton Harris, Research Coordinator of the Penn Jersey Transportation Study and Mr. M. Dutta, Department of Economics, University of Rhode Island, provided helpful discussions. Responsibility remains with the writer.

[2]A comparison of total federal, state, and local government expenditures as reported by the Bureau of the Census, by the National Income Division, and by the cash budgets of these government units can provide one illustration.

[3]By sources, I mean the Department of Commerce in the case of the income and product accounts and the Department of Labor (for example) in the case of input-output tables.

[4]Historical Summary of Governmental Finances in the United States, (U.S. Bureau of the Census: Washington) Vol. IV, no. 3 of the 1957 Census of Governments. In a letter dated September 1, 1960, Mr. George Jaszi, Ass't. Director of the Office of Business Economics, U.S. Department of Commerce, writes that "Government expenditures by function, in the national income and product framework, have not been prepared for the period prior to 1952."

A MULTI-STAGE PROBABILITY SAMPLE FOR TRAFFIC SURVEYS*

Leslie Kish, Survey Research Center, University of Michigan
Warren Lovejoy and Paul Rackow, The Port of New York Authority

Introduction

Origin-destination surveys of vehicular traffic serve to obtain data on the patterns of vehicular travel for use in planning the location, size configuration and competitive effect of future vehicular facilities. Since 1930, The Port of New York Authority has conducted such surveys periodically at the tolls plazas of its tunnels and bridges across the Hudson River on a "one shot" basis, i.e., as many interviews as possible were obtained over a period of one to three days and the results were assumed to be representative of the entire year's traffic.

Because of the deficiencies in this method, a study of the problems and the feasibility of using a probability sample was followed in January 1958 by the adoption of a new sampling technique, dubbed "Continuous Sampling." Briefly, this technique is based on a carefully designed and controlled probability sample which builds up enough interviews over time to give sufficient reliability for the desired estimates.

Estimating the characteristics of vehicular traffic on the six facilities poses some interesting problems in survey sampling. The traffic volume varies greatly among the different facilities as well as over different time intervals. This dimension of time has several identifiable sources of variation: seasonal over months, the pattern over days of the week and the diurnal fluctuation. But we wanted to avoid reflecting these extreme variations in the sampling probabilities for vehicles because the survey objectives made it advantageous to aim for equal probabilities.

Eight hour periods over a single facility, for which estimates of the variable traffic volume are available, served as primary sampling units. Despite these variable volumes we designed roughly equal sampling workloads to make the demands on the enumerator both reasonable and efficient. Equal over-all selection probabilities and equal sampling loads can both be achieved by first selecting the primary sampling units with probabilities proportional to estimated traffic and then sampling within those units with rates inversely proportional to the same probability.

Many other traffic problems possess similar essential features. For example, the traffic over roads crossing the boundaries of a state (or other defined area) has similar dimensions of variations over space and of the various types of variation over time. In still other applications, instead of space and time, other dimensions can be treated similarly. Hence, a description of the solution of such a problem may be of general interest.

Advantages of the Continuous Probability Sample

A number of important gains are achieved by this new traffic survey technique.

(1) By virtue of the fact that the laws of probability hold for this type of sample, we could set confidence limits on the estimates made from the results. In contrast the judgment samples used before could not yield valid inferences about the population in the statistical sense.

(2) Spreading the sampling over time avoids the possibility inherent in "one shot" surveys that seasonal or other variations might make traffic patterns on the day or days surveyed quite unrepresentative of the total population.

(3) This new technique provides up-to-date information on origin and destination patterns at all times because the survey data is continuously processed, tabulated and analyzed. In particular, this method of sampling provides measurements of seasonal variation which has been unavailable heretofore.

(4) The design of the survey eliminates the necessity of hiring large numbers of unskilled, temporary employees as interviewers and coders. For the current survey, four regularly employed interviewers did all of the interviewing as well as the coding. These men were trained in interviewing methods, had a thorough knowledge of the geography of the New York Metropolitan area, and were skilled in the coding of the data.

(5) Actual over-all costs for a year's survey using a probability sample is less than that for the judgment survey taken on three days in the past.

(6) Intensive sampling at any facility at any one time is avoided, thereby reducing the instances of congestion caused by obtaining origin and destination information. Furthermore, skilled interviewers reduce the time required for each interview, reduce the proportion of nonresponse, and obtain more complete information on each interview without undue delay to the motorist.

Data Obtained

During the year 1958, some 92,300 interviews were obtained at the six Port Authority vehicular facilities representing about 0.1 per cent of total annual traffic involved in the survey. A response rate of nearly 99 per cent was achieved which was a considerable improvement over the level obtained in the "one shot" surveys conducted in the past.

*Presented at the 1960 Annual meeting of the American Statistical Association, but did not appear in the 1960 Proceedings.

While the basic purpose of the survey was to obtain information as to the origins and destinations of autos, trucks, and tractor-trailers using the Port Authority tunnels and bridges, supplementary information was obtained concerning purpose of trip, license plate of vehicle, number of passengers in vehicles and time, day-type and direction. Buses were not included in the survey.

The survey was continued through 1959 and is in current use in 1960 with some modifications in the sample design to provide greater statistical reliability for certain types of peak period data.

Survey Requirements

The sample design had to be both economical and practical in order to meet the necessary restrictions inherent in this type of field survey especially as regards reasonable workloads, conditions of work and cost.

There were a number of specific requirements which the survey design had to accomplish:

(1) Sample estimates had to be produced for a great variety of specified lines of travel of varying volumes. A line of travel is defined as traffic moving between two geographically specified zones. These estimates had to be provided not only for the six Port Authority vehicular crossings combined, but for each facility separately on an annual basis as well as for each of the four seasons of the year.

(2) The sampling frame included all revenue autos, trucks and tractor-trailers using the six Port Authority vehicular crossings on all types of days (weekdays, Saturdays, Sundays and holidays). Buses were not included in the frame.

(3) The sample was to be a probability sample requiring that every vehicle trip made during the year have a known probability of selection, achieved by randomized selection procedures at every stage of the sample design.

(4) There had to be built into the sample design a replicative procedure so that proper measurements of the standard errors of the various sample estimates could be made.

(5) The statistical reliability required in the survey specified that with the expected over-all annual sample of 100,000 interviews, a sample estimate of 1 per cent, or 1,000 interviews, would have a coefficient of variation of the order of 3 to 6 per cent.

The design also had to be fitted to the practical problems inherent in the actual conduct of this type of survey, and these considerations necessitated the use of the following procedures:

(6) The survey budget permitted the assignment of four permanent field interviewers to this project. These interviewers were to be responsible for both the collection and the coding of the data obtained in the field. This limitation of available manpower restricted the number of tours of duty in the field to eleven per week.

(7) Reasonable working hours and the desirability of reducing variations in the hourly loads within a shift were used as criteria for designating the tours of duty at a facility. Thus, the shift, an eight hour tour of duty at one facility on one of the seven days of the week, became the primary sampling unit. The tours were designated as, a) 11 p.m. - 7 a.m. b) 7 a.m. - 3 p.m. c) 3 p.m. - 11 p.m.

(8) To improve coverage or different origin and destination patterns, directional flows, vehicle types and vehicular volumes at each facility during an eight hour period, the interviewer moves from one traffic lane to another each hour in a prescribed pattern of rotation. The design includes specified relief periods each hour and it is assumed, with good reason, that the traffic patterns during these relief periods are not different from those sampled during the interviewing periods.

(9) Because of the distances involved in covering the entire roll plaza of some facilities, it was necessary to subdivide some facilities into two or more locations. The locations were selected in such a manner as to allow interviewers to count the number of toll lanes open within a location during each hour of interviewing. The interviewer rotates among the open lanes at one prespecified location for four hours and then moves on to another location.

(10) Experience with previous "one shot" surveys supplemented by field testing indicated that an average workload of forty interviews per hour could be achieved by an interviewer. This is an average rate that can be maintained for long periods and still permitting considerably greater rates for occasional peak hours.

(11) Available data on hourly traffic volumes via each facility on the different days of the week were utilized in selecting uniform sampling rates during each shift. The probability of selecting a shift at a facility was made roughly proportionate to the average hourly density of traffic expected during that shift. Then the selection rates within the shifts were made inversely proportional to that same probability, thus producing equal final probabilities and roughly equal hourly workloads.

(12) Flexibility is required to handle such

factors as the ratio of actual to expected lanes open, loss of interviewing time due to bad weather, unusual variations in volume and nonresponse. The self-weighting sample was maintained by introducing a procedure of balancing actual against expected interviews by duplicating or omitting some as required.

Four-Stage Probability Sample

The sampling plan of the survey follows a stratified multi-stage probability design. The sample was selected in four stages with the probability of selection of any vehicle crossing a facility as the product of the probabilities in the several stages. These probabilities were controlled to maintain the equal chance of selecting all vehicles at any facility at any time. The four stages of selection were:

(1) Shift - an eight-hour tour of duty at one of the four facilities on a given day of the week.

(2) Location - a geographically contiguous group of toll lanes at a facility at which the interviewer remains for four hours.

(3) Lane - one of the toll lanes within the location.

(4) Vehicle - an auto, truck, or tractor-trailer passing through a lane.

The primary sampling units (shifts as defined above) were grouped into 84 strata consisting of four facilities by seven day types by three tours of duty. Each stratum was represented by the measure of size fj, a probability proportional to the expected hourly density of traffic flow for that stratum. Each fj, the approximate thousands of vehicles per hour, was an integer for ease of selection and balance. The sum of the fj's for all strata was 286 (26 weeks x 11 shifts) for each half year period, divided into 143 for two seasons. A table of selection was made up of the fj for each of the 84 strata and systematic random sampling was used to make up 13 weekly workloads of 11 shifts each with each workload roughly spread across the strata. Then the workloads were assigned at random to the 13 weeks in the season.

Probabilities for the Four Stages of Selection

The probability of selecting a vehicle crossing a facility may be represented as:

$$\text{Pr (Selecting a Vehicle)} = \left(\frac{11\ fj}{286}\right)\left(\frac{1}{24\ fj}\right)\left(\frac{3}{4}\right) = \frac{1}{832}$$

The first term on the right denotes the probability of selecting a shift. This is the first stage of selection. Eleven shifts are selected for any week with probability fj/286.

The third term, the fraction 3/4, reflects the fact that interviewing takes place only 44 minutes out of each hour $\left(\frac{45}{60} = \frac{3}{4}\right)$, the remainder of each hour being given to relief.

The middle term represents a composite of the last three stages of selection: location within shift which is the second stage, lane within location which is the third stage, and finally, vehicle within lane which is the fourth or ultimate stage of selection. Thus,

$$\frac{1}{24\ fj} = \text{Pr (Selection within a shift)}$$

Pr (Location within shift) x Pr (Lane within Location) x Pr (Vehicle within Lane)

For example, at a facility where there are two locations and four lanes are expected to be open during a specified hour at the selected location, the sampling interval for that hour would be determined as follows:

$$\frac{1}{24\ fj} = \left(\frac{1}{2}\right)\left(\frac{1}{4}\right)\left(\frac{1}{K}\right)$$

where K is the sampling interval. Thus,

$$K = \frac{24\ fj}{2 \times 4} = 3\ fj\ .$$

Hence, if fj for this particular shift were 4, the interviewer would select every 12th vehicle.

