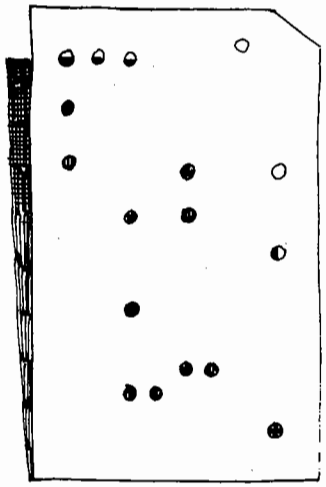

How To Count Better:

Using Statistics to Improve the Census

Morris H. Hansen *Westat, Inc.*
Barbara A. Bailar *Bureau of the Census**

THE IMPORTANCE OF CENSUS RESULTS

In 1790, Thomas Jefferson gave to George Washington, our first President, the results of the first census of the United States of America. Every 10 years since then, as provided in the Constitution, the decennial census has determined for the nation essential information about its people.

The basic constitutional purpose of the census is, of course, to apportion the membership of the House of Representatives among the states. From the beginning, the census has had many important purposes beyond the constitutional one. The development of legislative programs to improve health and education, alleviate poverty, augment transportation, and so on, is guided by census results. They are used also for program planning, execution, and evaluation.

*Now with the American Statistical Association.

Now the distribution of billions of dollars a year from the Federal Government to the states, and from the states to local units, is based squarely on census counts. Private business uses the census for such purposes as plant location and marketing. Much social and economic research would be essentially impossible without census information.

The importance of the census to current problems such as poverty, health, education, civil rights, and others brought about requests for shifting to a 5-year, rather than a 10-year, census in order to keep information more nearly up to date. Those requests came from governors of states, mayors of large cities, scientists, business people, and many others. In 1976, legislation was passed that called for a 5-year census, but Congress has never appropriated the funds to carry it out.

THE JOB OF PLANNING AND TAKING THE CENSUS

The need for a fast and reasonably accurate census is fairly obvious. What may not be so apparent are the major problems involved in taking a census and making the results available promptly, and in adapting the census questions to serve current needs. In each census, some questions have been changed in response to new needs, but certain basic information has been consistently required. Questions in the 1980 population census included age, sex, race, marital status, *household* relationship, education, school enrollment, employment and unemployment, occupation and industry, migration, *travel time to work, persons in carpool,* country of origin, *current language and English proficiency, ancestry,* income, and other subjects. (Items in italics were new in the 1980 census.)

The job of organizing and taking the census is a major administrative and technological undertaking. Even though most of the questionnaires are filled out by the respondents themselves in a "census by mail" and modern computers and other advanced technology have eliminated a lot of paperwork, the taking of a census requires the recruiting and training of hundreds of thousands of people, most of them for only a few weeks of work. The 1980 census, for example, recruited some 458,000 persons, with 270,000 persons active during the peak period of operations. Once specific goals are set in terms of questionnaire content and desired statistical results, the massive job of organization and administration begins.

The system for canvassing and for collecting, receiving, processing, and summarizing the vast numbers of completed questionnaires must be planned. Specialized electronic and paper-handling machines designed and built at the Census Bureau automatically read the information recorded by respondents or enumerators. These complex machines first photograph on microfilm and then scan and read microfilm copies of the census forms that have been filled out, in most cases, by the respondents themselves. The magnitude of the job is difficult to comprehend. For the 1980 census, a little more than a quarter of a billion pages (counting each side of a relatively large sheet as a page) were

handled. After being clerically reviewed, the results were recorded on magnetic tape, and then computers examined these results. In the process, the forms were edited for certain types of incompleteness or inconsistency, and adjustments were made automatically or special problem areas were identified for further manual investigation. The later steps of tabulation and printing for publication also were accomplished on electronic computers.

The approach to the job of taking a census differs from that of designing a totally new system, in that the census has been taken many times before, and the background and experience of the past serve to guide current efforts.

The availability of extensive past experience, however, has disadvantages. There may be long traditions that have come to be regarded as essential but that, in fact, only represent ways in which the job has been done in the past. For example, the tradition of taking the census by an enumerator canvassing an area and personally asking the questions of any responsible member of the household had been long established. This approach was regarded as proven by long use to be the only reasonable way to take the census. Because the concepts in some census questions are difficult, it was thought that only a trained enumerator could ask the questions and elicit the proper information. But this view did not recognize the difficulties in training and controlling an army of temporary interviewers. Nor did it recognize that the responses obtained in the census interview situations were sometimes based on a misunderstanding of the questions and that the interview process did not allow time for a considered reply. Furthermore, the interviewers' conceptions or misconceptions could importantly influence the response.

