DATA QUALITY OF THE SURVEY OF INCOME AND PROGRAM PARTICIPATION

Karen E. King, Rita J. Petroni and Rajendra P. Singh

Bureau of the Census

I. INTRODUCTION

This paper discusses sources of error found in the Survey of Income and Program Participation (SIPP). The discussion primarily focuses on potential sources of nonsampling error and what, if anything, is known about their effect on SIPP estimates. A brief discussion of reliability for subgroup analysis is also presented.

This paper summarizes a comprehensive document [9] on the potential sources of error in SIPP. The document unifies and summarizes many reports and memoranda developed during the last several years. While seeking to inform users of the SIPP data, it also helps staff at the Census Bureau to review the understanding of major error sources in SIPP and helps focus the SIPP's evaluation and testing activities to guide survey redesign activities in the future.

The design of any large scale, complex survey involves many decisions on the combination of methods to be used. These decisions are based on considerations of the costs and errors associated with alternative methods, and are interdependent in two ways. First, with a given budget, increased resources to reduce one source of error must be balanced by decreased resources and increased error elsewhere. Secondly, a change in methods to reduce one source of error may lead to an increase in another source of error. As with other large scale surveys, the design of the SIPP is the result of attempts to balance conflicting objectives and constraints. The SIPP design minimizes total survey error for a given budget. For example decisions were made between: 1) more interviews with smaller reference period or fewer interviews with longer reference period, and 2) more interviews with a smaller sample or fewer interviews with a larger sample.

The examples noted above are indicative of the kinds of design trade-offs which must be confronted in the development of any survey. Although the Census Bureau and the Department of Health and Human Services conducted an extensive development program prior to the collection of data in SIPP [18], many SIPP design decisions depended on the general knowledge and expertise of the Census Bureau staff.

II. GOALS AND OBJECTIVES OF SIPP

The overriding goal of SIPP is to provide policy makers with more accurate and comprehensive information about the economic situation of persons and households affected by government policy. This information is vital for improving the capability of federal agencies to formulate and evaluate their policies and programs in the areas of income and social welfare. The information is also important for social scientists to improve their understanding of the economic behavior of the U.S. noninstitutional population. For these purposes, both cross-sectional and longitudinal estimates are relevant and valuable.

To achieve this goal, three objectives were set: (1) to collect a wide array of information about characteristics such as income, program participation, labor force, etc.; (2) to make such data available in microdata files for simulation and other studies; and (3) to inform policy makers and others through a continuing series of publications. [16]

III. DESCRIPTION OF SURVEY

SIPP is a nationwide survey designed to provide comprehensive information that reflects the financial situation of persons, families, and households in the United States (except persons in institutions). The survey population includes persons living in group quarters, such as dormitories and religious group dwellings, but excludes persons living in military barracks, correctional facility inmates and nursing home residents, etc. [15]

A new sample panel of roughly 12,000 interviewed households (HHs) is introduced each year for 1986 and later. (The 1984 and 1985 panels started with larger samples--20,000 HHs and 14,300 HHs, respectively but were reduced later due to the budget cuts [11].) Persons in HHs interviewed in the first visit are contacted once every four months for two and a half years for a total of eight interviews. Thus, the design allows cross-sectional and longitudinal estimates from a combined sample from multipanels. If sample persons move, they are interviewed at the new address. "New" persons living with sample persons are considered part of the sample while living with these sample persons.

To provide smooth and steady workload for data collection and processing and to reduce operational problems, each panel is divided into four approximately equal subsamples, called rotation groups. These rotation groups are interviewed over four months, one each month. (In general, one cycle of four interviews covering the entire sample, using the same questionnaire, is called a Wave.) Persons interviewed in a given month

provide data for the previous four months. For more detailed information on the SIPP design see [4], [6], and [12].

The initial interview takes an average of about 30 minutes for the first person and 15 minutes for each additional person 15 years of age or older in the household. An interview is divided into the three main groups of questions: the control card, the core questions and the topical modules. The data is collected on basic demographic characteristics, labor force participation, program participation, amount of income by source, occupation, work history, assets and liabilities, etc. [5]

IV. SOURCES OF ERROR

Many factors contribute to error found in survey estimates. Section A lists the more important sources of SIPP error while Section B lists sources of minimal or trivial error. A detailed discussion of these errors can be found in [9].

