

Thomas J. Flewes, Bureau of Labor Statistics

Good planning of sample survey programs, perhaps more than is the case on other fields, requires a close and constant interaction between consumer and survey sampler. As Leslie Kish points out in his text on survey design, the dialogue is needed because survey objectives and sample design are a two-way street. While objectives should determine the design, problems of sample design often influence and change the objectives. However, in this age of functional specialization, user and sampler are often housed under different roofs.

This paper describes one such planning process, in which the Bureau of Labor Statistics and the Bureau of the Census are now engaged. The plans will lead to a redesigned Current Population Survey, and at least three concrete results: implementation of the recommendations of the National Commission on Employment and Unemployment Statistics; incorporation of information from the 1980 Census of Population and Housing; and improving the survey design so as to enhance the reliability of the data and do so in a cost-effective manner.

This is an undertaking of unprecedented magnitude. There have been significant sample redesigns at the completion of each decennial Census, and, in the mid-1960's, there was a major questionnaire and concept redesign to implement the findings of the Gordon Committee. Still, never before in the nearly 40 year history of the program has there been an attempt to simultaneously redesign and to expand the sample, introduce revisions in concept and questionnaire content, and improve the basic quality of the data.

The process which will lead to the redesign was started in early 1978. That was probably late by a year or two. It will continue through introduction of the new labor force concepts (which is tentatively scheduled for January 1983) and the introduction of a redesigned, enlarged sample in January 1984.

It is a process which has, until now, largely been in the stage of intra- and inter-agency discussion will set the stage, insure the potential, and then define the limitations of this important survey through the 1980's and into the 1990's. It is now high time to shed a little light on the process, so as to solicit the interest and input of others in the statistical community.

The Redesign Process. The process of survey redesign is an exciting and vital, if largely uncharted activity. Unlike sample design, about which numerous texts provide excellent step-by-step procedures, there is little guidance when approaching the redesign of an existing survey. In application, redesign is governed as much by the politics of compromise as by theory. In a redesign, the optimal often gives way to what is feasible.

There remain many parallels between design and redesign. Both activities involve the formulation of concepts, the development of selection and of estimation procedures, and the provision of means for processing and analysis of the data. In some critical respects, however, redesign differs from the design criteria to which statisticians were introduced in their statistics courses:

- o Redesign pertains to an on-going program, in which significant investments of skill, resources, and emotion have been made over the years;
- o Redesign often takes place in an atmosphere characterized by both inertia and demand for a positive change, and these goals conflict;
- o Redesign suffers from lack of resources. It is frustratingly true that product improvement seems harder to justify than product development, and on-going program requirements can compete overwhelmingly for the attention of the limited technical expertise available to a program.

These forces make the process of survey redesign a much more difficult task, in many respects, than design of a survey from scratch. Nonetheless, if approached properly, survey redesign is an invaluable learning activity. It forces a refocus on the entire program, and an examination of the conventional wisdom. It is a healthy activity.

Redesign Responsibility. The Bureau of Labor Statistics is the principle, but not the sole consumer of data obtained from the Bureau of the Census' Current Population Survey. Since 1959, the Bureau of Labor Statistics has had primary responsibility for the content of the basic CPS, and for analysis and publication of the labor force data collected. The BLS is a purchaser of collection and tabulation services from the Bureau of the Census. The Census Bureau has responsibility for all other aspects of the program, to include, importantly, design of the sample and the quality of the statistics produced.

While this division of responsibility creates the potential for conflict, the relationship between the agencies has, for years, been a model of inter-agency cooperation. There have been conflicts, but surprisingly few considering the fact that one Federal agency holds the purse strings, and another has ownership of the sample. This underlying spirit of cooperation set the stage for the give-and-take that invariably accompanies a redesign process.

The two agencies initially took different views of the redesign. A Census Bureau task force looked at the technical and methodological aspects of the survey. The redesign offered the opportunity to consider such topics as rotation group pattern; interviewing techniques and their best mix; the appropriate mix of self-versus proxy respondents; and the reliability criteria to be associated with geographic and demographic data.

Concurrently, a Bureau of Labor Statistics task group concentrated on aspects of the survey that related to the data that are published: period; timing and coverage; stratification criteria; the content of the questionnaire; estimating procedure used; and data processing. All the while, both agencies were forced to consider a number of concurrent events and trends which served to delimit the future requirements, including:

- o The deliberations of the National Commission on Employment and Unemployment Statistics, which led to the recommendations summarized in this session, and which will have a critical impact on the future concepts and methods;
- o Legislation proposed in Congress which mandates the use of statistical series in the allocation of Federal funds to States and areas;
- o The plans for the 1980 Census, which will be the basis for the sample redesign, but which also will place a burden on the Census staff and in the process, limit the resources that can be used in the CPS redesign;
- o The plans for the 1985 mid-decade Census;
- o Requirements for more and better data on the employment situation of minority, ethnic and demographic groups;
- o Improvements in analytical techniques, among which are better methods of exploiting microdata files and longitudinal analysis;
- o Increased concerns over the reporting burden generated by Federal government data collection programs; and
- o Closer scrutiny of the costs and effectiveness of government programs, generally.