Locations were selected by random procedures and balanced for relevant factors over each season. Lane rotation patterns at each location within a facility were specified in advance for the interviewer. Fractional intervals K for selection of vehicles within lanes were made into integers by a balanced process of randomization. at all stages of selection random procedures were always carefully specified and used.

It was mentioned previously that the design provided for a self-weighting sample, which is desirable for a number of reasons: (1) Cost of tabulation is reduced; (2) Simplified the analysis of data for meaningful and desirable subclasses; (3) The variance is at a minimum. One reason for adjusting the number of interviews arises from the following facts: (1) The third stage of the sample design is lane within location in which one lane is selected at random from among those open with probability 1/L, where L is the number of lanes assumed open in the location under concern for a specific hour; this number is supplied for each hour operating personal at each facility. (2) The number of lanes actually open at the selected location is noted by the interviewer for the hour interviewing of motorists is taking place.

The adjustment, if it be necessary, comes from the ratio of actual to assumed lanes open and it is needed to provide the correct actual probabilities in the third stage of selection. Then, the over-all probability of selecting a vehicle will remain the same at all facilities, location, etc., and the sample will be self-weighting. (An actual to expected ratio of more than one results in a increase of interviews and a ratio of less than one in a deficit.)

Similar corrections are made for loss of interviewing time; in very bad weather or in other emergencies the interviewer is permitted to make specified reductions of the working hour and these are noted and adjusted. Non-response also calls for an adjustment, either in case of refusals or if the rushed interviewer must pass a selected vehicle by. Where the need for adjusting the results for the hour arises, this is done by balancing the deficit against possible excess and the adjustment is made by duplicating or omitting cards at random.

Estimates and Sampling Errors

The objectives of the survey consists of many estimates of traffic volume, each estimate for a specified origin-destination couple. Often further specification refers to a defined facility or season, or day or hour, etc. In any case, the estimate is of an aggregate volume of traffic which can be estimated simply as $X'' = Kx$ where K is the inverse of the over-all sampling rate and $\underline{x}$ is the self-weighting card count in the sample with the specified characteristics. But instead of the above estimator, we use ratio estimators of the form $X' = Nx/n = Np$ where n is the card count for the entire sample or some specified subclass and N is the known total traffic for same; x is defined above. That is, the sample proportion $p = x/n$ for some specified characteristic and subclass is projected to a total by utilizing the information about the aggregate count (N) available from automatic counters.

We used ratio estimates throughout for estimating vehicles or passengers in any characteristic under consideration. They were used because of their ease of application to the data available; because the variance estimates are generally less for ratio estimates than for the simple estimates; and because this eliminates the bias due to the simple "slippage" or "non-coverage" which is bound to creep into field work.

Sampling Error

The two basic types of errors encountered to a greater or lesser degree in all sample surveys affecting the reliability of the results are sampling and non-sampling errors. Non-sampling error stems primarily from errors of response in collecting and in processing the survey results and from any bias in the sample due to non-response. Sampling error arises from the fact that the characteristics as pictured by the sample do not exactly coincide with the characteristics which would emerge from an equal complete coverage of the entire frame.

For computing the sampling error we used a model in which the entire selection for the season consisted of two random and independent halves. (For the sake of simplicity we recommend that this be actually done.) Actually, by "collapsing" strata we created the two computing units; by disregarding some further stratification this results in slight over-estimation of the variance.

The two self-weighting halves for the j-th season, may be represented as follows in estimating the proportion for some characteristic:

$$p_j = \frac{x_j}{n_j} = \frac{x_{j1} + x_{j2}}{n_{j1} + n_{j2}}$$

For the entire year of four seasons the similar estimator is the ratio of the sums of the seasons

$$p = \frac{x}{n} = \frac{\sum_{j=1}^{4}(x_{j1} + x_{j2})}{\sum_{j=1}^{4}(n_{j1} + n_{j2})}$$

The "relvariance" (the square of the coefficient of variation) of p can be estimated by:

$$C_p^2 \doteq C_x^2 + C_n^2 - 2\,C_{xn}$$

$$\text{where } C_x^2 = \frac{1}{x^2}\sum_{j=1}^{4}(x_{j1} - x_{j2})^2$$

$$C_n^2 = \frac{1}{n^2}\sum_{j=1}^{4}(n_{j1} - n_{j2})^2$$

$$C_{xn} = \frac{1}{xn}\sum_{j=1}^{4}(x_{j1} - x_{j2})(n_{j1} - n_{j2})$$

Thus from the sum of four seasonal contrasts for each year's estimates the relvariance can be computed with four degrees of freedom. These computations are made for a large number of items. These are then plotted and average values, subject to smaller variations, are used for estimating standard errors.

ASYMPTOTIC EQUIVALENCE OF EFFICIENCY FOR THREE PROCEDURES OF UNEQUAL PROBABILITY SAMPLING WITHOUT REPLACEMENT

J. N. K. Rao
Iowa State University

I. Introduction

Let y_j denote a characteristic attached to the j^{th} unit of a finite population of N units with population total $Y = \Sigma y_j$. It is well known that the use of unequal probabilities in selecting a sample may bring about considerable reduction in variance as compared to equal probability sampling. For example, such a procedure may be useful when a 'measure of size' x_j is known for all units in the population and it is suspected that these known sizes x_j are correlated with the characteristic y_j. One method (though by no means the only method) of utilizing x_j is to draw units with probabilities proportional to sizes x_j (p.p.s.), a technique frequently used in surveys, particularly for selecting primary units in multistage designs. Now the theory of sampling with unequal probabilities is equivalent to multinomial sampling provided the units are drawn with replacement. On the other hand, we know from equal probability sampling that selection with replacement results in estimators which are less precise than those computed from samples selected without replacement, the proportional reduction in the variance being given by the sampling fraction n/N. It is, therefore, natural to investigate the efficiency of unequal probability sampling without replacement as compared to unequal probability sampling with replacement.

A general theory of unequal probability sampling without replacement is given by Horvitz and Thompson (1952). Their estimator of Y is

$$\hat{Y} = \sum_{1}^{n} y_j/P_j \qquad (1)$$

with variance

$$V(\hat{Y}) = \sum_{1}^{N} \frac{y_j^2}{P_j} + 2 \sum_{i' > i = 1}^{N} \frac{P_{ii'}}{P_i P_{i'}} y_i y_{i'} - Y^2, \qquad (2)$$

where P_j is the probability for the j^{th} unit to be in a sample of size n and $P_{ii'}$ is the probability for the units i and i' both to be in the sample.* Now, when the P_j are proportional to the y_j, $\hat{Y}$ is constant and hence $V(\hat{Y})$ is zero which suggests that, by making P_j proportional to the known sizes x_j, considerable reduction in the variance will result if the x_j are approximately proportional to the y_j. Special devices are needed to satisfy this condition, namely $P_j = np_j$ where $p_j = x_j/\Sigma x_j$, when sampling with p.p.s. and without replacement. Yates and Grundy (1953) suggest an iterative procedure to obtain revised sizes x_j^* and draw the sample by selecting the first unit with probabilities proportional to the revised sizes, the second unit with probabilities proportional to the remaining (revised sizes), and so on, such that $P_j = np_j$. Let us call their method procedure 1. One can also use the iterative procedure of obtaining revised sizes for another well known sampling scheme which is as follows: n units are drawn with probabilities proportional to revised sizes with replacement. If any unit is selected more than once in the sample, reject the n selections and make further n selections with replacement and with probabilities proportional to revised sizes, the process being continued until n different units are selected in the sample. Let us call this method procedure 2 (e. g., Durbin (1953)). In order to draw a sample by procedure 1 or 2 it is convenient to use a method suggested by Lahiri (1951).* The application of Lahiri's method leads to the following scheme of drawing the sample for procedure 1: Let us consider, for illustration, n=2. Let x_o^* be a number not smaller than the largest of x_j^*. Select a random integer between 1 and N, say α where $1 \leq \alpha \leq N$. Select similarly a random number β, subject to the condition $0 < \beta \leq x_o^*$. If $\beta \leq x_\alpha^*$, then unit α is selected for the sample. If $\beta > x_\alpha^*$, unit α is not selected (at this 'draw') and the whole process has to be repeated with all the units until a selection is made. The probability of selecting a unit by this procedure is proportional to the revised sizes. After a unit is selected, repeat the whole process with the remaining x_j^* until another selection is made. For procedure 2, after a unit is selected by Lahiri's method, the whole process is repeated with all the x_j^* for the selection of the second unit. If the second unit selected happens to be the same as the first unit, reject both selections and repeat the whole process with all the units until two different selections are made.

There is, however, a well known procedure of drawing a sample with p.p.s. and without replacement which insures that $P_j = np_j$ with the original sizes x_j. This procedure, which we denote as procedure 3, is as follows: The N units in the population are listed in a random order and their x_j are cumulated and a systematic selection of n elements from a random start is then made on the cumulation (e.g., Goodman and Kish (1950)). It may be noted that for all the three procedures all x_j's should satisfy the necessary condition $P_j = np_j \leq 1$. Therefore, if there are some units for which np > 1, either select them automatically in the sample or use some other device like stratification or splitting a unit into smaller units, etc.

Recently (Hartley and Rao (1959), (1962)), the mathematical difficulties involved in evaluating the probabilities $P_{ii'}$, for procedure 3, are resolved with the help of an asymptotic

*It may be pointed out that in unequal probability sampling there are several different classes of linear estimators. The estimator (1) belongs to one of these classes and we concentrate here only on the estimator (1). For a detailed discussion of different classes of linear estimators the reader is referred to Koop (1957).

*Professor Leslie Kish suggested that I point out how Lahiri's method can be used to draw a sample by procedure 1 or 2.

theory, and compact expressions for the variance of the estimate of the population total together with variance estimates have been obtained. The following approach is used in developing the asymptotic theory: In sampling with p.p.s. and with replacement we have

$$\hat{Y}^1 = \sum^n y_j/np_j, \tag{3}$$

where $\sum^n$ denotes the summation over the n units selected with replacement, as the estimator of Y. Also the variance of $\hat{Y}^1$ is

$$V(\hat{Y}^1) = \sum_1^N P_j\left(\frac{y_j}{P_j} - \frac{Y}{n}\right)^2 \tag{4}$$

where $P_j = np_j$. Therefore, by assuming that each P_j is of the order of N^{-1} for large N, it is seen that $V(\hat{Y}^1)$ is of order N^2. In sampling without replacement the term of order N^2 will be the leading term in $V(\hat{Y})$ and hence the next lower order terms, namely terms of order N^1, will represent the gain in precision due to sampling without replacement. Therefore, Hartley and Rao evaluated (for their procedure 3) $V(\hat{Y})$ to order N^1. This is equivalent to evaluating $P_{ii'}$ to order N^{-3} and substituting it in the variance formula (2). Also, for the benefit of smaller size populations, they evaluated $P_{ii'}$ to order N^{-4} and hence $V(\hat{Y})$ to order N^0.

