Overhauling the Current Population Survey

Evaluating changes in the estimates

A 1–1/2 year overlap sample is being used to gauge differences in measures from the old and revised surveys; the effects of questionnaire changes, advanced computerization, and centralization of the interview process will be subject to special scrutiny

Chester E. Bowie,
Lawrence S. Cahoon,
and
Elizabeth A. Martin

The Current Population Survey (CPS) is the cornerstone of the U.S. labor market information system. It provides monthly statistics that serve as measures of both current labor force utilization and the overall performance of the economy. The data are used for both cyclical and secular trend analysis and also form the basis for the official U.S. labor force projections. The CPS is also used for a program of special inquiries on particular characteristics of the population and labor force, such as income and poverty, work experience and migration, school enrollment and educational attainment, and fertility. In addition, it is a widely used microdata source for research on a variety of labor market and social science topics.

The survey's most well-known statistic—the monthly national unemployment rate—often is used as a prime barometer of the health of the economy. Monthly unemployment rates for States, which are based either directly (for 11 large States) or indirectly (for the remaining States and the District of Columbia) on the CPS, are used in the allocation of Federal funds to local areas.

Recapping the reasons for change

For decades, the CPS has been the worldwide standard for household surveys. Its design, concepts, and operational procedures have served as a model for many other such surveys. Over the past few years, however, the household surveys of some other countries have surpassed the CPS in the use of more modern and innovative survey methods.

The current CPS labor force questionnaire has remained essentially unchanged since the last major revisions in January 1967, which were based in part on recommendations of the 1962 Gordon Committee. Additional revisions were proposed in the late 1970's and 1980's, most notably by the Levitan Commission. No major changes in the questionnaire have been implemented until now, due to the lack of funding for a large overlap sample necessary to assess their effect on the CPS labor force data series.

Current efforts to redesign the questionnaire, which began in 1986, resulted from joint Census Bureau and Bureau of Labor Statistics plans for a major redesign of all aspects of the CPS. The CPS redesign plan calls for the introduction of a new labor force questionnaire in January 1994, following a period of field testing and using a 1–1/2-year national overlap sample to estimate the effect of the changes on the labor force estimates. Concurrent with this initiative, attempts are being undertaken to eliminate paper and pencil data collection by adopting integrated computer-assisted interviewing methods. Finally, the redesign involves the selection of new sample areas and housing units from a sample frame developed from the

Chester E. Bowie is Assistant Chief of the Demographic Surveys Division; Lawrence S. Cahoon is Assistant Chief of the Demographic Statistical Methods Division, and Elizabeth A. Martin is Chief of the Center for Survey Methods Research, U.S. Bureau of the Census.

1990 Decennial Census, in order to account for changes in the population that have occurred since the preceding census. The redesigned sample will be phased in gradually starting in April 1994, 4 months following the introduction of the new questionnaire and computerization of the interviewing process.

Description of the modernization

Most of the objectives of the CPS labor force questionnaire redesign have been described earlier in this issue in articles by John E. Bregger and Cathryn S. Dippo and Anne E. Polivka and Jennifer M. Rothgeb, and so will not be discussed further here. However, the major points are summarized in the box below.

It is important to restate that efforts were made to enable consistent application of classification criteria for labor force concepts, and to incorporate the use of dependent interviewing. Dependent interviewing—using information from previous interviews to identify "real" change—was investigated to reduce the incidence of spurious change in gross flow and longitudinal data.

Another objective of the survey redesign that merits further discussion is the use of the capabilities of computer-assisted interviews for improving data quality and reducing respondent burden. The survey redesign strategy requires that all interviewing, and therefore all data capture, will be computer-based. This will involve both computer-assisted telephone interviewing (CATI) and computer-assisted personal interviewing (CAPI). Computer-assisted telephone interviewing takes place in either a central location where the questionnaire is administered by interviewers under direct supervision of a superior or from the home of a field representative who makes use of a laptop computer. Computer-assisted personal interviewing involves field representatives conducting interviews in the respondents' homes using a laptop computer. Consistent with current interviewing strategy, most CPS interviews will be conducted by telephone.