At the Bureau of Labor Statistics, the process was educational. It required asking some very basic questions: What data are really required to fulfill the BLS obligation to provide timely, comprehensive, and analytically-sound information on the labor force? What are the real needs of the community we serve -- the policy makers, the analysts, the researchers and the general public? The Bureau fortuitously had the outside, on-going evaluation of the National Commission to provide guidance, as well as a series of papers prepared by subject matter experts on various aspects of user needs.

Redesign Trade-offs. The notion of trade-offs set the tone for considering competing user needs. At the national level, for example, the importance of the unemployment rate and employment estimates as economic indicators dictates that priority be given to maximizing reliability of estimates of month-to-month change and change over short spans of time. Year-to-year change and comparative differences between demographic groups are important, but secondarily so. For State and area data, however, we would seek to maximize reliability of between-area comparisons over the year for better fund allocation.

One of the points at which these objectives are destined to conflict is in consideration of the optimal sample rotation pattern. The present 4-8-4 pattern yields more precise estimates of change between adjacent months and years than, say the 6-0 pattern employed in the Canadian household survey, which would improve precision of month-to-month change at the expense of over-the-year movements. Likewise, a 1-11-1-11 pattern would help the precision for year-to-year change and annual averages at the expense of monthly and quarterly change estimation. What is best for national economic indicator needs is suboptimal for the allocation of funds to States and areas, and visa versa. After considering these options, a half-way compromise position is being explored. One option receiving some support is a 3-9-3 pattern, which would reduce respondent burden, better the estimates of quarterly change, and preserve some precision for monthly and yearly change estimates.

Trade-offs are also evident in terms of periodicity of the survey. The CPS now collects data once each month, based on activity or status reported for the calendar week including the 12th of the month. Over the years, a number of serious proposals have been made to replace the reference week concept with weekly collection and monthly averaging. The idea was comprehensively examined in a 1970 report of the subcommittee on weekly enumeration of the Office of Management and Budget's technical group on labor force statistics. In this study, a good number of advantages were suggested, including:

- o Data would be more representative of the actual behavior of the population during all weeks of the month rather than for one week;
- o Averaging weekly data would have some positive dampening effect on the irregularity of the series;
- o Some data collected could be useful as weekly economic indicators; and
- o Problems associated with unusually early or late timing of the reference week during the month would be eliminated.

But the list of disadvantages was also weighty, and served to sway the argument toward maintaining the status quo:

- o Loss of comparability in reference period with other major labor force series, such as those in the survey of establishments;
- o Seasonal adjustment problems involved in such a dramatic shift from single week to monthly average;
- o Loss in timeliness of publication of the survey findings by as much as 2 weeks;
- o Problems of data collection during certain periods, such as over holidays; and
- o The awkwardness of interpreting data based on 4 and 5 week months.

Communicating Consumer Needs. Once requirements are identified, the consumer of survey results faces a new challenge -- that is, communicating those requirements to the samplers, who in turn must balance and blend the needs into a sample design. In this process, the Bureau of Labor Statistics staff found that we were forced to speak the sampler's language. It was necessary, in view of perceived reliability trade-offs, to assign priorities to long shopping lists of data needs, then to translate those needs into statements of the desired precision for the sample estimates. This is a discipline highly recommended for other consumers. A certain sympathy for the task confronting the sampler results from the process.

One example might well illustrate this process. There is a pressing need for better data on the employment and unemployment situation of Hispanic persons. Both Census and the BLS have been trying to improve data for this group. Indeed, we have mandated to do so in Congressional legislation that demands more comprehensive labor force measures for Hispanics.

We have recognized for some time that there is not sufficient sample to support more intensive exploitation of data for this population. The present coefficient of variation on month-to-month change for the Hispanic unemployment rate is 7.2 percent. But what should be done to improve these data? If additional sample is the answer, how much of an increase is warranted? In the final analysis, what coefficient of variation is acceptable?

In addressing this problem, both consumer and sampler were forced to a reconsideration of its basic nature. The problem, restated, is that we are unable to publish Hispanic data with the same detail as that for Whites and Blacks because, in the process of seasonally adjusting the data, the irregular component is so large that it clouds identification of the seasonal and trend-cycle components.

The restated requirement is not simply to expand the sample by "X" amount. Research is needed to come to a better understanding of the relationship between sampling error and the irregular component, which we have come to understand as an amalgam of both sampling and non-sampling errors. This relationship is not clear, but by restating the issue, we have a framework for approaching a solution. Research on better specification of the linkage between sampling error and the irregular component should provide the consumer and sampler with better guidance in drawing the sample and analyzing the results.

Through this process of exploration, founded in compromise, a framework for the future of the Current Population Survey is emerging. It will be an improved product, though not radically or even perceptively different from the current survey. Some important design changes, such as the proposal to base sample selection on the States rather than on a national basis, will give stronger, more reliable data, but may well be transparent to all but those most intimately involved.

There are many lessons to be learned from this redesign process. One important implication that may be drawn is that it is a process which needs an early start, and a forum for continued and meaningful dialogue between sampler and data consumer. The relationship between the consumer objectives and the eventual sampler's design is one of interaction, compromise and on-the-spot refinement.

The next stage in this process is the integration of feasible recommended improvements by the National Commission on Employment and Unemployment Statistics. One core activity will be the testing of suggested refinements in concepts and definitions by use of an enlarged Methods Test Panel, which will be discussed next by one of our Census Bureau colleagues.