

PRELIMINARY FIELD RESULTS OF AN ALTERNATIVE MEASUREMENT DESIGN FOR THE SURVEY OF INCOME AND PROGRAM PARTICIPATION

Karen Bogen, Nola G. Krasko, Jeffrey C. Moore, Kent H. Marquis, U.S. Bureau of the Census¹
Karen Bogen, Center for Survey Methods Research, Census Bureau, Washington, D.C. 20233-4700

KEYWORDS: response rate, costs

1. INTRODUCTION

The Survey of Income and Program Participation (SIPP) is one of the Census Bureau's major, continuing demographic surveys. The SIPP Cognitive Research Project (SIPP-CR) is an effort to reduce important measurement errors in SIPP using alternative procedures and field materials. This paper briefly reviews the overall research plan, summarizes the components of the experimental procedures, and then focuses on results of the field work to date. This paper presents the working hypotheses (based on production reports, field observations, and debriefing of interviewers) about why the experimental response rate was so low and the costs were so high. The paper also briefly describes some of the plans for addressing these problems in the future.

2. OVERALL RESEARCH PLAN AND FEATURES OF THE EVALUATION STUDY

Previous research has indicated that there are important errors in SIPP data. One persistent problem has been the "seam bias" (e.g., Burkhead and Coder, 1985) - the tendency for month-to-month change estimates measured across the seam between two interviews to be many times higher than estimates of change taken within the reference period of a single interview. The SIPP-CR was designed to address the seam bias as well as to address errors identified by Marquis and Moore (1990) and Marquis (1990).

The SIPP-CR consisted of two pretests of about 100 households each, followed by a full-scale Evaluation Study of about 750 households in Milwaukee, Wisconsin. The results presented in this paper are

from the Evaluation Study, which was conducted between September 1992 and May 1993. This was a longitudinal survey, with two interviews four months apart. Half the cases were conducted under the redesigned procedures and half using standard SIPP procedures and instruments. Sample cases for both treatments were selected from administrative records. Eventually the survey reports will be compared with the records, to determine the relative accuracy of income and program participation reporting under the two treatments.

The alternative measurement design used in the SIPP-CR includes new field materials and new procedures. All of the key new procedures were derived from findings from the SIPP Record Check Study (Marquis and Moore, 1990) and some subsequent exploratory cognitive research (Marquis, 1990). Following are the major components of the new procedures:

A. Use Personal Income Records - In an effort to preempt respondents' use of simple strategies for "recalling" their 4-month income history, the revised procedures insist that respondents use their personal records to report their individual, to-the-penny income payments, not monthly totals or summary amounts. The procedures also call for the interviewers to make callbacks to obtain income reports from records that are not available at the time of the interview.

B. Use More Realistic Recall Models - When respondents do not have records, which is often the case in initial interviews, interviewers are trained to guide respondents to use more realistic models to recall the dates and amounts of their income payments during the reference period.

¹ The authors thank Robert Williams of the Census Bureau's Field Division and Doug Lampe of the Census Bureau's Demographic Surveys Division for their important contributions to this research, and we thank our reviewers, Lynn Weidman and Cleo Jenkins, for their helpful comments. In addition, we acknowledge the work of the other members of the Census Bureau's SIPP Cognitive Research Committee: Don Fischer, Bob Kominski, Tom Palumbo, Gary Shapiro, Kathy Short, Raj Singh, Lynn Weidman, and Steve Willette. The views expressed are attributable to the authors and do not necessarily represent those of the U.S. Census Bureau.

C. Use an Unstandardized Approach to Collect Income Data - The core interview starts with a "free recall" section in which respondents are simply told the goals of the section, and asked to get their records and start reporting their income in the order in which it comes to mind. The interviewer asks whatever questions are necessary to get a complete, accurate accounting of the household's income.

D. Use a Non-Distracting Interview Setting - The new procedures require an appropriate, non-distracting interview setting. It is not acceptable to conduct a "doorstep" interview. The setting allows respondents to think and to have access to their records, and also demonstrates a clear commitment to high quality data.