A. Important Sources of Error

Data Collection Procedures: Data collection procedures were designed to provide the best quality data possible under given constraints. For example, SIPP uses both self and proxy responses. Insisting on self response increases survey costs and can also increase both person and household nonresponse. However, proxy responses may not be as accurate as self responses. Proxy responses have significantly higher nonresponse rates for some important items [3] and report more transitions from one state of condition (e.g. economic or labor) to another. [17]

Interviewers also play an important part in the collection of data. Interviewers may miss questions, rephrase, and, thus, change the meaning of the questions, or record information incorrectly.

Nonresponse Error: Every survey includes individuals who respond partially, or not at all, to the questionnaire. This can greatly affect data quality because nonrespondents may differ in some systematic way from respondents. For technical analysis, we call these "errors" of nonresponse, and divide them into several types. First, unit or household nonresponse occurs when no household members respond to any questions. The total eligible household sample loss for the 1984 panel was 22.3% [1]. Second, person or within unit nonresponse occurs when no responses are obtained for an individual in an interviewed household. It is estimated that after five interviews, roughly 20% of persons interviewed at Wave 1 of the 1984 panel missed one or more interviews. [13] Thirdly, item or question nonresponse occurs when an item on a questionnaire is not answered, but should be. Item nonresponse rates vary among characteristics. For example, item nonresponse rates for income amounts by types for the 1984 panel ranged from about 5% to 17% for the first quarter of 1984. [9] When responses for an individual are obtained for some but not all waves of a survey (wave nonresponse) a gap occurs in the longitudinal data and causes problems for analysts. For example, an individual may respond in the third and fifth waves, but not in the fourth. During the first five waves, about 5% of those who responded in Waves 1 and 5 of the 1984 panel, did not respond in one or more of the remaining three waves. [13]

Measurement Error: This occurs when what the survey records does not reflect the respondent's experience. Some sources of this error include memory loss, deliberate distortion, and misplacement of events in time. Individual sources of this error are difficult to detect and measure since the effects of one may cancel another. However, measurement error increases transitions or gross flows and the bias of estimates. In addition, it adversely affects the covariance structure of the data.

Sampling Error: This error occurs because data is collected from a sample of the population for purposes of calculating population estimates instead of from a complete census. Coefficients of variation (CVs) are a measure of the extent to which the results of the sample differ from the value being estimated. Based on the 1984 panel, for the fourth quarter of 1984, the CVs for mean monthly income of all households, nonfamily households, and single person female households are respectively 1.2%, 4%, and 8%. [9] Corresponding CVs for smaller subgroups will be larger. For example, the CV for single person female aged 20-24 households will be greater than 8%. Because of reduction in sample size from 20,000 interviewed households at Wave 1 of the 1984 panel to 12,300 households in February 1986, CVs for estimates obtained from the 1986 panel are expected to increase by roughly 30%. [9]

B. Other Sources

Construction and Maintenance of Sampling Frames: SIPP uses a multiple frame for sample selection. Errors arise from duplication of addresses and noncoverage. For example, persons entering the SIPP universe will have no chance of inclusion in the sample after the panel is introduced unless they move into a household which is being interviewed. [9]

Implementation of Sampling Scheme: Operations used to obtain the sample introduce what is believed to be a small amount of error. The errors come from missing units within addresses, listing addresses in wrong areas, interviewing wrong addresses, or because boundaries of

areas from which addresses are to be sampled are unclear. [9]

Data Preparation: In preparation for weighting and estimation, SIPP data is edited and keyed. Additionally, some coding and imputation is required. Each of these processes introduces some error. For example, individually imputed values will be different from true values and could adversely affect microlevel multivariate analysis. [9]

Estimation: Both the weights and variance estimates have some error associated with them. Weighting procedures used by SIPP are similar to those used by other large scale surveys such as The Current Population Survey (CPS). Thus, SIPP weights are expected to have similar kinds and degrees of bias as experienced by other surveys. In addition to bias resulting from the weighting methods used, errors may result from programming mistakes. [9] Variance estimates used to obtain generalized variance parameters are subject to a small degree of mean square error. However, SIPP's parameters are reasonable compared to CPS's parameters. [9]

Data Dissemination: SIPP data are released in the form of Public Use Data files or published reports. It is believed that errors introduced in this process are also minimal. [9]

V. CONTROL AND AVOIDANCE OF ERROR

Many procedures are used to control error and its effects. Some of the more important are explained below.