In the nineteenth century, many potential advantages of prior experience were lost, for the Census Bureau was not created as a permanent and continuing agency until 1902. It then became far more feasible, with a continuing staff, to benefit from lessons and experience of prior censuses. The situation for the 1900 and earlier censuses is illustrated in the following quotation from *The History of Statistics* (Cummings, 1918, pp. 678–679).

> Mr. Porter gives the following account of his experience, which must have been essentially that of every Superintendent of the Census.
>
> The Superintendent in both the last two censuses [1880 and 1890] was appointed in April of the year preceding the enumeration, but when I was appointed I had nothing but one clerk and a messenger, and a desk with some white paper on it. I sent over to the Patent Office building to find out all I could get of the remnants of ten years ago, and we got some old books and schedules and such things as we could dig out. . . . I was not able to get more than three of the old men from the city. . . . I knew most of the old census people. Some of them were dead and some in private business. . . . But little over sixty days were allowed for the printing of 20,000,000 schedules and their distribution, accompanied by printed instructions to the 50,000 enumerators all over the country, many of them remote from railroads or telegraph lines. . . . Now to guide us in getting up these blanks, we had only a few scrapbooks that someone had had the forethought to use in saving some of the forms of blanks in the last census. He had taken them home, a few copies at a time, and put them into scrapbooks. The government had taken no care of these things in 1885, when the office was closed up. Some of them had been sold for waste paper, others had been burned, and others lost.

In addition to showing the potential gains from continuity and learning from the past, this quotation suggests the great complexity of the census in the latter half of the nineteenth century. At that time there was little in the way of a continuing statistical program in the Federal Government and as a consequence the decennial census was loaded with a range of questions that proved difficult if not impossible to collect through decennial census inquiries—hence the great number and variety of questions and forms. Many of these types of information are now collected through sample surveys or compilations from administrative records.

THE USE OF STATISTICS IN PLANNING AND TAKING THE CENSUS

Statistical concepts and methods have provided fundamental improvements to the census over the past 50 years. We might say that these improvements form a technological explosion. Part of this explosive change has come from the introduction of the large-scale electronic computer, but even more of the change has come from statistical advances and the application of statistical studies. (In fact, development of the high-speed computer and modern statistical methods were both substantially motivated by census problems and were in part carried out by census scientists or with Census Bureau support.)

The Introduction of Sampling as a Tool for Census Taking

Sampling was first used in collecting census information in the 1940 census. A series of questions was added for a 5% sample of the population. [Roughly speaking, this meant asking every twentieth person the additional questions (Waksberg and Pearl, 1965).] This was a major advance, as the tradition for a century and a half had been universal coverage for every question. The questions asked of the 1940 sample included one on wage and salary income—income had not previously been a census question—one on usual occupation (as distinguished from current occupation; the 1940 census was planned during the Depression, when unemployment was very high and there was frequently a difference between a person's usual occupation and the occupation at which he or she was currently working), and several other questions.

The art and science of modern statistical sampling were evolving at the time of the 1940 census, and at the same time there was increasing public acceptance of sampling. It was possible to proceed with greater knowledge and confidence about what a sample would produce than would earlier have been possible. Thus, for a particular size and design of sample, statisticians could establish a reasonable range for the difference between a sample result and that obtained from a complete census. Suppose, for example, that a city had a total population

of 100,000 persons of whom 30,000 were employed and earned wages or salaries, and suppose that 10,000 of these received wages and salaries of less than $2,000 in 1939 (the year preceding the census). There would be approximately 5,000 people in a 5% sample from that city. The estimate of the number receiving less than the $2,000 wage and salary income would, with a very high probability, when estimated from the 5% sample, lie within the range 9,300 to 10,700. This kind of accuracy was sufficient to serve many important purposes and, in fact, was as much accuracy as could be justified in the light of the less-than-perfect accuracy of the responses to the question on income. Not only could estimates be prepared from the sample of what would have been shown by a complete census, but, in addition, the range of probable difference between the sample estimate and the result of a complete census could be estimated from the sample! Sampling theory also guides in designing samples to achieve a maximum precision of results per unit of cost.

