

Current Population Survey: Three Reprints from the *Monthly Labor Review*



U.S. Department of Labor
Bureau of Labor Statistics

Overhauling the Current Population Survey:

1. Why is it necessary to change?
2. Redesigning the questionnaire
3. Evaluating changes in the estimates

Overhauling the Current Population Survey

Why is it necessary to change?

*In recent years, the CPS has been
hard pressed to keep pace with the changes
in U.S. economic and social life;
as a result, a revamped survey,
geared to the 1990's,
will be initiated in January 1994*

John E. Bregger
and
Cathryn S. Dipbo

The Current Population Survey (CPS)—the most important source of information on the labor force, employment, and unemployment in the United States—is 53 years old this year. It yields perhaps the most eagerly awaited economic statistic—the monthly unemployment rate—and a host of other key data related to trends in employment. Data from this sample survey of 60,000 households across the entire Nation give policymakers, business persons, and academic analysts the key grist for their varied mills.

But the post-World War II years have seen vast changes in the Nation's socioeconomic life. The economy, buffeted by developments at home and on the international scene, has weathered periods of recession and shortages alternating with years of robust growth. The responses to these changes have included serious efforts to meet stiff competition in international markets; the streamlining of many industries and the loss of some; and the rapid growth of the services sector relative to the traditional goods-producing sector. On the social scene, there have been changes in the structure of, and number of workers in, American families; a massive movement of women into the labor force; the arrival of millions of immigrants from poorer nations, who must be assimilated into the society; and a general shift in the places where people work and the types of work they do.

Does the Current Population Survey, as it is currently structured and administered, allow us to fully capture the effects of these developments? The answer appears to be that it could do better. Mounting evidence of the problems prompted yet another wholesale examination of the survey, beginning in 1986. Changes to the survey as a result

of that review will be implemented and tested with the introduction of a new labor force questionnaire in January 1994.

This article presents an overview of the Current Population Survey, and generally discusses the areas slated for change. Anne E. Polivka and Jennifer M. Rothgeb next address specific points of the 1994 revision in greater detail. In the final article in this issue, Chester E. Bowie, Lawrence S. Cahoon, and Elizabeth A. Martin describe the plans to assess the effects on labor force estimates of the new methodological changes to the CPS.

Origins of the CPS

In 1940, with the effects of the Great Depression on economic activity in general and employment in particular still very much in evidence, the U.S. Government launched a monthly sample survey in an attempt to provide timely estimates of the number of employed and unemployed persons. The survey was rather small, involving only 8,000 households. Some experts judged it a shaky proposition because the use of sample surveys to detect trends in a complete population was not fully accepted at the time.¹ Moreover, no one had really agreed on the best way to identify and count the unemployed.

The Works Projects Administration (WPA), an agency created to help drag the Nation out of the depression, had attempted to count the unemployed 3 years earlier, in 1937. This Enumerative Check Census of Unemployment was a postcard survey covering the entire country.² Such a survey presented numerous problems for economists and statisticians alike, because no statistical controls—

John E. Bregger is Assistant Commissioner for Current Employment Analysis, and Cathryn S. Dipbo is Assistant Commissioner for Survey Methods Research, Bureau of Labor Statistics, U.S. Department of Labor.

such as accounting for nonresponse—were possible, but it certainly whetted appetites for more and better data. By late 1939, some forward-looking statisticians with the WPA had designed a national sample and a short questionnaire, based on the notion that one should be measuring employment and unemployment according to the actual *activity* of each member of the household. Following a brief testing period, a survey called the “Monthly Report of Unemployment” was initiated in March 1940.

The timing of the new monthly survey was, of course, intentional. It would give analysts a chance to compare the results of the sample survey with those from the full-blown 1940 Decennial Census. Results from the small sample matched those from the much larger census count very closely, as it turned out. Indeed, the new survey, with its extensive training of interviewers and cautious methods of enumeration, appeared to provide more accurate estimates than did the census—a phenomenon that confounded many nonstatisticians for years to come.

The WPA did not survive many more years, and a permanent home for the new sample survey had to be found. The Bureau of the Census was selected, and what is now called the Current Population Survey—or the more familiar acronym, cps—has been conducted by the agency since 1942. The Bureau of Labor Statistics entered the picture in 1959, when it assumed responsibility for the analysis and publication of the monthly unemployment statistics. And so, in the intervening years, the survey has grown, from one covering 8,000 households with a very simple questionnaire to the present 60,000 sample with a larger set of more complex questions.

With the passage of time, the survey has increased not only in size but also in usefulness (the two are quite naturally related). Much of its value derives from the nature of data collected. Because it is a statistical sample survey, the information collected relates to the employment status of the entire population, not just a portion of it. There is full demographic coverage, so that data are obtained on a broad array of characteristics of people and families. With respect to the labor force (the employed plus the unemployed), data are obtained on industrial, occupational, and class-of-worker status. For the employed, there are data on hours worked, providing information on the full-time and part-time status of workers, and on their usual weekly earnings. By way of example, the availability of full-time/part-time data has been useful in examining the growth of so-called contingent work in the United States, while earnings data—available both on a weekly basis, and annually from supplements to the regular March survey—have

been used with great frequency in examining such things as female/male pay differences over time. For the unemployed, data routinely are collected on duration of unemployment, the respondent’s job status at the time that his or her jobless spell began (lost or left job or labor force entrant), and jobseeking methods used. And, finally, among those not in the labor force, data are obtained for that elusive group, “discouraged workers”—persons who would like to have jobs but do not feel that they can find any. Taken in its entirety, there are a virtually infinite number of analytical possibilities provided by the cps.

But even given the huge amount of data already collected, there is a strong demand for more. The last time that the cps was materially changed was in 1967, when there were marked revisions to the questionnaire. These changes resulted largely from recommendations made by the President’s Committee to Appraise Employment and Unemployment Statistics, known more familiarly as the Gordon Committee, in 1962.³ Not only were a number of marginal changes made in the way employment and, particularly, unemployment were measured, but there was also a significant expansion in the types of data collected, including the beginnings of a data series on labor market discouragement. But that was 26 years ago, and, over the intervening period, a considerable backlog of unfulfilled data requirements has been generated. Adding impetus for change was the fact that, in 1979, another Presidential Commission, the National Commission on Employment and Unemployment Statistics, made recommendations for improving labor force statistics, particularly regarding the measurement of labor market discouragement.⁴

Just as there has been increasing demand for more data over the last two decades, there has been a demand for better data. As statistical research on improving the efficiency of the cps sample design and estimation procedures proceeded in preparation for the sample redesign based upon results of the 1980 census, it became apparent that major reductions in the mean square error of cps-based estimates would have to be made by reducing nonsampling, rather than sampling, error. Although there had always been a recognition of the existence of nonsampling error, in general—and measurement error, in particular—the documentation of rotation group bias during the mid-1970’s provided a quantification not previously available.⁵ Further research on this bias and its causes was limited, in part, due to a lack of a theoretical model and appropriate research methods. Then, in the early 1980’s, the introduction of two new survey research methodologies provided the means for understanding and reducing measurement error. These included the application of behavioral

science theory and methods—more commonly referred to as the cognitive aspects of survey methodology⁶—and computer-assisted interviewing. It is through the blending of these two methodologies that a new collection procedure, which focuses on reducing measurement error, was made possible.

Recognizing an opportunity, the Bureaus of Census and Labor Statistics convened, in 1986, a high-level working group that included the top leadership of both agencies. These experts met for 2 years, during which they formulated an extensive plan to overhaul the cps from top to bottom during the 1990's. Among the undertakings specified were to: 1) Revamp the questionnaire in its entirety, taking into account cognitive interviewing techniques; 2) conduct the survey entirely by computer-assisted means; and 3) develop a state-of-the-art data processing system that would accommodate the first two. These proposals were submitted to the U.S. Office of Management and Budget (OMB) as part of a comprehensive budget request to redesign the cps and were approved for both agencies by OMB and the Congress for the period 1990–96.

Change in data collected

The Bureau of Labor Statistics and Census Bureau researchers who undertook the development of the revised questionnaire considered a number of issues, needs, and recommendations. These included, but certainly were not limited to, the following:

- National Commission on Employment and Unemployment Statistics (Levitan Commission) recommendations;
- Demands for more data that would help resolve or explain certain ambiguities in labor force behavior or trends;
- Better longitudinal data;
- More reliable earnings data;
- Data that would assist in providing information about emerging issues of the times, such as the incidence of contingent work, the “glass ceiling” as a limit on women’s job advancement potential, availability of child care, and so forth;
- Changes in perceptions or meanings of certain words or concepts over time.

These are discussed in greater detail below.

Levitan Commission recommendations. The most recent examination of the Nation’s labor force statistics, including their underlying technical and conceptual foundation, was conducted by the Presidentially appointed National Commission on Employment and Unemployment Statistics (NCEUS) in 1978–79. The report that the NCEUS

issued on Labor Day 1979 had no specific recommendations for changing the concepts of employment and unemployment but did recommend changes in the way labor market discouragement was measured.

At present, discouragement is determined by asking persons who are not working at all in the survey reference week, and who have not looked for work during the prior 4-week period, if they would like to work “now,” either full or part time, with those answering in the affirmative being asked the reason why they are not seeking a job. Those who respond with a job market-related reason, such as their belief that work is not available, the fact that they gave up their search, or the effect of some personal factor (age, race, lack of skill, and so forth), are then counted as “discouraged workers.” Recognizing the inherent subjectivity in such a concept, the NCEUS recommended a measurement based on evidence of prior search and current desire and availability for a job. (This would be in keeping with the way discouragement is measured in the Canadian labor force survey.) The NCEUS recommendation was accepted by the Secretary of Labor 2 years later,⁷ but could not be implemented without the use of a separate overlap sample, which would determine if the changed concept had an impact on other measures, such as the unemployment rate. Therefore, any change had to await the availability of funding for questionnaire testing and a separate panel of households (overlap sample).

Other NCEUS recommendations for questionnaire development included the expressed need for more detailed information on persons not in the labor force, the “terms and conditions under which the unemployed are seeking work,” data on usual versus actual hours, and more information on the job-changing patterns of employed persons. The Commission also devoted an entire chapter of their report to the desirability of “measuring labor market related hardship,” implying the need for more and better data on earnings and on marginal labor force activities.

Other data needs. A commonly cited issue in the labor force field has been the large, often inexplicable, differences in employment levels and trends emanating from comparisons of data from the cps and those from the BLS survey of nonfarm establishments. These differences exist even after attempts to account for measurable differences through “reconciliation exercises,” which account for known variations, particularly coverage. Historically, employment estimates from the two surveys did tend to track fairly well over time, if not on a month-to-month basis. However, even that relationship broke down in the late 1980’s, when employment growth as measured by the establish-

ment survey significantly outstripped that measured in the household-based cps.

One factor that clearly played a role in the observed divergence was multiple jobholding, which is measured infrequently in cps supplements. (In the monthly measurements, employed persons are counted only once in the cps, regardless of the number of jobs they hold, whereas, in the establishment survey, persons with multiple jobs are counted for each job they hold.) It was estimated that increases in the number of persons working at two or more jobs accounted for as much as 65 percent of the difference in employment growth measured by the two surveys over the May 1985–89 period—May being the month in these 2 years during which data on dual jobholding was collected in special cps supplements.⁸ Thus, regular collection of data on multiple jobholding could prove quite useful, not only for shedding light on the survey differences in measuring employment, but also for determining once and for all the total counts of full-time and part-time jobs in the United States—a very important data need.

The measurement of self-employment in the United States is of continuing interest to data users concerned with the extent of business formation in this country. Also entering in are issues surrounding marginal work, such as unpaid family work (which must be at least 15 hours a week to be counted as employed in the cps). This is another area in which cps measures have been somewhat deficient, and in which improvements thus are desirable.

Better longitudinal data. Because of its 4–8–4 rotation scheme, wherein households enter the survey for 4 months, exit for 8 months, and then return for another 4 months, the cps provides the potential for tracking persons over a total of 16 months. In practice, there are difficulties with this methodology, due to such problems as rotation group bias, faulty recall, response bias, and the like. There is therefore strong interest in making a number of improvements that would reduce the likelihood of spurious, or otherwise inaccurate, changes being reported in the survey from month to month.

The use of dependent interviewing would go the furthest in solving this problem. With respect to the measurement of an employed person's occupation, for example, it has been determined that presently about a third of the cps participants who are employed in 2 consecutive months change their occupation. Testing suggests that a high proportion of these changes are spurious, that is, the persons did not change their occupations; what in fact happened is that different descriptions of their duties and responsibilities were reported, resulting in their being classified in a different

occupation. A more accurate estimate of the "true" monthly change in occupation would be about 10 percent from month to month.⁹ This problem could be "corrected" if the prior month's information could be made available to the interviewer, and only those persons who indicate that they have changed employers or that their duties and responsibilities have changed would be asked about their current jobs; otherwise, the prior month's status would be carried forward.

