

PLANS TO ASSESS THE EFFECTS OF METHODOLOGICAL CHANGES TO THE CPS ON LABOR FORCE ESTIMATES

Chester Bowie, Larry Cahoon, Elizabeth Martin
Chester Bowie, Bureau of the Census, Washington, D.C. 20233

BACKGROUND

Purpose of the CPS

The CPS is the cornerstone of the United States labor market information system. It provides monthly statistics that serve as a measure of both current labor force utilization and the overall performance of the economy. CPS data provide an indication of how well the labor market is functioning in providing jobs to those who are seeking work. These 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 segments, or 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--is often used as a prime barometer of the health of the economy. Monthly unemployment rates for states, which are based both directly (11 largest states) and indirectly (remaining states and the District of Columbia) on the CPS, are used in the allocation of Federal funds to local areas.

Purpose of a Major Modernization

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 household surveys. Over the past few years, however, some other countries' surveys have surpassed the CPS in utilizing 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 1970s and 1980s, 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 in questionnaire redesign, which began in 1986, resulted from joint Census Bureau and Bureau of Labor Statistics planning 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 a 1 1/2-year national overlap sample to estimate the effect of the changes on the labor force estimates. Concurrent with this initiative, we also set out on a course that will 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 most recent decennial census. It updates the sample to improve efficiency by accounting for changes in the population which have occurred since the preceding census. We will implement the redesigned sample starting in April 1994, 4 months after the introduction of the new questionnaire and computerization of the interviewing process.

Description of Modernization

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

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

- o To operationalize those concepts that are not well defined in the current questionnaire.

Examples are:

- part-time workers
- status of persons not in the labor force
- o To introduce revised concepts. An example is:
 - discouraged worker
- o To reduce respondent burden. Examples are:
 - retired persons
 - unable to work or disabled persons
 - duration of unemployment for unemployed
- o To reduce spurious change in certain CPS data series. Examples are:
 - industry and occupation
 - duration of unemployment
- o To increase available data on topics of analytic importance. Examples are:
 - dual jobholding
 - usual hours worked
 - earnings detail (for example, tips, commissions, overtime pay)
 - additional categories (for example, child care for part-time workers)
- o To reduce dependence on volunteered information. For example:
 - business in a household
 - full-time/part-time status and reason

Efforts were also 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 is to utilize 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 include computer-assisted telephone interviewing (CATI) and computer-assisted personal interviewing (CAPI). Computer-assisted telephone interviewing takes place either in a central location involving interviewers under direct supervision or out of a field representative's home on a laptop computer. Computer-assisted personal interviewing involves field representatives conducting interviews in the respondents' home using a laptop computer. Consistent with our current interviewing strategy, most CPS interviews are conducted by telephone.

The single most important dimension 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 it could not be conducted using a paper questionnaire. With the computer doing all the complicated work, the actual interview is simplified for the respondent and interviewer. The computer automatically brings the appropriate question to the screen. The computer can also be programmed to perform editing function and identify inconsistent answers. Another potentially important feature of computerized 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 aspects of the CPS by facilitating matching of household members between adjacent months.

Evaluation and Benchmarking the Changes

Whenever significant changes are made in an ongoing survey operation there is always the expectation that those changes will effect the data. It is important that we measure any such effects. We have designed an overlap sample for the CPS which will run from July 1992 through December 1994 for just this purpose.

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 precise estimates of overall differences due to the redesigned CPS, and less precise estimates for certain major subgroups of the population. Secondary goals are to measure the effect of individual changes. We view changes as occurring along three dimensions:

- Questionnaire changes
- Computerization of the interviewing processes
- Centralization of a portion of the interviewing

We expect to observe interaction effects among these three dimensions. We have designed special features within the overlap sample for specifically measuring 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 changes in the estimates. The Bureau of Labor Statistics and the Census Bureau want to be able to explain to the public why changes in 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 why differences occur is to use the information diagnostically to improve the data collection process during the overlap period (e.g., by improving training) to ensure a smooth transition from the overlap to full implementation of the redesigned CPS in 1994. If changes are made, however, this will reduce the effectiveness of the data comparisons.

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 of explaining the changes is limited. For the most part, the overlap design does not provide for experimental comparisons which 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. This is followed by a summary of the types of analysis we plan to evaluate the changes.

Design of the Overlap Sample

We based the overlap sample on the National Crime Victimization Survey design. This design was chosen since we are attempting to measure national-level effects only. The CPS is a state based design. None of the changes being made treat states differently. There was no need to design an overlap sample to measure effects at the state level.

The design is a stratified multistage sample. The larger metropolitan areas are included in the sample with certainty. The remaining areas are stratified with one Primary Sampling Unit (PSU) selected per stratum to represent the other PSUs in the stratum. The sample size for the overlap survey will be approximately 15,000 eligible housing units within the selected PSUs per month.

We will compare estimates from the overlap sample to those from the ongoing CPS. The overlap sample will provide annual average estimates with a standard error of approximately .11 percent for the unemployment rate and approximately .2 percent for the Labor Force participation rate.