In considering the advantages of the use of sampling, it may appear to some that the main work involved in taking a census is the time it takes in going from door to door and that, once some questions have been asked at a household, the cost of adding questions would be small whether they were added for a 5% sample or for the total population. Such a presumption is far from true. Suppose, for example, that an additional question about whether a person has a chronic illness adds an average of 20 seconds of work for each person counted in the census. With some 200 million people in the population, this would add more than 1 million hours of work and perhaps $4 million to the cost of the census. Thus obtaining the added information for, say, five questions from only a 5% sample instead of from all persons can produce needed and highly useful results at a fraction of the cost for complete coverage. Finally, the use of a sample permits tabulation and analysis much sooner than complete coverage does.

Starting in the 1940s sample surveys on a wide range of subjects were introduced so that continuing and up-to-date information would be available between the censuses. For example, the Current Population Survey is a sample of the population conducted monthly by the Bureau of the Census. The Survey collects information each month on employment, unemployment, and other labor force characteristics and activities of the population. It also serves to collect information on other subjects, with different supplemental questions in various months. In one month each year almost the full range of population census questions is asked. In other months information may be requested on recreational activity, housing, disability, or other subjects. (See the essay by Leon and Rones for further discussion of the Current Population Survey.) Sample surveys cover health, retail trade, business and personal services, the activities of governmental units, and so on. These surveys have large enough samples to provide national information, some information for regions of the nation, and even information for large states and metropolitan areas. They cannot, however, provide information for the many individual cities and counties and for relatively small communities within the cities and counties. To obtain that kind of fine detail, very large samples are needed, samples such as those taken as part of the decennial census.

In the 1950 census, the use of sampling was extended to some questions that in earlier censuses had been collected from all persons. For these questions, a 20% sample was used. This sampling was extended in the 1960 census to most of the items of information. In the 1960, 1970, and 1980 population censuses, only the basic listing of the population, with questions on age, sex, race, marital status, and family relationship, was done on a 100% basis.

The following question is often raised: If sampling is so effective a tool, why not take the whole census with a sample? Isn't it a waste of effort to do a complete census? One must remember that a primary purpose of the census is not to obtain national information but to provide information for individual states, cities, and counties, and for small areas within these. The results obtained by converting the whole census to a relatively large sample (perhaps including 20% of the population) would be adequate for some purposes. Such a sample, however, would not apportion representatives in the state legislatures in the same way as a complete census. Similarly, there could be important differences in the distribution of vast amounts of funds to thousands of individual small areas. Also, in some states, the legal status of many communities depends on the exact size of the official population count; for example, a city with 10,000 or more people can issue bonds. Hence a complete census is needed for total population counts and for some other basic population data, but the great bulk of the information may be collected from a relatively large sample.

Income—An Example of Sampling in the Census

When the 1940 census was planned, after a long depression, questions on wage and salary income were put on the census questionnaire to guide the nation on government and private sector programs. The income question was asked of everyone in the census, and a storm of protest was raised. There was some congressional opposition, and a New England senator led a campaign to persuade the public not to respond. Issues of privacy and confidentiality were given much play by the press. Congressional opposition and newspaper editorials focused on these issues of invasion of privacy and increased government intervention. There was support from the administration to ask the question. To allay fears of lack of privacy, the Census Bureau printed 20 million forms that individuals could fill out and mail directly to Washington, so that no enumerator could see the response. Only 200,000 such forms were ever used, and the nonresponse rate was less than 2%.

Since 1940, however, income has been asked routinely on census questionnaires for a sample of the population. In 1950 the expanded income questions were asked only of a 25% sample of the population. The practice of sampling for income questions has continued but it caused a problem after the 1970 census. Revenue sharing began in 1972, allocating billions of dollars of Federal funds to state and local governments, based on formulas that included local income estimates. For small areas, the 20% sample used in 1970 gave too much variability in sampling error between smaller and larger places, sampled at the same rate. Thus, in 1980, two sampling rates were used. A one-in-two rate was used for counties, incorporated places, and minor civil divisions with a population

of 2,500 or less. The rest of the country was sampled at a one-in-six rate. This differential sampling satisfied the data needs for small areas.

The uses of income data are so sweeping that it is now seen as one of the most important census items. There are over 20 Federal programs where the use of census data on income is legislated.

