DISCUSSION

Myron Straf, National Academy of Sciences

I enjoyed this set of very interesting papers and commend them to your reading. From them you will appreciate how federal government agencies, in seeking to describe and analyze social, economic, and biomedical problems, contributed to advances in statistical methods. I would also refer you to other sources of information on the history of the development of statistical methodology in the federal government, in particular the excellent monograph, Revolution in United States Government Statistics: 1926-1976, by Joseph Duncan and William Shelton (1978), Ingram Olkin's conversation with Morris Hansen as reported in Statistical Science (Olkin, 1987), Morris Hansen's reminiscences on survey sampling (Hansen, 1987), and the manuscript on the history of survey methods and data collection technology by Stephen Fienberg and Judith Tanur (1989) that appears elsewhere in this volume. Fienberg and Tanur refer to other historical perspectives, including Jean Converse's history of survey research (Converse, 1987), Steve Stigler's history of statistics (Stigler, 1986), and others.

BAILAR'S PAPER

Barbara Bailar has talked about three major statistical innovations from the Census Bureau: sampling, seasonal adjustment, and disclosure avoidance. By far the introduction of sampling was the most revolutionary innovation. I will later present my thesis for how it was that sampling took hold of the Census Bureau.

Bailar sketches some of the theory that was developed at the Bureau: multistage sampling, sampling with probability proportionate to a measure of size, composite estimation, and nonsampling errors. She credits much work to Morris Hansen at the Census Bureau. We should note, however, that Hansen also made major contributions to statistical theory while he was working on problems for the Bureau of Labor Statistics, including sampling from highly skewed distributions of business establishments and how to select samples of establishments and items to price within establishments for the Consumer Price Index.

Looking to the future, Bailar sees more work with models, for time series, for small-area data, and for missing observations, among other topics. I would add that the future will see further development and greater use of computer-intensive techniques, such as bootstrap estimation; obtaining statistics from administrative records; and matching and other linking of data files. There are also important developments to come in the field of cognitive aspects of survey methodology. The National Center for Health Statistics, the Bureau of Labor Statistics, and the Census Bureau have each established laboratories devoted to study in this field and have a formal cooperative agreement.

Bailar's focus on the Census Bureau credits much to it and deservedly so. Hansen indicated that his work there was made possible through team efforts. His work, for example, on redesigning the unemployment survey based on the new ideas of multistage probability sampling and cluster sampling involved many others, including Jerome Cornfield, Lester Frankel, William Hurwitz, and J. Stevens Stock.

It is important to note also that leading researchers moved not only among federal government agencies, but also to and from academia and private research and surve organizations. This movement led to increasing communication and cross-fertilization of ideas and methods and to building bridges across disciplines, institutions, and areas of application. The teams formed by drawing from different perspectives and experiences were a powerful effective force for the development of statistical methodology to solve the problems of the federal statistics agencies. There is a lesson here for the federal statistical system today.

NORWOOD AND KLEIN'S PAPER

Janet Norwood and Deborah Klein trace the historic development of three important concepts: industrial classification, race and ethnicity, and wages. At first blush, the concepts seem to have a precise economic, anthropological, or administrative definition. But, as the authors point out, concepts are often much more complex that what they seem to be. Should the process of production carry more weight than the output in classifying an industry? How does a respondent's self-classification of race change over time? Should compensation be measured by what employers pay or by what employees receive, and should it include noncash benefits?

Another major point of the paper is aptly stated:

> . . . phenomena underlying government statistics keep changing, the country's view of the concepts underlying data also changes, and those responsible for the measurement of these phenomena in official statistical series need to take account of the change in the definitions used in the conduct of surveys.

Our federal statistical system has not kept up with the times and Norwood and Klein direct our attention to the need for more attention to the identification and delineation of the underlying concepts. I applaud their call for a comprehensive reexamination of the concepts underlying the structure of the Standardized Industrial Classification and a modernization of the entire system. Another area in need of reexamination to which I would give priority is the concept of productivity.

The authors do point out the necessary trade-off between relevance to new conditions and continuity of time series. The balance, however, has too often favored continuity and not just for historical purposes. Even with the same concepts, comparisons over time may still be difficult. A dollar spent on a loaf of white bread today is not the same as it was decades ago, for many reasons beyond inflation. Production and transportation of wheat and baked goods have changed. Bread plays a different role in our diet. The quality and variety of bread has changed. And the cost of a loaf is a different portion of our income.

GREENHOUSE'S PAPER

Not included in these proceedings is Samuel Greenhouse's paper on the contributions to statistics from the National Institutes of Health. Greenhouse illustrates the prodigious development of methodology derived for application to the fields of health and medicine. The applications spanned both experiments and observational studies. Medical applications were the source of much methodology that later found its way into other fields. From the National Halothane Study (Bunker, Forrest, Mosteller, and Vandam, 1969) grew log-linear methods for the analysis of contingency tables that would find their way to the social sciences. From the need to analyze truncated medical observations grew survival analysis that later found application in other fields. The field of medicine and health was rich with data and, as Greenhouse points out, rich with creative biostatisticians and epidemiologists.

HOW ARE STATISTICAL INNOVATIONS INTRODUCED?