The purpose of the present paper is to present compact expressions for the variance and estimated variance together with simplified formulas for revised sizes in terms of the original sizes for procedures 1 and 2, obtained by using the asymptotic theory. The mathematical derivations are not given here and will be published elsewhere. It is shown that the three procedures have exactly the same value of $P_{ii'}$ to order N^{-3} and hence identical $V(\hat{Y})$ to order N^1. Since the terms of order N^1 are the important terms in $V(\hat{Y})$, which contribute to the gain in precision of sampling without replacement over sampling with replacement for large N, it follows that the three procedures have practically the same efficiency. However, with procedure 3 there is no need to compute the revised sizes, and this procedure, therefore, circumvenes an operation which may be cumbersome as N becomes large. It is also shown for the case n = 2 that, to order N^0, the estimator from procedure 1 has smaller variance than that from procedure 2 and that from procedure 2 has smaller variance than that from procedure 3. However, as N increases, the contribution from terms of order N^0 becomes negligible.

II. The case n = 2

Let $p_j^* = x_j^*/\Sigma x_j$. Then, for procedure 1, we find

$$p_j^* = p_j\left[1 + \frac{1}{4}(2p_j - 2\sum_1^N p_t^2 + 4p_j^2 - 3p_j\sum_1^N p_t^2 + 3(\sum_1^N p_t^2)^2 - 4\sum_1^N p_t^3)\right] \tag{5}$$

to order N^{-3}. If the sample is selected by procedure 1 using the revised sizes from (5), we obtain

$$\begin{aligned} V^{(1)}(\hat{Y}) = {} & \sum_1^N P_j\left(1 - \frac{P_j}{2}\right)\left(\frac{y_j}{P_j} - \frac{Y}{2}\right)^2 \\ & - \frac{1}{2}\sum_1^N \left(P_j - \frac{P_j^2 \sum_1^N P_t^2}{4}\right)\left(\frac{y_j}{P_j} - \frac{Y}{2}\right)^2 \\ & + \frac{3}{32}\left(\sum_1^N P_t y_t - \frac{Y\sum^N P_t^2}{2}\right)^2 \end{aligned} \tag{6}$$

and

$$\begin{aligned} v^{(1)}(\hat{Y}) = {} & \Big[1 - (P_i + P_{i'}) + \frac{1}{2}\sum_1^N P_t^2 \\ & - \frac{1}{2}(P_i^2 + P_{i'}^2) + \frac{5}{8}P_iP_{i'} - \frac{3}{32}(\sum_1^N P_t^2)^2 \\ & - \frac{1}{16}(P_i + P_{i'})\sum_1^N P_t^2 + \frac{1}{2}\sum_1^N P_t^3\Big] \\ & \cdot\left(\frac{y_i}{P_i} - \frac{y_{i'}}{P_{i'}}\right)^2 \end{aligned} \tag{7}$$

to order N^0, where $P_j = 2p_j$. Here $V^{(1)}(\hat{Y})$ and $v^{(1)}(\hat{Y})$ denote the variance and the estimated variance respectively for procedure 1, and i and i' are the two units included in the sample. For procedure 2 we have

$$p_j^* = p_j\left[1 + (p_j - \sum_1^N p_t^2 + 2p_j^2 - 2p_j\sum_1^N p_t^2 + 2(\sum_1^N p_t^2)^2 - 2\sum_1^N p_t^3)\right] \tag{8}$$

to order N^{-3}. If the sample is selected by procedure 2 using the revised sizes from (8), we find

$$\begin{aligned} V^{(2)}(\hat{Y}) = {} & \sum_1^N P_j\left(1 - \frac{P_j}{2}\right)\left(\frac{y_j}{P_j} - \frac{Y}{2}\right)^2 \\ & - \frac{1}{2}\sum_1^N \left(P_j^3 - \frac{P_j^2 \sum^N P_t^2}{2}\right)\left(\frac{y_j}{P_j} - \frac{Y}{2}\right)^2 \\ & + \frac{1}{8}\left(\sum_1^N P_t y_t - \frac{Y\sum^N P_t^2}{2}\right)^2 \end{aligned} \tag{9}$$

and

$$\begin{aligned} v^{(2)}(\hat{Y}) = {} & \Big[1 - (P_i + P_{i'}) + \frac{1}{2}\sum_1^N P_t^2 \\ & - \frac{1}{2}(P_i^2 + P_{i'}^2) + \frac{1}{2}P_iP_{i'} - \frac{1}{8}(\sum_1^N P_t^2)^2 \\ & + \frac{1}{2}\sum_1^N P_t^3\Big]\left(\frac{y_i}{P_i} - \frac{y_{i'}}{P_{i'}}\right)^2 \end{aligned} \tag{10}$$

to order N^0, where $V^{(2)}(\hat{Y})$ and $v^{(2)}(\hat{Y})$ denote the variance and the estimated variance respectively for procedure 2. Finally, for procedure 3, denoting the variance and the estimated variance by $V^{(3)}(\hat{Y})$ and $v^{(3)}(\hat{Y})$ respectively, we find

$$V^{(3)}(\hat{Y}) = \sum_1^N P_j\left(1 - \frac{P_j}{2}\right)\left(\frac{y_j}{P_j} - \frac{Y}{2}\right)^2 - \frac{1}{2}\sum_1^N \left(P_j^3 - \frac{P_j^2 \sum_1^N P_t^2}{2}\right)\left(\frac{y_j}{P_j} - \frac{Y}{2}\right)^2 + \frac{1}{4}\left(\sum_1^N P_t y_t - \frac{Y \sum_1^N P_t^2}{2}\right)^2 \qquad (11)$$

and

$$v^{(3)}(\hat{Y}) = \Big[1 - (P_i + P_{i'}) + \frac{1}{2}\sum_1^N P_t^2 - \frac{1}{2}(P_i^2 + P_{i'}^2) - \frac{1}{4}\Big(\sum_1^N P_t^2\Big)^2 + \frac{1}{4}(P_i + P_{i'})\sum_1^N P_t^2 + \frac{1}{2}\sum_1^N P_t^3\Big] \cdot \left(\frac{y_i}{P_i} - \frac{y_{i'}}{P_{i'}}\right)^2 \qquad (12)$$

to order N^0. From (6), (9) and (11) it follows that $V^{(1)}(\hat{Y}) < V^{(2)}(\hat{Y}) < V^{(3)}(\hat{Y})$, to order N^0. However, the three procedures have exactly the same variance to order N^1, namely

$$V(\hat{Y}) = \sum_1^N P_j\left(1 - \frac{P_j}{2}\right)\left(\frac{y_j}{P_j} - \frac{Y}{2}\right)^2. \qquad (13)$$

Also, the three procedures have exactly the same estimated variance to order N^1, namely

$$v(\hat{Y}) = \Big[1 - (P_i + P_{i'}) + \frac{1}{2}\sum_1^N P_t^2\Big] \cdot \left(\frac{y_i}{P_i} - \frac{y_{i'}}{P_{i'}}\right)^2. \qquad (14)$$

Equation (13) when compared with the variance in sampling with replacement, namely (4), clearly demonstrates the reduction in variance achieved by sampling without replacement through the 'finite population corrections' $(1 - P_j/2) < 1$.

Since the three procedures have exactly the same variance to order N^1 and since the terms of order N^1 are the important terms that contribute to the reduction in the variance achieved by sampling without replacement for large N, one can conclude that in most of the practical situations there is little to choose between the three procedures on the basis of efficiency alone. However, since there is no need to compute revised sizes with procedure 3, this may be preferred to other procedures. If one is using procedure 1 and is satisfied with the variance to order N^1, then the revised sizes x_j^* are obtained from the simplified formula

$$p_j^* = p_j\left[1 + \frac{1}{2}\Big(p_j - \sum_1^N p_t^2\Big)\right] \qquad (15)$$

to order N^{-2}. Similarly, for procedure 2, the revised sizes x_j^* are obtained from

$$p_j^* = p_j\left[1 + \Big(p_j - \sum_1^N p_t^2\Big)\right] \qquad (16)$$

to order N^{-2}.

An example for the evaluation of the variance formulas

Horvitz and Thompson (1952 , Table 2) give the data of a population of N = 20 blocks in Ames, Iowa. Here y_i and x_i denote the number of households and the 'eye-estimated' number of households respectively in the i^{th} block. Using their data the following values are obtained:

$$V(\hat{Y}^1) = 3{,}241 \text{ using } (4)$$
$$V(\hat{Y}) = 3{,}025 \text{ to order } N^1 \text{ using } (13).$$

Incidentally, for this example, $\sum^{20} P_t y_t = 49.63$ and $Y \sum^{20} P_t^2/2 = 49.95$. Therefore, we find that, to the nearest integer and to order N^0, $V^{(1)}(\hat{Y}) = V^{(2)}(\hat{Y}) = V^{(3)}(\hat{Y}) = 3{,}007$. It may be of interest to exhibit the nature of convergence of the various approximations to the variance by regarding the variance formula for sampling with replacement as an approximation to order N^2 as set out in Table 1 below.

Table 1. Terms in the approximations to the variance formula

Order of approximation	Sampling procedure	Variance	Difference
$O(N^2)$	with replacement	3,241	216
$O(N^1)$	procedures 1, 2 and 3	3,025	18
$O(N^0)$	procedures 1, 2 and 3	3,007	

The convergence in this example appears to be satisfactory although the population size (N = 20) is much smaller than those usually encountered in survey work. The variance in equal probability sampling without replacement for this example is 16,219. Therefore, all these procedures of p.p.s. sampling are vastly superior compared to equal probability sampling. It must not be forgotten, however, that there are other devices of decreasing the variance in the latter case with the help of the known x_i values (e.g., ratio estimation). The gain in precision through unequal probability sampling without replacement as compared to sampling with replacement is about 7% (235/3241).