E. Conduct Self-Response, Group Interviews - The new procedures call for maximum self-response, preferably in a group interview format, in which all eligible household members are interviewed together. These procedures maximize the likelihood that records will be available, and allow other household members to help each other recall income sources and details.

F. Revise Interviewer Evaluation Criteria - The alternative procedures put less emphasis on high response rates and high efficiency and more on quality interviewing behavior. All interviews are taped; a sample of the tapes is monitored and interviewers are evaluated on the extent to which their performance is consistent with the primary goal of obtaining complete and accurate information from respondents.

Other changes to the procedures include the use of reference periods that overlap across waves, resolution of discrepant income sources across waves, reconciliation of sources and amounts reported in the overlap period, the simplification of skip patterns in the questionnaire, and the clarification of sections by adding transitions and explanations for the respondents.

3. EARLY RESEARCH FINDINGS

While definitive evidence must await the complete matching of survey data with the administrative data, several earlier sets of analyses suggest that the revised procedures and instruments resulted in improved data quality. Results from the two pretests showed that respondents in the experimental treatment used their income records much more often than

respondents in a standard SIPP interview (Bogen, Moore, and Marquis, 1992). Also, pretest data showed rates of underreporting program participation and a seam bias which were much lower than in regular SIPP (Moore, Bogen, and Marquis, 1992). The fact that there were signs of data quality improvement is the motivation to look for ways to overcome the high nonresponse and costs.

4. EVALUATION STUDY RESPONSE RATES

Although measurement quality is the main focus of the Evaluation Study, a great deal of attention has been paid to field production issues since, from the start of the initial pretest, the response rate and costs for the experimental treatment have been clearly out of line with Census Bureau standards.

Table 1 shows response rates for Wave 1 of the Evaluation Study: 83% for the experimental treatment versus 95% for the control treatment. Most of the nonresponse cases were refusals, which were steady over all four months of Wave 1 interviewing. However, about 20% of the nonresponse on the experimental treatment was a result of no-one-home cases. In contrast, on the control side, there was only one such case over all 4 months of Wave 1. Table 2 shows that the response rate in Wave 2 was also higher for the control treatment.

TABLE 1.
WAVE 1 ASSIGNMENT OUTCOMES

	Oct.-Jan. Total*	
	Experi- mental	Control
Total Workload	648	648
Interviews	493	573
Noninterviews	104	31
> Refusal	84	28
> No-one-home	17	1
> Other	3	2
Ineligibles	51	44
Response Rate	83%	95%

* The response rates exclude the interviews conducted in September, which was designated as a practice month.

TABLE 2.
WAVE 2 ASSIGNMENT OUTCOMES

	<u>Experi- mental</u>	<u>Control</u>
Total Workload	425	434
Interviews	369	416
Noninterviews	40	9
> Refusal	29	6
> No-one-home	6	2
> Other	5	1
Ineligibles	16	9
Response Rate	90%	98%
Longitudinal Response Rate	75%	93%

5. HYPOTHESES ABOUT HIGH NONRESPONSE IN THE EXPERIMENTAL TREATMENT

Hypotheses regarding the causes of the high nonresponse in the experimental treatment can be divided into several major categories: the role of the experimental procedures, the role of the field staff and management, or a combination.

A. The Role of the Experimental Procedures - The first hypothesis is that the experimental procedures are burdensome and invasive and so directly affect respondents' willingness to participate. The procedures make respondents work harder to report the information, for example, using records to report individual payments or reconstructing income details in the absence of records. Other procedures that have been suggested as contributing to the nonresponse problem are tape recording the interviews, the group interview requirement, and the requirement for an appropriate interview setting.