I would like to suggest a thesis for how statistical innovations are successfully introduced in federal government agencies and illustrate it with the example that Bailar recounts of the introduction of sampling at the Census Bureau.

Statistical research developed in the federal statistical agencies because of three reasons:

(1) a societal need for information to guide the social and economic programs of the government;
(2) innovations in statistical methodology; and
(3) institutions through which the theory could be applied, adapted, and extended to meet the needs for information.

For an institution to succeed in this goal, however, there had to be a confluence of three types of leadership:

(1) in communication to disseminate the statistical innovations and to bring people together across institutions, disciplines, and areas of application;
(2) in scientific intellect to understand the the new methods and to adapt them or to develop further ones by which to obtain and analyze the information; and
(3) in management to recognize the value of applying the statistical innovations and to influence others, especially senior government officials, to adopt the new methods.

These three types of leadership are akin to Hansen's reasons for the success of sampling at the Census Bureau, which were referred to in Bailar's paper: team-work approaches, a core of sampling experts, and support from the top. Each of these aspects, however, required leadership. The introduction of sampling is, at least for me, the best example of how the setting and leadership led to adopting, adapting, and extending a statistical innovation. Let me embellish on this story that Bailar, Hansen, Duncan and Shelton, and others tell.

It was 1937, during President Roosevelt's New Deal, and society had great need to guide the many recovery programs. It was also three years after the appearance of Neyman's classic paper (Neyman, 1934), a Copernican revolution in statistics. Neyman derived the sampling error of the mean of a stratified random sample from a general finite population, showed how to allocate the sample among the strata. He debunked purposively selected samples and launched the theory of confidence intervals.

W. Edwards Deming, known as a leader in disseminating statistical innovations, arranged for Neyman to come to Washington in April of 1937 to present a series of lectures. Among those who attended was Morris Hansen. Hansen had been brought into the Census Bureau's Statistical Research Division just the previous year by Calvert L. Dedrick, who thought there was a future for the Bureau in sampling and assigned Hansen to explore the field.

In the summer of 1937, Congress passed a law requiring a voluntary census of unemployment. Dedrick had testified that collecting data in such an uncontrolled manner would not be useful, but his arguments did not prevail on the Congress. President Roosevelt set up a temporary agency headed by John D. Biggers to conduct the census of the unemployed. Dedrick, with the help of Frederick F. Stephan and Samuel A. Stouffer, convinced Biggers that voluntary reporting would lead to a very large undercount and that an inexpensive probability sample could provide a check on the results.

The voluntary census was conducted in November of 1937 and the sample household canvass was conducted by postal carriers in the week ending December 4, 1937. The preliminary results were available on New Year's Day. Working on producing those results in his office at midnight on New Year's Eve, Morris Hansen called his wife to wish her a Happy New Year.

The sample survey, dubbed "The Enumerative Check Census of 1937," was a success. It produced larger and more realistic measures of the unemployed. The success of this sample survey laid the groundwork for sampling to be used in the decennial census for the first time in 1940, and the questionnaire procedures that had been developed were employed by the Works Project Administration in 1940 for a national sample survey of unemployment, still continuing to this day as the Current Population Survey. Hansen went on to extend Neyman's work, in particular to sampling with probability proportionate to measures of size and to sampling

in multiple stages with varying probabilities.

The need for information was present: Congress wanted to know the extent of unemployment. The statistical innovations were there in Neyman's 1934 paper. The institution at which it all took place was the Census Bureau. Deming took the lead in seeing that Neyman's results were communicated. Hansen understood the importance of the results and was able to apply, adapt, and extend them. And Dedrick convinced administration officials to adopt the innovation of sampling.

It is my pleasure in discussing these three papers today to note that the presenters have themselves been effective leaders in communication, in scientific intellect, and in management.

REFERENCES

Bunker, John P.; Forrest, William H., Jr.; Mosteller, Frederick; and Vandam, Leroy D. (1969). The National Halothane Study, Report of the Subcommittee on the National Halothane Study of the Committee on Anesthesia, Division of Medical Sciences, National Academy of Sciences--National Research Council. Washington, D.C.: U.S. Government Printing Office.

Converse, Jean M. (1987). Survey Research in the United States: Roots and Emergence, 1890-1960. Berkeley: University of California Press.

Duncan, Joseph W. and Shelton, William C. (1978). Revolution in United States Government Statistics: 1926-1976, Office of Federal Statistical Policy and Standards, U.S. Department of Commerce. Washington, D.C.: U.S. Government Printing Office.

Fienberg, Stephen E. and Tanur, Judith M. (1989). Some history of survey methods and data collection technology. Manuscript to appear in this proceedings volume.

Hansen, Morris H. (1987). Some history and reminiscences on survey sampling, Statistical Science. Vol. 2, 180-190.

Neyman, J. (1934). On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection, Journal of the Royal Statistical Society, Vol. 97, 558-625.

Olkin, Ingram (1987). A conversation with Morris Hansen, Statistical Science, Vol. 2, 162-179.

Stigler, Stephen M. (1986). The History of Statistics: The Measurement of Uncertainty Before 1900, Cambridge, Mass.: Harvard University Press.