III. The general case $n \geq 2$

Most of the published literature on unequal probability sampling deals only with the case n=2 and does not have anything to offer for n > 2 due to difficulties in evaluating $P_{ii'}$, and hence $V(\hat{Y})$ and $v(\hat{Y})$. Though the case n = 2 is important, particularly in stratified designs, situations often arise when n is greater than 2. A striking feature of our asymptotic approach is that it permits an easy evaluation of $V(\hat{Y})$ and $v(\hat{Y})$ for n > 2. It can be shown that the three procedures have exactly the same variance and estimated variance to order N^1, namely

$$V(\hat{Y}) = \sum_1^N P_j \left[1 - \frac{(n-1)}{n} P_j\right] \left(\frac{y_j}{P_j} - \frac{Y}{n}\right)^2 \quad (17)$$

and

$$v(\hat{Y}) = (n-1)^{-1} \sum_{i>i'}^{n} \left[1 - (P_i + P_{i'}) + \frac{1}{n}\sum_1^N P_t^2\right] \left(\frac{y_i}{P_i} - \frac{y_{i'}}{P_{i'}}\right)^2, \quad (18)$$

where $P_j = np_j$. The assumption that P_j is of order N^{-1} implies that n is relatively small compared to N. For procedure 1 the revised sizes x_j^* are obtained from

$$p_j^* = p_j \left[1 + \frac{(n-1)}{2} p_j - \frac{(n-1)}{2} \sum_1^N p_t^2\right], \quad (19)$$

and for procedure 2 the x_j^* are found from

$$p_j^* = p_j \left[1 + (n-1) p_j - (n-1) \sum_1^N p_t^2\right] \quad (20)$$

to order N^{-2}.

The extension of the theory, presented here for simple sampling, to multi-stage sampling, etc., is fairly straightforward and will not be discussed here.

References

Durbin, J. (1953). Journal of the Royal Statistical Society, Series B, p. 253.

Goodman, R. and Kish, L. (1950). Journal of the American Statistical Association, p. 350.

Hartley, H. O. and Rao, J. N. K. (1959). Proceedings of the Social Statistics Section, American Statistical Association, 1960, p. 56.

Hartley, H. O. and Rao, J. N. K. (1962). Annals of Mathematical Statistics (in press).

Horvitz, D. G. and Thompson, D. J. (1952). Journal of the American Statistical Association, p. 663.

Koop, J. C. (1957). Institute of Statistics, Mimeo Series No. 296, North Carolina.

Lahiri, D. B. (1951). Bulletin of the International Statistical Institute, Part II, p. 133.

Yates, F. and Grundy, P. M. (1953). Journal of the Royal Statistical Society, Series B, p. 235.

EMPIRICAL SEQUENTIAL TESTS

Lester S. Adelman, Martin Miller and Peter Nemenyi
State Univ. of N. Y., Downstate Medical College, Brooklyn

0. Summary
1. General Description of the Proposed Method
2. Comparison with Fixed Sample Size and S.L.R. Tests (General)
3. Example: Binomial p
4. Example: Normal Mean, Variance Known
5. A Nonparametric Example
6. A Modification and Refinement

Summary

Sequential likelihood ratio tests [1] require a knowledge of the power function of a test statistic. It often happens, especially in nonparametric procedures, that the power function of a statistic has not been tabulated.

A multi-stage procedure is proposed which can be used with any statistic whose Null-distribution is tabulated. The method is not optimal but offers appreciable average savings in sample size, at least in certain Normal situations.

In a modification, the tail probabilities P_1 at each step are converted to normal deviates and fed into a Wald sequential scheme.

1. General Description of the Proposed Method

Consider any statistic U designed to test a null hypothesis H_o against an alternative H_1. Assume that U tends to be big when H_1 is true and that the full distribution of U under H_o is known, so that any percentage point can be found.

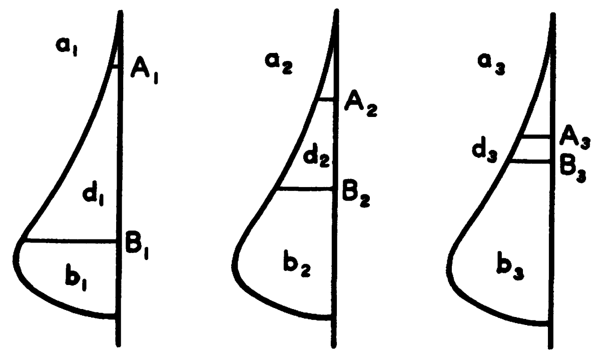
The usual fixed sample size procedure is to calculate U from a sample of size n, reject H_o at the α level if $U \geq U_\alpha$ and "accept H_o" or "reserve judgment" if $U < U_\alpha$. Here U_α is defined by $\Pr\{U \geq U_\alpha | H_o\} = \alpha$ if U is a continuous statistic (e.g. a sample mean) or $\leq \alpha$ if it is discrete.

If H_1 is simple or is in the form of a prior distribution, and if the distribution of U under (each) H_1 is also known, then the power of the test, Pr $\{\text{rejecting } H_o | H_1\}$ can also be calculated.

The fixed sample size procedure can always be modified as follows to yield sequential or multi-stage tests:

Take a sample of size n_1 and calculate the value of the statistic, which will be called U_1. If $U_1 \geq A_1$, reject H_o; if $U_1 < B_1$, accept H_o; if $B_1 \leq U_1 < A_1$, take a small (independent) sample of size n_2 and proceed as before with subscripts 2 in place of 1. Continue in this way until the first time that $U_i \geq A_i$ or $< B_i$. At each step the probabilities $a_i = \Pr\{U_i \geq A_i | H_o\}$ and $b_i = \Pr\{U_i < B_i | H_o\}$ are known, A_i and B_i having been chosen in advance. The probability that sampling will be continued after a given step is $1-a_i-b_i$, which will be called d_i for short.

The level of the test will be $\leq \alpha$ if the constants n_i and $A_i \geq B_i$



are chosen in any manner such that

$$\alpha \geq a_1 + d_1 a_2 + d_1 d_2 a_3 + \dots \qquad (1).$$

The power is

$$1-\beta = a_1' + d_1' a_2' + d_1' d_2' a_3' + \dots \qquad (2).$$

The expected sample size is

$$\text{ASN} \doteq \begin{cases} n_1 + d_1 n_2 + d_1 d_2 n_3 + \dots & \text{under } H_o \\ n_1 + d_1' n_2 + d_1' d_2' n_3 + \dots & \text{under } H_1 \\ n_1 + d_1'' n_2 + d_1'' d_2'' n_3 + \dots & \text{under } H \end{cases}$$

where primed probabilities are calculated under the assumption that H_1 is true, and double prime refers to probabilities based on any other assumption H. The summations are carried to i = M, the point of truncation where A_i is set = B_i and sampling terminates automatically ($d_i = 0$).

β is known if the distribution of U under H_1 is known for fixed sample sizes; (the null distribution of U for fixed n is assumed to be known). The ASN can also be calculated under any assumption for which fixed sample size distributions of U are known.

It is possible to devise non-truncated strategies, for example by setting $a_i = a$ = constant, and $b_i = b$ = constant $(i = 1, 2, \ldots, \infty)$, $d = 1-a-b$, $\alpha = a/(1-d)$. This particular infinite strategy is poor, and others have not been explored. Perhaps the method proposed in this paper should therefore be called "multi-stage" rather than "sequential". In particular, we have given most attention to three-stage strategies, finding that these can do a good deal better than two-stage strategies but that we were unable to obtain substantial further gains by employing four stages with this method.

2. Comparison with Fixed Sample Size and S.L.R. Tests (General)

When α, β and the number of steps M have been specified, the problem arises how to choose the constants n_i, a_i and b_i $(i = 1, 2, \ldots, M)$ from all possible sets satisfying (1) and (2) in such a manner as to minimize the ASN under H_o or H_1 (or some intermediate H). Alternatively it might be desired to specify α, M (or $\sum_1^M n_i$) and the ASN under some H, and minimize β.

This kind of problem cannot be solved in general terms, as the solution depends on the sampling distributions of the statistic used. Even in the simpler special cases an optimum may be very difficult to find. But it can be shown that:

(a) the proposed M-step method can be designed so as to yield an ASN appreciably smaller than the fixed sample size necessary for the same level and power, at least in certain Normal situations, and

(b) no choice of constants can make the ASN as small as a sequential likelihood ratio test [1]. This is clear because the proposed method treats all outcomes in the zone of indifference $B_i \leq U_i < A_i$ alike while the Wald test takes full account of the actual value of U_i within this interval. (The multi-stage tests frequently used in sampling inspection [2] are like Wald sequential tests in this respect).

Thus the proposed method is not optimal but can be recommended as a worthwhile improvement over fixed sample size methods when Wald-like sequential methods are not available, as for example in many nonparametric problems. The point of view is that if The Best is not available, more modest improvements should not be shunned.

The M-stage method is rather flexible. The sample sizes used in any one step, or from step to step, need not be equal. Also it does not matter if external conditions (e.g. variances) change from step to step; and if conditions call for it, different test statistics for the same hypothesis may be used at different stages so long as the decision to change is not made after inspection of the data with an eye on the outcome. And only the null distribution of test statistics needs to be known to make the procedure possible.

In making comparisons between the proposed multi-stage tests and either fixed sample size or s.l.r. strategies, it is desirable to compare their performance not only at two points specified by H_o and a simple H_1 but over a whole range of parameter values, since various values may occur in practice. Thus we have set up comparisons in such a way as to juxtapose the ASN of a given multi-stage test at each parameter point with that fixed sample size that would yield the same power at this point; alternatively we compare the power of a given strategy with the power of a test using a fixed sample of size equal to the ASN calculated at that point. This type of comparison of two curves seems to be more realistic than the more usual comparison of a curve with the horizontal straight line based on a simple alternative.

In most problems, the values of β and ASN are not readily available for s.l.r. strategies at parameter values other than H_o and H_1. We therefore chose to make the comparison with s.l.r. tests in a simple binomial problem where we were able to get information on the s.l.r. test for at least one intermediate parameter value.

3. Example: Binomial p

Multi-stage - H_o: $p = .1$, H_1; $p > .1$.

$M = 3$, $n_1 = n_2 = n_3 = 48$.

U_i = number of "successes" observed at i-th stage.

Step 1: Reject H_o if $U_1 \geq 9$,
accept H_o if $U_1 < 7$,
continue if $7 < U_1 \leq 9$.

Step 2: Reject H_o if $U_2 \geq 8$,
accept H_o if $U_2 < 7$,
continue if $U_2 = 7$.

Step 3: Reject H_o if $U_3 \geq 8$,
accept H_o if $U_3 < 8$.

In other words,

$$\begin{Bmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \end{Bmatrix} = \begin{Bmatrix} 9 & 8 & 8 \\ 7 & 7 & 8 \end{Bmatrix}.$$

From Bureau of Standards binomial tables,

$$\begin{Bmatrix} a_1 & a_2 & a_3 \\ 1-b_1 & 1-b_2 & 1-b_3 \end{Bmatrix} = \begin{Bmatrix} .046 & .102 & .102 \\ .200 & .200 & .102 \end{Bmatrix}.$$

d's by subtraction, (and $d_3 = 0$).

$$\alpha = a_1 + d_1a_2 + d_1d_2a_3 = .0575.$$

At $p = .2$,

$$\begin{Bmatrix} a_1' & a_2' & a_3' \\ b_1' & b_2' & b_3' \end{Bmatrix} = \begin{Bmatrix} .642 & .771 & .771 \\ .129 & .129 & .229 \end{Bmatrix},$$

and $\beta = b_1'+d_1'b_2'+d_1'd_2'b_3' = .164$.