Interviewer error is controlled by intensive training, observation, and reinterview. Training is a continuous process and consists of initial training, refresher courses given twice a year and supplemental training in weak areas. The objective of this training is to minimize human error, improve understanding of the concepts and questions, and improve data quality. In addition to observation under the training program, interviewers are observed at least once a year and provided feedback on their general performance. The reinterview program is designed to check performance in the areas of household coverage, unit coverage, and income sources and helps to identify areas where improvement is needed either in interviewer performance or field procedures. For the reinterview program, a household is reinterviewed by telephone by a field supervisor shortly after a regularly scheduled visit using a condensed version of the questionnaire.

Quality control is used in the processing stages of bureau surveys to monitor quality of work and to identify work needing to be reprocessed and staff needing to be retrained. For example, keying data is under strict quality control. Each batch of a keyer's work is verified and errors are detected and corrected. If the keyer's error rate is above an acceptable limit, his/her work is checked 100 percent until the rate is 0.43 percent or less. Usually, the current error rate for SIPP is about 0.11 percent [9].

Feeding back income data collected in the previous interview is used in SIPP to help with recall. At each subsequent interview, the respondents are asked if they still have the income types reported four months earlier or if there have been any changes. Also, an experiment with feeding back assets and liabilities reported in the previous year was conducted in Wave 7 of the 1984 panel to determine if feedback could provide more accurate response. The results, however, do not show a noticeable impact on the associated estimates.

Several methods are used to control and avoid nonresponse error. First, enumerators receive training to help reduce nonresponse and are motivated to diligently seek responses since their evaluations depend upon obtaining low household nonresponse rates. Secondly, in Wave 1 of the 1987 panel, a sample of households were offered token gifts of appreciation to determine if these help decrease household attrition. Additionally, nonresponse is compensated for by imputation and weighting. For cross-sectional estimates the weighting process inflates the weights of interviewed households to compensate for similar non-interviewed households and imputation is used to compensate for both person and item nonresponse and replaces unanswered questions by values obtained from similar respondents. At present there is no longitudinal imputation procedure to compensate for wave nonresponse; persons who miss interviews are zero weighted and persons who respond in all waves have their weights increased. The success of these complex techniques in avoiding bias is unknown.

In addition, many other things are done or will be in the near future to help avoid, measure, or adjust for errors. Some examples of these are listed below.

To avoid errors in estimates, a thorough review of the weighting is done by mathematical statisticians and subject matter specialists. It is believed that most errors are detected and corrected. Additionally, research to improve estimation is being conducted. First, administrative data, such as IRS records, are being investigated as possible sources for controls in SIPP weighting to reduce mean square error of income and related estimates. Second, research is planned on the best method for combining data from more than one panel to produce more reliable estimates. Also, reduction of mean square error in variance estimation will be accomplished by using a new estimation procedure.

Administrative records studies currently under way include matching individual records on recipiency and developing a model of SIPP response and imputation errors in measure of program participation and amounts received. Such studies may assist in the improvement of gross flow estimates. [14]

SIPP data released in the form of Public Use Data files or published reports are reviewed by Census Bureau staff and meet all statistical standards set by the Census Bureau.

A SIPP User's Guide explaining concepts, limitations of the data, universe, time frames, interviewing techniques, and other survey differences from year to year is published to help data users avoid errors.

VI. SUMMARY

A. Overview

This paper discusses sampling and nonsampling errors associated with SIPP estimates. The magnitude of sampling errors can be estimated, but quantification of the various sources of nonsampling error and their impact on estimates is difficult, if not impossible. A sense of the overall effect of nonsampling errors can be obtained by comparing SIPP estimates to those from independent sources. Tables 1 through 3 provide estimates from SIPP and other sources which, with other information in this document and [9], can be used by analysts to determine the quality of SIPP estimates. The acceptable quality will differ according to the particular use to be made of the data.

Table 1 compares SIPP and CPS income estimates with independently derived estimates for a selected group of income types. These comparisons have been made based on the aggregate income received by the population and indicate some variation in the ratios within a year between different income surveys. Typically the estimates of income from the two surveys fall short of those derived from independent sources. The shortfalls in the SIPP estimates for monthly figures are, in most cases, less than the CPS shortfalls for annual amounts. [9] Table 2 presents a few selected characteristics derived from the SIPP longitudinal research file and compare them to 1983 and 1984 CPS estimates.

SIPP sampling and nonsampling errors can be reduced by additional expenditures. Sampling errors for SIPP can be improved by increasing sample size and/or by combining the sample from two panels. Nonsampling errors can be reduced by improving various aspects of the survey. Described below are areas in which improvement may have an important impact on meeting SIPP objectives.

