

SURVEY OF INCOME AND PROGRAM PARTICIPATION EXPERIMENTS AND THEIR RESULTS

Thomas J. Carmody, Donald P. Fischer, and Andrea J. Meier¹
U.S. Bureau of the Census

I. Introduction

The Survey of Income and Program Participation (SIPP) is a national household-based survey designed to provide improved data on income and program participation. Because of the SIPP's longitudinal design, persons in sample households are interviewed at regular intervals over a 2½ year period. A number of methodological issues are involved in this unique survey, such as adjusting for missing interviews in longitudinal estimates, gaining the cooperation of respondents, and supplementing SIPP data with information from other sources. In order to improve the SIPP's methodology and the quality of SIPP estimates, and to increase the efficiency of the survey, a series of research and evaluation experiments have been undertaken by the Bureau. Six experiments conducted for these purposes are discussed here. They are: the Missing Wave experiment, the Gift experiment, a Debriefing of SIPP Respondents, the Asset and Liability Feedback experiment, the Maximum Telephone Interviewing experiment, and the Employer Provided Benefits experiment. Both the methodologies and results of these experiments are described. All statistical tests in the experiments were performed at the ten percent level of significance, unless mentioned otherwise. In addition, future plans for each experiment are provided where applicable.

II. SIPP Design

The SIPP is a panel survey in which a new sample (or panel) of households is introduced each year. Each eligible person in the sample is interviewed once every four months, with approximately one quarter of the sample being interviewed each month. The panel is interviewed at four month intervals for about 2½ years (generally, eight interviews). All persons age 15 and over in a sampled household at the time of the first interview (called original sample persons) remain in sample for the duration of the panel, even if they move to a new address within the United States. In addition, any individuals age 15 or over who move in with original sample persons (or vice versa) after the first interview are also interviewed for as long as they are living with an original sample person.

To even out field and processing work loads, households in sample are divided into four subsamples of equal size called rotation groups, and one rotation group is interviewed each month. Thus, the entire sample is interviewed over a four month period. One cycle of four interviews covering the entire sample, using the same questionnaire, is called a wave. The reference period of a questionnaire

for a given wave is the four month period prior to the interview month.

The questionnaire contains a "core" of income, program participation, and labor force questions which are asked at each interview, and a set of "topical module" questions which change from interview to interview. Topical modules include subjects such as assets and liabilities, taxes, and work history. For more information on the SIPP see Nelson, McMillen and Kasprzyk (1985).

III. The Experiments

A. Missing Wave Experiment

In a longitudinal survey such as the SIPP, respondents often miss one or more interviews, causing a gap in the longitudinal data. This gap may be filled by collecting retrospective data for missed interviews during later waves. However, the longer reference period needed to collect retrospective data may introduce bias arising from the increased recall period. As a potential tool for improving longitudinal imputation and noninterview adjustment, a set of questions was developed for persons who missed one interview. This set of questions was referred to as the Missing Wave Section. A primary goal of the Missing Wave experiment was to determine how well the Missing Wave Section could detect wave transitions (changes) in receipt of income, government programs, and assets. A preliminary analysis of the missing wave data was completed, but it did not include statistical testing.

The Missing Wave Section became part of the SIPP questionnaire in Wave 4 of the 1984 Panel and was part of the questionnaire through Wave 9. Analyses were initially completed using data collected during Wave 9 for persons who were not interviewed in Wave 8. The section was administered to individuals who granted an interview two waves earlier and for the current wave, but not for the previous (missing) wave. Whereas the current wave questionnaire asks about the preceding four months, the questions in the Missing Wave Section ask about the period five through eight months ago. The Missing Wave Section was placed at the end of the questionnaire so that the current wave's data would not be affected by data collected in the Missing Wave Section. In addition, some of the information obtained during the current interview is used in the Missing Wave Section.

There are basically four parts to the Missing Wave Section of the SIPP questionnaire. The questions in the first part attempt to find out about the person's labor force status. The second part of the Missing Wave Section concerns

the receipt of various income sources, and the third part deals with asset income reciprocity. The fourth part of the Missing Wave Section attempts to obtain retrospective data about specified programs such as Medicare, Medicaid, and any types of educational assistance.