However, there is little evidence that Wave 1 respondents refused because they objected to the new procedures (such as record use or reconstruction of income details), since the refusals generally occurred at the doorstep, before the interviewer even began to explain the purpose and procedures of the survey. There is little evidence of Wave 1 interviews being terminated once the interview requirements were

revealed. The other way that procedures may have affected response rates was that the required interview arrangements (self-response, group interview, proper setting) did not leave the interviewer much room to negotiate the conditions for the interview with the respondent. In the control interview, the interviewer is prepared to give up certain arrangements in order to get the interview. For example, control interviewers might conduct a doorstep proxy interview, as quickly as possible, without insisting on record use. The experimental interviewers were instructed that this was not acceptable and were also cautioned against rushing through the interview.

B. The Role of the Field Staff - A second hypothesis is that the difference in response rates was a direct result of the differences in the field staffs and management. Both staffs contained some experienced interviewers. However, in the control treatment, there were fewer interviewers, each completing a larger assignment. As a result, the experienced interviewers in the control treatment were assigned and completed a much larger proportion of the total cases in that treatment. Experienced interviewers typically are better at both avoiding and converting refusals.

Another hypothesis regarding the response rate differences is that the field staffs on the two treatments handled follow-up of reluctant respondents in different ways, particularly that the control staff conducted follow-up earlier and more often, thereby increasing the possibility of obtaining those interviews. There is some evidence to support this in the noninterview report forms, but it may be an artifact of the way the noninterview report forms were completed by the follow-up staffs on each treatment.

Another contributing factor to the response rate differences could be the crew leaders for the two treatments. The crew leader's job was to motivate and guide new interviewers in their work, as well as to complete their own assignments. The crew leader for the control treatment averaged about 38 completed interviews per month, including a high percentage of conversions of the refusals generated by the staff. In stark contrast, the crew leader on the experimental treatment averaged about 9 interviews per month. There are two opinions about the reasons for this huge difference. One opinion is that the control staff crew leader had a smaller staff, fewer problem cases, and an easier interview to convert, so the performance difference was due to the procedure differences. The other opinion is that the control staff crew

leader was generally a much more capable interviewer and a far better motivator and leader for her staff.

C. The Combined Role of the Procedures and the Field Staff - A third general hypothesis is that the effect of the field procedures on the field staff caused the high costs and nonresponse. The notion is that the new procedures contributed to refusals in an indirect way: the procedures affected the interviewers' burden and expectations and, as a result, affected the interviewers' behaviors at the door. This behavior in turn affected the respondent's decision to refuse.

There may be a number of ways this happens. The first may be that the high level of interviewer burden leaves the interviewer concerned about all the procedures she must remember to do once inside the door. The second may be that the high level of respondent burden makes the interviewer apprehensive about all the potentially objectionable things he or she must ask the respondent to do. Fears about the work to be done once in the interview could distract the interviewer from focusing on getting in the door in the first place. These fears could make the interviewer more tentative and less persuasive at the door.

Typically, in a standardized, scripted interview, an interviewer's biggest hurdle is getting in the door. Once he or she has achieved that, the hardest work is done. The interviewer reads the interview that has already been scripted. They do not put any extraordinary burdens on the respondent. In the alternative interview, this is not the case.

Once in the door, the interviewer has many more things to think about. First, he/she must ask for permission to tape record the interview. In the early months of interviewing, interviewers may have been concerned about how this request would be received. The next big step is to try to arrange a group interview, which interviewers have reported to be quite difficult to achieve. The interviewer also knows that the setting must be appropriate and may need to ask the respondent to turn down the TV or may need to offer to come back later when the respondent is not preoccupied by another task. The interviewer then begins the Free Recall section of the interview, which is an unscripted section in which he/she must think about what information is needed and set about getting it. There is no script to follow, and the interviewer must be focused on the goals of the section in order to complete it successfully. Through-

out the income portion of the interview, the interviewer must ask respondents to consult their income records to report the dates and amounts of income received. Likewise, the interviewer must collect to-the-penny income amounts for all sources, even payment that seem negligible, such as small lottery winnings. So, the procedures may have indirectly affected response rates through interviewers' apprehension about both the burden on herself and the burden on the respondent. The only evidence to support this hypothesis is the concurrence of some of the interviewers.

