EFFECTS OF A COGNITIVE INTERVIEWING APPROACH ON RESPONSE QUALITY IN A PRETEST FOR THE SIPP

Kent H. Marquis, Jeffrey C. Moore and Karen Bogen[1], U.S. Bureau of the Census
Kent Marquis, U.S. Bureau of the Census, CSMR, WPB 1, Room 433, Washington DC 20233-4700

KEYWORDS: Response error, Record check, Cognitive interviewing

1.0 INTRODUCTION

The Survey of Income and Program Participation is a major vehicle for producing information about the economic situation of people and families in the United States. Recent research, however, has indicated the presence of important amounts of measurement (or response) error (Marquis and Moore, 1990a) and that the measurement error can result in substantial distortions in subject-matter estimates from the survey (Bollinger and David, 1993, Marquis and Moore, 1990a).

As a result, the Census Bureau has designed a new, cognitive interviewing approach for SIPP that may reduce measurement errors importantly. In this paper we mention the primary features of the new interview, discuss the design of the pretest, describe the design of the associated check of administrative records and present some results concerning measurement errors, response rates and costs. We conclude by raising some issues to be addressed in later research.

2.0 METHODS

2.1 Features of the New Interview

The goal of the new interview procedures is to get the respondent and interviewer working conscientiously to accurately reconstruct the family's income history for the past several months. To preempt the human tendency to use simple heuristic strategies in reconstructing income, we ask for reports of each income payment rather than for monthly, quarterly or yearly summaries. The core of the approach is the emphasis on the use of personal income records, such as pay stubs, bank statements, and food stamps receipts. During the first interview, we wait while family members retrieve their receipts and also train them in how to keep records for future interviews in the panel. Training covers the saving of receipts and writing down the dates and amounts of income payments that do not come with receipts.

To assure complete reporting, we rely on a combination of free recall and recognition procedures. To assure good coverage and correct linkage of information from one interview to the next, we employ a reconciliation procedure that reviews income sources reported in each interview and allows any needed corrections to be made on the spot.

Other new features include reporting income received up to the day of the interview, a simplified questionnaire, initial self-response in a group interview setting, and quality control based on tape recording all interviews and monitoring a sample of them.

2.2 Pretest Sample Design

We conducted this pretest in Milwaukee, Wisconsin from December 1991 through March 1992. It immediately followed an initial feasibility study.

We drew a sample of 100 households from a special frame consisting of addresses in administrative records. The frame consisted of approximately 400 randomly selected addresses, 100 from each of 4 government programs: Aid to Families with Dependent Children (AFDC), Food Stamps, Supplemental Security Income, and Unemployment Compensation. The frame also included a subset of employees from a large employer in the area.

Our realized sample consisted predominantly of people in households that had been receiving government income assistance, the economically underpriveleged in the city of Milwaukee. Interviewers view such assignments as particularly difficult to complete well and on time.

2.3 Pretest Fieldwork Design

The sample was divided into two rotations and we conducted two interviews with each household.

Households in the first rotation were interviewed in December about income received since August 1st. They were interviewed again in February about income received since December 1st, a two-month reference period that we were forced to use because of schedule constraints.

2.4 Pretest Record Check Design

A record check compares a person's information from the survey to information in administrative records in order to assess the quality of the survey response. There are several ways to design a record check, each yielding a different kind of information about survey response error.

PROGRAM PARTICIPATION DATA--

--In Administrative Record

--In Survey

	Record Yes	Record No
Survey Yes	A	B
Survey No	C	D
	A + C	

Figure 1: Record Check Definitions

The AC design is our basic record-check procedure. It involves estimating cells A and C in Figure 1 by drawing a sample of people from the administrative records and interviewing them. The value of cell A is the number of people who are participating in the program and report it in the survey. The C value is the number of people who are participating but do not mention it in the survey. The ratio C/(A+C), which we call the underreporting rate, is the probability that a person who is truly participating in the program is not reported, in the survey, to be participating. For this statistic, one generalizes to the population of true program participants. AC design estimates are less expensive to obtain than full-design estimates of net response bias and response error variance but they tell only part of the response error story. Nevertheless, we feel that the underreporting response error estimates are a useful evaluation tool at this stage of development.