Projected estimates of the number of transitions that we expected to see for Wave 8 were compared to the actual number of transitions detected using the missing wave form during Wave 9. A wave transition, in this case, is a change in status of reporting receipt of an asset or income type between two interviews. When comparing the missing wave number of transitions for receipt of income types to the expected number of transitions (78), a large discrepancy was found. Only 51% (40) of the transitions that were expected to be reported were reported on the missing wave form. A similar situation occurred with receipt of asset types. Only 26% (70) of the 258 expected transitions were reported. Based on certain assumptions made, these expected numbers of transitions are subject to error. However, preliminary analysis shows that our estimates of transitions are not greatly improved by using the Missing Wave Section. (Huggins, 1987).

Another observation is also relevant to the Missing Wave experiment. It was determined that a maximum of 0.3% of the SIPP sample households eligible for interview in Wave 8 would experience an improvement in their noninterview cell assignment if missing wave data obtained during Wave 9 were used to assign households to noninterview cells for the missing wave, Wave 8. (Huggins, 1987).

Based on these analyses of one wave of missing wave data and the belief that other analyses would yield similar results, researchers suggested that implementation of the Missing Wave Section of the SIPP questionnaire be discontinued. The justification for this recommendation was based on indications that data collected during the missing wave portion of the interview would not significantly improve SIPP estimates. It was felt that noninterview adjustment and imputation utilizing previous waves and waves subsequent to the missing wave would provide comparable results. Thus, the extra respondent burden was not deemed justifiable. (Huggins, 1987).

B. Gift Experiment

Since SIPP sample persons are to be tracked and interviewed every 4 months over two and one half years, there is sample attrition over the life of the panel as some persons move and others no longer agree to participate in the survey. It was thought that giving gifts to respondents might lower the attrition rate. To evaluate this hypothesis the Census Bureau designed an experiment for

1987 Panel households. It was decided to begin the experiment in Wave 1 of the 1987 Panel since earlier panels had shown that Wave 1 has the highest rate of new Type A noninterviews (Type A noninterviews are households which are eligible to participate in the survey but do not). The largest percentage of Type A's are households which refuse to participate.

Interviewing for Wave 1 of the 1987 Panel was conducted in February through May of 1987. Since noninterview rates are calculated on a rotation group basis, one rotation group (April sample households) was chosen to be the treatment group and receive the gift during the Wave 1 interview. The other three rotation groups made up the control group.

For the treatment group, after finding a "responsible" person at the sample household, the interviewer followed the usual introduction procedures and explained the purpose of the visit. The interviewer then presented the gift, a small solar-powered calculator, to the respondent and explained that it was a token of the Census Bureau's appreciation for the household's participation in this important survey. Respondents were not required to keep the calculators, and interviewers did not ask to have the calculators back in situations where persons still refused to participate.

After conducting interviews for April of 1987, the interviewers were asked to complete an evaluation form for the experiment. The respondents seemed to like the calculators; however, only 41 of 352 interviewers who completed forms felt that giving gifts helped them gain cooperation from the respondents. About 18 percent (65) of the interviewers believed that interviewers' skill had more of an impact on how cooperative the respondents were. Forty-three of the interviewers cited a specific instance in which the gift helped them avoid a type A noninterview. (Jackson, 1987).

Comparisons were made at both the national and regional office levels between 1987 Panel Wave 1 type A noninterview rates for the gift recipient households (April, rotation 4) and the nonrecipient households (February, March, May, rotations 1-3). For nine of the twelve regional offices, the type A noninterview rates appeared lower for the recipient households. However, this difference was only significant for the Charlotte regional office. This difference was also significant at the national level, with the type A rate for the two groups differing by one percentage point. Looking back at both the 1985 and 1986 Panels, the national level April type A rate appeared lower than the combined February, March, and May rate. However, the differences were not significant. (Gbur, 1987).

Further comparisons were also made between the reciprocity and nonreciprocity

groups' cumulative type A rates for Waves 2 and 3 at the national level. For both Wave 2 and 3 of the 1987 Panel, the difference between the reciprocity group and the nonreciprocity group cumulative type A rate was significant. (Gbur, 1988).

