The Survey of Income and Program Participation (SIPP) is a new, ongoing nationally representative household survey program of the Bureau of the Census. The SIPP is designed to improve the measurement of information related to the economic situation of households and persons in the United States, and is the culmination of a large-scale development program, the Income Survey Development Program (ISDP), which examined concepts, procedures, questionnaires, recall periods, and the like (Ycas and Lininger, 1981).

This paper discusses specific methodological, survey design and statistical issues of concern to the program, reviewing in each case work completed or in progress. Topics to be discussed include: (1) questionnaire design; (2) respondent rules; (3) data collection mode; (4) length of reference period; (5) following movers rules; (6) sampling for special subpopulations; (7) nonresponse and imputation; (8) accessing SIPP microdata; (9) longitudinal concepts; (10) statistical estimation for longitudinal concepts; (11) program record check study; (12) composite estimation; and (13) SIPP gross flow data.

Income Survey Development Program

The need for a survey like SIPP arose because of the limitations of the March Income Supplement of the Current Population Survey, the principal source of information on the distribution of household and personal income in the United States. These limitations are inherent in the survey design, survey instrument, and survey procedures and can not be easily modified. As a consequence the Income Survey Development Program (ISDP) was established in 1975 by the U.S. Department of Health and Human Services to develop methods to overcome the shortcomings of the CPS--1) the underreporting of property income and other irregular sources of income; 2) the underreporting and misclassification of participation in major income security programs; and 3) the lack of information necessary to analyze program participation and eligibility.

The collection of data in the ISDP centered around several field tests conducted to examine different concepts, procedures, questionnaires, recall periods and the like. Several features distinguished these field tests from other data collections, particularly the CPS. They include: 1) interviews obtained at regular intervals within a year; 2) most types of income reported on a monthly basis; 3) income reported on an individual basis; 4) individuals followed over the survey period to obtain data on changes in income and family composition; and 5) information collected on special topics such as disability, child care, fertility, net worth, and taxes paid to provide insight into the context of program benefits, program receipt, and economic well-being.

Because the ISDP was the predecessor to SIPP, many ISDP characteristics can be seen in the SIPP design, including the survey design, content, and questionnaire format.

Survey of Income and Program Participation

SIPP began in October 1983 as an ongoing survey program with one sample panel selected to represent the noninstitutional population of the United States; about 21,000 households were occupied and eligible for interview. Each household is interviewed once every four months for approximately 2 1/2 years; the reference period for the principal survey items is the 4 months preceding the interview. This interviewing plan results in eight interviews per household. Each year a new, slightly smaller panel is introduced. This design allows cross-sectional estimates to be produced from a combined sample of approximately 25,000 households. Information concerning the SIPP design, content, and operations can be found in Nelson, McMillen, and Kaspzryk (1984). In the following, design and methodological issues are discussed; in some cases the importance of the issue has its origin in the early days of the SIPP development program.

Questionnaire Design

The principal effort of the ISDP, improved questionnaire design, was directed to overcoming CPS income measurement problems which lead to underreporting and misclassification of income. In an ISDP field test two questionnaires approaches were developed. For simplicity, one version may be referred to as the "short" form and the other as the "long" form.

The short form approach attempted to gather income data directly while keeping respondent burden at a moderately low level. The general strategy of the long form approach was to isolate events, experiences, and other attributes associated with the receipt of specific types of income. The long form contained an extensive set of probes about the receipt of income and lengthy questions to ascertain income amounts.

The hypothesis tested was that the long form approach produces more complete and accurate reporting of income. Olson (1980) provides a summary of the analysis conducted on the two questionnaire formats. Several analytical approaches were implemented and are discussed in Olson's summary. The long form was perceived by both interviewers and respondents as less burdensome and also was shown to have higher income reporting rates.

An experiment on alternative questionnaire formats was included in the ISDP; this experiment contrasted a household screening format with a person-based format which had evolved from prior field tests. The household screening approach was based on a revised version of the questionnaire used in the April 1978 CPS Income Supplement Test. This approach was expected to reduce interview time without reducing data quality.