The government program agencies provided us with a complete history of payments they made to each sample person during the 6-month survey period. We compared the survey responses to these records to make our underreport error estimates.

2.5 Comparison Estimates

To gain some interpretive perspective, we compare the pretest underreport estimates to underreport estimates obtained from a record check study of regular SIPP (Marquis and Moore, 1990b). We will call this the 1984 Record Check Study. Let us describe the essential features of the 1984 Study so we can exercise the necessary caution when comparing the results.

The data are from up to about 10,000 people who were in the 1984 SIPP panel and residing in Wisconsin, Pennsylvania, New York or Florida. We compared their survey information about several programs to the program administrative records. We excluded people for reasons such as a refusal to report their social security number or because they did not report for the full 8-month reference period. Sometimes we were unable to get good quality administrative records for a program in a particular state. This resulted in excluding additional people from some analyses. Despite the large sample sizes, the n's used to estimate underreports are relatively small, in part because they are conditional on a true "yes" in the record, which is a rare event.

3.0 RESULTS

3.1 Indirect Indicators of Response Quality

Before discussing the underreport estimates, let us review some indirect indicators of improved response quality that we have presented in more detail elsewhere (Bogen, Moore and Marquis, 1992, Moore, Bogen and Marquis, 1992).

Because a major goal of the new interview is to encourage households to use their income

records, we are particularly interested in how often records are used. This is of real concern because, when we describe these procedures to new audiences, the universal reaction is that, "You'll never get people to do that." We can look at record use in three ways, as in Figure 2:

Measure	Wave 1	Wave 2
HOUSEHOLDS using at least one record	83%	85%
INCOME SOURCE reports using at least one record	64	78
	NATIONAL SIPP = about 20%	
PAYMENT reports using at least one record	59	75

Figure 2: Personal Record Use Rates

The percentage of households that used at least one record in the pretest was 83% in the first interview (wave 1) and 85% in the second interview. We find this encouraging.

The percentage of income sourcesy for which at least one record was used was 64% in Wave 1 and 78% in wave 2. The comparable figure for National SIPP during this time was about 20% (Singh 1991, 1992). So we are encouraged both by the high level of source record use and by the apparent increase from the first to the second interview (indicating that record keeping training was having an effect).

A third approach is to ask, What percent of individual income payments are reported with the aid of records? Based on reports of the interviewers, the payment record use rate was 59% in Wave 1 and 75% in wave 2. We are similarly encouraged by the levels and the change over time.

3.2 Underreporting Response Error

Next we discuss the underreporting error rates for the 4 programs in the pretest and compare them to estimates made from the larger 1984 Record Check Study of regular SIPP. Results in Figure 3 suggest that the pretest underreporting error rates, represented by the white bars, are considerably lower than the rates from the 1984 study, represented with dark bars. Although the two samples are very different, we are encouraged by the lower levels in the pretest.

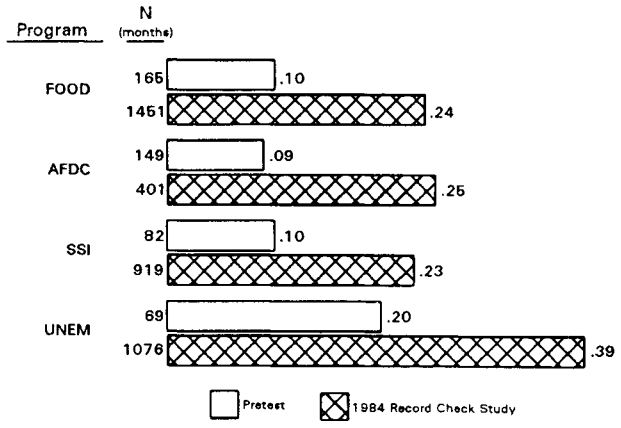

Figure 3: Underreport Rates

3.2.1 Food Stamps. The 165 in the N(months) column is the total number of months, according to records, that the sample people participated in the food stamps program during the reference period . Respondents reported 149 of these participation months correctly and failed to report 16. The pretest underreport rate for food stamps, then, is 16/165 or 10 percent, as indicated by the white bar.