Similar problems exist with measuring the duration of unemployment. Norman Bowers and Francis W. Horvath found a significant negative relationship between the unemployment duration reported during the first interview and the reported change between months.¹⁰ They concluded that the average duration of unemployment may be overstated by at least 2 weeks. Estimates might be improved for persons found to be unemployed in both the current and previous months if the number of weeks the persons had been jobseeking were to be carried forward automatically from the previous month's count.

More earnings data. Whereas most cps data users are familiar with the annual March supplement, with its full array of income, earnings, and work experience data covering the previous calendar year, fewer users are aware that data on the "usual weekly earnings" of wage and salary workers, and the hourly rate of pay for those paid by the hour, have been collected monthly in the cps since 1979 (in the two out-going rotation groups, 4 and 8). These data (tabulated and published quarterly) have been used, for example, to study trends in the female/male earnings ratio, which was up to 76 percent for full-time workers in the second quarter of 1993.¹¹ Given the considerable potential value of these data, there is interest in furthering their usefulness—first, by improving the accuracy of the measures, and then by expanding to some degree the amount of data available.

Data on other important current issues. How many contingent workers are there in the United States? Does the glass ceiling still exist, or alternatively, is it being gradually pierced? How many workers have child care problems? People look to the cps to answer these and many other timely questions. If items were added to the cps questionnaire relating to weekly hours of work and if more detail became available on the reasons why people work less than full time (defined as fewer than 35 hours per week), the resulting data could prove useful in the examination of both the contingent work force and the need for child care. Better estimation of occupational and earnings data (referred to above) would be helpful (but certainly not all-serving) in answering the glass-ceiling

question. There are undoubtedly other issues, not yet on the national (or local) agenda, on which survey improvements of this sort could shed light.

Changes in perceptions or meanings. An excellent example of this phenomenon involves the measurement of persons "on layoff." In the past, most persons defined layoff as a temporary separation, whereby there was an expectation of recall as soon as business conditions improved. More recently, we have come to recognize that the term layoff has taken on a much broader meaning. For example, in recent focus group exercises, when persons were asked "Do you have a job from which you are temporarily laid off?", some respondents who had permanently lost their jobs answered in the affirmative.¹² In other words, when used in isolation, the term now signifies a degree of permanence to many respondents. This recognition is crucial to proper measurement, if only to get a better fix on how many people are truly laid off, in order to differentiate this group from those who are unemployed for other reasons—especially those jobless due to permanent discharge. In the present questionnaire, the distinctions are insufficiently clear, and thus revised wording is necessary to ensure the precise meaning. A likely outcome would be fewer persons enumerated as being on layoff and more persons assigned to the "other job losers" category.

The data collection process

Whereas the 1967 questionnaire changes were prompted by a shift in the concepts of employment and unemployment, changes in the mode of collection have, historically, been the result of the need to cut costs (for example, the increase in telephone interviewing during the 1970's) or reduce processing time (as in the case of the switch to FOSDIC¹³ in 1961). For the 1994 revisions, improving data quality is the primary reason for changing both the questions and the mode of collection.

Despite the existence of what theoretically may appear to be very concise definitions of labor force status, the conversion of concepts into a series of questions for determining status is not straightforward.¹⁴ While, for most people, the measurement of labor force status is robust and reliable, there is a small minority for which it is not.

To reduce measurement error, it is necessary to investigate the processes of question asking, question answering, and the interaction between respondent and interviewer.¹⁵ The question asking task has been characterized as encompassing question reading or recall and question formulation. In fact, the interviewer's task goes beyond question asking to include categorizing responses

to open-ended questions when the questionnaire provides only a small set of response categories, and converting sometimes lengthy descriptions of kind of work or important activities or duties into responses that can be entered on one 5-inch line.

Question sequencing has long been recognized as having an effect on data quality.¹⁶ In a cps-based example, questions for persons not in the labor force were moved from the first- and fifth-month interviews to the fourth- and eighth-month interviews in January 1970, because there was evidence that the original timing of the questions resulted in unacceptably high reporting of additional unemployed persons as well.¹⁷

To minimize interviewer error associated with question wording, cps interviewers are trained to pose the questions exactly as worded. Yet, a study using 1969 and 1971 data showed that the question "Did . . . have a job or business from which he was temporarily absent or on layoff last week?" was reworded by the interviewer in 18 percent of the observed interviews.¹⁸ Interviewers participating in focus groups indicate that this question and others are difficult to ask and often are too complex for respondents to understand.¹⁹

Under certain circumstances, interviewers are given instructions to deviate from the script and to formulate their own questions. For example, interviewers are to probe about unpaid work if the household is involved with a farm or business. Yet, the current questionnaire contains no explicit question to identify such households, implying that the use of such probes can be "iffy" at best.

Instructions to interviewers on how to proceed—termed "question sequencing"—are given using two methods, directional arrows or italicized instructions. In the absence of either method, the interviewer's default procedure is to proceed in numerical order. As a result, questions are sometimes skipped; the missing data rate for some questions can be as high as 4 percent.

The respondent's task of question answering is not always easy either. One model of question answering includes four tasks: question comprehension, knowledge recall, judgment, and response verbalization.²⁰ Comprehension may be a problem for respondents, because some of the words and terms used in cps questions have multiple meanings (such as "job"); some are given a broader interpretation by respondents than intended by BLS (for example, "on layoff"); and some have diverse meanings (as in "looking for work").²¹

Recall of the many nonsalient activities of everyday life is complicated by their storage in memory as *schema* rather than as specific episodes. For example, people have a typical workweek. To answer accurately a question such as "How many hours did you work last week?", they would have to access their episodic memory to re-

Overhauling the cps: Why Change?

construct the actual days worked, beginning and ending times, and periods of absence. Motivation on the part of respondents to search their memories for specific episodic events varies. Search strategies also vary. In the case of proxy respondents—persons who respond on behalf of other household members—the specific information may not be known at all.

Respondents must judge the adequacy and relevance of recalled information. They may recall incorrectly but judge what they recall as accurate. Or they may engage in "satisficing behavior," such as providing an adequate but incomplete answer when a more substantial cognitive effort is needed to answer a question completely and accurately.²² Thus, instead of trying to recall specific hours worked last week, many people respond with an estimate of their usual hours worked. They may also recall correctly but judge the information as irrelevant. For example, there is evidence that adults may use a different standard for judging an activity, such as looking for work, when the subject is a child (teenager) rather than an adult.²³

A number of factors can influence a respondent's decision at the communication stage, including the sensitivity of the question, the social desirability of the answer, and the expectation of the interviewer. Respondents may feel a need to give a favorable impression of themselves. Both respondent and interviewer characteristics interact with social desirability. For example, it has been discovered during laboratory studies of proxy reporting that there are respondents who answer the cps questions as if they went to work last week but who, if their spouses are to be believed, are actually unemployed or not in the labor force.²⁴ In one case, the proxy even commented that her spouse was ashamed of not having a job and often misrepresented his labor force status.

Interviewer processing of a response parallels the question answering process, consisting of comprehension, recall, judgment, and recording. Judgment plays an important role in classifying responses to open-ended questions. Interviewers may judge some responses as inadequate and probe for details; others may check the category "other" and write in the response. (Interviewers are trained to include only legitimate job search activities in "other," but a recent study showed that 32 percent of "other" responses should have been classified into one of the prespecified categories.²⁵)

The physical form of the questionnaire is not an insignificant factor in the information exchange between respondent and interviewer. Inconsistencies in instructions to the interviewers, inadequacies in question sequencing, and incompleteness of questions for certain circumstances are in part due to the space limitations of the current FOSDIC

booklet. The separation of the household roster forms, which are retained in the Census Bureau regional offices, from the labor force questionnaires, which are new each month and contain no pre-printed information, requires interviewers to re-enter basic demographic information each month on the labor force questionnaires.

The mode of interviewing (personal visit, telephone, computer-assisted telephone contact) also can affect response.²⁶ After 24 months during which centralized computer-assisted telephone interviewing (CATI) was used in the cps, the unemployment rate for the CATI test panel (not all data for sample households in the CATI-designated panel are actually collected using CATI) was found to be 0.8 percentage point higher than for the control panel.²⁷

The new cps questionnaire has been developed with an eye toward minimizing these and other problems. It is recognized, of course, that they cannot be eliminated altogether. The guiding principle has been to reduce measurement error by aiding the cognitive processing of both respondents and interviewers. Comprehension is enhanced by including definitions and by dividing questions involving multiple concepts. Recall is aided by providing respondents with a strategy for remembering their activities. Judgment is augmented by including probes in the survey instrument and using the computer to do calculations. Better communication is achieved with the improved flow of the interview through the use of the computer. Interviewers feel more professional, and respondents feel less threatened.

The use of a computer as a survey instrument provides a powerful tool that allows greater flexibility in question wording and sequencing. The computer permits the tailoring of question wording and complex branching among questions to meet specific situations, thus relieving interviewers of a significant burden. It allows for information collected in a previous interview to be used in the current interview to lighten the task of both respondents and interviewers. Automated collection procedures also provide the means for extending the types of data collected in order to meet the demands of the many users without a significant increase in respondent burden.

A QUARTER OF A CENTURY has passed since the last major revision to the Current Population Survey. During that time, our society has changed greatly and the demand for information has exploded. Increasingly, legislators and policymakers are looking to the Federal statistical system to provide them with accurate and timely data on a wide variety of issues. While trying to meet some of these ever-increasing demands for data, we must not lose sight of our primary measurement goal: to

provide a monthly snapshot of the activities of our population of working age—how many people are working, how many are looking for work, and how many are out of the labor force altogether. Moreover, we have a commitment to providing the best quality data possible, using a measurement process that takes advantage of the

significant improvements in survey methodology. It is believed that the introduction in 1994 of the "new" cps—the culmination of a multi-year project undertaken by BLS and the Bureau of the Census to revise and modernize the survey—will go far toward meeting the Nation's needs in this regard. □

Footnotes

¹ J.W. Duncan and W.C. Shelton, "U.S. Government Contributions to Probability Sampling and Statistical Analysis," *Statistical Science*, vol. 7, no. 3, 1992, pp. 320–28.

² G. Bancroft, *The American Labor Force* (New York, John Wiley and Sons, Inc., 1958), p. 183.

³ President's Committee to Appraise Employment and Unemployment Statistics, *Measuring Employment and Unemployment* (Washington, U.S. Government Printing Office, 1962).

⁴ National Commission on Employment and Unemployment Statistics, *Counting the Labor Force* (Washington, U.S. Government Printing Office, 1979).

⁵ B.A. Bailer, "The Effects of Rotation Group Bias on Estimates from Panel Surveys," *Journal of the American Statistical Association*, March 1975, pp. 23–30; and C.A. Brooks and B.A. Bailer, "An Error Profile: Employment as Measured by the Current Population Survey," Statistical Policy Working Paper 3 (Washington, Office of Federal Statistical Policy and Standards, 1978).

⁶ T. Jabine, M. Straf, J. Tanur, and R. Tourangeau, eds., *Cognitive Aspects of Survey Methodology: Building a Bridge Between Disciplines* (Washington, National Academy Press, 1984).

⁷ "Final Report of the Secretary of Labor on the Recommendations of the National Commission on Employment and Unemployment Statistics" (U.S. Department of Labor, October 1981).

⁸ J.F. Stinson, Jr., "Multiple jobholding up sharply in the 1980's," *Monthly Labor Review*, July 1990, pp. 3–10.

⁹ C.L. Collins, "Comparison of Month-to-Month Changes in Industry and Occupational Codes with Respondent's Report of Change: cps Job Mobility Study," Response Research Staff Report 75-5 (Bureau of the Census, May 15, 1975).

¹⁰ N. Bowers and F.W. Horvath, "Keeping Time: An Analysis of Errors in the Measurement of Unemployment Duration," *Journal of Business and Economic Statistics*, April 1984, pp. 140–49.

¹¹ "Usual Weekly Earnings of Wage and Salary Workers: Second Quarter 1993," *News* (Bureau of Labor Statistics, July 21, 1993).

¹² M. Palmisano, "Respondent Understanding of Key Labor Force Concepts Used in the cps," Paper presented at the Annual Meeting of the American Statistical Association: Section on Survey Research Methods, 1989.

¹³ FOSDIC stands for Film Optical Sensing Device for Input to Computers. "The Current Population Survey: Design and Methodology," Technical Paper No. 40 (Bureau of the Census, January 1978), p. 3.

¹⁴ B.A. Bailer and N. Rothwell, "Measuring Employment and Unemployment," in C.F. Turner and E.A. Martin, eds., *Surveying Subjective Phenomena* (New York, Russell Sage Foundation, 1984), pp. 129–42.

¹⁵ J.E. Sander, F.G. Conrad, P.A. Mullin, and D.J. Herr-

mann, "Cognitive Modelling of the Survey Interview," *Proceedings of the American Statistical Association: Section on Survey Research Methods*, 1992, pp. 818–23.

¹⁶ H. Schuman and S. Presser, *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context* (New York, Academy Press, 1981).

¹⁷ Bailer, "The Effects of Rotation Group Bias."

¹⁸ A. Rustemeyer and N.D. Rothwell, "Observation on cps," *Bureau of the Census Memorandum* (Bureau of the Census, Dec. 8, 1969); and A. Rustemeyer and N.D. Rothwell, "Observation of cps Interviews in January 1970," *Bureau of the Census Memorandum* (Bureau of the Census, Dec. 29, 1971).