Analysis of Data Effects

We designed the overlap sample to measure directly the effects of all of the changes. We have embedded in the overlap sample and in the current CPS sample a number of split-panel designs to measure the effects of some components of the change.

The cube pictured in figure 1 shows the types of changes which could be analyzed. The historical system is represented by the lower right hand corner with the current questionnaire, no computer, and no centralized interviewing. One goal is the diagonally opposite corner in the upper left with the new questionnaire, using a computer, and centralized interviewing.

Each of the lines along the edge of the cube represent a dimension for which we would like to obtain a measure of the effects represented by that change. For example, from point 1 to point 5 represents 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.

The new questionnaire is sufficiently complex that we feel it is unreasonable to attempt to construct and use a paper version of the questionnaire. For this reason, some corners of the cube represent unrealistic situations. These are corners which would require the use of the new questionnaire without computerization.

In order to evaluate the effect of all of the changes, we have focused our efforts in these directions:

- 1) Analysis of the overall effect
- 2) Analysis of questionnaire effects
- 3) Analysis of mode effect. Computerization and Centralization

These areas are discussed in the following sections.

New versus Old Questionnaire

Numerous changes were made in the questionnaire to better operationalize 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. For the few labor force concepts for which definitions were changed (consistent with the recommendations of Levitan and Gordon Commissions), substantial differences between the old and new questionnaire are expected, in particular, declines in the number of economic part-time workers and in the number of discouraged workers. Finally, we expect dependent interviewing to greatly reduce month-to-month changes in industry, occupation, and class of worker classifications. The hypothesized effect is to reverse the direction of the bias in the current data and reduce it: a large overreporting bias will be replaced by a much smaller underreporting bias.

We will not attempt to conduct paper interviews using the new questionnaire. It incorporates complex branching patterns and dependent interviewing that are not feasible on paper; therefore, we will not know the effect of the new versus the old questionnaire on paper or the effect of automation on the new questionnaire.

We will be able to tabulate the effects of the questionnaire change using MIS 2-4 and 6-8 CATI cases by comparing CATI cases in the overlap sample with CATI cases in the current sample across common PSUs. For MIS 1 and 5 cases under the current design where the old questionnaire is conducted on paper and the new on CAPI, we only can get an overall measure of the

effects of computerization and the new questionnaire. Since these cases are all done by personal visit, there is no effect of centralization.

Computer-Assisted versus Paper Administration

Automation ideally makes it possible to achieve greater control over how the survey is actually administered, resulting in greater standardization. Automation necessarily reduces interviewer errors in following skip patterns or asking questions out of order, and very likely reduces variability in how 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 reduce rapport or have other unintended effects on the interview.

As noted above, for the new questionnaire, it is not possible to measure the effects of automation separately from the effects of the questionnaire, because it is not feasible to implement the new questionnaire on paper. A variety of qualitative and quantitative measures 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 versus Decentralized Interviewing

The Census Bureau's field staff is highly experienced and generally well-trained. Many CPS field representatives have years of experience conducting the survey. In contrast, interviewers in the Census Bureau's Hagerstown facility have many fewer years of experience and less training, and the staff in the newly-opened Tucson centralized facility will have even less training and experience. These differences in field staff training and experience are hypothesized as sources of differences in data quality in the old and new CPS, which may result in differences in results obtained in centralized versus decentralized modes of interview.

In order 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 measures to be collected include:

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

Weaknesses identified will be addressed through supplementary training.

Centralization is also hypothesized to affect CPS results because it permits more communication among interviewers, and more monitoring of them by supervisors, than can occur 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 classification of job search methods, and obtaining job titles rather than occupation information.

In the overlap sample, cases will be randomly assigned for interviewing by Hagerstown and Tucson versus decentralized field staff. This will make it possible to estimate the effect of centralization, which is recognized as a possibly important source of variation. CATI interviewing,

however, and therefore the experimental assignment, will only be implemented in multi-interviewer PSUs, not in single-interviewer PSUs. Multi-interviewer PSUs tend to be urban and suburban areas. Hence, we will not know the effects of centralization versus decentralization for rural respondents, who will only be interviewed by decentralized field staff. We also will not measure the effects of centralization on the new questionnaire in rural PSUs.

These illustrate some of the comparisons being made, and there are many others. We are very concerned with examining, to the extent possible, the effects of each change. Budget constraints forced a design of the overlap sample to measure primarily the total effect of the changes. The sample sizes, therefore, will not always permit us to make firm conclusions, especially for small estimates and small changes.

Conclusion

The planned redesign and modernization of the CPS is an extraordinary 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 new modern data collection system.

At the time of the redesign implementation, we must be able to estimate what the effects of the new questionnaire and the use of automation (CATI/CAPI) have on the published CPS labor force estimates. In addition, we need to explore reasons for these changes.

The design and implementation of an overlap sample and the various analytical effects described above should provide the information required to address these objectives.

FIGURE 1