To gain some perspective on the meaning of the pretest underreport rate, the dark bar shows the underreport rate from the larger, 1984 record check study. Using record-checked data from the 1984 SIPP sample in 3 states (excluding New York), we obtained a food stamps participation underreporting rate of 24% per month. Although the two studies are different, we are encouraged to see a pretest underreport rate that is numerically less than half of the rate from the 1984 study.

3.2.2 Aid to Families with Dependent Children. In the pretest, there were 149 true months of participation in AFDC, 13 of which were not reported in the survey. This yields an underreport rate of 9% which is numerically smaller than the 25% underreport rate we obtained in the 1984 SIPP study. For the 1984 study estimate, we have included data for only two of the 4 states. (We

were unable to obtain satisfactory administrative record data for New York so we have omitted those people from the analysis. Also, many people in Pennsylvania refer to their AFDC program as general assistance and that is the way that many interviewers recorded it. This caused a pretty substantial underreport rate. Therefore, for the comparative analysis here, we have excluded sample people who lived in Pennsylvania.)

3.2.3 Supplemental Security Income. The Supplemental Security Income Program provides extra income to the blind, disabled and elderly poor. In Wisconsin, there is both a federal and state component to many of the monthly SSI payments. For this analysis we have combined them, defining a match whenever participation is reported in either (or both) components for the sample person. For the combined SSI then we measured 82 participation months according to the records and survey respondents reported 74 of those months. (6 of the 8 omissions were for just one person.) This yields a 10% underreport rate for the pretest compared to a 23% rate for the 1984 Study. For this 1984 estimate we used administrative record information for all 4 states.

3.2.4 Unemployment Compensation. According to Figure 3, the pretest underreport rate for Unemployment Compensation is 20%. Unemployment compensation checks usually come weekly. For this analysis, we defined reported participation as the report of any unemployment compensation for a person at any time during the month. Using this definition, interviewed sample people had 69 months of unemployment compensation in the administrative records and reported 55 of those months. The pretest underreporting percentage is about half of the 39% rate obtained in the 1984 study (excludes the New York sample).

3.3 Response Rates and Costs

The disappointing results from the pretest concern response rates and costs (see Figures 4 and 5). Wave 1 response rates in the pretest were 73% compared to about 90% for National SIPP. Wave 2 response rates were 88% compared to about 95% for National SIPP. These are very big differences and we view the pretest levels as unacceptable for a Census Bureau Survey.

Study	Response Rates	
	Wave 1	Wave 2
Pretest	73%	88%
Nat'l SIPP	about 90%	about 95%

Figure 4: Response Rates

Comparison	Interview Cost Ratio
Pretest : National SIPP	1.6
Pretest : K.C. Region	1.9

Figure 5: Costs

Pretest interviewing costs were extremely high. Compared to National SIPP, the Milwaukee Pretest costs were about 1.6 times as much. Compared to the average costs for SIPP in the Kansas City Region, the pretest costs were almost double.

Preliminary analyses by the Census Bureau's Field Division suggest that the low response rates and high costs can be traced, in part, to inexperienced interviewers, their general lack of skills in handling reluctant respondents, and their lack of sophistication in optimizing their calls to get interviews and to complete their monthly assignments.

Later assessments have also implicated the new procedures: our practice of allowing unlimited callbacks to get either non-distracting interviewing conditions or complete income information is clearly adding substantially to travel and time costs. The sheer difficulty of this interview, compared to others in an interviewer's assignment, may work in subtle ways to increase refusals and other kinds of noninterviews.

4.0 DISCUSSION

Some underreports are due to discrepancies between the survey and administrative records concerning the month of a payment. For example,

the program may have mailed a check at the end of one month and the respondent reported receiving it at the beginning of the next month. We conducted some sensitivity analyses focusing on adjusting the record payment date to account for time delays between sending and receiving checks (due to the mail, holidays, etc.). The results were pretty uninteresting due partly to the small sample sizes but mainly to the fact that time discrepancies are not a major cause of underreporting in the pretest. While some underreports are artificially generated because the program mailed the check one month and the household reported receiving it the next month, the effects tend to "wash out" for long term recipients.

The majority of underreporting for these programs tended to occur from only a small number of people who failed to report any of their participation months.

