

EFFECTS OF METHODOLOGICAL CHANGE IN THE CURRENT POPULATION SURVEY

Cathryn S. Dippo, Donna L. Kostanich, Anne E. Polivka

Cathryn S. Dippo, Bureau of Labor Statistics, 2 Massachusetts Ave. NE, Washington, DC 20212

Key Words: Parallel testing, Mode effects, Questionnaire effects, Computer-assisted interviewing

In January 1994, a new measurement system was introduced into the Current Population Survey (CPS). The introduction of a new questionnaire and data collection procedures was the result of a research and development process that began in 1986. Two important goals of the redesign were to improve data quality and to increase the availability of information on labor market activity on a monthly basis.

By modifying question wording and sequencing to better operationalize existing definitions, reduce the reliance on volunteered information, and embed explicit and implicit recall strategies, data quality was improved. Changes in the definition of discouraged workers recommended by the National Commission on Employment and Unemployment Statistics were implemented, along with increasing the availability of discouraged worker data by moving the questions from the quarter-sample being interviewed for the fourth and eighth times to the full sample. Questions on multiple job holding were also added.

The new questionnaire was developed for a fully-automated collection environment with field representatives using laptop computers for both their personal visit and telephone interviews (CAPI) and with a centralized staff of interviewers from a telephone facility (CATI). Thus, it was possible to use complex skip patterns to tailor questions to respondents' situations and use information from prior interviews to reduce respondent burden.

To measure the effects of the new measurement system, a Parallel Survey (PS) of 12,000 occupied housing units was conducted from July 1992 through December 1993. While the PS was designed to replicate the collection methodology and procedures as they would be incorporated into CPS in January 1994, it was only one-fifth the size of CPS, had a national rather than a state-based sample as in CPS, and used a simpler post-stratification estimation procedure. Since the percentage of sample units to be interviewed from the centralized telephone facilities was to double in 1994, the PS had about twice the percentage of its sample units being assigned to centralized telephone interviewing. In order to test for the effects of different

modes of collection, subpanels were embedded in the PS (and CPS) for testing the following hypotheses:

- H₁: No centralized telephone interviewing effect
- H₂: No centralized and computer-assisted interviewing effect
- H₃: No questionnaire effect given centralized telephone interviewing
- H₄: No questionnaire designed for computer-assisted interviewing effect.

An important factor in the measurement process that could not be controlled in an experimental design sense was the interviewing staff. Recognizing the importance of interviewers to the measurement process, the decision was made to try to replicate in the PS the level of interviewer experience that would exist in CPS in January 1994. To that end, some experienced CPS and other Census current survey interviewers were transferred to the PS (only 25% of PS interviewers were new hires), and all PS interviewers were given up to six months for field experience with the new questionnaire and collection methodology. Still, the possibility existed of interviewer's prior experience affecting their knowledge and behavior and, consequently, the results.

Based on the 1993 annual average estimates from CPS and the PS, the new measurement system was expected to yield 0.5 percentage point higher estimates of the overall unemployment rate and 0.7 percentage point higher estimates of the unemployment rate for women. While no significant difference was found between estimates of the overall employment-to-population ratio from the PS and CPS, the estimate from the PS for men was 0.6 percentage point lower than CPS and 0.7 percentage point higher for women. The PS also yielded 20.5 percent lower estimates of the number of persons working part time for economic reasons and 62.4 percent fewer discouraged workers.

When the embedded panels were used to test the above hypotheses, evidence of possible effects was found for the following.

Centralized interviewing: Total unemployment rate (marginally); total black and black male unemployment rate (UER).

Centralized and computer-assisted interviewing: Total, total female, and total male UER; total white, white female, and white male UER; black female UER.

Questionnaire given centralized telephone interviewing: Total female UER (marginally), total black and black female UER.

Questionnaire designed for computer-assisted interviewing: Total, total male, and total female UER; total white, white male, and white female UER; total black, black male, and black female UER; total female and white female labor force participation rate.

In order to determine if various design differences between the PS and CPS could be affecting the results, the CPS was reweighted using PS post-stratification procedures (no significant difference in the UER), and the PS was reweighted to reflect the percentage of centralized interviewing to be used in January 1994 (.08 percentage point lower). Also, since the CPS would begin using 1990-census-based population controls in January 1994, it was reweighted to determine the effect of switching from 1980 to 1990 controls (.10 percentage point higher). It was not possible to evaluate the effects of not using the composite estimator in January 1994 or the effects of using seasonal adjustment factors based on old CPS with the new.