The Use of Sampling and Experimental Studies to Evaluate and Improve Census Methods and Results

Substantial steps to evaluate and improve census methods begun in the 1940s have been continued and greatly extended since then. Statistical studies have been made of various aspects of the census. One such study was made by repeating the census enumeration, in a well-designed sample of areas, shortly after the original census enumeration, using the same procedures as in the initial census. Thus we were able to see something of how much alike two censuses taken under the same conditions and procedures would be.

Studies of these types show high consistency and accuracy of response for questions such as sex, age, race, and place of birth, but they show higher degrees of inconsistency and inaccuracy in responses to the more difficult questions relating to occupation, unemployment, income, education, and others. The information from such studies helps both in improving the questions in the next census (by, for example, showing which questions cause trouble and need rewording) and in interpreting the accuracy of census results when the questions are put to specific use.

Studies that compare alternative methods and procedures within the framework of well-designed and randomized experiments have been exceedingly important for learning about the effectiveness of various procedures and in comparing their cost and their accuracies. Such comparisons have been made, for example, of various types of questionnaire design and other variations in procedures. For example, one study reexamined the census coverage and questions in a sample of areas, but used more highly trained enumerators, more detailed sets of questions, and other such expensive improvements.

These rather wide-ranging studies led generally to the conclusion that some of the methods earlier regarded as the way to achieve major improvements would not be effective in relation to their cost (although some worthwhile improvements in questionnaire design and procedures were accomplished). We find, for example, that simply spending much more time and money on training an army of temporary enumerators would add to cost but probably not lead to corresponding improvements in accuracy. A problem in the 1970 and 1980 censuses has been the turnover of enumerators. Many are trained but some quit before starting work and others do the easy cases and then quit. Supplemental pay experiments in the test censuses leading to 1990 have been encouraging. Employees who meet standards qualify for supplemental payments. In the test censuses, the increase in production and the resulting savings in enumerator wages have more than compensated for the cost of the supplemental payments. Continued experimentation is planned before the final decision on 1990 pay plans.

The Surprising Effect of Enumerators

One study, however, led to surprising conclusions and then to a basic improvement in census procedures. It had long been known that enumerators can and do influence the answers they obtain—presumably unconsciously most of the time. But the magnitude of this enumerator effect was not known. Hence a large statistical study was carried out as part of the 1950 census to measure the magnitude of enumerator effect.

The study plan (in a simplified description) was based on areas divided into 16 work assignments (areas small enough so that 2 work assignments could be canvassed by a single enumerator) and 8 enumerators. Two of the 16 work assignments, chosen at random, were given to the first enumerator; two of the remaining ones, also chosen at random, were given to the second enumerator, and so on. Of course, the whole 16-fold experiment was repeated many times.

The random choice of work assignments was essential here in order to interpret the results in a useful way. (Random choice means choice by a method equivalent to writing the numbers 1, 2, . . . , 16 on identical cards, shuffling or mixing them thoroughly, and then picking first one, then another, and so on.)

Another essential feature of the plan was that each enumerator had 2 work assignments and that there were a number of enumerators. That way, good estimates could be obtained for the variability introduced by a single enumerator (roughly, the differences in performance by the same enumerator that were not attributable to differences in areas) as compared with the variability stemming from differences among the enumerators.

Further details of this path-breaking analysis cannot be given here, but we can summarize the results. Far greater differences between enumerators were found than had been anticipated, not so much on items such as age and place of birth, but on the more difficult items such as occupation, employment status, income, and education. For those items, in fact, a complete census would have as much variability in its results (because of enumerator effect) as would a 25% sample if there were no enumerator effect!

What should we do about this? One approach might be to expend far more resources on the selection, training, and supervision of enumerators. But the other studies mentioned above have shown that this is not feasible under the conditions of a national census.

Another possible answer is to eliminate the need for the enumerator by leaving the carefully designed questionnaires with the respondents and asking them to fill them out and mail them. (Enumerators would be needed only when respondents' returns were incomplete or where the respondents asked for help.) This method was tested in further studies and found to work quite well. It was used in the 1960 census for the longer census forms and was a great success, in terms of both cost and added accuracy. Hence, in 1970 and 1980, this method was used still further: most of the population received and returned the census forms by mail.

Thus, in the process of using statistics to improve the census, the completeness of coverage has been improved, and the accuracy of the items of information collected has been increased.