It was contrasted with a person-based approach. Under this approach, questions on all sources of income were asked of the first household member, then repeated for the second, and so on. A separate form was filled out for each adult in a sample household, but extensive use was made of skip instructions and check items to reduce the number of questions asked of any one respondent.

Differences in data quality obtained with the two questionnaire formats and differences in the interview times appeared slight. Large differences were not observed between the two approach-
es in estimates of income recipiency rates, and in the incidence of "don't know" and "refusals." Since the household screening format did not offer a significant improvement over the person-based approach, this person-based format, with modest improvements and refinements, was adopted for SIPP.

Questionnaire design issues and discussions concerning data collection procedures continue to be a part of the SIPP program. The general issue which now occurs is whether interviews conducted without the use of responses from previous interviews (the so-called independent approach) produce better estimates than interviews conducted using the previous interview responses to remind respondents of earlier statuses (the so-called dependent interview approach). In the SIPP, a dependent approach is used to update income receipt patterns, but the approach has not been evaluated.

A similar dependent approach to data collection is also possible with the data collected on personal net worth. These data are obtained at two points-in-time, one year apart. Specifically, data on asset and liability values are collected in both Wave 4 and Wave 7 of the 1984 Panel. To examine differences between the dependent and independent approach one half the sample in Wave 7 was provided information collected on asset and liability values collected in Wave 4, while the other half was not provided the previously reported information. An evaluation of the two approaches will be conducted in 1987.

The same questionnaire design issue, the dependent versus independent interview, has also occurred in the repeated measurement of industry and occupation data. During the 1984 and 1985 SIPP panels these data were collected independently during each interview even though the individual had not changed employers.

The independent collection of industry and occupation data has, however, several problems. Undue variation in occupation classification can result when respondent descriptions of duties vary slightly or when the interpretation of the written description varies between the clerical staff members assigning the classification codes. Research into this problem has provided some estimates (which appear unusually high) of the number of times occupation and industry classifications change from interview to interview for persons with the same employer.

As a result, a modification was made in the 1986 Panel to reduce changes in occupation and industry codes resulting from random response error and clerical interpretation, and to reduce interview time. The modification introduces a "screener" question that asks if activities or duties have changed during the past 8 months. A negative response eliminates the detailed occupation and industry questions. The occupation and industry classifications would then be derived from responses given in the previous interview. It is important to note that while this change is being made for the 1986 Panel, industry and occupation data from the 1985 Panel collected during the same time period will be collected independently each wave, giving rise to a natural experiment embedded in the two panels.

Respondent Rules

When interviewing households with more than one member, a problem which must be addressed is the extent to which proxy responses are acceptable. Since not everyone may be present at the time of the interview, both time and money can be saved by asking another household member about the person who is not present. The difficulty with this is that along many dimensions of the survey instrument, the proxy report may result in less accurate data than the self-report (Kalton, McMillen, and Kasprzyk, 1986).

A formal test of respondent rules, conducted in the ISDP, compared the quality of reporting in a treatment group where proxy interviews are accepted from any household member who feels qualified to answer for a missing person with a treatment group where interviews must be in person except for extreme situations (respondent physically or mentally incapable, unable to speak English, etc.). Refusal rates were slightly higher for the self-response treatment and the percent of households interviewed was slightly higher for the proxy response treatment. The differences, however, were too small to give insight into which rule should be preferred. Differences between treatment groups in reported income recipiency rates also appeared to be small and unaffected by the response rule, and combined "don't know" and "refusal" rates for income amounts of various income types were not consistently lower under the self-response mode.

Under self-response rules, records were used more often by persons when answering wages and salary questions, and response rates on hourly wage rates were higher, but in general the evidence for either set of response rules was not conclusive. Thus, as a result of these findings, estimated costs for using a self-response rule, and the implementation of a "call back" procedure to obtain critical information unavailable at the time of the interview, the SIPP respondent rules now allow proxy interviews to be taken.