The single most important dimension that the computer brings to the interviewing environment is the ability to simplify the process for the field representative. The redesigned CPS labor force questionnaire has become so complex that it could not be conducted using a paper questionnaire. However, with a computer doing all the complicated work, the actual interview is simplified for both the respondent and interviewer. The computer automatically brings the appropriate question to the screen. It can also be programmed to perform editing functions and to identify inconsistent answers. Another potentially important feature of computerized data collection systems is the ability to store and display data from earlier interviews, so as to permit dependent interviewing. In addition, CATI/CAPI enhances the longitudinal as-

An overview of the modernization

The objectives of the CPS labor force questionnaire redesign are as follows:

- To improve the measurement of those concepts which, although well defined, are not measured precisely in the current questionnaire. Examples are:
  - — employment/unemployment status
  - — layoff
  - — hours worked
  - — self-employment/unpaid family work
  - — earnings

- To define those concepts that are not well explained in the current questionnaire. Examples are:
  - — part-time workers
  - — status of persons not in the labor force

- To introduce revised concepts. A prime example is the need to introduce a new concept for measuring discouragement.

- To reduce respondent burden. Examples include revising the approach to:
  - — retired persons
  - — persons unable to work or disabled persons
  - — inquiries concerning duration of unemployment for unemployed persons

- To reduce spurious changes in certain CPS data, such as those related to:
  - — industry and occupation
  - — duration of unemployment

- To increase available data on topics of current analytic importance, such as:
  - — dual jobholding
  - — usual hours worked
  - — earnings detail (for example, tips, commissions, overtime pay)
  - — child care for part-time workers

- To reduce dependence on volunteered information about such topics as:
  - — the existence of a business in a household
  - — full-time/part-time work status and the reason for that status

pects of the CPS by facilitating matching of household members between adjacent months.

Evaluating the changes

Whenever significant changes are made in an ongoing survey operation, there is always the expectation that those changes will affect the characteristics of the data obtained. It is important to measure any such effects. For just this purpose, an overlap sample—a sort of control group—has been designed for the CPS, to run from July 1992 through December 1993.

The primary objective of the overlap sample is to provide a reference point for the transition of the main labor force series from the "old" to the "new" CPS. The main measurement objective of the overlap is to obtain precise estimates of overall differences due to the redesign, and less precise estimates for certain major subgroups of the population. Secondary goals are to measure the effect of specific types of changes:

- questionnaire changes;
- computerization of the interviewing processes; and
- centralization of a portion of the interviewing.

It is highly likely that interaction effects among these three types of change will be observed. The overlap sample has been designed with special features to specifically measure the effects of some individual components of the change.

A large number of survey design features are being changed in the new CPS, and a number of them, alone or in combination, could result in significant changes in the estimates. The Bureau of Labor Statistics and the Census Bureau want to be able to explain to the public why differences between the new and old series occur, and to comment on whether any changes reflect improvements in the quality of the data. The two agencies also need to understand from a scientific point of view the effect of different design features on labor force estimates. A third reason for wanting to know the reasons for differences between the two surveys is to use the information diagnostically to improve the data collection process during the overlap period (for example, by improving training) to ensure a smooth transition from the overlap to full implementation of the redesigned CPS in 1994.

The overlap sample is designed to meet the first objective of calibrating the new and old CPS estimates, *but* its ability to meet the second objective—explaining the differences—is limited. For the most part, the overlap design does not provide for comparisons that would permit estimation of the effects of different design features on overall estimates.

The following section provides an overview of the design of the overlap sample. It is followed by a summary of the types of analysis that are planned in order to evaluate the changes between the "old" and "new" surveys.

Overlap sample design

The overlap sample design was based on that used in the National Crime Victimization Survey, which is conducted by the Census Bureau for the Bureau of Justice Statistics. This design was chosen because the principal intent is to measure national-level effects. Although the CPS is a State-based design, none of the changes being made to the survey treats States differently. Cost constraints mitigated against designing an overlap sample that would measure effects at the State level.

The design is a stratified multistage sample. The larger metropolitan areas are included in the sample with certainty—that is, with a probability of 1. The remaining areas are stratified with one Primary Sampling Unit (PSU) (or locality) selected per stratum to represent the other PSU's in the stratum. The sample size for the overlap survey is approximately 14,000 eligible housing units within the selected PSU's per month.