¹⁹ "Response Errors on Labor Force Questions Based on Consultations with Current Population Survey Interviewers in the United States," Paper presented before the OECD Working Party on Employment and Unemployment Statistics (Bureau of Labor Statistics, 1988).

²⁰ C.F. Cannell, P.V. Miller, and L. Oksenberg, "Research on Interviewing Techniques," in S. Leinhardt, ed., *Sociological Methodology* (San Francisco, Jossey-Bass, 1981), pp. 389–437; and R. Tourangeau, "Cognitive Science and Survey Methods," in Jabine and others, eds., *Cognitive Aspects of Survey Methodology*, pp. 73–100.

²¹ E.A. Martin, "Some Conceptual Problems in the Current Population Survey," *Proceedings of the American Statistical Association: Section on Survey Research Methods*, 1987, pp. 420–24; and C.S. Dipbo and J.L. Norwood, "A Review of Research at the Bureau of Labor Statistics," in Judith M. Tanur, ed., *Questions about Questions* (New York, Russell Sage Foundation, 1992), pp. 271–90.

²² J.A. Krosnick, "Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys," *Applied Cognitive Psychology*, no. 5, 1991, pp. 213–36.

²³ J.M. Tanur and J.K. Lee, "Youth Unemployment Measured by the cps," *Proceedings of the American Statistical Association: Section on Survey Research Methods*, 1992, pp. 667–72.

²⁴ T.J. Demaio, "Social Desirability and Survey Measurement: A Review," in Turner and Martin, eds., *Surveying Subjective Phenomena*, pp. 257–82; and P.A. Mullin and B.E. Tonn, "Factors Associated with Proxy Knowledge of Employment-Related Information," Paper presented at the Annual Winter Meetings of the American Statistical Association, Fort Lauderdale, FL, January 1993.

²⁵ M.P. Fracasso, "Reliability and Validity of Response Categories for Open-ended Questions in the Current Population Survey," *Proceedings of the American Statistical Association: Section on Survey Research Methods*, 1989, pp. 481–85.

²⁶ Bailer and Rothwell, "Measuring Employment."

²⁷ L.S. Cahoon, "CATI Phase-In Analysis: Results of cps CATI Phase-In Analysis: January 1991–December 1992," *Memorandum for Distribution* (Bureau of the Census, Mar. 8, 1993).

Overhauling the Current Population Survey

Redesigning the CPS questionnaire

The wording and order of questions, the survey instrument, and the interaction of interviewers and respondents all are vital to the success of a survey; modifications addressing these factors should greatly reduce the labor force misclassification of individuals in the "new" CPS

Anne E. Polivka
and
Jennifer M. Rothgeb

Probably the single most politically sensitive number published by the Federal Government is the seasonally adjusted monthly unemployment rate. This measure, along with other information about the U.S. labor force, such as earnings, number of hours worked, and job search intentions of those not in the labor force, is calculated using data collected through the Current Population Survey (CPS). Nevertheless, despite the importance of the statistics derived from the survey, and the changing American economy, the CPS has remained virtually unchanged since 1967.

In 1986, the Bureau of Labor Statistics (BLS), in collaboration with the Bureau of the Census, began a program to modernize the CPS. An integral part of this effort was evaluating and redesigning the survey questionnaire. The result is a completely redesigned questionnaire, which will be implemented in January 1994. This article briefly elaborates on the history of and concepts underlying the CPS questionnaire, as discussed by John E. Bregger and Cathryn S. Dippo on pages 3-9. Its chief focus, however, is the new questionnaire: the need for the redesign; methods used to test alternative versions; comparisons of the revised questionnaire with the current one; and the extent to which labor force misclassification appears to be reduced through the redesign.

Background in brief

Since its inception as a national survey of sample households in 1940, the CPS has based its measurement of employment and unemployment on individuals' activities. However, the implementation of these activity-based measurements (and the auxiliary information collected with the CPS) has undergone some alterations throughout the years.

The most fundamental changes to the CPS questionnaire occurred in 1945 and 1967. In 1945, four standardized questions were incorporated to ascertain whether individuals were employed, unemployed, or not in the labor force. Previously, enumerators had assigned labor force classifications during the interview by following a complicated prioritization scheme. However, special studies conducted at the time demonstrated that the lack of specifically worded questions resulted in the exclusion from the labor force of a large number of part-time and intermittent workers, and created inconsistencies among individuals' labor force classifications. The introduction of specific questions ensured uniformity in data collection and relieved enumerators of the burden of applying complicated prioritization schemes.¹

In the 1967 revision, these four standardized labor force questions were maintained. However, based on the recommendations of the President's Committee to Appraise Employment and Unem-

Anne E. Polivka is an economist with the Office of Research and Evaluation, Bureau of Labor Statistics. Jennifer M. Rothgeb is a social science statistician with the Center for Survey Methods Research, Bureau of the Census.

ployment Statistics (the Gordon Committee), a 4-week timeframe was inserted in the question to determine if an individual had looked for work within the preceding 4 weeks. Separate questions about methods of job search and ability to take a job also were added.² Since 1967, to be classified as unemployed, an individual has to have *actively* looked for work within the last 4 weeks.

Other changes based on the Gordon Committee's recommendations included: raising the cut-off age for exclusion from the labor force from under 14 years to under 16 years; inclusion of two questions about extra hours worked and time taken off by persons working between 35 and 48 hours per week; addition of probing questions about the duration of unemployment, and about self-employment; and inclusion of questions asking if individuals who had not recently tested the job market wanted to work and, if so, why they had not searched for jobs.³

Since 1967, there have been no major changes to the cps questionnaire, although earnings questions were added in 1979, questions about union membership and union coverage were incorporated in 1984, and inquiries about school enrollment were included in 1985.

In the late 1970's and 1980's, revisions to the survey were proposed, most notably by the National Commission on Employment and Unemployment Statistics (the Levitan Commission). Changes based on these recommendations were tested by the Bureau of the Census in the early 1980's through the Methods Development Survey, jointly developed by BLS and the Census. No major changes were implemented, however, due to the lack of funding for a large overlap sample necessary to assess the effects of questionnaire changes.

The need for change

The U.S. socioeconomic environment has changed radically since 1967, and modifications to the cps questionnaire to account for these changes are long overdue. The primary goal in revising the questionnaire is to obtain more precise and consistent measures of the labor force. An auxiliary, but still important, reason for the change is to collect more information about the U.S. population of working age.

The major way to obtain more precise measures of the labor force is to reduce the misclassification of individuals' labor force status that results from interviewer or respondent error or problems with the questions asked. In the recent redesign of the cps, four ways were identified to reduce misclassifications: 1) clarification of existing definitions, 2) incorporation of definitional changes proposed by various commissions, 3) revision of question wordings and

sequencings, and 4) computerization of the survey collection process.⁴

Clarification of existing definitions. Identification of labor force activities frequently requires cps respondents to understand specific concepts. However, several of these concepts are not explicitly defined in the questionnaire. Better implementation of existing definitions entails clarifying these concepts, rather than assuming that all survey participants employ a common definition. Problematic concepts that were targeted for clarification, include: "work"—because previous research indicated that marginal work activities, such as unpaid family work and intermittent-type work were missed, while volunteer work sometimes was mistakenly captured in the survey; "on layoff"—because evidence indicated that between 30 to 50 percent of "laid off" respondents did not include the expectation of recall in their definition of layoff; and "looking for work"—because it was suspected that passive job search methods, such as reading newspaper advertisements, were inappropriately being coded as active search methods, causing some individuals to be falsely classified as unemployed.⁵

Existing definitions also could be improved by building definitions of phrases and terms directly into the questions. Two phrases specifically identified as needing to be explained were "last week" and "full-time work." Previous research indicated that only 17 percent of respondents used a definition of "last week" that matched the cps definition of Sunday through Saturday. The majority, 54 percent, defined last week as Monday through Friday.⁶ The Levitan Commission recommended defining the phrase "full-time work" in the questionnaire, because the cps definition of full-time work as 35 hours per week or more is not in accord with many legal definitions and may not match commonly held notions of such work.⁷

A third way of clarifying existing definitions is to ask explicitly for information rather than inferring it from answers provided. One example in which information currently is assumed rather than explicitly requested involves the classification of individuals as involuntarily part time (part time for economic reasons). Underlying the concept of part time for economic reasons is the assumption that individuals who are involuntarily working part time want, and are available, to work full time. However, in the current cps, individuals who are classified as part time for economic reasons are never explicitly asked about their desire or availability for full-time work. Rather, both are assumed from the reasons respondents provide for working part time. For example, individuals who say they are working part time because they are too busy with housework, school, or personal busi-

ness are assumed not to be available for or to desire full-time work. But, individuals who say they are working part time because they cannot find full-time work are assumed to be available for and to want such work.

Incorporation of recommended definitional changes. The Levitan Commission criticized the discouraged worker definition for its reliance on implicit information, as well as the subjectivity of its definition. The Commission recommended that the definition of "discouraged workers" be more precise, and that the possibility of misclassification be reduced by including direct questions about recent prior job search and current availability for work.

Revisions to question wording and question sequencing. A third way identified to reduce labor force misclassification involves improving question wording and resequencing questions within the questionnaire to reduce error in the interaction between respondents and interviewers. One means of doing so involves splitting long, complex questions into two or more separate items. Perhaps the most notoriously complex question in the current cps is the one asking about *both* temporary absences and layoffs—"Did you have a job or business from which you were temporarily absent or on layoff LAST WEEK?" Also, the follow-up question, "Why were you absent from work LAST WEEK?", could confuse individuals who may not consider their layoffs as absences.

A second means of restructuring the questionnaire to minimize error is to reduce reliance on volunteered information. An example of a case in which interviewers must rely on volunteered information to classify individuals is the counting of unpaid family workers. The current question about last week's work activities has parenthetical instructions telling the interviewer to ask about unpaid work if the household has a farm or business. The current questionnaire, however, does not provide a mechanism for interviewers to directly establish the existence within a household of a business or farm. The interviewer may not learn of a family business until a person is identified as self-employed through the industry and occupation questions. This may be after data collection has been completed for other household members who may also be involved in the business.⁸

A third technique for reducing errors is to incorporate memory aids into the questionnaire. For example, previous research indicates that the question about actual hours worked in the current cps is cognitively difficult for respondents to answer. The hours question is supposed to measure how many hours respondents actually

worked in the reference week, taking account of extra hours worked, time taken off due to illness or vacation, and hours worked on secondary jobs. It is suspected, however, that the hours reported as "actual" are a mixture of usual hours, exact actual hours, and some approximation of actual hours.⁹

A fourth means of reducing error in the interaction between interviewers and respondents is changing the wording of questions to reflect current sociological and economic conditions. Questions in need of updating occur in both the employed and the unemployed series. For example, current question wordings do not adequately reflect the growth in numbers of working mothers or the increased incidence of paid work done at home. Interviewers are instructed to tailor the first question in the labor force series to the age and gender of the respondent. Specifically, if the respondent "appears to be a homemaker," the interviewer is supposed to ask "What were you doing most of last week—keeping house or something else?" If the respondent seems relatively young, the interviewer asks "What were you doing most of last week—going to school or something else?" For all other respondents, interviewers are to ask "What were you doing most of last week—working or something else?" The next question currently is "Did you do any work at all LAST WEEK, not counting work around the house?" Because home offices and other work arrangements that involve individuals working in their homes have become more prevalent, the phrase "not counting work around the house" could be confusing. The combination of the interviewer's tailoring of the first work question and the phrase "not counting work around the house" is outdated, and could be offensive to many respondents.

A fifth way to reduce nonsampling error in the questionnaire involves revising the precoded response categories. In the current questionnaire, approximately half of the questions are "open-ended"—questions that do not include specific possible responses *in the body of the question* or that are not simple "yes/no" questions. For the majority of these questions, interviewers have available a fixed set of response categories to record respondents' answers. Research by Maria P. Fracasso indicates that the accuracy of coding in the cps is reduced where response categories are inappropriate, nonexhaustive, or not mutually exclusive.¹⁰ For example, examination of the coding of the reasons for being temporarily absent from work indicated that 15.5 percent of the responses were categorized as "other." The vast majority of these responses could have been classified as maternity leave if such a category had been provided in the survey instrument.

Computerization of the collection process. The final method identified to reduce labor force misclassification was the automation of the survey instrument so that all interview data could be collected using a computer. Starting in January 1994, all interviews will be conducted using a combination of laptop computers and centralized computer telephone interviewing facilities.

Computerization of interviews can reduce the incidence of misclassification errors in several ways. For example, the computer will automatically skip interviewers from question to question. It also will automatically tailor questions, inserting appropriate pronouns or names and changing verb tenses. This automation will relieve interviewers of the burden of correctly following instructions to skip certain questions on the basis of answers to earlier questions, and will ensure that respondents move smoothly through the questionnaire. Automated skip patterns also permit the use of more appropriate and detailed questions, geared towards specific groups or situations.