A related problem is the response rule for college students. Students are usually considered members of their parents' households until they establish a permanent residence elsewhere. Thus, the usual procedure for students living away from home while attending school treats them as household members temporarily absent and obtains proxy interviews from other members of their parents' household. In order to measure the accuracy of information taken from proxy interviews for students living away from home, one interview during a field test of the ISDP was first obtained by proxy at the parents' household and then by self-interview at the student's school residence. The results of this study are described in Roman and O'Brien (1984). The procedures adopted for SIPP were the usual Census Bureau procedures as described above.

Data Collection Mode

The SIPP has followed the practice of conducting most interviews (approximately 95 percent) in person (Kalton, McMillen, and Kasprzyk, 1986). Because of the rising costs of a personal interview the Census Bureau is considering the possibility of conducting a substantially larger number of SIPP interviews by telephone.
As a result, a SIPP National Telephone Test is planned from August to November 1986 to study the large-scale use of the telephone in SIPP. Households within 50 percent of the segments will be designated as maximum telephone interview cases; the remaining 50 percent will be maximum personal visit cases. Interviewers will conduct almost all of the telephone interviews from their homes.

Length of Reference Period

The ISDP focused on data collection techniques designed to improve the reporting of cash and noncash income, and as such the length of the reference period for most survey items was a critical design decision.

This issue was addressed twice during the ISDP. First, a single interview using a six month recall period was compared with two consecutive interviews, both using 3-month reference periods (Olson, 1980). Second, an experiment was conducted comparing reported property income using a three month recall with a six month recall period. Analyses of the latter experiment were not conducted due to the withdrawal of funding for the development program.

The results of the first experiment along with the additional ISDP experience led to a four month recall period in SIPP; this decision maintains cost at the appropriate budget level while trying to maintain satisfactory data quality.

Following Movers Rules

An important design feature in the ISDP and now the SIPP is that all persons in a sample household at the time of the first interview remain in sample during the 2 1/2 year period of the panel; this rule holds even if a person(s) should move to a new address—subject to a minor geographic constraint.

A sample of addresses was selected and individuals were identified at these addresses at the time of the first interview. After the first interview, the sample was no longer address-based but rather person-based, consisting of all individuals enumerated during the first interview. Thus, these people and anyone they share living quarters with are interviewed in subsequent interviews.

During the ISDP two issues concerning movers were important to 1) the production of cross-sectional estimates at each interview; and 2) the costs associated with following movers. Huang (1984) presents several unbiased base weights for cross-sectional estimates for the noninstitutionalized population when the sample contains movers.

The issue of costs was addressed by a study conducted as part of the ISDP. The purpose of "The Mover's Cost Study" was to shed some light on the data collection costs resulting from following movers to their new addresses. White and Huang (1982) describe the study and provide some results based on the movers procedures adopted for the field test.

Recently, Jean and McArthur (1984) discuss data collection issues in the SIPP as they pertain to movers and offer recommendations to improve coverage in future SIPP panels. They report that only 80 percent of movers between the first and second waves of the SIPP Panel were traced. Kalton and Lepkowski (1985) also discuss the procedures adopted in SIPP, and suggest a research program aimed at measuring the extent of noncoverage from various sources.

Sampling for Special Subpopulations

Early design goals of the ISDP emphasized a concern for improving the reliability of subpopulation estimates. This was exhibited in the emphasis placed in the ISDP on sampling from administrative program lists. Thus, samples were oftentimes drawn from lists of current participants of Federal or state administered programs. The problems experienced and the lessons learned are discussed in Kasprzyk (1985).

During 1984 a Census Bureau Working Group analyzed subsampling (screening) proposals for oversampling special populations. The issue under consideration was the reliability of estimates when different subsampling schemes are introduced. Results of that investigation are reported in a Census Bureau memorandum (1985).