The 1987 Panel type A rates for the recipient group computed so far were also found to be less than the rates projected from the 1985 and 1986 Panels. Although the type A rates for the recipient group for Waves 1 through 3 of the 1987 Panel are significantly lower than the rates for the nonreciprocity group, the projected rate comparisons suggest that these results should be viewed with caution. (Gbur, 1988).

Future plans for the Gift experiment include calculating type A noninterview rates for subsequent waves and comparing them to projected rates for the recipient group from previous panels and previous panels' rates. Although the differences between type A rates for recipients and nonrecipients have not been very large through Wave 3, further analyses will be done to determine if there is a consistent trend of lower type A rates for the treatment group for each wave. (Gbur, 1988).

C. Debriefing of SIPP Respondents

During reinterview for the last wave of the 1985 SIPP Panel, a debriefing of SIPP respondents was implemented primarily to determine reasons why respondents do and do not use records during interviews. Reasons for which respondents continued to participate in the SIPP and whether respondents could offer comments which could aid in the improvement of SIPP data collection were also of concern. The aim of analyzing the debriefing responses was to provide suggestions for improving the SIPP. Reasons for nonresponse, bias, and nonsampling error were also examined.

The debriefing sample was a group of 516 retiring SIPP respondents. Since the 1985 Panel reinterview sample for rotations 2-4 of Wave 8 was already in place, it underwent the debriefing. The questions asked during the debriefing regarded: reasons for continuing with the SIPP, reasons for not using records, suggestions for making record use easier, clarity of certain questions, learning of other government programs through the SIPP, and overall comments and suggestions. Since the questions were open-ended, the response categories were determined after the debriefing took place by categorizing them into what seemed to be distinct categories. Some of the response categories, however, were preprinted on the questionnaire, so it is conceivable that the respondents may have been prompted with possible responses in some cases. The overall response rate for the debriefing experiment was 89.5 percent. (Meier, 1988).

The tax and W-2 form record use rate for the SIPP obtained from debriefing data was compared to a similar rate estimated from 1984 Panel tax and annual roundup topical module data. This module is an extra set of questions concerning calendar year income and types of tax returns filed. One point to note is that the debriefing rate was determined by asking respondents if they had used tax and W-2 forms during the final SIPP interview. In contrast, the tax and annual roundup rate was determined from check item boxes in the questionnaire that were checked off by the interviewer indicating whether the respondent was referring to records. In all, 56.4 percent of the debriefing respondents claimed that they used tax and W-2 forms during collection of data at their final interview. Of the persons responding to questions in the tax and annual roundup module, only 30 percent were checked off as using W-2 forms and 34 percent as using their tax form during the interview. Based on the estimated overlap between these two groups, we would expect the percentage of persons using tax and W-2 forms during the annual roundup to be a maximum of 13.8 percent. With this in mind, the debriefing rate of 56.4 percent is much higher than would be anticipated. However, if the debriefing respondents interpreted the debriefing question as asking whether they used tax or W-2 forms (as hypothesized), then the small degree of overlap between the 30 and 34 percent would indicate a tax and annual roundup record use rate that would be fairly close to the debriefing rate of 56.4 percent. In any case, increased record use rates are needed. (Meier, 1988).

Based on debriefing responses, 60.6 percent of the respondents reported that they routinely referred to records such as bank statements and pay stubs during regular SIPP interviews as compared with the 56.4 percent who said that they used records during the final interview (Note that these two percentages are not significantly different). Although a large majority of the respondents who did not use records claimed that there was nothing we could do to encourage them to use records (about 80%), their reasons for not using records were variable. The major reason given for not using bank statements and pay stubs during regular interviews was that respondents felt they knew the information without referring to records (54.1%). However, the major reason given for not using tax and W-2 forms during the annual roundup part of the last interview was that records were not available (32.3%). (Meier, 1988).