In addition, computerization makes it possible for editing, verification procedures, range checks, and consistency checks to take place during the interview. The use of these procedures should improve data quality by detecting possible errors while respondents can still aid in the correction of the data. Without these automated procedures, inconsistent answers can be "corrected" only through imputation in the postinterview processing.

Computerization also improves the quality of the data by allowing the automatic transmission of household and demographic information for the cases interviewers are scheduled to interview. Currently, all demographic background data for a cps household must be kept within the survey supervisors' offices. This requires interviewers to transcribe all demographic and identifying information onto the paper questionnaires being used for the current interview. Furthermore, while a full accounting of an interviewer's assignment on a housing-unit basis is undertaken clerically at the supervising office, there is currently no accounting for each person listed on the household roster. Specifically, there is no attempt to ensure that information has been collected for each person listed on the survey supervisor's demographic control card for the household for the month. Both the hand transcribing of information and the lack of complete checks on respondent enumeration will introduce error into the data. The computerization of the transmittal of demographic information eliminates many of these constraints, and thus will improve the quality of the data.¹¹

Expansion of collected information. Although not a major reason for the questionnaire rede-

sign, the new cps questionnaire will expand the amount of data collected. One group about whom it was deemed necessary to collect additional information is multiple jobholders. Consequently, in the redesign, information about the number of multiple jobholders will be collected monthly. In addition, it was also decided to collect information on the number of hours multiple jobholders work on their main jobs and secondary jobs separately, along with information on the industry and occupation of the second job. With this information, it will be possible to track the number of dual jobholders over the business cycle; to obtain a count of the number of individuals in part-time jobs, as opposed to the number who are working part time (individuals who hold two or more part-time jobs which total more than 35 hours are currently classified as full-time workers); and to reconcile, at least partially, differences in employment measures between the cps and establishment surveys.

There also was a recognized need for information on: persons earning overtime pay, tips, or commissions; the usual hours for all workers, and not only for those who worked less than 35 hours in the reference week; and data on potentially discouraged workers for all those in the sample, rather than just those in outgoing rotation panels. Obtaining this additional information will make possible a more complete assessment of labor force dynamics.

Preparation for the redesign

Preliminary work. In 1986, a joint BLS and Bureau of the Census task force convened to identify problems with the current questionnaire, suggest possible solutions for those problems, and develop a plan for research to improve and test the design of the questionnaire along with related survey procedures.¹²

Beginning in 1986, the two agencies conducted a number of research projects to obtain information necessary to guide the development of a revised cps questionnaire. The projects included interviewer focus groups, respondent focus groups, respondent debriefings, a test of interviewers' knowledge of concepts, indepth cognitive laboratory interviews, response categorization research, and a study of respondents' comprehension of alternative versions of labor force questions.¹³

Evaluation and selection of new questions. Based on recommendations from the Levitan Commission, suggestions made by the 1986 Census-BLS Questionnaire Redesign Task Force, and the results of BLS-Census research projects conducted prior to 1990, two alternative versions of the cps questionnaire were developed. These versions

were compared to the current questionnaire in a test conducted from July 1990 to January 1991, during which data were collected for more than 72,000 individuals. The results of the test were used to develop a single alternative version of the questionnaire. This version was compared to the current version of the questionnaire in a second phase of testing conducted from July 1991 to October 1991, during which data were obtained for more than 32,000 individuals.

Both phases of testing used computer-assisted telephone interviewing (CATI) and a random-digit dialing (RDD) sample. Questionnaire versions were assigned randomly to households, and once a version was assigned, the household was interviewed for 4 consecutive months using the same questionnaire. The results of both phases of testing were used in development of the final revised CPS questionnaire.

Both quantitative and qualitative methods were used to select questions and to identify problems along with potential solutions. Analysis was based on behavior coding of interviewer-respondent interactions, interviewer debriefings, respondent debriefings, response distributions, and item nonresponse measures.

Behavior coding of interviewer-respondent interactions entails monitoring or tape recording actual interviews and keeping a quantitative record of those questions that interviewers read incorrectly or that respondents had difficulty answering. Using a specially developed form, researchers from BLS and Census coded interviewer-respondent interactions from July 1990 to December 1990 and from July 1991 to October 1991. Interviewer behaviors that were recorded included reading the questions exactly as worded, making minor or major changes in the question wording, posing of probing questions, and adopting appropriate verification procedures. Respondent interchanges that were recorded included providing an adequate answer, giving an inadequate answer, and requesting clarification of a question.¹⁴

Interviewer debriefings tap interviewers' opinions about, and experiences with, the questionnaires being tested. The CPS CATI/RDD research used two interviewer debriefing techniques: 1) self-administered debriefing questionnaires, and 2) focus group discussions. Interviewers were first asked to complete a self-administered questionnaire. Several weeks later, they were brought together into focus groups consisting of eight to ten people for a more free-form discussion of their opinions and impressions. Questions that were asked of interviewers included: 1) which questionnaire version flowed the best/worst, 2) which series of questions seemed most difficult for respondents to answer, and 3) which question did respondents refuse to answer most often. In addition, interviewers were

asked which concepts or terms they thought were most commonly misinterpreted or misunderstood by respondents. They also were invited to cite any other problems in the tested surveys.¹⁵

Field-based respondent debriefings use a series of question-specific probes to ascertain whether certain words, phrases, or situations are understood by respondents in the manner intended by the questionnaire designers. In the CPS CATI/RDD study, respondents were administered the respondent debriefing after their final monthly interview. The respondent debriefing consisted of two parts. Ninety percent of the respondents received two or three sets of followup questions that were keyed to responses given during the main labor force survey. Ten percent of the respondents were read short hypothetical vignettes and then asked to classify the scenario presented in each vignette. (For example, respondents were asked to classify an individual in a vignette as "working" or "not working.") Among other purposes, these respondent debriefing questions were designed to determine if key labor force concepts were being misunderstood and to ascertain whether certain questions in the main survey were superfluous.¹⁶

Analysis of both response distribution and item nonresponse was based on data collected directly with the test questionnaires. Response distribution analysis consisted of statistical comparisons of the distribution of answers with the distribution for comparable questions among the various questionnaires being tested. Item nonresponse analysis was conducted by calculating the "don't know" and "refusal" rates for each question. A high "refusal" rate was used to flag a question as potentially sensitive, while a high "don't know" rate was used to indicate a potentially confusing question.

No single method was employed to determine which version of a question was best; instead, decisions were based on the combination of information from various sources.¹⁷

Features of the new questionnaire¹⁸

Based on the results of the CATI/RDD tests, a revised questionnaire was developed. This section highlights many of these revisions. The discussion is divided into three parts, dealing with revisions related to 1) employment, 2) unemployment, and 3) persons classified as not in the labor force. Accompanying the description of each revision is a brief discussion of the problem with the current question that made the change necessary, and some of the CATI/RDD results used to support the revision. The empirical results were calculated using cumulative data. The reported standard errors were corrected using a Roa-Scott adjustment procedure¹⁹ to account for the clustering of indi-

viduals within households and the repeated observations of individuals over several months. The empirical results from the CATI/RDD tests are presented only as evidence to support decisions about the inclusion of certain questions and about question wording. The CATI/RDD tests results are not adequate for assessment of the impact of the questionnaire revisions on labor force estimates. Limitations of an RDD sample, along with the large variance of estimates due to the small size of some subsamples, preclude the measurement of significant differences in the labor force estimates.²⁰

Employment related revisions

1. "At work" questions. The first two labor force questions in the current CPS are: "What were you doing most of LAST WEEK?" and—for persons who do not say "working" or "unable to work"—"Did you do any work at all LAST WEEK, not counting work around the house?" As noted earlier, these questions may fail to classify as "working" those individuals involved in intermittent or casual work activities, while inappropriately classifying as "working" those persons involved in volunteer activities.

The wording of the current "at work" questions also is ambiguous, outdated, and potentially confusing. If respondents were to answer the first question literally, they might say "sleeping," while persons who operate businesses from their households might say "no," they did not do any work outside the house. To eliminate this type of problem, the first two labor force questions were replaced with a single question: "LAST WEEK, did you do any work for pay?"

To obtain a more complete measurement of unpaid family workers and the self-employed, an introductory question was added to the labor force series inquiring about the existence of household businesses. If a household business is reported, then the single "at work" question is expanded to include the phrase "or profit." Individuals who live in households that report a business, but say that they did not do any work in response to the "at work" question, are subsequently asked if they had done any unpaid work in the family business.

The CATI/RDD results for phase II indicate that approximately 15.1 percent ($n = 7,604$, where n is the sample size) of the households reported a business. (It should be noted, however, that the initial inquiry about the existence of a business was never intended to obtain a measure of the number of household businesses and therefore does not impose any criterion for establishing the legitimacy of a reported business. The new question was added merely to improve the measurement of unpaid family workers and work done for profit.)

The combination of the revised work question and the introductory business question was successful in that 25.2 percent of individuals in households with businesses who answered "no, they did not work" to the initial work question did report doing unpaid work in the family business later in the revised survey.

2. Multiple jobholder questions. Recognizing the need for timely information on multiple jobholders, the CPS was revised so that individuals are asked every month, "LAST WEEK, did you have more than one job, including part-time, evening, or weekend work?" Individuals who indicate that they are multiple jobholders will be asked a followup question to establish the number of jobs they have. In addition, multiple jobholders will be asked to report the hours usually and actually worked on their main job, separately from the hours for all their other jobs combined.²¹ In their fourth and eighth month interviews, multiple jobholders also will be asked about the industry and occupation of their second jobs. The latter information will be used to aid in the reconciliation of CPS data with data collected in the Current Employment Statistics Survey of establishments.

In the second CATI/RDD test, 7.5 percent ($n = 9,437$) of employed persons were multiple jobholders. This estimate is slightly higher than those obtained from the first CATI/RDD test and from the 1991 CPS supplement. In the first CATI/RDD test, when the two alternative versions are averaged, 6.8 percent ($n = 29,149$) of persons employed were multiple jobholders, and in the May 1991 supplement, 6.2 percent of those employed were multiple jobholders. The differences in the percentages could be due to changes in the number of multiple jobholders over the business cycle—a phenomenon that can be analyzed when monthly collection of data on multiple jobholders begins.

The combination of the questions about multiple jobholders, and the reporting of hours on main jobs separately from hours worked at all other jobs also allows a more complete examination of the characteristics of part-time jobs. For example, in the second CATI/RDD test, it was found that, based on usual hours, 4.5 percent of multiple jobholders had two full-time jobs of 35 hours or more per week, 69.0 percent had one full-time job and one part-time job, and 25.5 percent had two part-time jobs ($n = 599$).²²

3. Hours questions. In addition to providing more complete information on the number of part-time jobs, asking multiple jobholders for hours on their main job and all other jobs separately will also improve the quality of the hours data. In the current survey, working individuals are asked, "How many hours did you work last week at all

jobs?" However, interviewers reported during debriefings conducted in 1988 that at least some respondents do not seem to hear the phrase "at all jobs,"²³ which might result in the respondent's underestimation of mean hours worked. It may also falsely yield a higher estimate of the proportion of individuals working part time, if multiple jobholders report hours at only one of their jobs.

The current hours questions pose problems for other worker groups as well. Asking respondents to report the number of hours they actually worked, without a surrounding context, could result in a large amount of rounding in the reported hours. Currently, if individuals report working between 35 and 48 hours, they are asked followup probes to determine if they worked any extra hours or took any time off. Interviewers are instructed to correct the actual hours reported based on the answers to the probes. However, it is uncertain whether interviewers correct the actual hours information uniformly. Also, posing the probes after the initial hours question has been asked does not aid respondents in reporting actual hours. Finally, because the probes are not asked of everyone who is working, there could be a difference in the accuracy of the hours data between persons who work more than 35 hours and those who work less. Because BLS defines full-time workers as all individuals who actually worked more than 35 hours in the survey week, this lack of symmetry could translate into a biased measurement of full- and part-time workers.

To better measure actual hours and the incidence of various types of work schedules, an "anchor and recall" strategy was built into the revised questionnaire. Under this strategy, individuals are first asked "How many hours per week do you USUALLY work at your job?" Then, they are asked "LAST WEEK, did you lose or take off any hours from work, for ANY reason such as illness, slack work, vacation or holiday?" and "LAST WEEK, did you work any overtime or EXTRA hours that you do not USUALLY work?" Individuals who indicate that they have worked extra hours or have lost hours are asked about the number. Finally, the respondents are asked how many hours they actually worked, with the question "(So, for) LAST WEEK, how many hours did you ACTUALLY work at your job?"

Data from the second CAT/RDD test indicate that the anchor and recall strategy was slightly more sensitive in obtaining estimates of individuals working less than 35 hours per week. The proportion of employed individuals who worked full time was 69.3 percent ($n = 9,028$) as measured using the current wording and 66.5 percent ($n = 8,654$) when the revised wording was used. The mean of reported hours measured with the current

wording was 39.0 compared to 37.9 hours measured with the revised wording. (X^2 for the difference = 15.8, p -value = 0.000).