Estimates of Gross Flows: One of the major goals of SIPP is to provide longitudinal estimates of gross flow (transition from one state of economic or labor condition to another state i.e. income or recipiency). These estimates can be very useful in explaining social-economic events that are relevant to existing or new government policies. Preliminary analysis of unweighted data from SIPP suggests that the gross flows for pairs of two consecutive months reported in the same interview are considerably lower than those reported from two consecutive interviews. [2], [8], and [10] However, a study on food stamp transitions by Judkins [7] was encouraging. It showed that transitions based on the combination of months within a reference period and months between reference periods are very close to those derived from administrative sources. (Table 3). These analyses show that the microlevel estimates may be seriously affected by nonsampling errors and that, at least for food stamps, macrolevel estimates may not be. They also suggest that all or most transitions are reported, but are misplaced in time.

Coverage Within the Unit: Evaluation of SIPP coverage shows, like other surveys, a differential coverage by age, race and sex. (Coverage is the ratio of the SIPP estimate to an independent estimate.) This coverage is worst (64%) for black males of 22-24 years of age [9]. Longitudinal and cross-sectional estimates which are highly correlated with poor coverage groups may be seriously biased. For example, for households with black males aged 22-24, income estimates may be biased because of the low coverage of such persons.

Nonresponse: The sample loss due to household (unit) noninterviews increases as a panel ages. This rate starts at 5 to 7% and reaches over 20% by the last interview. Also, some members of an inter-viewed household may not be interviewed. In addition to unit and persons in unit nonresponse, item nonresponse is also present in SIPP. The item nonresponse varies by item. For some items, such as market value of stocks and mutual funds, the item nonresponse rate is as high as 41%. For other items, such as Aid to Families with Dependent Children in the fourth quarter of 1984, the item nonresponse rate is only 5.5%. After the overall item nonresponse rate is even higher.

Wave nonresponse creates a gap in the longitudinal data and causes problems for analysts. Using imputation to fill the gap for these cases affects gross flow estimates, while dropping these cases increases variances.

Time-in-Sample Effect: The number of times respondents are interviewed may affect their responses. Although more investigation is required, the percentages of table 1 suggest that this phenomenon may be affecting estimates of food stamp amounts , but not wage and

salary income estimates.

Combining Panels: Because more than one SIPP panel covers a given reference period, comparable data from two panels may be combined to produce more reliable estimates than can be obtained from one panel.

B. Conclusion

SIPP provides a wealth of data which could be used to serve various important goals of the data users. This paper, due to space constraints, provides a brief overview of the data quality. A detailed discussion is presented in [9]. While much research remains to be done, the information in this paper and in [9] will allow data users to judge the quality of the SIPP data for their analytical purposes.

ACKNOWLEDGMENTS

The preparation of this document involved the efforts of many people. Guidance was received from many subject matter specialists within the Census Bureau and the SRM/SSRC Work Group on Technical Aspects of SIPP. Special thanks also go to Kimberly Wilburn and Carol Burroughs for their secretarial support.

Table 1. SIPP and CPS Estimated Aggregate Income Amounts Received As Percents of Independently Derived Estimates.

	SIPP		CPS (1983)
Wage and Salary	3Q83	95.0	99.0
	4Q83	94.3	
	1Q84	93.2	
	2Q84	94.4	
	3Q84	95.2	
	4Q84	94.5	
Unemploy. Comp.	3Q83	100.9	75.5
	4Q83	105.8	
	1Q84	85.2	
	2Q84	83.1	
	3Q84	80.3	
	4Q84	100.9	
Food Stamps	3Q83	90.1	71.2
	4Q83	83.1	
	1Q84	85.2	
	2Q84	86.2	
	3Q84	84.6	
	4Q84	83.6	

Table 2. Mean Annual Income Amounts And Estimates of Persons Ever Receiving Benefits from Selected Programs from the March CPS and SIPP 1983-1984 Longitudinal Research File.

Income Sources	SIPP 1983 -1984	March CPS 1983	March CPS 1984
Social Sec.	$4,512 (34,122)	$4,583 (32,182)	$4,358 (31,731)
Fed. SSI	$2,248 (3,941)	$2,366 (3,568)	$2,221 (3,442)
AFDC	$2,980 (3,987)	$3,072 (3,561)	$3,034 (3,468)
Pensions:			
Federal	$10,115 (1,937)	$11,032 (1,555)	$11,013 (1,609)
Military	$11,586 (1,297)	$10,267 (1,493)	$10,538 (1,337)
Dividends	$1,427 (26,807)	$1,543 (19,858)	$1,459 (18,690)

Table 3. Start-Up Rates (%) for Food Stamp Participation from SIPP and the Urban Institute.