Faced with the possibility of a major break in series of an important economic indicator, the decision was made to continue parallel measurement, only with the old measurement system being used in the PS. While it would have been preferable to keep at least some portion of the measurement process constant from December to January, this was not possible, due to the differences in the sample designs. Namely, to provide monthly estimates for the 11 largest states and annual average estimates for the remaining states with the required reliability, the entire state-based CPS sample had to use the same collection methodology, and the PS with its national versus state-based sample could not be used.

In preparation for the conversion of CPS from old to new methodologies in January 1994, CPS interviewers began training on the new methodology in September 1993. They were given home studies, attended classroom training, and had several months of practice assignments in the field. Moreover, provisions were made for moving some PS interviewers to CPS and others to other Census-collected current surveys. Thus, when the decision was made to continue the PS, almost 65% of the required staff had to be hired and trained in less than two months. In both surveys, the respondents in the overlapping panels between December and January (75%) experienced a shift from one measurement system to another.

Monthly estimates of the overall unemployment rate and the employment-to-population ratio for women from the two measurement systems are presented in the figures at the end of this paper. (The PS was discontinued after May.) While the estimated unemployment rate from CPS was higher in January than December, it does not appear in the first figure as if the expected annual average increase of 0.6 percentage point (0.5 for methodology change and 0.1 for change in population controls) has been observed thus far in 1994. Moreover, while it was known that there could be problems with the PS in the first few months of 1994 due to new interviewers and switching respondents from new to old methodology, the monthly UER estimates continued to be higher than those from CPS. On the other hand, the expected effects of changing measurement systems on other labor force characteristics appear to have occurred. For example, in the second figure, it can be seen that the expected increase, based on the PS, in the employment-to-population ratio for women was observed in CPS after January 1994.

To help users of CPS time series data bridge the gap between the old and new methodologies, the parameters of a linear model with main effects for survey and collection method were estimated by generalized least squares under three sets of restrictions, necessitated by the model not being fully identifiable. While a primary goal of this modeling effort was to develop adjustment factors for analysts, significance tests for main effects were also conducted. Under the restriction that the old and new collection method effects were the same both before and after the switch in January, it was possible to estimate a new method effect and a PS effect. For some labor force items, the estimated new method effect is similar to that estimated by comparing 1993 annual averages from the two surveys. In other cases, most importantly, the overall unemployment rate, this is not true.

There are a number of lessons that can be learned from this experience of parallel testing. Primary among these is that parallel testing should be done if major changes are to be made in a measurement process. In this era of tight budgetary constraints, we were given a rare opportunity of conducting a separate survey. Although it was smaller, for a shorter period of time, and incomplete in an experimental design sense than desired, the PS yielded information that was invaluable in educating users about the possible effects of methodological change. Other ways of parallel testing are possible. For example, BLS and Census tested the

effects of changing the layout of the diary for collecting consumer expenditure information by using 20% of the regular sample, and Statistics Canada plans to use 20% of its Labour Force Survey sample for testing its new questionnaire. While there are problems with pulling out a portion of the survey itself for parallel measurement, just as there were problems with a separate parallel survey, one needs to consider the consequences of not conducting some sort of parallel testing. For example, what would have happened if the unemployment rate for January 1994 (and succeeding months) were estimated to be 0.6 percentage point higher than that for December 1993 and there had been no parallel testing?

It is just as important to remember, especially when analyzing the results for an ongoing rotating panel survey like CPS, that only one observation is available for each of two treatments. Either one or both of the estimates could be from the tail of the sampling distribution. Moreover, one must recognize that nonsampling error can be far more important in analyzing data than sampling error. While one can estimate the sampling variance, nonsampling error is really the focus, since, presumably, the purpose of changing the measurement process is to reduce nonsampling error. Therefore, to the extent possible, it is necessary to address all of the factors that could affect both sampling and nonsampling errors. With the exception of the methodological changes being tested, a parallel survey should mimic as much as possible the entire survey design environment of the main survey. Analysts should also be keenly aware of those features that cannot be mimicked and the resulting limitations on data interpretation. These features include interviewer training and experience, field management procedures, and sample design. If multiple changes are being tested, a sample design that allows for the testing of separate effects should be attempted in order to assess which of the changes might be affecting the nonsampling error structure.

For further details on the results of the CPS parallel testing, see the following reports that can be obtained by contacting the Office of Survey Methods Research, Bureau of Labor Statistics, at 202-606-7370 or fax 202-606-7426. A complete bibliography for the redesign is also available from the above.