In addition to improving the data, restructuring of the questionnaire will address one of the concerns of the Levitan Commission. Usual hours data will be collected for *all* employed individuals, rather than only for those who actually worked less than 35 hours. This will allow classification of employed individuals based on their usual full-time/part-time status, rather than on a combination of actual and usual hours as is currently the case. It will also permit all individuals, rather than only those who actually worked part time during the survey week, to be classified by reason (economic or noneconomic) for their part-time status.

4. Economic part time. In the CPS, persons who work part time (less than 35 hours a week) are divided into two groups: voluntary part time (those working part time for noneconomic reasons); and involuntary part time (those on part time schedules for economic reasons). To be classified as economic part time, persons must give such reasons as slack work, slower business, or inability to find full-time work.

The size and characteristics of the involuntary part-time work force have been closely watched as important cyclical indicators of economic activity. In addition, the composition and number of involuntary part-time workers have been examined to determine the economic hardship associated with being part time.²⁴ To adequately analyze the policy implications and labor market dynamics associated with part-time work, it is necessary to establish clearly the difference between its voluntary and involuntary components.

As previously noted, the concept of economic part time is not well defined in the current questionnaire. Individuals who usually and actually work less than 35 hours a week are asked, "What is the reason you USUALLY work less than 35 hours a week?" Based on the reasons individuals provide, their availability and desire for full-time work are inferred. There are no direct questions to obtain such information.

Even if the reason for working part time was all the information desired, the current open-ended CPS question would be inadequate. Previous research indicates that a basic problem with open-ended questions asking about "reasons" for doing or not doing something is that the survey designers and respondents sometimes do not share a common frame of reference.²⁵ If the survey designers' intended frame of reference is not explicit and answer categories are not presented, respondents with different frames of reference

may fail to spontaneously give a suitable answer because they do not know the intent of the question. It may also be that respondents would not think spontaneously of an answer that might be preferred once it is suggested.²⁶

A final problem with the current series, cited by the Levitan Commission, is that the reasons for working part time are obtained only for individuals who actually worked less than 35 hours during the reference week. Individuals who usually worked less than 35 hours, but actually worked more than 35 hours during the survey week, would not be included in estimates of part-time workers.

In developing the revised questionnaire, the concept of economic part-time status has been better implemented by including a direct question on the desire for full-time work for all individuals who usually work less than 35 hours a week. All individuals who want to work full time will be asked for the main reason they usually work part time. Those who do not want to work full time will be asked why they are not working full time, but they will not be classified as part time for economic reasons, and any reasons they provide for working part time will be classified as noneconomic. Those who provide an economic reason for working part time will be asked if they are available to work more than 35 hours a week.²⁷

To help respondents better understand what is being asked, the "reason for working part time" question has been reworded to include both economic and noneconomic reasons for working part time. Specifically, individuals who say they want to work full time are asked "Some people work part time because they cannot find full-time work or because business is poor. Others work part time because of family obligations or other personal reasons. What is your MAIN reason for working part time?" (Individuals who indicate that they do not want to work full time are asked "What is the main reason you do not want to work full time?") Results of the second CATIRDD test indicate that 26.6 percent ($n = 1,734$) of persons who usually work less than 35 hours a week want to work full time. More than 90 percent ($n = 195$) of those classified as usually part time for economic reasons reported that they were available for full-time work "last week."

Explicitly asking about desire and availability for full-time work slightly lowered the proportion of those who usually work part time who were classified as part time for economic reasons. Among those who usually and actually worked part time during the week preceding the interview, 14.4 percent of those who received the current wording were classified as part time for economic reasons, compared with 11.0 percent of those who received the revised wording. (The X^2

for the difference was 5.61, and the p -value was 0.028.)

5. *Industry and occupation questions.* Every month in the current CPS, respondents are asked:

- a. For whom did you work?
- b. What kind of business or industry is this?
- c. What kind of work were you doing?
- d. What were your most important activities or duties at this job?
- e. Were you:
 - An employee of a PRIVATE COMPANY, or business, or individual for wage or salary or commission
 - A FEDERAL Government employee
 - A STATE Government employee
 - A LOCAL government employee
 - Self-employed in own business, professional practice, or farm
 - Is this business incorporated?
 - Yes
 - No
 - Working WITHOUT PAY in a family business, or farm
 - NEVER WORKED

Going through these questions every month is quite burdensome for respondents and interviewers. Frequently, interviewers report that respondents complain that they had already provided the information in the previous month.

Collecting information repeatedly every month also results in spurious reports of industry and occupation shifts by workers. This problem occurs because it is sometimes difficult for a respondent to describe an occupation consistently from month to month, the household respondent who is providing answers for the industry and occupation questions may be a different person in subsequent months, and interviewers may change slightly the way in which the same answer is recorded. (For example, they may paraphrase longer responses.) The situation is aggravated by the fact that respondents' answers are coded using 3-digit industry and occupation codes. Because categories at the 3-digit level are quite detailed, very subtle wording changes can result in an incorrect report of change in industry or occupation. Also, when codes are assigned to respondents, it frequently is important to know if an individual's industry or business is in manufacturing, retail trade, or wholesale trade. The question "What kind of business or industry is this" does not always yield this information.

A 1974 study by Candace L. Collins showed that about 32 percent of the three-quarters of the sample that overlap between 2 consecutive months had a change in their 3-digit occupation classifications, and about 16 percent had a change in their 3-

digit industry classifications.²⁸ In the first CATI/RDD respondent debriefing, self-respondents who received the current question wording were asked if they had changed employers from the previous month. Those who had not were then asked if their usual activities or duties had changed. These responses were then compared to the actual change in industry and occupation codes between the previous month and the current month determined in the main survey. Of those who said "no, they had not changed employers or job duties" in the debriefing, 39.0 percent ($n = 1,014$) had changed industry codes between consecutive months in the regular survey, and 21.7 percent ($n = 1,064$) had changed occupation codes. These high rates of change reflect response and coding errors rather than true change, because by their own reports in the debriefing, respondents' employers and job duties had not changed.

Another problem with the current occupation question is that obtaining information about individuals' *most important* activities may not accurately reflect what individuals usually do or what their occupations are.

A final problem with the current questionnaire wording is that the class-of-worker information—the individual's status as a private wage and salary worker, a government worker, self-employed, or an unpaid family worker—is collected after other industry and occupation information has been recorded. This questionnaire structure, along with the question wording, frequently induces interviewers to fill in the answer to the class-of-worker question without consulting the respondent. Interviewers' "silent" coding could lead to an inaccurate class-of-worker distribution. Further, because the assignment of some industry codes varies depending on whether an individual is self-employed or a private industry worker, errors in an individual's class-of-worker information could translate into errors in assignment of industry and occupation codes.

Not knowing an individual's class-of-worker status prior to asking the other industry and occupation questions, also precludes rephrasing questions to make answering them less awkward. For example, self-employed individuals are asked "For whom did you work?" This question usually would prompt the response of "myself," when in actuality the name of the individual's business or company is desired.

To reduce the possibility of misclassification, the industry, occupation, and class-of-worker questions were rearranged in the revised questionnaire. Respondents are first asked: "Were you employed by government, by a private company, a non-profit organization, or were you self-employed?" Not only does putting the class-of-worker question first prevent interviewers from

filling in the answer without asking the question, but it also permits the wording of subsequent industry and occupation questions to be custom tailored. For example, rather than asking everyone for whom they work, the revised questionnaire asks government workers, "What is name of the government agency for which you work?" Private wage and salary workers are asked "What is the name of the company for which you work?", while the self-employed are asked "What is the name of your business?" To further aid in industry coding, the explicit probe "Is this business or organization mainly manufacturing, retail trade, wholesale trade, or something else?" will appear on interviewers' computer screens. Interviewers are instructed to ask this probe if necessary.

To improve the quality of the occupation data, the duties question has been altered to ask for usual activities and the phrase "that is, what is your occupation" was added to the question inquiring about the kind of work. The two revised occupation questions thus will read "What kind of work do you do, that is, what is your occupation?" and "What are your usual activities or duties at this job?"

To ease respondent burden, dependent interviewing—using information from the previous month's interview in the current interview—has been incorporated into the industry and occupation questions. After the industry and occupation data are collected in the first month, rather than being asked for the same information every month, individuals interviewed in successive months are asked:

- a. Last month, it was reported that you worked for (company name). Do you still work for (I-company name)?
Yes (Ask next question)
No
Don't know (Skip to independent industry/occupation question series)
Refused (Skip to independent industry/occupation question series)
- b. Have the usual activities and duties of your job changed since last month?
Yes (Skip to independent industry/occupation question series)
No (Ask c)
Don't know (Ask c)
Refused (Ask c)
- c. Last month you were reported as (previous month's occupation or kind of work performed) and your usual activities were (previous month's duties). Is this an accurate description of your current job?
Yes (End series)
No (Skip to independent industry/occupation question series)

Don't know (End series)
Refused (End series)

Results from the second CAT/RDD test indicate that the revised questions did result in data that were more easily coded, and that verbatim descriptions that had to be referred to expert supervisory coders or that were not assigned a code due to insufficient information were reduced significantly.

To obtain an up-to-date measure of true changes, expert coders analyzed a sample of individuals' jobs descriptions collected from the same respondent for three pairs of consecutive months. (The methodology used in this analysis was based on a 1975 job mobility study by Candace L. Collins.²⁹) The purpose was to determine if, in fact, a true change at the 3-digit level occurred.³⁰ Under contract to the Bureau of Labor Statistics, WESTAT, Inc., designed and analyzed the results of this test. As shown below, WESTAT found estimates of true change to range from 3.8 to 4.2 percent for industry codes and from 5.9 to 7.4 percent for occupation codes at the 3-digit level. Following are the average month-to-month changes in industry and occupation codes as determined by the two surveys and the expert coders (sample sizes in parentheses). Note that estimates of change for the revised questionnaire in the second CAT/RDD test include month-to-month change for both the dependent and independent measures of industry and occupation.

Percent change in 3-digit categories from—		
	Current cps	Revised cps
Industry.....	23 (1,426)	5 (1,361)
Occupation.....	39 (1,392)	7 (1,392)
WESTAT expert coders		
Industry.....	3.8-4.2 (256)	
Occupation.....	5.9-7.4 (406)	

If one accepts the expert coding results as a measure of the "true" change between interviews, the current questionnaire greatly overestimates the rates of gross flows of persons among industries and occupations, while those from the revised questionnaire are fairly close to those assumed to be rates of "true change."³¹

An added byproduct of using dependent interviewing for collecting industry and occupation data is that it permits the identification and analysis of individuals who have changed employers without a period of unemployment or who have changed occupations while remaining with the same employer. This analysis will partially meet the request of the Levitan Commission for more information on job mobility.

Data from the second CAT/RDD test indicate that 4.8 percent ($n = 5,366$) of persons eligible for dependent interviewing changed employers between months. Some 1.9 percent ($n = 5,081$) of those who did not change employers reported that their duties had changed, while another 2.3 percent ($n = 4,894$) reported that the previous month's descriptions of their duties was inaccurate.

6. *Earnings questions.* To obtain estimates of weekly earnings, wage and salary workers in the CPS currently are asked the following series of questions:

- How many hours per week do you USUALLY work at this job?
- Are you paid by the hour on this job? (if no skip to d)
- How much do you earn per hour?
- How much do you USUALLY earn per week at this job before DEDUCTIONS? Include any overtime pay, commissions, or tips usually received.

There are several problems with these questions. The current version of the earnings series forces everyone to report earnings on a weekly basis, even though that may not be the easiest way for respondents to recall or report earnings. Data from the first CAT/RDD test indicate that, in one test version, only 14 percent ($n = 853$) of the non-hourly workers were paid at a weekly rate, and in the other test version of the questionnaire, only 24.3 percent ($n = 1,623$) of such workers found it easiest to report their earnings as a weekly amount.

A second problem is the complexity of the current weekly earnings question. This complexity arises because the instructions to include any overtime pay, tips, and commissions and to report gross earnings (before deductions) all are embedded in one question. A final problem is that asking hourly workers how many hours they usually work, how much they earn per hour, and how much they earn per week could sound quite redundant if respondents are not listening carefully.

Given these concerns, the earning question series has been altered dramatically. In the revised questionnaire, the series first requests the periodicity for which it is easiest for respondents to report their earnings. Once a periodicity is identified, the wording of questions requesting the amount earned is tailored to this periodicity. Individuals also are asked a specific question to determine if they receive overtime pay, tips, or commissions. To improve the quality of the data, if individuals indicate that they do receive "extra" earnings, a lead-in is included in the earnings amount question, reminding respondents to include overtime pay, tips, and commissions when reporting gross earnings. The accuracy of the data also will be

improved through a series of range checks and verification procedures that have been programmed into the survey instrument.

Because, for some minimum wage studies, it is important to obtain an estimate of the number and rate of pay of hourly workers, additional questions are asked of individuals who do not initially find it easy to report earnings on an hourly basis.

Below is an example of the earnings question series an individual who chooses to report monthly would be asked:

"This month I have a few questions about earnings."