Apart from record use, debriefing found that the major reason respondents gave for continuing to participate in the SIPP was that they liked the interviewer(s) (27.8%). It was also determined that bias exists in SIPP estimates due to

SIPP respondents' learning of other government programs through the SIPP and then applying for them. About 2.2 ± 1.3 percent of the respondents claimed that they learned of a government program through the SIPP and then applied for it. If we assume that these persons are receiving benefits from the program they learned of through the SIPP, then it appears that an estimate of persons on government programs who learned of them through the SIPP is 12.1 percent. [A 90 percent confidence interval for this is given by (0%,26.4%). A more accurate measure of the range of reliability is currently being computed.] This bias may also introduce further biases since the skip pattern of the questionnaire and answers to other questions are dependent upon the answers to government program questions. (Meier, 1988).

Since the respondents were receptive towards the debriefing and valuable information was learned, it was recommended that further research be done as a result of the debriefing experiment. Some possible methods for increasing record use were to have interviewers call before the interview to remind respondents to gather records, or possibly experiment with adding a statement to the SIPP questionnaire itself asking the respondents to retrieve records. A cognitive lab approach was also suggested for determining which portions of the SIPP questionnaire are unclear, what might aid interviewers in eliciting more accurate responses, and what could be done to increase record use. (Meier, 1988).

D. Asset Feedback Experiment

Topical module questions about assets and liabilities are asked twice during each panel; during the 1984 Panel these questions were asked in Waves 4 and 7. The reference periods for these two waves were exactly one year apart. One of the primary purposes of administering these sets of questions twice in one panel is to measure a person's savings, that is, change in net worth, between the two interviews. The purpose of the Asset Feedback Experiment was to determine whether respondents would be influenced in their reporting of asset amounts in Wave 7 if they knew the amounts which they had reported in Wave 4.

To conduct the experiment, respondents in half of the eligible households in Wave 7 were provided specific information which they had supplied in Wave 4. Respondents in the remaining households were interviewed without previously supplied information. In order to comply with the Census Bureau's assurance of confidentiality, interviewers used the feedback form only for self respondents and in situations where the proxy respondent was the same one who provided the information in Wave 4.

Use of the feedback form did not appear to have a significant effect on estimates of year-to-year change in median net worth. First, changes in median net worth from Wave 4 to Wave 7 were considered separately for the feedback and control groups, then the two groups were compared on the basis of these changes. In the feedback group, median net worth of the SIPP universe was estimated to be \$32,048 in Wave 4 and \$30,890 in Wave 7 (the change in estimates of \$1158 was not significant), while in the control group the figures were \$32,944 in Wave 4 and \$32,357 in Wave 7 (again, the change in estimates of \$587 was not significant). In addition, the SIPP universe was divided into 50 subpopulations (based on age, race, education, family type, labor force status, asset ownership, and income) and the same comparisons between Wave 4 and Wave 7 were made in each subpopulation. In only 12 of these 50 subpopulations were there significant changes between Wave 4 and Wave 7 for either the feedback or control group. (McNeil and Lamas, 1987).

The numerical difference between the change reported in the feedback group and the control group was \$571 (\$1158-\$587) and was not significant. Also, the numerical difference between groups was not significant for any subpopulation. Thus, it appears that use of the feedback form did not affect the reporting of changes in net worth. However, it should be noted that our ability to detect differences between the two groups in the reporting of year-to-year changes was reduced because there appeared to be few changes to report (i.e., few significant changes between Wave 4 and Wave 7 were found in either group). If there had been large real changes between waves, then differences between the feedback and control groups may have been more apparent. All tests were performed at the 5 percent level. In some tests a correlation of 0.55 between Wave 4 and Wave 7 estimates in both groups was assumed. Also, all estimates were in 1984 constant dollars.

The numerical differences between groups may have been artificially reduced by the imputation procedure because missing information from a respondent in one group was replaced by information from a respondent with similar characteristics who may have come from the other group. In Wave 7, a relatively large percentage of the value of some asset types was imputed. For example, 19 percent of the value of IRAs was imputed as were 39 percent of the value of stock and mutual fund shares (McNeil and Lamas, 1987).

E. Telephone Interviewing Experiment

After conducting personal visit interviews for about one year, the Census Bureau decided to evaluate telephone interviewing as a way to save money. A

successful pretest, conducted in June 1985 in two of the Bureau's regional offices, indicated that the SIPP questionnaire could be given over the telephone. A national test was then conducted in two phases. Phase I cases were interviewed from August to November of 1986, and Phase II cases were interviewed from February to April of 1987.