Kostanich, Donna L. and Cahoon, Lawrence S. (1994), *Effect of Design Differences Between the Parallel Survey and the New CPS*, CPS Bridge Team Technical Report 3, March 4.

Miller, Stephen M. (1994), *What Would the Unemployment Rate Have Been Had the Redesigned Current Population Survey Been in Place from September 1992 to December 1993?: A Measurement Error Analysis*, CPS Bridge Team Technical Report 1, March 3.

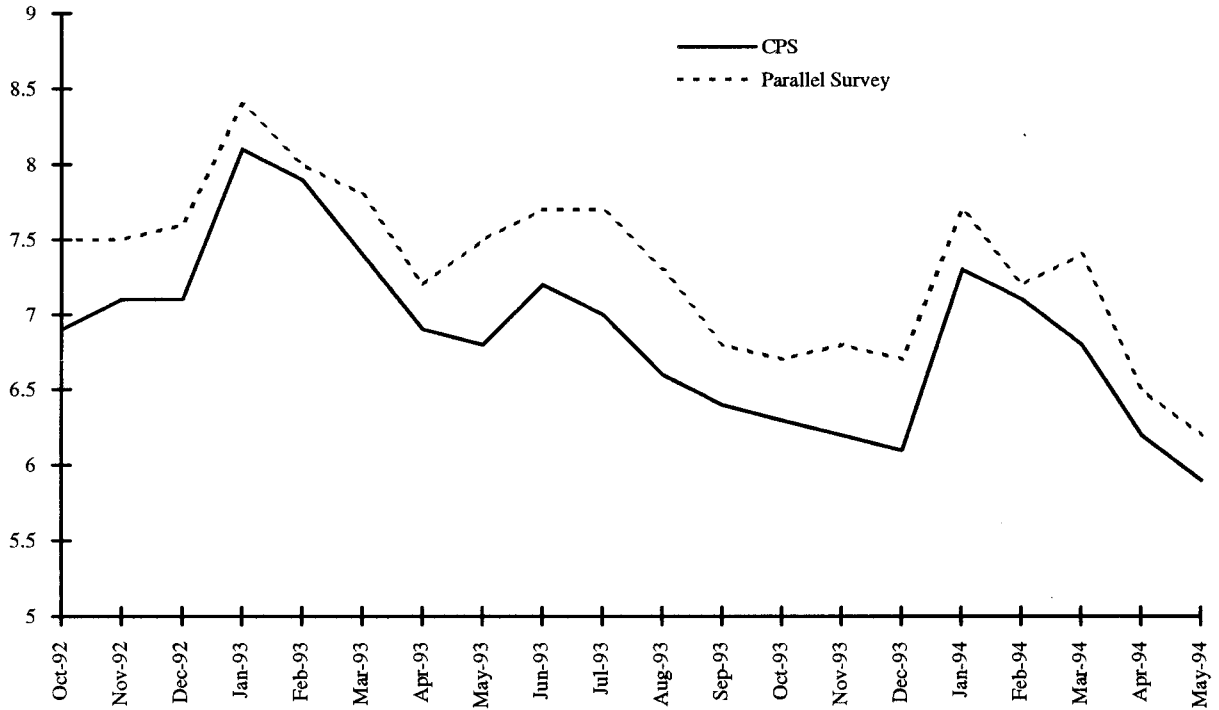
Polivka, Anne E. (1994), *Comparisons of Labor Force Estimates from the Parallel Survey and the CPS During 1993: Major Labor Force Estimates*, CPS Overlap Analysis Team Technical Report 1, March 18.

Polivka, Anne E. and Miller, Stephen M. (1994), *The CPS After the Redesign: Refocusing the Economic Lens*, paper presented at NBER/CRIW Conference, December.

Rothgeb, Jennifer M. (1994), *Revisions to the CPS Questionnaire: Effects on Data Quality*, CPS Overlap Analysis Team Technical Report 2, April 6.

Thompson, Jenny (1994), *Mode Effects Analysis of Major Labor Force Estimates*, CPS Overlap Analysis Team Technical Report 3, April 14.

Alternative Measures of Overall Unemployment Rate
 October 1992 to May 1994 (Not seasonally adjusted, using 1990 weights)



Alternative Measures of Employment-to-Population Ratio for Women
 October 1992 to May 1994 (Not seasonally adjusted, using 1990 weights)