- a. For your job, what is the easiest way for you to report your total earnings BEFORE taxes or other deductions: hourly, weekly, annually, or on some other basis?
- b. Do you usually receive overtime pay, tips, or commissions?
- c. (Including overtime pay, tips and commissions,) What are your usual monthly earnings on this job, before taxes or other deductions?
- d. I have recorded your total earnings as . . . MONTHLY before taxes or other deductions. Is that correct?
- e. Even though you told me it is easier to report your earnings monthly, are you PAID AT AN HOURLY RATE on this job?
- f. (EXCLUDING overtime pay, tips and commissions) What is your hourly rate of pay on this job?

The parenthetical phrases are added if the answer to question b is "yes."

Results from the second CATI/RDD test indicate that the revisions to the earnings series did lead to significant decreases in the incidence of "don't know" responses for calculated weekly earnings. With the current wording, "don't know" responses were received from 18.6 percent ($n = 3,467$) of the eligible individuals, in sharp contrast to the 12.8-percent ($n = 2,985$) "don't know" rate among those who were administered the revised questionnaire (the X^2 for the difference = 39.1, the p -value = 0.00). Refusal rates were 8 percent for both versions. The drop in the "don't know" rate illustrates how much more information can be obtained from respondents if they are provided with reporting options instead of being asked to compute their earnings in terms of a specified periodicity.

While data from the second CATI/RDD test suggest that the revised earnings series produces a lower nonresponse rate than the current series, the data do not indicate any significant differences in earnings estimates. In the second CATI/RDD test, average weekly earnings for persons who responded to the current wording was \$465 ($n =$

2,823), compared with \$456 ($n = 2,355$) for those who responded to the revised series ($f = 0.89$, p -value = 0.355). The respondent debriefing did indicate, however, that a larger percentage (85.1 versus 70.4 percent) of those receiving the revised wording were reporting *gross* (rather than net) earnings, compared to those who received the current wording. The phase II test also demonstrated the importance of asking additional questions of those who do not find it easy to report earnings on an hourly basis in order to obtain the number of hourly workers.

Without the additional question, only 38.3 percent of those who received the revised questionnaire would have been classified as hourly workers, compared with 61.7 percent of respondents who received the current questionnaire. With the additional question, the estimate of the number of hourly workers was significantly higher in the revised version (66.9 percent, $n = 2,800$) than in the current version (61.7 percent, $n = 3,402$, X^2 for the difference = 15.5, f -value = 0.00).

Unemployment-related revisions

1. *Unemployed—looking for work.* To be classified as unemployed, an individual has to have looked for work within the last 4 weeks, or be on layoff. In the current questionnaire, the two key questions used to classify an individual as "unemployed—looking for work" are "Have you been looking for work during the last 4 weeks?" and if yes, "What have you been doing in the last 4 weeks to find work?"

In response to the latter question, interviewers are instructed to place each *active* job search method reported into one of the prespecified response categories. An active job search method is defined as any effort that could have resulted in a job offer without any further action on the part of the jobseeker. In other words, an active job search is one that directly brings potential employers and employees into contact. The prespecified active search methods categories on the current CPS are: "checked with public employment agency," "checked with private employment agency," "checked with employer directly," "checked with friends or relatives," "placed or answered ads," or "other." (Interviewers are instructed to code all *passive* job search methods, such as reading the newspaper, attending job training courses and practicing typing, into the "nothing" category.)

Research indicates that interviewers have difficulty correctly recording active and passive job search methods using current response categories. A 1989 observational study of interviewers conducted by Fracasso found that only 39 percent of the "looked at ads" answers were correctly placed into the "nothing" category, while 45 percent were

incorrectly put into the "placed or answered ads" category, and 16 percent wound up in the "other" category.³² The interviewer study also found that 64 percent of the "attended job training" answers were incorrectly placed in the "other" category. As the results of the interviewer study revealed, coding a passive search method as "nothing" seems illogical to many interviewers, because the respondent actually did something. A 1986 examination of verbatim entries conducted by Martin also established the fact that there was confusion among interviewers in the use of the "other" category.³³

In addition to the difficulty noted in correctly distinguishing between active and passive job search, there also is concern that interviewers are not consistently probing for cases in which more than one job search method was used. Correctly distinguishing between active and passive job search, along with completely enumerating all search methods used can be crucial, because only individuals who report at least one *active* search method are classified as unemployed. All others who say they have "looked for work," but who have not used at least one active job search method, are classified as not in the labor force. Therefore, failing to record all active job searches could lower the unemployment rate. Conversely, incorrectly classifying passive job search methods as active could artificially raise the unemployment rate.

There also was concern that the initial "looking for work" question might be interpreted by respondents as requiring intensive job search over an extended period. If the initial "looking for work" question were too restrictive, individuals would be incorrectly excluded from being unemployed before the interviewer even reached the job search methods question. In redesigning the questionnaire, the intention was to use the broadest initial "looking for work" question, and later screen out passive jobseekers in the job search methods question.

To broaden the group of individuals who could be classified as unemployed, the initial "looking for work" question was changed to "Have you been doing anything to find work during the last 4 weeks?" To ensure that interviewers ask about the possibility that more than one job search method was used, the computer displays the probe "anything else?" after the response to the initial job search method has been entered.

To address concerns about the distinction between active and passive job search methods, several response categories were added or had their labels changed. To further reduce confusion, the response categories also were reformatted so that active and passive search methods appear in two, separately labeled columns. On the interviewer's

computer screen, the search method questions will appear as follows:

"What are all of the things you have done to find work during the last 4 weeks?"

<i>Active</i>	<i>Passive</i>
Contacted:	Looked at ads
Employer directly/interviewed	Attended job training programs/courses
Public employment agency	Other passive
Private employment agency	
Friends or relatives	
School/university/employment center	
Sent out resumes/filled out applications	
Placed or answered ads	Nothing
Checked union/professional association registers	
Other active	

The active response categories that were added are "contacted school/university/employment center," "sent out resumes," "filled out applications," and "checked union/professional association registers." The passive categories added are "looked at ads," and "attended job training programs/courses." In addition, the "other" category was split between "other passive" and "other active."

The differences resulting from the use of the "other" and "nothing" categories, as measured in the second CATI/RDD test, most clearly exemplify the impact of the additional response categories. If all the passive search methods in the revised questionnaire are collapsed into a single "nothing" category, 23.6 percent ($n = 706$) of the responses would have been coded as "nothing," compared with only 4.7 percent ($n = 773$) of the responses to the current question. It should be noted, however, that the effect of the additional response categories on the classification of individuals as unemployed is not as dramatic as the differences in the "nothing" category might imply. In the revised questionnaire, all of the passive methods can be used in combination with active methods. Only individuals who actually said they did "nothing" or used passive methods exclusively would be classified as not in the labor force instead of unemployed. In the revised questionnaire, 5.9 percent ($n = 706$) of those stating that they had looked for work in the last 4 weeks used only passive methods, compared to 4.7 percent ($n = 773$) of those responding affirmatively to the search question in the current questionnaire.

The additional response categories also reduced the use of the "other" category. In the second CATI/RDD test, 25.9 percent ($n = 773$) of the responses to the current wording were classified as "other," compared with only 10.3 percent ($n = 706$) of those with the revised wording. Concur-

Overhauling the cps: Redesigning the Questionnaire

rent with the reduction of the use of the "other" category was a reduction in the misclassification of passive job search methods as "other."

As shown in the tabulation below, analysis of verbatim responses collected in the second CATI/RDD test indicates that 31.5 percent ($n = 133$) of the "other" responses measured with the current wording were actually passive methods that should have been coded as "nothing." By comparison, 8.3 percent ($n = 48$) of the "other active" methods measured with the revised questionnaire should have been coded as passive.³⁴

In the revised questionnaire, there was some error in the other direction. Thirty-four percent of the "other passive" methods should have been coded as active methods. The effect of these misclassifications should be smaller than the errors measured with the current wording, however, because only 5.4 percent ($n = 706$) of the responses in the revised questionnaire were classified as "other passive." The probability of individuals inappropriately being excluded from being unemployed based on these misclassifications is further reduced by the fact that active job search methods can be mentioned in conjunction with job search methods classified as "other passive." In the second CATI/RDD test, "other passive" job search methods were mentioned by themselves only 0.8 percent of the time.

Following are the percentages of respondents who were misclassified in the specified categories as described above, suggesting the potential error in job search methods classified as "other":

	"Other"	"Other active"	"Other passive"
Total.	25.9	6.8	5.4
Sample size (n) . . .	(773)	(706)	(706)
False active in current version	31.5	—	—
n	(133)		
False active in revised version	—	8.3	—
n		(48)	
False passive in revised version	—	—	34.2
n			(38)

2. Unemployed on layoff. In addition to being a jobseeker, the other major way to be counted as unemployed in the current cps is to be classified as on layoff. According to the official cps definition, individuals on layoff must expect to be recalled. However, as mentioned previously, the current cps does not verify whether individuals on layoff expect to be recalled. The measurement of individuals on layoff in the current questionnaire is further complicated by the fact that the layoff in-

quiry is part of a complex question that also inquires about temporary absences. Previous research determined that this question is long, awkward, and frequently misunderstood.³⁵ To reduce respondent error, the temporary absence and layoff question was split into two separate items. In the revised questionnaire, respondents are first asked "LAST WEEK, did you have a job either full or part time? Include any job from which you were temporarily absent."³⁶ Respondents who say "no" are then asked, "LAST WEEK, were you on layoff from a job?"

The recall expectations of those who say they are on layoff are ascertained in the revised questionnaire using a series of two questions. Individuals who say they are on layoff are first asked "Has your employer given you a date to return to work?" Individuals who have not been given a date are then asked "Have you been given any indication that you will be recalled to work within the next 6 months?" In the second CATI/RDD test, of those who initially said "yes, they were on layoff," only 36.3 percent ($n = 300$) expected to be recalled. Of these persons, 57.8 percent ($n = 109$) had been given a date to return to work.

Individuals who initially said "yes, they were on layoff," but did not indicate that they expect to be recalled are asked the "looking for work" questions. Consequently, even those who do not expect to be recalled have the opportunity to be classified as unemployed. In the second CATI/RDD test, 78.5 percent ($n = 191$) of those who did not expect to be recalled in the revised version were still classified as unemployed because they had actively looked for work within the last 4 weeks.

The overall effect of the revised layoff series on the unemployment rate is unclear. The direct layoff question may increase the number of individuals who say "yes," they are laid off. However, the expectation of recall question should screen out many of those initially reported to be laid off in response to the direct question, thus lowering the unemployment rate.

On the other hand, if the increase in the number of individuals who initially said they were on layoff in response to the direct question was large enough, or the layoff series increased the proclivity of those who go on to the job search question to say "yes," they had looked for work, then the revised layoff series could actually increase the unemployment rate.³⁷

Results of the second CATI/RDD test indicate that the revised layoff series will have little impact on the composition of the unemployed. Of those classified as unemployed in the second CATI/RDD test, 13.0 percent ($n = 722$) of persons responding to the revised wording were classified as on layoff, compared to 14.5 percent ($n = 841$) of those who

responded to the current wording (X^2 for the difference = 0.721, p -value = 0.396).

3. *Duration of unemployment.* Economists examine the duration of unemployment to test theories about declining reservation wages, the effects of unemployment insurance, and distinctions between those who are unemployed and those who are not in the labor force. In addition, the duration of unemployment is considered an indicator of economic hardship and of relative position within a business cycle. Nevertheless, despite the interest in the duration of unemployment, several studies have established that the current CPS duration data are prone to a great deal of measurement error. In the current CPS, respondents are forced to report how many *weeks* they have been looking for work or on layoff. Individuals who have been looking for work or who have been on layoff in consecutive months are also asked to report their jobless durations independently each month. Norman Bowers and Francis W. Horvath established that forcing respondents to report durations of joblessness in terms of weeks resulted in overreporting of duration for those with shorter spells of unemployment and underreporting of duration for those with spells of unemployment lasting 26 weeks or longer.³⁸ Research by Horvath also established that, although respondents were requested to report duration in weeks, there was a great deal of clustering of durations at the 4-week intervals.³⁹

Analysis of the data from the first CATI/RDD test established that durations were not consistently reported from month to month. In the first CATI/RDD test, when data were collected independently each month, only 22.3 percent ($n = 269$) of those looking for work in consecutive months increased their reported time of unemployment by 4 weeks plus or minus a week. Only 12.6 percent of those looking for work in consecutive months reported an increase in the length of time they had been looking by exactly 4 weeks.

Analysts also identified another potential source of measurement error in the duration of unemployment.⁴⁰ They speculated that, although the reported duration is supposed to be the time an individual is without work, it also could include weeks during which individuals were still working. This overestimate could occur because the current CPS asks individuals how long they have been looking for work, not how long they have been looking while without a job. CPS interviewers are instructed to include only the time an individual spends looking for work while not employed. In the current survey, there are questions asking when an individual last worked at a full-time job and, if it was within the last 12 months, what the last month was in which an individual worked. The answers to these questions are sup-

posed to provide interviewers with the basis for correcting reported durations. However, for respondents who do not volunteer additional clarifying information and for those who have most recently worked part time, the interviewers' instructions are irrelevant or misleading.