(Usually the probabilities a_i and b_i would be chosen first and cutoff points A_i and B_i derived; but with discrete statistics, integer A's and B's must be chosen, of course with the probabilities in mind, as these really constitute the strategy).

<u>Wald strategy</u>: H_o: $p = .1$, H_1: $p = .2$.

$\alpha = .0575$, $\beta = .164$;

i.e. the s.l.r. test is equivalent to our strategy at these points as far as level and power is concerned. For this particular s.l.r. test the value of β at $p = .1462$ can be found in Dixon and Massey [3]; it is .6049.

A rough comparison is carried out in the table below:

COMPARISON OF THREE-STEP, S.L.R., AND FIXED-N STRATEGIES IN A BINOMIAL EXAMPLE

OPERATING CHARACTERISTIC

p	3-step	s.l.r.
(0)	(1.0)	(1.0)
.1	0.9425	0.9425
.12	0.3273	
.1462		0.6049
.15	0.5582	
.18	0.2280	
.20	0.1640	0.1640
.30	0.0041	
1.0	0	0

AVERAGE SAMPLE NUMBER

p	3-step	s.l.r.	equiv. fixed-N
(0)	(48.0)	(14.7)	
.1	56.101	39.847	74.69
.12	60.980		70.70
.1462		56.902	
.15	65.044		73.90
.18	63.231		74.753
.20	60.074	44.663	68.48
.30	49.109		37.77
1.0	48.0	3.861	

Equivalent fixed sample sizes were calculated from the normal-approximation formula based on values of β actually calculated for the three-stage test at the various values of p.

Rough graphical interpolation of power functions indicates that the OC curves are fairly close together at least for $0 < p \leq .3$. Hence a point-for-point comparison of expected sample sizes is meaningful. Expected sample sizes are compared in fig. 1. For values of p up to about .24, the ASN of the three-stage test is intermediate between Wald's ASN and the fixed sample sizes calculated to produce the same OC. For larger values of p our strategy becomes more expensive than a fixed sample, due to the requirement that at least the first 48 individuals be drawn.

4. Example for a Normal Mean, Variance Known

Comparison with fixed-sample-size strategies is easiest in the problem of testing hypotheses about the mean of a normal distribution with known variance σ^2. We have tried out a number of three-step strategies and a few two- and four-step strategies, seeking to find out empirically what the better strategies look like and how they behave. The table below shows calculations and results for the best three-stage strategy we found so far:

CALCULATION OF THE OC AND ASN

THREE-STEP TEST FOR A NORMAL MEAN

H_o: $\mu = \mu_o$. H_1: $\mu = \mu_o + \delta$ ($\delta > 0$). $\alpha = .05$

$a_1 = .035$ $1-b_1 = .177857$ $d_1 = .142857$

$a_2 = .07$ $1-b_2 = .141429$ $d_2 = .071429$

$a_3 = .49020$

$\delta\frac{\sqrt{n}}{\sigma}$	0.1	0.5	2.0	4.0
A_1	1.712	1.312	-0.188	-2.188
B_1	0.824	0.424	-1.076	-3.076
A_2	1.376	0.976	-0.524	-2.524
B_2	0.974	0.574	-0.926	-2.926
A_3	-0.051	-0.451	-1.951	-3.951
a_1'	.0434496	.0947618	.5745971	.9856646
$1-b_1'$	.2049726	.3357847	.8590355	.9989509
d_1'	.1615230	.2410229	.2844864	.0132863
a_2'	.0844134	.1645351	.6998585	.9941981
$1-b_2'$	.1650311	.2829862	.8227741	.9982830
d_2'	.0806177	.1184511	.1229156	.0040849
a_3'	.5203371	.6740045	.9744709	.9999611
$1-\beta$	.0638599	.1536610	.8077315	.9989281
ASN	1.1745	1.2696	1.3194	1.0133
N*	1.4811	1.5575	1.5808	1.3891
% saved	20.7	18.5	16.5	27.1

*N = fixed sample size required for $\alpha = .05$ and the same β as that of 3-step test for the same δ.

The following table of % saved by the three-step strategy was calculated in the same way

$\delta\frac{\sqrt{n}}{\sigma}$	0.1	0.2	0.3	0.4	0.5	1.0
% saved	20.7	19.9	20.2	19.4	18.5	14.1

$\delta\frac{\sqrt{n}}{\sigma}$	1.5	2.0	2.5	3.0	4.0
% saved	13.0	16.5	22.1	26.1	27.1

The comparison of equivalent sample sizes is illustrated graphically in Figure 2.

The results obtained so far suggest the following conclusions:

(1) Three-step strategies do substantially better than two-step strategies of the sort discussed in this paper (which yield only diminutive savings and frequehtly losses). Soon after three steps a point of diminishing returns is reached: We have been unable to find a four-step strategy which does much better than a good three-step. (A computer will be used to investigate this further).

(2) The more efficient strategies are of the "converging" kind for which the upper and lower cutoff points move closer together at successive steps. (If the evidence in earlier steps has been suggestive enough to warrant continued sampling, it seems logical to reach a decision on the basis of weaker evidence subsequently than would have been required at the first step.) Thus even a strategy as extreme - (in the sense that the third step tests at the 72% level) - as the following, is quite good:

$a_1 = .03 \quad 1-b_1 = .1966667 \quad d_1 = .1666667$

$a_2 = .06 \quad 1-b_2 = .1433334 \quad d_2 = .0833334$

$a_3 = .71994 = 1-b_3.$

$\delta\frac{\sqrt{n}}{\sigma}$	0.1	0.5	1.0	1.5	2.0	4.0
% saved	24.6	19.9	13.2	11.7	15.2	29.2

(3) The choice of a_1 is rather crucial in determining a good strategy, as a_1 is "undiluted" by factors d_1, d_2 in the formulae for α and β. Similarly n_1 must be quite important, being undiluted in the formula for the ASN. (However we have only tried out strategies with $n_1 = n_2 = \ldots$.)

(4) As might be expected, well-chosen multi-stage strategies are most economical for parameter values very close to those specified by H_o and values quite far from them, and not very helpful for moderate deviations. However, for values extremely far from H_o, any given multi-stage strategy becomes uneconomical because of the requirement that no less than n_1 readings be taken; (in practice this merely means that the n_i's should be made small if extremely large deviations from H_o are to be anticipated).

5. A Nonparametric Example

To illustrate the use of the multi-stage sampling plans with distribution-free statistics, we consider a rank test.

H_o: Two continuous populations are identical

H_1: Translation.

A sample of 5 is drawn from each population at each step. The statistic "U" is Wilcoxon's rank sum statistic, (not in the form of a Mann-Whitney U, although critical values were obtained by transformation from the Mann-Whitney table).

Step 1: Reject H_o, accept H_o or continue according as $U_1 \geq 37$, $U_1 < 32$ or $32 \leq U_1 < 37$;

Step 2: Reject H_o, accept H_o or continue according as $U_2 \geq 36$, $U_2 < 33$ or $33 \leq U_2 < 36$;

Step 3: Reject H_o or accept H_o according as $U_3 \geq 24$ or $U_3 < 24$.

From a table of the Wilcoxon distribution (or rather, by conversion from the Mann-Whitney table), we find that the a's are .028, .048 and .726, the (1-b)'s are .210, .155 and .726 and hence the d's are .182, .107 and 0. The idea was to approximate what seemed like a good three-stage strategy with $\alpha = .05$ as closely as possible with the discrete distribution at hand. From the a's and d's we calculate $\alpha = .05087$.

The power function for this rank test cannot be computed without a knowledge of the power function of Wilcoxon's two-sample statistic for fixed sample sizes 5. (The Dutch school worked it out for sizes up to about 3). Therefore we can only study the performance of the strategy - i.e. a strategy using the same a_i's and b_i's - in the parametric comparison of two means with (let us say) known equal variances. The table below shows some percentage savings in the parametric case, calculated exactly as in the example of Section 4:

d/σ	% Saved by 3-Step
0.06324	27.14
0.3162	20.77
0.6324	12.90
0.6956	11.82
0.7905	10.66
0.8854	10.17
0.9486	10.25
1.0118	10.61
1.1067	11.68
1.2016	13.26
1.2648	14.55
1.8972	27.74

6. A Modification and Refinement

A possible refinement would be to subdivide the zone of indifference at a given step i into subintervals. Sampling is continued whenever U_i falls into any of the subintervals of the middle zone, but the boundaries at the next step are made to depend on the subinterval.

In practice such a subdivision at several successive steps would lead to hopelessly complicated calculations. But the problem can be simplified by going to the limit:

Let the probability of at least the observed U_i at the i-th step given H_o be P_i. (In effect we are doing the probability integral transformation). Let $Z_i = \Phi^{-1}(P_i)$ where Φ is the cumulative standard normal distribution function. Then Z_i at each step is standard normal and is independent of the previous Z's (except that the existence of Z_i (i.e. of an i-th sample) is conditional on earlier Z's being medium-sized). Stop sampling and reject H_o at step M if $\sum_1^M Z_i/\sqrt{M}$ gets big enough. Stop sampling and accept H_o if $\Sigma Z_i/\sqrt{M}$ gets small enough. For intermediate values of $\Sigma Z_i/\sqrt{M}$ continue sampling. For example, the boundaries ("big enough" and "small enough") could be set by an s.l.r. procedure.

Although the method has now been Waldianized in a sense, it cannot generally be made equivalent to Wald's test. This would mean using n_i's = 1 (at least from a certain point on), which sometimes leads to absurdities. E.g. in sign tests, taking one pair at a time contributes a probability p = 1/2 or 1 at each step, thus $Z_i = 0$ (no contribution) or $-\infty$, so that H_1 is almost inevitably accepted even if H_o is true. It is thus necessary that several readings be taken at each step, enough to yield something of a probability distribution each time.

. . . .

References

1. Wald, A., Sequential Analysis, Wiley, 1947.

2. Hastay, W.; Wallis, W. A.; and Eisenhardt, C. Selected Techniques of Statistical Analysis, McGraw-Hill, 1947.

3. Dixon, W. J., and Massey, F. J., Introduction to Statistical Analysis, McGraw-Hill, 2nd Ed., 1957. (See p. 305).

4. National Bureau of Standards, Tables of the Binomial Probability Distribution, U. S. Gov't Printing Office, 1949.

5. Davies, O. L., Design and Analysis of Industrial Experiments, Hafner, 1956.

6. Armitage, P., Sequential Medical Trials, Charles C. Thomas, Springfield, 1960.

7. Chilton, N. W.; Fertig, J. W.; and Kutcher, A. H., "Studies in the Design and Analysis of Dental Experiments, III Sequential Analysis, (Double Dichotomy)", J. Dental Research, 40, No. 2, 1961, 331-340.