SIPP 84 Panel

Ref. Months:	1-2	2-3	3-4	4 -5	Avg
Start-Up	4.9	4.7	4.5	10.9	6.2
(s.e.)	(.8)	(.8)	(.7)	(1.1)	(.5)
Exit Rate	3.3	3.5	3.1	12.8	5.7
(s.e)	(.7)	(.7)	(.6)	(1.2)	(.5)

Urban Institute

Cal. Months:	Jun to Jul	Jul to Aug	Aug to Sep	Sep to Oct	Oct to Nov	Nov to Dec	Avg
Start -Up	6.7	6.9	6.1	6.2	6.7	5.0	6.3
(s.e)	(.6)	(.6)	(.5)	(.5)	(.6)	(.5)	(.3)
Exit Rate	7.3	5.8	6.7	7.0	6.1	5.1	6.3
(s.e)	(.6)	(.5)	(.6)	(.6)	(.5)	(.5)	(.3)

REFERENCES

[1] Bowie, C., "SIPP Operational Statistics, Report #6," Internal Census Bureau Memorandum from Bowie to Distribution List, January 7, 1987.

[2] Burkhead, D. and J. Coder (1985), "Gross Changes in Income Recipiency from the Survey of Income and Program Participation," Proceedings of the Section on Social Statistics. American Statistical Association.

[3] Coder, J., and A. Feldman (1984), "Early Indications of Item Nonresponse on the Survey of Income and Program Participation," Proceedings of the Section on Survey Research Methods, American Statistical Association.

[4] David, M. (1985); "Introduction: The Design and Development of SIPP," Journal of Economic and Social Measurement, Vol. 13, No.3 and 4.

[5] Frankel, D., "Summary of the Content of the 1984 Panel of the Survey of Income and Program Participation," SIPP Working Paper Series No. 8504, Bureau of the Census.

[6] Herriot, R., and D. Kasprzyk (1984), "The Survey of Income and Program Participation," given at the 1984 American Statistical Association, published with other papers given in The Survey of Income and Program Participation and Related Longitudinal Surveys: 1984, U.S. Department of Commerce, Bureau of the Census.

[7] Judkins, D.R., "SIPP Gross Flows: Validation of Food Stamp Turnover", Internal Census Bureau Memorandum from Judkins to Singh, October 16, 1986.

[8] Kalton, G., D. McMillen, and D. Kasprzyk (1986), "Nonsampling Error Issues in the Survey of Income and Program Participation," Proceedings of the Second Annual Research Conference, Bureau of the Census.

[9] King, K., R. Petroni, and R. Singh, "Quality Profile for the Survey of Income and Program Participation," SIPP Working Paper Series, No. 8708, Bureau of the Census.

[10] Moore, J., and D. Kasprzyk (1984), "Month-to-Month Recipiency Turnover in the ISDP," Proceedings of the Section on Survey Research Methods, American Statistical Association.

[11] Moore, T., "Changes in the Survey of Income and Program Participation (SIPP) Sample," Internal Commerce Department Memorandum From Butz to Ortner, April 1987.

[12] Nelson, D., D. McMillen, and D. Kasprzyk, "An Overview of the Survey of Income and Program Participation," SIPP Working Paper Series No. 8401, Update, Bureau of the Census.

[13] Short, K., and E. McArthur (1986), "Life Events and Sample Attrition in the Survey of Income and Program Participation," to be published in the Proceedings of the Section on Social Statistics, American Statistical Association.

[14] Singh, R., L. Weidman, and G. Shapiro (1986), "Gross Flow and Spell Estimates from SIPP," presented at the International Symposium on Panel Surveys, Washington, D.C., November 1986.

[15] U.S. Bureau of the Census (1978), "The Current Population Survey: Design and Methodology," Technical Paper 40, Washington, D. C., U. S. Government Printing Office.

[16] U.S. Bureau of the Census, "SIPP Combined Objectives," Internal Census Bureau memorandum from R. Herriot for SIPP Off-Site Conference Attendees, February 24, 1986.

[17] Weidman, L. (1986), "Investigation of Gross Changes in Income Recipiency from the Survey of Income and Program Participation," Proceedings of the Section on Survey Research Methods, American Statistical Association.

[18] Ycas, M. and C. Liniger (1981), "The Income Survey Development Program: Design Features and Initial Fundings," Social Security Bulletin, November 1981, Vol. 44, No.11.