Analysts will be comparing estimates from the overlap sample with those from the ongoing CPS. The overlap sample will provide annual average estimates with a standard error of approximately 0.11 percent for the unemployment rate and approximately 0.2 percent for the labor force participation rate.

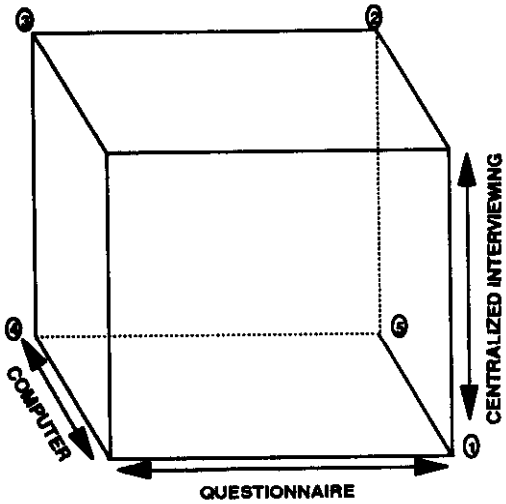
Analysis of data effects

As indicated earlier, the overlap sample was designed to measure directly the effects of all of the changes to the survey. Embedded in the overlap sample and in the current CPS sample are a number of split-panel designs to measure the effects of some individual components of the change to the new system.

The cube pictured below shows the types of changes that could be analyzed. The historical system is represented by the lower right-hand front corner, and consists of the current questionnaire, no enhanced use of computer technology, and no centralized interviewing. The goal is the diagonally opposite corner in the back upper left, depicting a system using the new questionnaire, advanced computer technology, and both centralized and decentralized interviewing.

Each of the lines along the edge of the cube represents a dimension of change, the effects of which analysts would like to obtain a measure. For example, the segment from point 1 to point 5 rep-

resents the use of computers with the present questionnaire and no centralized interviewing. By gaining an understanding of the effects of each of the changes individually, we hope to gain a better understanding of the reasons for any overall effects.



1. Old questionnaire; decentralized; no computer (historical)
2. Old questionnaire; centralized; computer
3. New questionnaire; centralized (and decentralized); computer (goal)
4. New questionnaire; decentralized; computer
5. Old questionnaire; decentralized; computer

The new questionnaire is sufficiently complex that it is almost certainly unreasonable to attempt to construct and use a paper version. For this reason, some corners of the cube represent unrealistic situations. These are scenarios that would require the use of the new questionnaire without computer assistance.

To evaluate the implications of all the changes to the survey, we have focused our efforts in these directions:

1) Analysis of the overall effect;
2) Analysis of questionnaire effects; and
3) Analysis of mode effect—that is, the joint effects of computerization and centralization of data collection.

These areas are discussed in the following sections.

New vs. old questionnaire

Numerous changes were made in the questionnaire to better define CPS concepts, improve respondents' understanding of the intent of questions, reduce reliance on volunteered information, and improve the reliability of classification by interviewers. The effects of these changes are hypothesized to be improvements in data quality and more consistent labor force classifications, but few net differences in estimates between the old and the new surveys. For the few labor force concepts for which definitions were changed (consistent with the recommendations of the Levitan and Gordon Commissions), substantial differences in estimates between the old and new questionnaire are expected, in particular with regard to declines in the number of economic part-time workers and in the number of discouraged workers. Finally, it is expected that dependent interviewing will greatly reduce reported month-to-month changes in industry, occupation, and class-of-worker categories. In short, the hypothesized effect is that the direction of the bias in the current data will be reversed and reduced: a large overreporting bias will be replaced by a much smaller underreporting bias.

We will not attempt to conduct paper-based interviews using the new questionnaire. It incorporates complex branching patterns and dependent interviewing techniques that are not feasible to incorporate in a paper survey instrument; therefore, we will not know how the absence of a paper questionnaire will affect survey results, or be able to gauge the effect of automation on the new questionnaire.

We will be able to tabulate the effects of the questionnaire change using MIS (number of a group's months in the sample) 2–4 and 6–8 CATI cases by comparing CATI cases in the overlap sample with those in the current sample across common PSU's. For MIS 1 and 5 cases, using the current design with the old questionnaire administered on paper and with the new one on CAPI, we get only an overall measure of the effects of computerization and use of the new questionnaire. Because these cases all involve personal visit, there is no effect of centralization of data collection.