Nonresponse and Imputation

For longitudinal surveys such as those of the ISDP and the SIPP, the problems of refusal and selective nonresponse are compounded by cumulative losses in responses over the course of the panel. Therefore, an important aspect of both the ISDP and SIPP work is the study of methods for compensating for nonresponse. To that end, Kalton (1983) reviews procedures currently used in survey research. Imputation procedures were also discussed by Kalton and Kasprzyk (1982), where bias and variance properties for several classes of procedures are summarized.

SIPP data can be treated as both cross-sectional and longitudinal. Procedures to compensate for unit nonresponse in SIPP as well as other Census Bureau surveys are described in Bailey, Chapman, and Kasprzyk (1985). They also describe the complications arising in the treatment of unit nonresponse in a multi-interview survey.

Due to complete interview nonresponse patterns, data gaps can be treated through weighting adjustments as one typically treats unit nonresponse or the gaps can be assigned data through an imputation system and treated as item nonresponse. The Survey Research Center/U. of Michigan has begun to study this topic (Kalton, 1985; Kalton, Lepkowski, and Lin, 1985; Kalton and Miller, 1986). Preliminary results, taking into account the longitudinal relationship between the variable and simple prediction models, show imputation can be more efficient than weighting. However, the practical realities of developing good imputation models for a missing wave is a significant undertaking.

The complexity of the missing data problems exists even when unit/person response is complete, because data at the item level can be present, missing, or not applicable for each interview. The treatment of item nonresponse in a longitudinal manner has been a topic of substantial interest for the SIPP (Kalton and Lepkowski, 1983). Other work has occurred recently which lays the foundation for future evaluations of the SIPP longitudinal imputation procedures.

An investigation of the feasibility of using model-based imputations has been conducted (Huggins and Weidman, 1986a; Huggins and Weidman, 189
1986b). Models which impute missing response patterns based on the frequency distribution of response patterns have also been investigated (Samuhel, and Huggins, 1984; Huggins, Samuhel, and Weidman, 1985), and several imputation procedures for continuous data were compared in a small simulation study (Huggins, 1986).

Accessing SIPP Microdata

To facilitate the processing of SIPP data, the Census Bureau is proceeding along two paths: 1) developing software for creating, extracting, and analyzing longitudinal data, and 2) looking at alternative methods of accessing large complex data sets.

One problem presented by SIPP is that many of the statistics to be calculated from SIPP data are not options in any tabulation or statistical software. Within the Census Bureau general software has been written to calculate and develop extract files of spell characteristics such as duration, and beginning and ending dates of spells as well as the characteristics of individuals during spells in a particular state. Census Bureau staff has also continued research into alternative methods of accessing large complex data sets. Following internal experimentation with SIPP and the Scientific Information Retrieval (SIR) data base management system, the feasibility of using RAPID to access SIPP data is under evaluation. RAPID is a data base management system developed by Statistics Canada for processing Canadian Census and survey data (Jeays, 1985). Two features of RAPID make it attractive for SIPP processing. First, RAPID works with a transposed data structure which implex data sets. Second, RAPID has interfaces with SPSS, SAS, and TPL. These user friendly interfaces are a central focus of the investigation.

Longitudinal Concepts

Household and family level analysis in a longitudinal survey is complicated by the fact that the composition of households and families can change over time, since original sample persons leave to join other households or families, or to set up new ones. The principal issue is the development of definitions of households and families which account for survey measurements at two or more points in time and which do not create serious conflicts with the traditional cross-sectional household and family constructs. Definitions have been refined through numerous internal discussions (McMillen and Herriot, 1985; Citro, 1985). Research examining alternative definitions of longitudinal households and measures of annual income status and family type has been reported by Citro, Hernandez, and Herriot (1986) and Citro, Hernandez, and Moorman (1986). This research has indicated that the choice of definition does not affect annual measures of low income status or of households by type. If this finding does not change after additional research, considerations, such as ease of implementation and operational simplicity, will be the determining factor in choosing a longitudinal household definition.