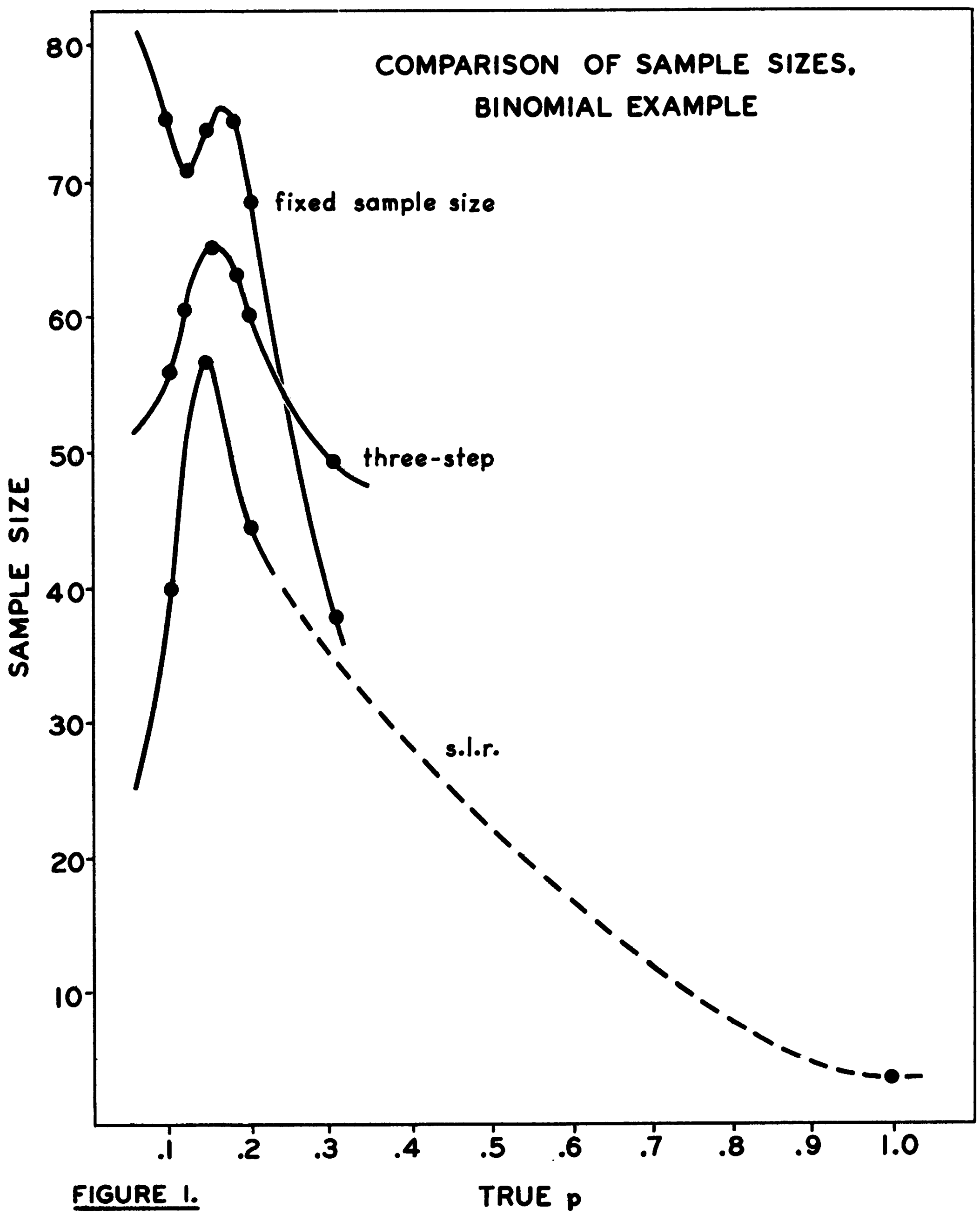

COMPARISON OF SAMPLE SIZES,
BINOMIAL EXAMPLE
fixed sample size
three-step
s.l.r.
SAMPLE SIZE
80
70
60
50
40
30
20
10
.1
.2
.3
.4
.5
.6
.7
.8
.9
1.0
TRUE p

FIGURE 1.

COMPARISON OF 3-STEP ASN AND EQUIVALENT FIXED SAMPLE SIZE IN TESTS FOR A NORMAL MEAN

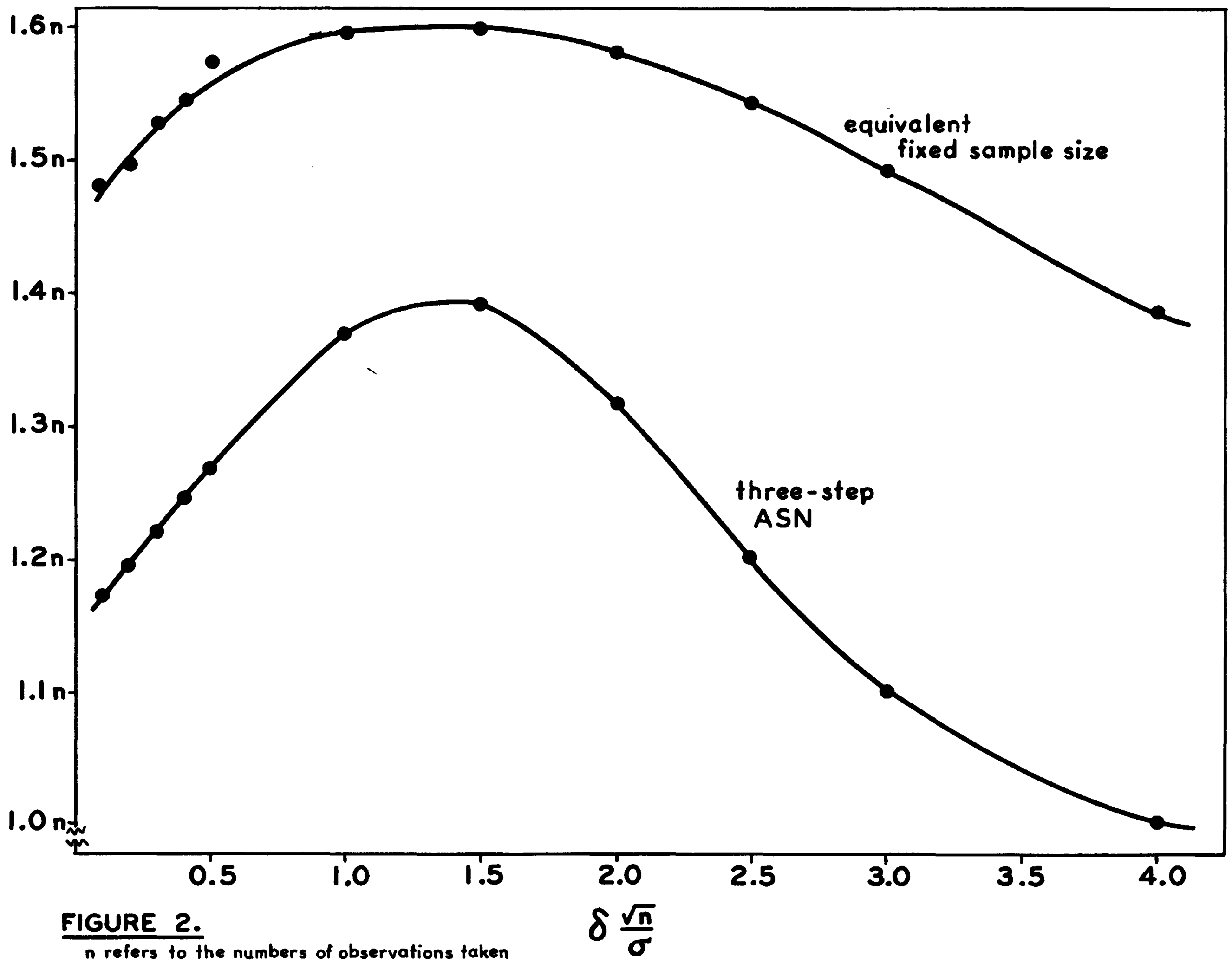


FIGURE 2.
n refers to the numbers of observations taken at each step in the 3-stage test.

ON THE DERIVATION OF OPTIMUM ALLOCATION FORMULAS IN STRATIFIED MULTI-STAGE SAMPLING BY THE USE OF THE CAUCHY INEQUALITY

J. C. Koop, North Carolina State College, Raleigh, North Carolina

INTRODUCTION

Without using the differential calculus, Neyman [1] and Sukhatme [2], in his well-known textbook on statistical sampling, have derived optimum allocation formulas for their sampling problems by purely algebraic arguments. Stuart [3] applied Cauchy's inequality to simplified variance and cost functions to obtain such formulas and illustrated the basic simplicity of the method.

This well-known inequality may be stated as follows. If a_i and b_i $(i = 1,2,....r)$ are real numbers then[1]

$$(a_1^2 + ... + a_r^2)(b_1^2 + ... + b_r^2) \geq (a_1 b_1 + ... + a_r b_r)^2 \qquad (1)$$

and equality is only attained if

$$a_1/b_1 = a_2/b_2 = ... = a_r/b_r . \qquad (2)$$

When equality in (1) is attained, under the conditions stated at (2), then the left hand side of (1) is necessarily a minimum.

In this expository paper the application of (1) to variance functions and certain cost functions appropriate to more involved sample designs in stratified multi-stage sampling, and some related, will be considered. By letting $\sum a^2$ play the role of the variance function, and $\sum b^2$ that of the cost function, it will appear that the set of conditions specified by (2) in respect of the unknown sample sizes for each stratum and for each step in sampling will lead to equations from which these unknowns can be solved for each of the situations when either the over-all cost of survey or the variance (for a character of interest) is fixed. These solutions give the optima.

The presentation of such a paper appears to be warranted as so far the simplicity of this method together with its ease of application for more ramified sample designs is still not so well known.

STRATIFIED SAMPLING IN TWO STAGES

We shall consider two cases here, namely (i) when the first- and second-stage units are sampled with equal probabilities and without replacement at each stage, (ii) when the first-stage units are sampled with unequal probabilities (e.g. probabilities proportional to certain measures of size) and with replacement and the second-stage units with equal probabilities and without replacement. In all problems the estimation of population totals will be considered. This will be sufficient to illustrate the method. For theoretical development of formulas reference is made to textbooks on statistical sampling.