6. EVALUATION STUDY COSTS

Table 3 shows some of the Wave 1 cost comparisons for the two treatments. The experimental treatment showed many more hours charged, many more miles travelled, and overall about twice the dollar cost per case compared to the control treatment.

TABLE 3.
EVALUATION STUDY WAVE 1
FIELD COSTS PER CASE

	<u>Experi- mental</u>	<u>Control</u>
Interview Hours	3.3	1.6
Travel Hours	2.6	1.0
Miles	34.2	20.0
Interviewer Costs per Case	\$51	\$24

7. HYPOTHESES ABOUT HIGH COSTS IN THE EXPERIMENTAL TREATMENT

The working hypotheses regarding the causes of the high costs in the experimental treatment can be divided into two categories: the role of the experimental procedures and the role of the field staff.

A. The Role of the Experimental Procedures - This hypothesis suggests that the self-response, group interview, record use, and a proper interview setting requirements resulted in many more visits to the households and, thus, are the major contributors to the high cost of the experimental treatment. Table 4 shows that the experimental treatment averaged

almost 5 personal visits per household. The control treatment averaged about 3 visits per household. The experimental procedures made it very difficult for the interviewer to complete an interview in just a single visit. Some callback was required, either to collect self-response data, or to arrange a group interview, or to collect income data from records.

TABLE 4.
EVALUATION STUDY WAVE 1
COST FACTORS

	Experi- mental	Control
Average Personal Visits per Case	4.9	2.9
Average Phone Calls per Case	1.3	1.4
Average Minutes per Interview	94	66

Another result is that the actual in-house interviewing time was higher for the revised procedures. Preliminary data show the experimental treatment averaged just over an hour and a half per household while the control treatment averaged just over an hour. While it is likely that this difference contributed to the higher costs on the experimental treatment, it seems unlikely that it was the main cause.

B. The Role of the Field Staff - Staff differences may have played some small role in the higher costs on the experimental treatment in two ways. First, a much higher percentage of the experimental interviews were completed by new interviewers (more than 75% of the October - January interviews on the experimental treatment versus less than 50% on the control treatment). To the extent that more experienced interviewers are more efficient, this could affect costs. Likewise, there is likely a cost advantage for the control treatment as a result of the economies of scale resulting from larger assignment sizes. For example, travel time to cases is lower when the interviewer has more addresses to visit.

8. CONCLUSIONS AND NEXT STEPS

In the effort to collect more accurate data, the new procedures were deliberately designed to not accept the quick and dirty, and presumably lower-quality, interviews. Likewise, somewhat higher costs were

anticipated as a result of the decision to not always accept interviews on the first visit. However, it is clear that the Census Bureau would be highly unlikely to adopt these new procedures, no matter how much measurement error was reduced, at the current levels of nonresponse and costs.

Assuming that the Evaluation Study record check results continue to show large improvements in data quality, another research panel will be implemented to address the operational issues. The operations panel will seek the appropriate mix of procedures that can reduce costs and nonresponse to acceptable levels without sacrificing the quality gains that have been made. Cost reductions may be achieved by putting strict limits on the number of callback visits for some of the procedures, while maintaining the emphasis on record use. With regard to reducing the nonresponse, the plan is to use only experienced interviewers and to improve training so that interviewers are more prepared when they go to the field.

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 May 1992 meetings of the American Association of Public Opinion Research.
- Burkhead, D. and J. Coder (1985). "Gross Changes in Income Reciprocity from the Survey of Income and Program Participation." Proceedings of the Social Statistics Section, American Statistical Association, Washington, D.C.
- Marquis, K. and J. Moore (1990). "Measurement Errors in SIPP Program Reports." Proceedings of the Annual Research Conference, Census Bureau, Washington, D.C., pp. 721-745. Also available as Report No. 9008 in the SIPP Working Paper Series (June 1990).
- Marquis, K. (1990). "Report of the SIPP Cognitive Interviewing Project." Unpublished report, U.S. Bureau of the Census, August 1990.
- 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, Ottawa, Canada.