Computers vs. paper

Ideally, automation makes it possible to achieve greater control over how a survey is actually administered, resulting in greater standardization. Automation necessarily reduces interviewer errors in following instructions to skip certain items or in asking questions out of order, and very likely reduces variability in the way questions are asked. Standardized probes are programmed, which contribute to greater uniformity in how problem situations or "don't know" responses are handled. On the other hand, automation involves reliance on machines, which can break down or malfunction in ways that can disrupt the interview. In addition,

there is the possibility that CAPI interviewing, which involves bringing the computer into the respondents' homes, may inhibit rapport or have other unintended effects on the interview.

As noted above, it is not possible to measure the effects of automation separately from the effects of the new questionnaire, because it is not feasible to implement the new questionnaire on paper. However, a variety of qualitative and quantitative information will be collected to assess interviewers' and respondents' reactions to CAPI data collection. These include item nonresponse measures, response distributions, respondent and interviewer debriefing data, and behavior coding of interviewer/respondent interactions.

Centralized vs. decentralized interviewing

The Census Bureau's field staff is highly experienced and generally well-trained. Many CPS field representative have years of experience conducting the survey. In contrast, interviewers in the Census Bureau's Hagerstown, MD, facility have many fewer years of experience and less training, and the staff in the newly opened Tucson, AZ, centralized facility have even less training and experience. These differences in training and experience may turn out to be sources of differences in data quality between the old and new CPS, which may result in differences in results between the centralized and decentralized modes of interview.

To assess and monitor possible effects of interviewer training and experience on the quality of data, a number of measures will be collected, primarily to use as tools for diagnosing and correcting problems. The measures are intended to identify problems with the implementation of CAPI and/or the new questionnaire, which would be addressed primarily during training. The means of assessing interviewer performance will include:

- interviewer focus groups;
- monitoring (in CATI) and taping (in CATI and the field) of interviews; and
- capture of data on frequency of interviewer backups and corrections in CATI and CAPI.

Weaknesses identified will be addressed through supplementary training.

It is also believed that centralization will affect CPS results, because it permits more communication among interviewers, and more monitoring of workers by supervisors, than is possible within a decentralized field staff. Greater communication means that interviewers in a centralized facility can, and do, develop their own agreed-upon interpretations of survey procedures and questions. This is beneficial when interviewers' interpretations agree with standard procedures, but this is not always the case. Past experience has suggested that interviewers in Hagerstown have their own idiosyncratic ways of handling certain situations, such as classifying job search methods in other than the intended way, and obtaining job titles rather than occupation information.

In the overlap sample, cases are being randomly assigned for interviewing by either the Hagerstown and Tucson offices or the decentralized field staff. This makes it possible to estimate the effect of centralization, which is recognized as a possibly important source of variation between the old and new surveys. However, CATI interviewing, and thus the experimental assignments, is only being implemented in multi-interviewer PSU's, not in single-interviewer PSU's. Multi-interviewer PSU's tend to be urban and suburban areas. Hence, analysts will not know the effects of centralization versus decentralization for rural respondents, who are being interviewed only by decentralized field staff. The effects of interview centralization on the new survey also will not be measured in rural PSU's.

The above cases illustrate some of the comparisons being planned, and there are many others. Under ideal circumstances one would wish to examine, to the extent possible, the effects of each change. However, budget constraints forced a design of the overlap sample to measure primarily the total effect of the changes. The resulting small sample size will not always permit formation of firm conclusions about the effects of the survey redesign, especially for small estimates and small changes.

THE PLANNED REDESIGN and modernization of the CPS is an extraordinarily important and ambitious undertaking. The result of planning and testing since 1986 will culminate in the replacement of the current CPS operation with a revised questionnaire and a modern data collection system beginning next year. At the time of the redesign implementation, it must be possible to estimate what the effects of the new questionnaire and the use of automation (CATI/CAPI) have on the published CPS labor force estimates, and to explore the reasons for these changes. The design and implementation of an overlap sample and the various analytical efforts described above should provide the information required to address these objectives. □