Case (i). There are L strata and it is desired to estimate the total

$$T = \sum_{h=1}^{L} \sum_{i=1}^{N_h} \sum_{j=1}^{N_{hi}} x_{hij} . \qquad (3)$$

where x_{hij} is a variate of interest of the j^{th} second stage unit $(j = 1,2,...N_{hi})$ of the i^{th} first stage unit $(i = 1,2,...N_h)$ of the h^{th} stratum $(h = 1,2,...L)$. The best unbiased linear estimate based on n_h first-stage units $(h = 1,2,...L)$ and n_{hi} second-stage units $(i = 1,2,...n_h)$ is given by

$$T' = \sum_{h=1}^{L} \frac{N_h}{n_h} \sum_{i=1}^{n_h} \frac{N_{hi}}{n_{hi}} \sum_{j=1}^{n_{hi}} x_{hij} . \qquad (4)$$

The variance of T' is given by

$$V(T') = \sum_{h=1}^{L} N_h^2 \left[\frac{S_h^2}{n_h} (1 - \frac{n_L}{N_L}) + \frac{1}{N_h n_h} \sum_{i=1}^{N_h} N_{hi}^2 \frac{S_{hi}^2}{n_{hi}} (1 - \frac{n_{hi}}{N_{hi}}) \right] \qquad (5)$$

where in the h^{th} stratum S_h^2 is the variance of the totals for the first-stage units and S_{hi}^2 is the variance of the second-stage units in the i^{th} first-stage unit each with divisor $N_h - 1$ and $N_{hi} - 1$ respectively. The expected cost C appropriate to variance function (5) is given by

$$C = C_o + \sum_{h=1} (c_h n_h + \sum_{i=1}^{N_h} \frac{C_{hi}}{N_h} n_h n_{hi})$$

or

$$C - C_o = \sum_h c_h n_h + \sum_h \sum_i \frac{c_{hi}}{N_h} n_h n_{hi} \qquad (6)$$

where C_o is the fixed over-all cost, c_h is the cost per first-stage unit, and c_{hi} is the cost per second-stage unit in the i^{th} first-stage unit in stratum h. The above cost function (6) is sufficiently general. Others more involved do not lead to easy solutions.

To apply Cauchy's inequality, (5), after rearrangement of terms, is rewritten as

$$V(T') + \sum_{h=1}^{L} N_h S_h^2 = \sum_{h=1}^{L} \frac{N_h^2 Z_h^2}{n_h} + \sum_{h=1}^{L} \sum_{i=1}^{N_h} \frac{N_h N_{hi}^2 S_{hi}^2}{n_h n_{hi}} , \qquad (5.1)$$

where $Z_h^2 = S_h^2 - \sum_{i=1}^{N_h} N_{hi}S_{hi}^2/N_h$. It will be noted that the left hand sides of both (6) and (5.1) do not involve the unknowns explicitly. Let us assume that all Z_h^2 are positive. Now applying the inequality we have

$$(V(T') + \sum_h N_h S_h^2)(C-C_o)$$

$$\geq \left[\sum_h \frac{N_h Z_h}{\sqrt{n_h}} \cdot \sqrt{c_h n_h} + \sum_h\sum_i \frac{\sqrt{N_h} \cdot N_{hi}S_{hi}}{\sqrt{n_h n_{hi}}} \cdot \sqrt{\frac{c_{hi}}{N_h} n_h n_{hi}}\right]^2 \quad (7)$$

$$\text{i.e.} \geq \left[\sum_h (\sqrt{c_h}\ N_h Z_h + \sum_i \sqrt{c_{hi}}\ N_{hi}S_{hi})\right]^2 . \quad (7.1)$$

It should be noted that

$$N_h Z_h/\sqrt{n_h}\ ,\quad h = 1,2,..L$$

and

$$\sqrt{N_h}\ N_{hi}S_{hi}/\sqrt{n_h n_{hi}}\ ,\quad h=1,2,..L \text{ and } i=1,2,...N_h\ ,$$

play the role of the a's. Similarly $\sqrt{c_n n_h}$ and $\sqrt{c_{hi}n_h n_{hi}/N_h}$ for relevant h and (h,i) play the role of the b's. On the right hand side of (7), the value of the expression in square brackets does not involve the unknown sample sizes. This value is the minimum value and is attained when

$$\frac{\frac{N_h Z_h}{\sqrt{n_h}}}{\sqrt{c_h n_h}} = \frac{\frac{\sqrt{N_h}\ N_{hi}S_{hi}}{\sqrt{n_h n_{hi}}}}{\sqrt{\frac{c_h n_h n_{hi}}{N_h}}} = \lambda, \quad (8)$$

where λ is a constant, for all relevant h and (h,i) noted above. There are

$$L + \sum_{h=1}^{L} N_h \text{ such equations.}$$ From (8) we find the optimum values to be

$$n_{hi} = \sqrt{\frac{c_h}{c_{hi}}} \cdot \frac{N_{hi}S_{hi}}{Z_h}, \quad (9)$$

$$(h = 1,2,..L,\ i=1,2,...N_h)$$

and

$$n_h = \frac{1}{\lambda} \cdot \frac{N_h Z_h}{\sqrt{c_h}} \cdot (h = 1,2,..L). \quad (10)$$

From (9) we see that for a given stratum the n_{hi} are independent of the parameters relating to other strata. By fixing either the variance or the cost, the constant $1/\lambda$ can be determined from (5.1) or (6) as the case may be by substituting for n_{hi} and n_h given by equations (9) and (10). Let $V(T')$ be fixed at V_o, i.e. $V(T') = V_o$. Then from (5.1) it will be found that

$$\frac{1}{\lambda} = \sum_{h=1}^{L} \sqrt{c_h}\ (Z_h + \frac{1}{\sqrt{c_h}} \sum_{i=1}^{N_h} \sqrt{c_{hi}}\ N_{hi}S_{hi}) / (V_o + \sum_h N_h S_h^2) \quad (11)$$

and thus

$$n_h = \frac{N_h Z_h}{\sqrt{c_h}} \sum_h \sqrt{c_h}\ (Z_h + \frac{1}{\sqrt{c_h}} \sum_i \sqrt{c_{hi}}\ N_{hi}S_{hi}) / (V_o + \sum_h N_h S_h^2). \quad (12)$$

The conclusions which emerge from an inspection of the formulas for n_{hi} and n_h in (9) and (12) respectively (barring the exceptional situations for which Z_h^2 may be negative) are quite familiar and will not be repeated. When the over-all cost $C = C'$, then from (6) we will find

$$\frac{1}{\lambda} = (C'-C_o)/\sum_h (\sqrt{c_h}\ N_h Z_h + \sum_i \sqrt{c_{hi}}\ N_{hi}S_{hi}),$$

and thus

$$n_h = \frac{C'-C_o}{\sqrt{c_h}} \cdot \frac{N_h Z_h}{\sum_h (\sqrt{c_h}\ N_h Z_h + \sum_i \sqrt{c_{hi}}\ N_{hi}S_{hi})}. \quad (13)$$

Finally, from (7.1) the minimum expected variance or cost is obtained when either cost or variance is fixed respectively.

Case (ii). In this sampling plan the first-stage units are sampled with unequal probabilities and with replacement. Let p_{hi} be the probability of selection of the i^{th} first-stage unit in the h^{th} stratum where $\sum_{i=1}^{N_h} p_{hi} = 1$ for $h = 1,2,....L$. In this case a simple unbiased estimate T'' of T will be given by

$$T'' = \sum_{h=1}^{L} \frac{1}{n_h} \sum_{i=1}^{n_h} \frac{N_{hi}}{p_{hi}n_{hi}} \sum_{j=1}^{n_{hi}} x_{hij}\ , \quad (14)$$

with variance

$$V(T'') = \sum_{h=1}^{L}\left[\frac{\sigma_h^2}{n_h} + \frac{1}{n_h} \sum_{i=1}^{N_h} \frac{N_{hi}^2}{p_{hi}} \cdot \frac{S_{hi}^2}{n_{hi}}(1-\frac{n_{hi}}{N_{hi}})\right], \quad (15)$$

where

$$\sigma_h^2 = \sum_{i=1}^{N_h} p_{hi}(\frac{x_{hi}}{p_{hi}} - x_h)^2, \quad (h = 1,2,...L)$$

in which

$$\left.\begin{aligned} x_{hi} &= \sum_{j=1}^{N_{hi}} x_{hij}, \\ \text{and} \quad x_h &= \sum_{i=1}^{N_h} \sum_{j=1}^{N_{hi}} x_{hij} \end{aligned}\right\} (h = 1,2,...L)$$

The remaining notation is the same as for Case (i). The variance function (15) can be rewritten as

$$V(T'') = \sum_{h=1}^{L}\left[\frac{U_h^2}{n_h} + \sum_{i=1}^{N_h}\frac{N_{hi}^2 S_{hi}^2}{p_{hi} n_h n_{hi}}\right], \tag{16}$$

where $U_h^2 = \sigma_h^2 - \sum_{i=1}^{N_h}(N_{hi}S_{hi}^2)/p_{hi}$, to facilitate the application of the inequality. It will be assumed that all U_h^2 are all positive. An expected cost function C_1 appropriate to (16) is

$$C_1 = C_o + \sum_{h=1}^{L}(c_h n_h + n_h \sum_{i=1}^{N_h} p_{hi} c_{hi} n_{hi}). \tag{17}$$

Proceeding on the same lines as in Case (i) the optima n_{hi} are given by

$$n_{hi} = \sqrt{\frac{c_h}{c_{hi}}} \cdot \frac{N_{hi}S_{hi}}{p_{hi}U_i} \cdot (h=1,2,..L; i=1,2,..N_h) \tag{18}$$

For fixed cost C_1' or fixed variance V_o the optima n_h are given respectively by

$$n_h = \frac{C_1'-C_o}{\sqrt{c_h}} \cdot \frac{U_h}{\sum_h(\sqrt{c_h}\, U_h + \sum_i \sqrt{c_{hi}}\; N_{hi}S_{hi})}$$

and

$$n_h = \frac{1}{\sqrt{c_h}} \cdot \frac{U_h}{V_o} \sum_h(\sqrt{c_h}\, U_h + \sum_i \sqrt{c_{hi}}\; N_{hi}S_{hi}) \quad . \ (h = 1,2,..L) \tag{19}$$

SAMPLING IN THREE OR MORE STAGES

The extension of the problem considered under Case (i) to three or more stages will be given. Let x_{hijk} be the variate of interest of the k^{th} third-stage unit ($k = 1,2,...N_{hij}$). The remaining notation is as before. Let n_{hij} third-stage units be sampled without replacement and with equal probabilities in the j^{th} second-stage unit of the i^{th} first-stage unit of the h^{th} stratum. The total T will be given by

$$T = \sum_h \sum_i \sum_j \sum_k x_{hijk}, \tag{20}$$

the best unbiased linear estimate of which will be given by

$$T' = \sum_h \frac{N_h}{n_h} \sum_i \frac{N_{hi}}{n_{hi}} \sum_j \frac{N_{hij}}{n_{hij}} \sum_k x_{hijk} \tag{21}$$

with variance

$$V(T') = \sum_{h=1}^{L} N_h^2 \left[\frac{S_h^2}{n_h}(1 - \frac{n_h}{N_h}) + \frac{1}{N_h n_h}\sum_{i=1}^{N_h} N_{hi}^2 \cdot \left\{\frac{S_{hi}^2}{n_{hi}}(1 - \frac{n_{hi}}{N_{hi}}) + \frac{1}{N_{hi}n_{hi}}\sum_{j=1}^{N_{hi}} N_{hij}^2 \cdot \frac{S_{hij}^2}{n_{hij}}(1 - \frac{n_{hij}}{N_{hij}})\right\}\right] \tag{22}$$

The variance function (22) can be rewritten as

$$V(T') + \sum_{h=1}^{L} N_h S_h^2 = \sum_h \frac{N_h^2 Z_h^2}{n_h} + \sum_h \sum_i \frac{N_h N_{hi}^2 Z_{hi}^2}{n_h n_{hi}} + \sum_h \sum_i \sum_j \frac{N_h N_{hi} N_{hij}^2 S_{hij}^2}{n_h n_{hi} n_{hij}}, \tag{22.1}$$

where

$$Z_h^2 = S_h^2 - \frac{1}{N_h}\sum_i N_{hi}S_{hi}^2, \qquad Z_{hi}^2 = S_{hi}^2 - \frac{1}{N_{hi}}\sum_j N_{hij}S_{hij}^2. \tag{22.2}$$

The Z_h^2 of (22.1) is not the same as that in (5.1) and this notation is maintained only for the sake of consistency. It will be assumed that Z_h^2 and Z_{hi}^2 are all positive. An expected cost function C_2 appropriate to (22) is

$$C_2 = C_o + \sum_{h=1}^{L} n_h c_h + \sum_{h=1}^{L}\frac{n_h}{N_h}\sum_{i=1}^{N_h} c_{hi}n_{hi} + \sum_{h=1}^{L}\frac{n_h}{N_h}\sum_{i=1}^{N_h}\frac{n_{hi}}{N_{hi}}\sum_{j=1}^{N_{hi}} n_{hij}c_{hij}, \tag{23}$$

where in stratum h, c_{hij} is the cost per third-stage unit in the j^{th} second-stage unit of the i^{th} first-stage unit.