COVERAGE ERROR IN THE CENSUS

Since 1950, the Census Bureau has been measuring the coverage it attains in the census. The rates are very high, between 98% and 99%. However, the rates vary over population subgroups. About 5% of the Black population is missed, with rates approaching 20% for Black males in certain age groups. Hispanics are also counted at lower rates.

After the 1980 census, the Census Bureau was sued by several localities. Because of the concentration of the undercount in minority groups, cities with large numbers of minorities were convinced that they had been undercounted. They asked for a statistical adjustment of the census results. In the court cases, the Census Bureau argued that though statistical methods existed that were accurate enough to measure the undercount for evaluation purposes, the methods were not accurate enough to adjust the census counts. Other statisticians, representing the jurisdictions affected, said the population distributions would be improved by adjustments. One of the lawsuits—that filed by the state and city of New York in 1980—was decided after seven years. On December 8, 1987, Federal Judge John Sprizzo issued an opinion in favor of the government. At the time this essay was written, however, many other suits remain to be decided.

The Census Bureau instituted an intense, focused research program on methods of measuring the undercount and distributing it to all levels of geography. Several demonstrations have taken place in test censuses. However, on October 30, 1987, the Department of Commerce announced that it had decided against making an adjustment for any undercount for 1990.

PROBLEMS

1. Give several reasons why a census is taken.
2. Why is it desirable in some instances to take a 5% sample as opposed to a complete census? What does the 5% sample lose?
3. Comment on the advantages and disadvantages of using enumerators in a census.
4. Answers to questions regarding sex, age, race, and place of birth seem to be more reliable than those regarding occupation, unemployment, income, and education. Give two possible reasons for this. Do you think answers to mailed questionnaires would be more accurate than interviews by enumerators on these questions? Why or why not?
5. (For group work.) Take a small random sample of people and administer a questionnaire on some issue, for example opinions of the local newspaper. Report on the problems and situations that arise.
6. Suppose the Census Bureau decided to add a battery of six questions on energy consumption to the 1990 census. About how much money would the Census Bureau save if, instead of asking everyone the questions, they asked them of only a 1% sample?
7. Suppose that a census uses mailed questionnaires instead of enumerators, and 40% of the returned questionnaires have no answer marked to a question on mental health. Would the Census Bureau be justified in counting only the marked responses and publishing the result as a 60% sample? Justify your answer.
8. A scientist in the Census Bureau suggests using telephone interviews instead of mailed questionnaires for families with telephones, thinking that this technique might lead to a more accurate census.
   a. What arguments can you think of which support or refute this idea?
   b. Design an experiment to determine whether this idea is correct.
   c. What other considerations (beside accuracy) would determine whether or not the telephone technique would be adopted?

REFERENCES

Barbara A. Bailar. 1983. "Counting or Estimation in a Census—A Difficult Decision." *1983 Proceedings of the Social Statistics Section of the American Statistical Association.* Washington, D.C.: American Statistical Association, pp. 42–49.

Barbara A. Bailar. 1985. "Comment." *Journal of the American Statistical Association* 180: 109–114.

John Cummings. 1918. "Statistical Work of the Federal Government of the United States." *The History of Statistics,* John Koren, ed. New York: Macmillan, pp. 678–679.

Morris H. Hansen. 1987. "Some History and Reminiscences on Survey Sampling." *Statistical Science* 2(2): 180–190.

Morris H. Hansen, Leon Pritzker, and Joseph Steinberg. 1959. "The Evaluation and Research Program of the 1960 Censuses." *1959 Proceedings of the Social Statistics Section, American Statistical Association.* Washington, D.C.: American Statistical Association.

Morris H. Hansen and Benjamin J. Tepping. 1969. "Progress and Problems in Survey Methods and Theory Illustrated by the Work of the United States Bureau of the Census." In *New Developments in Survey Sampling,* Norman L. Johnson and Harry Smith, Jr., eds. New York: Wiley (Interscience), pp. 3E(1)–(11).

Ingram Olkin. 1987. "A Conversation with Morris Hansen." *Statistical Science* 2(2): 162–179.

Ann Herbert Scott. 1968. *Census, U.S.A.: Fact Finding for the American People 1790–1970.* New York: Seabury.

Joseph Waksberg and Robert B. Pearl. 1965. "New Methodological Research on Labor Force Measurements." *1965 Proceedings of the Social Statistics Section, American Statistical Association.* Washington, D.C.: American Statistical Association, pp. 227–237.