Statistical Estimation for Longitudinal Concepts

Research on estimation has proceeded along two dimensions--longitudinal person estimation and longitudinal household (family or recipient unit) estimation. The work on person estimation includes the calculation of selection probabilities to yield unbiased longitudinal estimates of individual characteristics and the use of controls in additional stages of estimation (Judkins et al., 1984). A refinement of this work and a description of the method proposed to produce longitudinal weights for person analysis covering the first three SIPP interviews has been reported by Kobilarcik and Singh (1986).

The topic of longitudinal household (family or recipient unit) estimation is also under study. Some approaches to this issue were presented at the 1984 meeting of the American Statistical Association (Ernst, Hubble, and Judkins, 1984). As more precise statements of various longitudinal household definitions were developed, research on obtaining unbiased weights and developing weighting adjustments continued (Ernst, 1986). Recent empirical research by Citro, Hernandez, and Moorman (1986) will considerably simplify estimation for longitudinal households.

Program Record Check Study

A record check project has been initiated to investigate response quality issues in SIPP through a case-by-case comparison of SIPP data with administrative record information. The research questions to be addressed in the project include: 1) the quality of recording report quality; 2) the effects of recall period length on report quality; 3) the nonexperimental effect of self/proxy respondent status on report quality; 4) the extent of misclassification errors; 5) the between-wave turnover effects; and 6) demographic correlates of report quality.

Four State administered programs and six Federally administered programs are included in the study; some results are expected in 1987.

Composite Estimation

Composite estimation is a technique that combines estimates from the current and previous time periods with the goal of improving the precision of survey estimates by taking advantage of the correlations between responses for the same analytic units at different time periods. Composite estimation is particularly effective when the correlations are high, which is likely to be the case for many important data items in SIPP. A preliminary review of types of composite estimators appropriate for the SIPP data structure is available (Chakrabarty, 1986).

SIPP Gross Flow Data

Analysis of program data on a month-to-month basis in ISDP revealed a tendency for reported program turnover to occur between waves of Interviewing more often than within the wave (Moore and Kasprzyk, 1984). Further analysis of this topic was deferred until the availability of three observations from the SIPP. The analysis (Burkehead and Coder, 1985) which covered month-to-month changes in recipiency of income benefit amounts for a 12-month period focused on changes occurring between the last month of one reference period and the first month of the succeeding reference period (between months 4 and 5 and between months 8 and 9 in sample.) The results using SIPP data are similar to the ISDP results, where an uneven pattern of change is observed and
that this pattern is clearly associated with the interviewing scheme. The main causes for the problem seem to be questionnaire wording/design, respondent recall error, and the interaction between these two factors.

The concern about this topic has increased since the first data were collected in the SIPP, and several research projects have begun to study the potential biases in gross flow estimates due to response error. Staff have developed a model to estimate biases in gross flow estimates that result from response error. They have developed a methodology for estimating the parameters of the model using SIPP response error rates and the ratios of within-wave and between-wave flow estimates. Using this model and methodology, the bias, using several assumptions can be computed. Preliminary results are now available (Hubble and Judkins, 1986).

Other staff are also involved in an effort to determine the causes for bias in SIPP gross flow estimates. Exploratory data analysis to determine variables related to over/under reporting and if imputed responses are responsible for a disproportionate number of between-wave transitions has begun. An empirical examination did not detect any real relationships between gross change distributions and key demographic variables (Weidman, 1986).

Conclusion

The topics discussed above summarize issues and interests of the SIPP program through August 1986. During the summer of 1986, an internal Research and Evaluation Committee was formed to identify research and evaluation projects to be initiated in 1987. Projects to be identified fall under two general areas: 1) projects which assist the Census Bureau in the production of a long-term household estimation for the proposed longitudinal definition. SIPP Working Paper Series No. 8605, Washington, D.C.: U.S. Bureau of the Census.


REFERENCES