The values of n_{hij}, n_{hi} and n_h which minimize cost or variance can be found from the following set of equations analogous to (2), viz.

$$\frac{N_h Z_h}{n_h\sqrt{c_h}} = \frac{N_h N_{hi} Z_{hi}}{n_h n_{hi}\sqrt{c_{hi}}} = \frac{N_h N_{hi} N_{hij} S_{hij}}{n_h n_{hi} n_{hij}\sqrt{c_{hij}}} = \lambda \text{ a constant.} \tag{24}$$

$$(h = 1,2,...L;\ i = 1,2,..N_h;\ j = 1,2,..N_{hi})$$

Thus we find

$$n_h = \frac{1}{\lambda} \cdot \frac{N_h Z_h}{\sqrt{c_h}} \tag{25}$$

$$n_{hi} = \sqrt{\frac{c_h}{c_{hi}}} \cdot \frac{N_{hi}Z_{hi}}{Z_h} \tag{26}$$

$$n_{hij} = \sqrt{\frac{c_{hi}}{c_{hij}}} \cdot \frac{N_{hij}S_{hij}}{Z_{hi}} \tag{27}$$

As demonstrated in the foregoing discussion, $\frac{1}{\lambda}$ can be determined either for the condition of fixed cost or fixed variance from (23) or (22.1), as the case may be, by substituting for n_h, n_{hi} and n_{hij} given at (25), (26) and (27) respectively.

In the multi-stage case when sampling is with equal probabilities and without replacement all the way, and with the same type of expected cost functions and linear estimator the equations for optima will be as in (24), only more extended, e.g.

$$\frac{N_h Z_h}{n_h\sqrt{c_h}} = \frac{N_h N_{hi} Z_{hi}}{n_h n_{hi}\sqrt{c_{hi}}} = \dots = \frac{N_h N_{hi} N_{hij} N_{hijk} Z_{hijk}}{n_h n_{hi} n_{hij} n_{hijk}\sqrt{c_{hijk}}} = \dots\dots = \lambda \qquad (28)$$

The Z-functions can be identified from the variance function and the first two will be those given at (22.2). At the last stage of course, as can be seen from (22.1), only the variance internal to the unit will be involved so that the last expression in (28) will have a symbol of the type S_{hijklm} representing the square root of this internal variance for the unit in question.

Further for other sampling plans (e.g. sampling with unequal probabilities and with replacement at one or more intermediate stages) and with similar type of expected cost functions and linear estimators such as those considered above, formulas similar to (25), (26) and (27) will be obtained.

When non-response is taken into account, the optima for first-stage, second-stage sample sizes, etc. and the resampling rate or successive resampling rates for those units at the last stage (suffering from non-response) can be determined by exactly the same method. The structure of such sample designs may be described as multi-stage multi-phase. Unfortunately, the formulas are too long and unwieldy for presentation in an expository paper of this kind. These problems have also been considered recently by Mr. G. T. Foradori [4] in an unpublished work.

OPTIMUM ALLOCATION FOR SEVERAL ITEMS

This problem can be solved by the classical method of Lagrange but the equations obtained do not lead to algebraic solutions and can only be solved by iteration. Yates [5] has considered the problem for two variates. For more than two variates the computational work can become very tedious.

In one alternative approach, which is appealing because of its simplicity, the weighted sum of the relative variances of the variates of interest (corresponding to the items) is minimized for a fixed expected cost or the expected cost is minimized fixing the said weighted sum. The weights of course sum to one. The items are weighted according to their importance. The relative variance rather than the variance is considered because by so doing each component measure of precision is reduced to the same dimension and thus their linear compound is meaningful.

We shall investigate this problem in the light of the latter statement considering only stratified one-stage sampling with equal probabilities and without replacement. As will be seen subsequently, the problem for more than one stage of sampling, and with more involved probability systems, can be solved along the same lines.

Consider the estimation of the totals for r items (e.g. in a demographic survey, numbers employed, numbers unemployed, etc. in a region or country) using the best linear unbiased estimator for each.

The weighted sum of the relative variances will be

$$V = \sum_{t=1}^{r} w_t \frac{1}{X_t^2} \sum_{h=1}^{L} N_h^2 \frac{{}_tS_h^2}{n_h}\left(1 - \frac{n_h}{N_h}\right), \qquad (29)$$

where w_t is the weight assigned to the t^{th} item whose variance in the h^{th} stratum is ${}_tS_h^2$ and X_t is the total for the t^{th} item in the entire universe. The cost of the survey is given by

$$C = C_o + \sum_{h=1}^{L} c_h n_h \qquad (30)$$

The sumbols for which no explanation has been given here bear the same meaning as under Case (i), but restricted to one-stage sampling. (29) can be rewritten as

$$V' = V + \sum_{t=1}^{r} \frac{w_t}{X_t^2} \sum_{h=1}^{L} N_h\, {}_tS_h^2 = \sum_{h=1}^{L} \frac{N_h^2}{n_h} \sum_{t=1}^{r} w_t \frac{{}_tS_h^2}{X_t^2} \qquad (29.1)$$

We find

$$(C-C_o)V' \geq \left(\sum_{h=1}^{L} \frac{N_h}{\sqrt{n_h}}\sqrt{\sum_t w_t \frac{{}_tS_h^2}{X_t^2}} \cdot \sqrt{c_h n_h}\right)^2, \qquad (31)$$

and the minimum is attained when

$$\frac{\frac{N_h}{\sqrt{n_h}}\sqrt{\sum_t w_t \frac{{}_tS_h^2}{X_t^2}}}{\sqrt{c_h n_h}} = \lambda \text{ a constant.} \qquad (h = 1,2,\dots L) \qquad (32)$$

From (32) we obtain

$$n_h = \frac{1}{\lambda} \cdot \frac{N_h}{\sqrt{c_h}} \sqrt{\sum_{t=1}^{r} w_t \frac{{}_tS_h^2}{X_t^2}}, (h=1,2,\dots L) \qquad (33)$$

Fixing the over-all relative variance in (29.1) at v_1 and eliminating λ between the resulting equation and (33) we obtain

$$n_h = \frac{N_h}{\sqrt{c_h}} \sqrt{\sum_t w_t \frac{{}_tS_h^2}{X_t^2}} \left(\sum_{h=1}^{L} c_h N_h \sqrt{\sum_t w_t \frac{{}_tS_h^2}{X_t^2}}\right) \Big/ \left(v_1 + \sum_t \frac{w_t}{X_t^2} \sum_h N_h\, {}_tS_h^2\right) \qquad (33.1)$$

Similarly when the total cost of the survey is fixed at c_1 we obtain

$$n_h = \frac{c_1 - C_o}{\sqrt{c_h}} \; N_h \sqrt{\sum_t w_t \frac{S^2_{th}}{X^2_t}} \Big/ \sum_h \sqrt{c_h} N_h \sqrt{\sum_t w_t \frac{S^2_{th}}{X^2_t}} \, . \tag{33.2}$$

In the light of the above discussion the interpretations to be made for the optima n_h given by (33.1) and (33.2) are readily apparent.

COMMENTS

Firstly, a question arises as to what should be done when one or more values such as the $Z^2_h = S^2_h - (\sum_i N_{hi} S^2_{hi}) / N_h$ are negative. There is no guarantee that they will always be positive since the underlying quadratic form in the variates concerned is not positive definite. Thus the square roots of the values concerned are no longer real but imaginary numbers. In any situation where the various parts of the variance function exhibit this refractory tendency the general method given above no longer holds, but, the basic procedure on which it rests, that is multiplying the cost and the variance function together, and determining the sample sizes which minimize the product, still holds. A discussion on this question is found in Sukhatme's book [2] on pp. 312-314, and will not be repeated here.

Secondly, the solutions for sample sizes presented here are formal. In practice integer values or such nearest values as are convenient to the survey organization will be used. Also they may usually be based on poor estimates of the unknown values involved. It can be demonstrated that "moderate" deviations from the optima do not increase the desired variance or alter the expected cost to an appreciable extent.

Finally, when sampling with unequal probabilities, as in Case (ii) above, the number of first-stage units may be fixed because variation in respect of the first stage is already controlled by the device. In such circumstances, the optima may also be determined as above for the remaining stage or stages. The only change is that those parts of the variance function and its corresponding cost function for example on the left hand sides of (16) and (17) respectively will be augmented by

$$-\sum_{h=1}^{L} \frac{U^2_h}{a} \quad \text{and} \quad -\sum_{h=1}^{L} c_h a$$

respectively, where a is the number of first-stage units per stratum selected on the basis of other considerations. The reasons behind such considerations may be of real importance in the conduct of sample surveys.

REFERENCES

[1] Neyman, J.. "Contribution to the Theory of Sampling Human Populations." J. Amer. Statist. Ass., 33 (1938) 101-116.

[2] Sukhatme, P. V. Sampling Theory of Surveys with Applications. New Delhi, India. The Indian Society of Agricultural Statistics. Ames, Iowa. The Iowa State College Press, 1953.

[3] Stuart, A. "A Simple Presentation of Optimum Sampling Results." J. Roy. Statist. Soc., B16 (1954), 239-241.

[4] Foradori, G. T. "Some Non-Response Sampling Theory for Two Stage Designs." Unpublished so far (1961).

[5] Yates, F. Sampling Methods for Censuses and Surveys. New York, Hafner Publishing Company, (1953, 2nd edition), 338-342.