In theory, this pattern could result from record error, matching error, processing error, curbstoning the interview or some other reason out of the respondent's control. However, we view any of these possibilities as very unlikely and suspect that there are underlying "cognitive" explanations to be discovered.

There were 3 people in the unemployment analysis and one person each in the AFDC, Food Stamps, and SSI analyses who underreported all their participation. (The AFDC and Food Stamps underreporter was the same person!) One can speculate that these respondents may have forgotten their program benefits or may have been aware of their program benefits but decided not to report them for some reason. In all but the SSI case, these people reported income from jobs. Perhaps they did not believe the interviewer's confidentiality assurance and feared that negative consequences might follow from reporting both program and job income to a representative from the government.

While we have designed the current pretest procedures to address a number of human cognitive strengths and weaknesses, we probably have not addressed the complete forgetting and deliberate underreporting issues well enough. In theory, complete forgetting can be overcome by using recognition and other memory priming techniques, but only if the respondent is willing and able to pay attention to them. In theory, motivated lying can be overcome through appropriate persuasive techniques that reassure a person that he has nothing to lose by reporting and, perhaps, a little something to gain. But unless deliberate liars can be identified and given some extra persuasion, it is unlikely that we can remove underreporting error entirely.

5.0 SUMMARY

In summary, we have described a pretest of a method for reducing underreporting response error in the Survey of Income and Program Participation. The evaluation of the method involved sampling known program benefit recipients from records, interviewing people in beneficiary households about their income, and comparing responses of all adults in those households to administrative records.

We calculated pretest underreporting rates for 4 programs and compared them to rates for SIPP obtained several years ago in a larger study that was less geographically restricted. In general, pretest underreporting error was lower than in the prior study but pretest costs and noninterview rates were unacceptably high.

We have just finished applying this record check method to a larger experimental study in Milwaukee, where we assigned households at random to either the standard SIPP or the new interviewing procedures. If the new interviewing methods achieve substantially higher response quality, SIPP will continue this research program, giving added attention to bringing response rates and costs under better control.

REFERENCES

Bogen, K., J. Moore and K. Marquis (1992), "Pretest Results of an Alternative Measurement Design for the Survey of Income and Program Participation," Paper presented at the American Association for Public Opinion Research, May, 1992.

Bollinger, C. and M. David, 1993, "Modeling Food Stamp Participation in the Presence of Reporting Errors," Proceedings of the 1993 Annual Research Conference, Bureau of the Census, Washington DC.

Marquis, K. and J. Moore, 1990a, "Measurement Errors in SIPP Program Reports," Proceedings of the Sixth Annual Research Conference, Bureau of the Census, Washington DC.

Marquis, K. and J. Moore, 1990b, "Measurement Errors in SIPP Program Reports," SIPP Working Paper No. 9008, U.S. Bureau of the Census, Washington DC, June 1990, especially Appendix Table 1.

Moore, J., K. Bogen and K. Marquis, 1992, "A 'Cognitive' Interviewing Approach for the Survey of Income and Program Participation: Development of Procedures and Initial Test Results," Proceedings of Statistics Canada Symposium 92, Design and Analysis of Longitudinal Surveys, November 1992.

Singh, R., 1991, "SIPP 91: Wave 1 Results of the Record Check Study," Unpublished U.S. Bureau of the Census memorandum for the SIPP Research and Evaluation Steering Committee, December 19, 1991.

Singh, R., 1992, "SIPP 91: Wave 2 Results of the Record Check Study," Unpublished U.S. Bureau of the Census memorandum for the SIPP Research and Evaluation Steering Committee, June 15, 1992.

ACKNOWLEDGMENTS

We gratefully acknowledge the advice and assistance from our headquarters advisory group (Don Fischer, Nola Krasko, Bob Kominski, Doug Lampe, Tom Palumbo, Gary Shapiro, Kathy Short, Raj Singh, Lynn Weidman, Steve Willette and Robert Williams), from the Kansas City Regional Office Staff (especially Richard Taegel, the field supervisor), from the Milwaukee field staff, from the ASA-Census Advisory Committee and from the ASA/SRM Working Group on the Technical Aspects of SIPP.

NOTE

1. The views expressed are attributable to the authors and do not necessarily represent those of the U. S. Bureau of the Census.