The objectives of the first phase of the national test were to compare data quality between the two modes of interviewing (telephone and personal visit). A secondary consideration was to compare cost data for the two modes. The test was split between Waves 2 and 3 to increase the amount of unaffected data that would be available for the topical modules if the telephone test had such a bad effect on data quality that the data could not be used.

For Phase I of the test, approximately half of the households were designated as cases to be interviewed by personal visit and the rest to be interviewed by telephone. Interviewers were instructed to use the designated mode for each case unless there were circumstances which might jeopardize the interview. As in the pretest, interviewers completed the self-study for telephone interviewing. Also, the regional offices mailed letters and appropriate flashcards to households which were designated as telephone interview cases. Most of the telephone interview cases were conducted from interviewers' homes, but some were conducted from the regional offices so that observers could monitor the calls.

In the following analysis, the personal visit and telephone groups are compared by designated interview mode, not the actual interview mode. Because both personal visit and telephone interviews will always be needed in the SIPP, the designated mode gives a truer picture of the effect of implementing that mode. In Phase I, 14 percent of the personal visit group were interviewed by telephone. Of those in the telephone group, between 21 and 46 percent received personal visit interviews (Gbur and Durant, 1987). The exact percentage is not known because some interviewers did not indicate the interview mode on the questionnaire. Probably, these interviewers thought that if a designated telephone interview were done by telephone, no indication of that was necessary. Therefore, we believe it likely that about 21 percent of interviews in the telephone group were personal visits.

Interviewers who participated in the telephone test were generally positive about telephone interviewing with 74.7 percent saying that it could be successful in SIPP, the most often cited reason being that respondents prefer telephone interviews. However, 29.3 percent of interviewers believed telephone interviewing would not work for SIPP; reasons

given included the sensitive nature of the questions and that respondent trust cannot be built and maintained over the telephone (the two percentages sum to more than 100 because some interviewers gave reasons for and against). Household size should be a factor in deciding whether to use the telephone according to 83 percent of interviewers, with a majority believing two persons are the maximum number that can be interviewed in one sitting. (Gbur and Durant, 1987).

It was thought that respondent refusal would be easier over the telephone than in person, leading to higher nonresponse rates. However, the household refusal rate was not higher in the telephone group (3.8 percent in the telephone group vs. 4.1 percent in the personal visit group, the difference was not tested). All results for the telephone test were based on unweighted data. (Gbur and Durant, 1987).

The overall household nonresponse rate in the personal visit group was 8.2 percent and was about the same in the telephone group (7.3 percent). The person nonresponse rate within interviewed households was 2.7 percent in the personal visit group and 2.9 percent in the telephone group (The difference was not significant). It was expected that asking sensitive questions over the telephone could increase item nonresponse rates. No significant differences were found for nonresponse rates on 32 items of income and labor force participation. However, the nonresponse rate was numerically higher in the telephone group for 29 of the 32 items. (Gbur and Durant, 1987). This raises the possibility that telephone use may increase item nonresponse rates.

In evaluating the nonresponse rates, it should be kept in mind that the telephoned respondents had already received one or two personal visits from their telephone interviewer. Thus, respondents had a chance to develop trust in their interviewer before the test began. If this trust had not been developed during the personal visits, telephone nonresponse rates may have been higher. Results from Phase II of the test may help to clarify the issue. Given all the extraneous factors that would enter in, it was felt that no reliable comparison of personal visit and telephone costs could be made from Phase I.

There were three objectives for Phase II of the telephone test: 1) to determine how well households would respond to telephone interviews in two consecutive waves; 2) to evaluate any cost differences between the two modes of interviewing; and 3) to allow observations of the effect upon telephone interviewing of having three SIPP panels in the field at once (Only two panels were in the field during the first phase of the test). Two of the Phase II test month's telephone

interviews were with households which were interviewed by telephone in Wave 3. The third month's telephone interviews were with households which were interviewed by telephone for Wave 2 and by personal visit for Wave 3. No results are yet available from Phase II of the telephone test.