Several changes have been made in the questionnaire to reduce response error and thus improve the consistency and quality of data collected. To reduce response error, respondents are permitted to report their duration of joblessness in any periodicity they wish. Respondents are asked "As of the end of LAST WEEK, how long had you been looking for work" (or for those who are on layoff, "As of the end of LAST WEEK, how long had you been on layoff?") However, to encourage more precision in the cases of shorter spells of unemployment, if respondents report a duration of 4 months or less, they are then asked to report their duration in weeks, with the question "We would like to have that in weeks, if possible. Exactly how many weeks had you been looking for work (or on layoff)?" To ensure consistency in reported durations for those who are unemployed in consecutive months, the duration of unemployment will be updated automatically by 4 or 5 weeks.

So that search for alternative employment while still employed is excluded from the duration of unemployment, the date that an individual last worked is collected for those who have worked in the last year. If that date results in a jobless duration that is shorter than the reported duration of jobseeking, the duration will be set equal to the length of time since an individual last worked during the postinterview edit of the data.

The exact set of questions included in the revised questionnaire was not tested in either of the CATI/RDD tests. However, in the first CATI/RDD test, the questionnaire that did have both the dependent interviewing and a followup probe requesting durations in weeks for those who reported being unemployed for 1 or 2 months did result in an average duration of jobseeking of 20.8 weeks ($n = 546$, $SD = 40.0$), compared with an average duration of 13.0 weeks as measured with the current wording ($n = 534$, $SD = 18.8$), a difference that was statistically significant. The followup probe tested in the first phase—"In terms of weeks, how long had you been looking for work," was modified because the results revealed that 74.7 percent ($n = 79$) of those reporting durations of 1 or 2 months gave a duration in weeks that was exactly 4 times their monthly report.

Not in labor force-related revisions

1. *Reduced burden for persons retired, disabled, and unable to work.* One of the most frequent complaints from respondents about the CPS is that

it is burdensome for retired and disabled individuals who have no attachment to the labor force. Currently, persons who say they are retired in response to the first major activity/status question still are asked if they did any work last week, if they were temporarily absent or on layoff from a job, if they had looked for work in the last 4 weeks, and—in their fourth and eighth month interviews—their work histories within the last 5 years. To reduce respondent burden, the response category of “retired” has been added to each question about labor force status. If individuals 50 years of age or older volunteer that they are retired in answer to any of these questions, they are skipped to a question asking whether they currently want a job, and, in outgoing rotation groups, when they last worked. If these individuals indicate that they do not want to work, they are classified as “retired—not in the labor force” and the interview is terminated. In the second CATI/RDD test, only 3.9 percent ($n = 1,790$) of those who volunteered that they were retired said that they wanted to work. The use of the “retired” response option in reducing respondent burden was demonstrated by the large proportion of individuals 50 years or older reported as retired on the first labor force question in the first CATI/RDD test. The data from both of the alternative versions of the questionnaire tested showed that, of all those reported as retired, 43 percent ($n = 5,985$) were classified as retired based on the response to their first question.

To further reduce the burden in subsequent months, the revised questionnaire was designed to use dependent interviewing for those reported as retired in the previous month. Initially, it is ascertained if individuals classified as retired in the previous month had done any work in the last week. If these individuals did not work, they are then asked to verify that they are retired through the questions “Last month you were reported to be retired, are you still retired?” and “Do you currently want a job?” For those who confirmed that they are still retired and do not want a job, the interview is concluded.

Similar changes were made to reduce the burden for those who volunteer that they are “unable to work” or “disabled.” If individuals are reported as “disabled” or “unable to work” in any of the major labor force classification questions, a followup question is asked to determine if they can do any gainful work in the next 6 months. If they cannot, they are classified as “not in the labor force—disabled/unable to work” and the interview is terminated. Data from the second CATI/RDD test indicate that 88.3 percent ($n = 248$) of those reported as “disabled” and 70.5 percent ($n = 105$) of those reported “as unable to work” were unable to do any kind of work in the next 6 months. In

subsequent months, they are asked if they worked in the last week. Individuals who have not are asked to verify their previous month’s status as disabled or unable to work. Almost 99 percent ($n = 1,284$) of those reported as retired, disabled, or unable to work in the previous month verified that their status had not changed.

It should be noted that the response categories of “retired” and “disabled” were not developed to obtain a complete measurement of retired or disabled persons. Classification as retired or disabled is dependent on individuals’ volunteering the status. No attempt is made to determine if those who say that they are retired ever worked at a paying job. Nor is there an attempt to distinguish between disability and retirement for older individuals. The categories of “retired,” “disabled,” and “unable to work” were added to the major labor force questions merely to reduce respondent and interviewer burden. Interviewer focus groups conducted during the first and second CATI/RDD tests indicate that the additional response categories and the resulting questionnaire flow were useful in meeting these objectives.

2. Discouraged workers. The current definition of a discouraged worker—someone who is not employed, wants a job, but is not looking for work because of perceived job market factors—has been widely criticized. The Levitan Commission faulted it as being too subjective because it was primarily based on the “desire for work,” rather than more objective measures of recent job search activities. The definition also has been criticized because individuals’ information about availability and desire for work is inferred from the reasons provided for not looking. The measurement of discouraged workers is important because, like the unemployed, these individuals represent underutilized resources. In fact, one of the unemployment rates published by BLS includes discouraged workers in the measurement of unemployment.⁴¹

To accommodate the recommendations of the Levitan Commission, the revised questionnaire includes questions to determine whether a person has searched for a job within the last 12 months, and whether an individual was available to work during the reference week. During the respondent debriefing for the first CATI/RDD test, it was found that a substantial proportion of those classified as discouraged had looked for work in the last year, found a job, lost or left that job, and had not looked for work since. Because the intent of asking about discouragement is to measure individuals who have given up looking, it was thought inappropriate to include those persons described above in the measurement. For this reason, the revised CPS also includes questions asking if individuals who have

worked in the past year have looked since they last worked.

The revised questionnaire continues to ask individuals who want jobs why they have not searched for work in the last 4 weeks. However, unlike the current questionnaire, the new instrument asks for the *main* reason individuals are not looking. The current question asking about reasons for not looking for work allows multiple responses. A complex priority scheme applied during processing determines which reason takes priority, and thus whether an individual can be classified as discouraged. By asking for the main reason, the revised questionnaire allows respondents to prioritize the reasons themselves.

The final change to the discouraged workers series in the revised questionnaire is that the series will be asked of eligible individuals every month, rather than only during fourth and eighth month interviews. Collection of data from all eligible individuals every month will permit BLS to publish information about discouraged workers on a monthly, rather than on a quarterly, basis, as is currently done.

Using the new criteria to define discouragement could potentially reduce the number of discouraged workers by more than half. When these criteria were applied in the second CATI/RDD test, the percentage of persons not in the labor force classified as discouraged was 0.3 percent ($n = 2,048$), compared with 1.1 percent ($n = 2,127$) as measured with the current questionnaire wording.

Assessing effects of the revisions

As noted at the outset of this article, one of the major objectives in redesigning the CPS questionnaire was to reduce labor force misclassifications. In this section, the effectiveness of the redesign in meeting this objective is evaluated using data from the second CATI/RDD test. The evaluation procedure primarily compared distributions of various labor force classifications and examination of answers to the respondent debriefing questions. Again, given the limitations of the CATI/RDD sample, the results presented here cannot be generalized to the population as a whole.

Work. One objective in redesigning the work questions was to avoid misclassifying individuals who had actually done work during the survey week. The effectiveness of the revised working series was evaluated during the respondent debriefing. The first individual who was reported as not working was asked "In addition to people who have regular jobs, we are also interested in people who may work only a few hours per week. LAST WEEK did you do any work at all, even for as little as one hour?"

Of those asked the "missed work" probe, 3.81 percent ($n = 1,524$) of those who received the current wording and 2.61 percent ($n = 1,263$) of those responding to the revised wording reported missed work activities (X^2 for the difference = 3.2, p -value = 0.07). These percentages represent the proportion of individuals classified as not working who should have been reported as working. However, to obtain a more accurate view of the effect of the revised series, a measure was constructed to analyze those who were identified as working only through the respondent debriefing as a proportion of all those measured as employed. To obtain a comparable sample between the respondent debriefing and the main survey, only the first person in each household who worked more than 1 hour during the survey reference week was included in the index. The constructed index showed that 2.9 percent of those working had their work activity missed with the current wording, compared to 2.0 percent of those who received the revised questionnaire. (This difference is also marginally significant: X^2 for the difference = 3.25, p -value = 0.071.)⁴²

These apparent differences in missed work activity did not translate into large differences in the proportion of the population reported as working. In the second CATI/RDD test, 57.7 percent ($n = 16,175$) of the population was reported as working with the current questionnaire, compared with 58.4 percent ($n = 15,609$) of those who received the revised questionnaire. The revised questionnaire, however, did appear to elicit more reports of work activities involving relatively few hours. The proportion of those working less than 20 hours in the reference week was 10.0 percent ($n = 8,654$) with the revised wording, compared with 8.8 percent ($n = 9,028$) with the current wording (X^2 for the difference = 8.5, p -value = 0.00).⁴³

Hidden part-time workers. Another purpose in redesigning the CPS questionnaire was to collect usual hours for all those who were employed to obtain a more complete count of part-time workers. Data from the second CATI/RDD test indicate that the current procedure of classifying individuals based on the number of hours actually worked in the survey week as opposed to the number of hours they usually work does result in an undercount of usual part-time workers. Phase II data indicate that 5.8 percent ($n = 1,656$) of those who usually worked part time actually worked full time in the survey week. These usually part-time workers represented 1.67 percent ($n = 5,734$) of those reported as working full time in the survey week. Similarly, 7.4 percent ($n = 215$) of those who usually worked part time for economic reasons worked full time in the survey week.

Missed unemployed. In redesigning the cps, an attempt was made to revise the "looking for work" questions in order to improve the classification of unemployed individuals. One method employed in the second CATI/RDD test to assess the success of the revision was to ask respondents who had been identified in the main survey as wanting a job (but not classified as unemployed) the following debriefing questions:

During the last 4 weeks, have you done any of the following to find work (READ LIST)

- a. Checked union or profession registers?
- b. Contacted an employer directly?
- c. Contacted an employment agency or center?
- d. Placed or answered ads?
- e. Contacted friends or relatives about jobs?

Using the responses to this debriefing question, a quasi-measure of those who were revealed as unemployed in the respondent debriefing as a proportion of all unemployed (main survey and debriefing) could be constructed. These measures indicated that the proportion of unemployed missed was not significantly different between the revised and current questionnaire. These constructed indexes were thought instructive for comparisons between versions of the questionnaire. However, because the respondent debriefing question was fairly leading and apt to elicit socially desirable responses, the magnitude of "missed unemployment" was judged to be unreliable.

Conclusion

Beginning last July (1992), the revised questionnaire has been administered to a national overlap

sample survey of 13,000 eligible households per month. This will continue until December 1993 to enable 18 months of data to be collected. (Households follow the same rotation pattern as that in cps, that is, they are in sample for 4 consecutive months, out for 8, and back in for 4 more.) The overlap sample is a test of the revised questionnaire in a fully automated environment of CATI/CAPI. The overlap data are being processed using the revised labor force edit and processing system which will be used to process "live" cps starting in January 1994.

The results from the overlap sample are being compared with data from the current cps to benchmark differences in the labor force estimates due to the new questionnaire, new modes of interviewing, and other design changes. During the analysis of the labor force estimates, special emphasis will be placed on changes which may differentially affect certain demographic groups, differences that may influence measurements of changes in social trends (such as the trend in the number of hours women work outside of the home), and differences due to definitional changes (for example, measurement of discouraged workers). Evaluation of the effect of the revised questionnaire on labor force classification error will continue. Methods of evaluation to determine questionnaire effects will be similar to those used during the CATI/RDD test.

The revised questionnaire will be fully implemented in the field in January 1994. Reports will be issued in February 1994 to document the effects of the redesigned questionnaire. Later reports will include analysis of the data to explain the differences that may be observed. □

Footnotes

¹ See Louis Ducoff and Margaret Jarman Hagood, "Labor Force Definitions and Measurement," in National Commission on Employment and Unemployment Statistics, *Counting the Labor Force, Readings in Labor Force Statistics*, appendix vol. III (Washington, U.S. Government Printing Office, 1979), pp. 32-43.

² Robert L. Stein, "New Definitions for Employment and Unemployment," *Employment and Earnings*, February 1967, pp. 3-27.

³ Other influential changes to the cps that did not have a direct effect on individuals' labor force classification include: 1) the July 1953 replacement of a consecutive 6-month interview schedule with the current 4-8-4 interview schedule, whereby households are interviewed for 4 consecutive months, are out of the sample for 8 months and then are interviewed for 4 more consecutive months; 2) the addition in May 1955 of questions concerning the number of hours worked and the reasons for working part-time; and 3) starting in January 1957, the transfer of individuals on temporary lay-off and those waiting to begin a new wage and salary job within 30 days from the employment to the unemployment classification. (This reclassification involved 250,000 indi-

viduals in January 1957 or 7.7 percent of the unemployed and 0.4 percent of the employed.)

⁴ See Kennon Copeland and Jennifer M. Rothgeb, "Testing Alternative Questionnaires for the Current Population Survey," *Proceedings of the American Statistical Association: Section on Survey Research Methods*, 1990; and Jennifer M. Rothgeb, Anne E. Polivka, Kathleen P. Creighton, and Sharon R. Cohany, "Development of the Proposed Revised Current Population Survey Questionnaire," *Proceedings of the American Statistical Association: Section on Survey Research Methods*, 1992.

⁵ See, for example, Elizabeth Martin, "Some Conceptual Problems in the Current Population Survey," *Proceedings of the American Statistical Association: Section on Survey Research Methods*, 1987, pp. 420-24; or E. Martin, P. Campanelli, and R.E. Fay, "An Application of Research Analysis to Questionnaire Design: Using Vignettes to Study the Meaning of 'Work' in the Current Population Survey," *The Statistician*, vol. 40, 1991, pp. 265-76.

⁶ See Pamela C. Campanelli, Elizabeth A. Martin, and Jennifer M. Rothgeb, "The use of respondent and interviewer

debriefing studies as a way to study response error in survey data," *The Statistician*, vol. 40, 1991, pp. 253-64.

⁷ National Commission on Employment and Unemployment Statistics, *Counting the Labor Force* (Washington, U.S. Government Printing Office, 1979).

⁸ Rothgeb and others, "Development of the Proposal."

⁹ Campanelli and others, "The use of respondent and interviewer debriefing."

¹⁰ Maria P. Fracasso, "Reliability and Validity of Response Categories for Open-Ended Questions in the Current Population Survey," *Proceedings of the American Statistical Association: Section on Survey Research Methods*, 1989.

¹¹ Cathryn Dippo, Kathrine Creighton, Anne Polivka, and Jennifer Rothgeb, "Redesigning a Questionnaire for Computer Assisted Data Collection: The Current Population Survey Experience," Unpublished paper prepared for the 1992 Field Technologies Conference, St. Petersburg, FL.

¹² William P. Butz and Thomas J. Plewes, "A Current Population Survey for the 21st Century," *Proceedings of the Fifth Annual Research Conference* (Bureau of the Census, 1989), pp. 3-13.

¹³ See Campanelli, Martin, and Rothgeb, "The use of respondent and interviewer debriefing studies"; W. Sherman Edwards, R. Levine, and Sharon R. Cohany, "Procedures for Validating Reports of Hours Worked and for Classifying Discrepancies between Questionnaire Reports and Validation Totals," *Proceedings of the American Statistical Association: Section on Survey Research Methods*, 1989; Fracasso, "Reliability and Validity"; Gregory Gaertner, David Cantor, and N. Gay, "Tests of Alternative Questions for Measuring Industry and Occupation in the cps," *Proceedings of the American Statistical Association: Section on Survey Research Methods*, 1989; Martin, "Some Conceptual Problems"; Mark Palmisano, "Some Results on the Application of Laboratory Research Techniques to Survey Methodology," Paper presented at the meeting of the American Association for the Advancement of Sciences, 1989; and Mark Palmisano, "Respondent Understanding of Key Labor Force Concepts used in the cps," *Proceedings of the American Statistical Association: Section on Survey Research Methods*, 1989.

¹⁴ See James L. Esposito, Pamela C. Campanelli, Jennifer M. Rothgeb, and Anne E. Polivka, "Determining Which Questions Are Best: Methodologies for Evaluating Survey Questions," *Proceedings of the American Statistical Association: Section on Survey Research Methods*, 1991.

¹⁵ *Ibid.*

¹⁶ *Ibid.*

¹⁷ *Ibid.*

¹⁸ Portions of this section were extracted from Jennifer M. Rothgeb and Sharon R. Cohany, "The Revised cps Questionnaire: Differences Between the Current and the Proposed Questionnaires," *Proceedings of the American Statistical Association: Section on Survey Research Methods*, 1992, forthcoming.

¹⁹ J.N.K. Rao and A.J. Scott, "On Chi-squared test for multiway contingency tables with cell proportions estimated from survey data," *Annals of Statistics*, vol. 12, 1984, pp. 46-60.

²⁰ Annual averages for 1992 from the cps show that 6.2 percent of all households in the United States do not have telephones. Black and other nonwhite households had lower coverage rates. Some 15.8 percent of black households did not have phones as did 14.2 percent of Hispanic households. Some 11.9 percent of all unemployed individuals lived in households without phones, while 4.2 percent of those currently employed and 6.4 percent of those not in

the labor force lived in such households. source: Unpublished tabulations from the cps, provided by the Demographic Survey Division of the Bureau of the Census.

²¹ A decision was made not to ask about hours on each job separately because a weighted average from the two alternative questionnaires in the first CATI/RDD test showed that 91.3 percent of multiple jobholders had only two jobs.

²² Technically, the second "job" could have been a combination of more than one job. However, because 92.2 percent of the multiple jobholders in the second CATI/RDD test had only two jobs, the discussion was framed with respect to two jobs for ease of explication.

²³ "Response Errors on Labor Force Questions Based on Consultations with Current Population Survey Interviewers in the United States," Paper prepared for the Organization for Economic Cooperation and Development, Working Party on Employment and Unemployment Statistics (Bureau of Labor Statistics, 1988).

²⁴ Rebecca M. Blank, "Are Part-time Jobs Bad Jobs?" in Gary Burtless, ed., *Future of Lousy Jobs* (Washington, The Brookings Institution, 1990), pp. 123-52; and Sar A. Levitan and Elizabeth A. Conway, "Part-time Employment: Living on Half Rations," *Challenge*, May-June 1988, pp. 9-16; and Elizabeth Conway, "Working Along the Edge," *Business and Health*, October 1988, pp. 12-15.

²⁵ H. Schuman and S. Presser, *Question and Answers in the Attitude Surveys: Experimentation in Question Form, Wording and context* (New York, Academy Press, 1981).

²⁶ J. Salancik, "Inference of One's Attitude From Behavior Recalled Under Linguistically Manipulated Cognitive Sets," *Journal of Experimental Social Psychology*, September 1974, pp. 415-27; and J. Morton-Williams, "Study of Interviewer Recording of Respondent Answers," *Survey Methods Newsletter*, 1986.

²⁷ There is one exception to the direct inquiring about availability. Individuals who usually work part time, but who actually worked full time in the reference week, are assumed to be available for full-time work.

²⁸ Candace L. Collins, "Comparison of Month-to-Month Changes in Industry and Occupation Codes with Respondent's Report of Change: cps Job Mobility Study," Response Research Staff Report no. 75-5 (Bureau of the Census, May 15, 1975).

²⁹ *Ibid.*

³⁰ David Cantor, "Draft Recommendations on Dependent Interviewing," Unpublished memorandum (Bureau of the Census, 1991).

³¹ See Cantor, "Draft Recommendations," for more detail on the expert coding test.

³² Fracasso, "Reliability and Validity."

³³ Martin, "Some Conceptual Problems."

³⁴ The revised version had several new "active" categories. In the comparison of "other" responses between versions, some of these new active categories were recoded to "other." However, in the discussion of the incorrect use of the "other" categories in the two surveys, only those answers actually coded in an "other" category were compared ("other" in the current version; "other" active and "other passive" in the revised version). Therefore, the percentage reported for the revised version as falling into the "other" category will be larger than the percentage reported as just "other active" in the revised survey.

³⁵ See Palmisano, "Same Results"; and Jennifer Rothgeb, "Summary Report of July Follow-up of the Unemployed," Unpublished Bureau of the Census Memorandum (1982).

³⁶ It should be noted that individuals who are on layoff can

Overhauling the CPS: Redesigning the Questionnaire

still say "yes, they have a job" and then provide "layoff" as their reason for being absent. In the second CAT/RDD test, 5.4 percent of those who said they were absent from work provided layoff as the reason.

³⁷ Using the numbers from the second CAT/RDD test, the potential for the revised layoff questions to increase the unemployment rate can be illustrated as follows. In response to the current questionnaire wording, 129 individuals said they were on layoff. In the revised questionnaire 300 individuals said "yes" they were on layoff in response to the initial layoff question. Of these, 109 expected to be recalled and thus were directly classified as unemployed. The remaining 191 were asked the "looking for work question." Of these, 150 had actively looked for work. Consequently, in the revised questionnaire, 259 individuals were classified as unemployed through the layoff series. It is impossible to make a complete assessment of the effect on unemployment estimates between the questionnaire versions, however, because it is impossible to ascertain how many of those who said "no" to the current temporarily absent/on layoff question and then went directly to the looking for work question would have said "yes" to the revised layoff question.

³⁸ Norman Bowers and Francis Horvath, "Keeping Time: An Analysis of Errors in the Measurement of Unemployment Duration," *Journal of Business and Economic Statistics*, April 1984, pp. 140-49.

³⁹ Francis Horvath, unpublished tables from the CPS (1991).

⁴⁰ Joseph Antos and Wesley Mellow, "The Youth Labor Market: A Dynamic Overview," Staff Paper 11 (Bureau of Labor Statistics, 1979); and Peter J. Mattila, "Job Quitting and Frictional Unemployment," *American Economic Review*, vol. 64, 1974, pp. 235-39.

⁴¹ BLS issues a table of eight alternative measures of unemployment each month, based on various definitions of the labor force. One of these measures includes discouraged workers in the count of the unemployed.

⁴² Elizabeth Martin and Anne Polivka, "The Effect of Questionnaire Design on Measurement of Work in the Current Population Survey," *Proceedings of the American Statistical Association: Section on Survey Methods*, 1992, forthcoming.

⁴³ *Ibid.*

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. Dipbo 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.

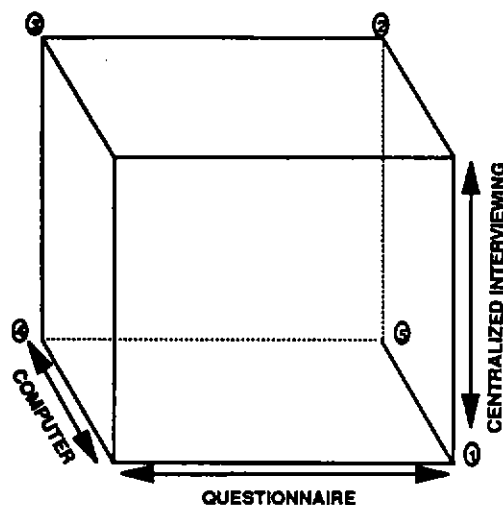
Analysis of data effects

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 ques-

tions, 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 CATI, 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 representatives 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.

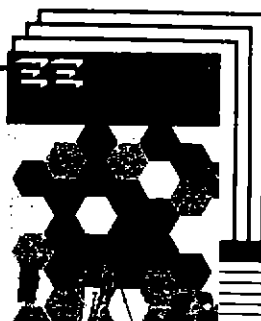
THE PLANNED REDESIGN and modernization of the CPS is an extraordinarily important and ambitious undertaking. The result of planning and testing since 1986 will culminate in the replacement of the current CPS operation with a revised questionnaire and a modern data collection system beginning next year. 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. □

Where to Find Information on Employment and Unemployment

Employment and Earnings:

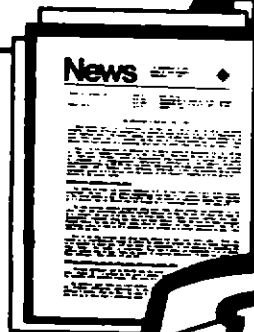
Monthly periodical containing labor force and establishment data. National, State, and area figures on employment, unemployment, hours, and earnings. Order

Employment and Earnings from Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402. Includes text, statistical tables, and technical notes.



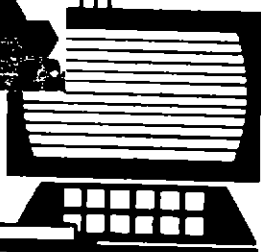
Employment Situation News Release:

Copies of this national statistical monthly release reach the public about a week after the release date. Write: Inquiries and Correspondence, Bureau of Labor Statistics, Room 2860, 2 Massachusetts Ave. NE, Washington, D.C. 20212-0001.



Electronic News Release:

Quickest. Accessible electronically immediately at release time through BLS news release service. Write to the Office of Publications and Special Studies, Bureau of Labor Statistics, Room 4110, 2 Massachusetts Ave. NE, Washington, D.C. 20212-0001, or call (202) 606-5902.



Telephone:

Quick summary on 24-hour recorded message. Key numbers, plus other BLS indicators and upcoming release dates. Call (202) 606-STAT.



Machine-Readable Form:

Labor force data from the household survey and employment, hours, and earnings data from the establishment survey are available on both computer tape and diskette. For

information, write to the Office of Employment and Unemployment Statistics, Bureau of Labor Statistics, Room 4675, 2 Massachusetts Ave. NE., Washington D.C. 20212-0001 or call (202) 606-6373.



Monthly Labor Review:

Employment and unemployment statistics included in monthly 47-page summary of BLS data and in analytical articles. Available from the Superintendent of Documents, U.S. Government Printing Office Washington, D.C. 20402.