F. Employer Provided Benefits Experiment

The Census Bureau tested the feasibility of having respondents sign a form which would allow the Bureau to contact the respondent's employer and collect supplemental information. The employer would be asked to provide information about employer and employee contributions to health insurance plans, pension plans, and life insurance plans. This information would be more useful than the limited employee benefits information currently collected from respondents. The test was conducted in rotation 4 of Wave 8 of the 1985 Panel. The Wave 8 interview would be the household's last interview, so the employer test would not affect future response rates for this rotation group. Employed persons 18 or over in one half of the rotation group were eligible to receive the authorization form.

If the interview was a proxy, the authorization form was left at the household with instructions on how to complete it and return it to the regional office. If the interview was conducted by telephone, the interviewer explained the test and mailed the form to the respondent.

There was no follow-up by the regional offices in situations where the interviewer left the form with the respondent or mailed it to the respondent. All forms which were signed and received in the regional offices by September 15 were sent to the employers. The employers were also sent a letter explaining the study and a franked envelope for returning the completed form.

In situations where the employer had not returned the form within three weeks, the employer was contacted twice to try to obtain a response. However, if the employer was unable or unwilling to provide the information by the end of October, there was no further follow-up.

Authorization forms were given to 1352 sample persons. Of those, 596 (44.1 percent) signed the forms. Of the 596 forms mailed to employers, 549 (92.1 percent) were mailed back. Thus, forms were received for 40.6 percent of eligible persons. (McNeil, 1988).

To obtain a more realistic response rate, the sample loss that has occurred through Wave 8 should also be considered. We calculated there would have been 1543 SIPP households in one half of rotation 4 given no sample loss and that there are 1.2 employed persons 18 or older per household. Thus, our base is 1852 eligible persons. Under these assumptions,

73.0 percent of possible eligible persons participated in the test; 32.2 percent signed the authorization form, and forms were received from employers for 29.6 percent of the 1852 persons.

Analysis of the data will continue and if nonresponse rates are found to be fairly uniform across population subgroups and can be lowered, then acceptable data quality may be attained. In that case, information from employers may be collected from the entire SIPP sample every two or three years (McNeil, 1988).

ACKNOWLEDGEMENTS

The authors would like to thank Thomas Moore, Rajendra Singh, Maurice Kniceley, and Dennis Schwanz for their comments. Thanks also goes to Nicoletta Palumbo, Kimberly Wilburn, and Sandra Carnegie who provided secretarial support.

REFERENCES

- [1] Gbur, P., "SIPP 87: Initial Results from the Gift Experiment", Internal Census Bureau memorandum from Gift Experiment Workgroup to Singh, October 1, 1987.
- [2] Gbur, P., "SIPP 87: Gift Experiment Results Through Wave 3, Internal Census Bureau Memorandum from Gift Experiment Workgroup to Singh, July 6, 1988.
- [3] Gbur, P. and S. Durant (1987), "Testing Telephone Interviewing in the Survey of Income and Program Participation and Some Early Results," Presented at the International Conference on Telephone Methodology, November 8-11, 1987.
- [4] Huggins, V. (1987) "Evaluation of Missing Wave Data from the Survey of Income and Program Participation (SIPP)", Proceedings of the Section on Survey Research Methods, American Statistical Association.
- [5] Jackson, D., "Interviewers' Evaluation of the Gift Experiment", Internal Census Bureau memorandum from Jackson to Bowie, August 5, 1987.
- [6] McNeil, J. (1988) "Progress Report on Employer Provided Benefits Study," Internal Census Bureau Memorandum from McNeil to Singh, April 18, 1988.
- [7] McNeil, J. and E. Lamas, (1987), "An Analysis of the SIPP Asset and Liability Feedback Experiment," Proceedings of the Section on Survey Research Methods, American Statistical Association.
- [8] Meier, A., "SIPP 85: Results of Debriefing Respondents", Internal Census Bureau memorandum from Meier to Singh, to be distributed, 1988.
- [9] Nelson, D., D. McMillen, and D. Kasprzyk, "An Overview of the SIPP: Update 1," SIPP Working Paper Series No. 8401, U. S. Bureau of the Census.

¹This paper reports the general results of research undertaken by Census Bureau staff. The views expressed are attributable to the authors and do not necessarily reflect those of the Census Bureau.