Survey Methodology, June 1988
Vol. 14, No. 1, pp. 45-58
Statistics Canada

Research Issues in the Survey of Income and Program Participation[1]

DANIEL KASPRZYK[2]

ABSTRACT

The Survey of Income and Program Participation (SIPP) is an ongoing nationally representative household survey program of the Bureau of the Census. The primary purpose of the SIPP is to improve the measurement of information related to the economic situation of households and persons in the United States. It accomplishes this goal through repeated interviews of sample individuals using a short reference period and a probing questionnaire. The multi-interview design of the SIPP raises methodological and statistical issues of concern to all panel surveys of families and persons. This paper reviews these issues as they relate to the SIPP. The topics reviewed are: 1) questionnaire design; 2) data collection, including respondent rules, data collection mode, length of reference period, and rules for following movers; 3) concepts, design, and estimation; and 4) response error.

KEY WORDS: Panel surveys; Questionnaire design; Survey design; Longitudinal estimation; Response error.

1. INTRODUCTION

The Survey of Income and Program Participation (SIPP) is an ongoing nationally representative household survey program of the U.S. Bureau of the Census. It provides comprehensive information on the economic resources of the American people and on how public transfer and tax programs affect their financial circumstances. The data from the SIPP provide government policy makers with an information base for studying the efficiency of government tax and transfer programs, for estimating future program costs and coverage, and for assessing the effects of proposed policy changes. 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).

The need for a survey like SIPP arose because of the limitations of the March Income Supplement of the Current Population Survey (CPS), 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 was established in 1975 by the U.S. Department of Health and Human Services to develop methods to overcome the principal 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 other types of information that people generally find difficult to report accurately (for example, monthly detail on income earned during the year);

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

[2] Daniel Kasprzyk, SIPP Research and Coordination Staff, United States Bureau of the Census, Washington D.C. 20233

and (3) the lack of information necessary to analyze program participation and eligibility. Several features distinguish field tests of the ISDP from other data collections, particularly the CPS. They include: (1) interviews for the same persons were obtained at regular intervals within a year; (2) most types of income were reported on a monthly basis; (3) income was reported on an individual basis; (4) individuals were followed over the survey period to obtain data on changes in income and family composition; and (5) information was 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 dependency, and overall economic well-being. Because the ISDP was the predecessor to SIPP, many characteristics of the ISDP can be seen in the SIPP, including the survey design, content, and questionnaire format.

The SIPP began in October 1983 as an ongoing survey program with one sample panel of 21,000 households selected to represent the noninstitutional population of the United States. Each household is interviewed once every four months for approximately 2½ 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 panel is introduced. This design allows cross-sectional estimates to be produced from the combined sample of 2 panels. Information concerning the SIPP design, content, and operations can be found in Nelson, McMillen and Kasprzyk (1985).

This paper reviews specific methodological, survey design, and statistical issues of concern to the program. The general categories of interest are: (1) questionnaire design; (2) data collection, including respondent rules, data collection mode, length of reference period, and rules for following movers; (3) concepts, design, and estimation; and (4) response error.

2. QUESTIONNAIRE DESIGN

The principal effort of the ISDP was directed to overcoming problems which resulted in underreporting and misclassification of income in the CPS March Supplement. In an ISDP field test, two questionnaire 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. For each household member, questions were asked directly about the receipt of certain income types. If income were received, the amount received during the reference period was determined before proceeding to the next source of income.

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. This form contained an extensive set of probes about the receipt of income and lengthy questions to ascertain income amounts. Amounts associated with specific income types were not obtained until all sources of income were determined.

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 approaches to the analysis were implemented and are discussed in Olson's summary: (1) staff observation of training and interviewing; (2) debriefing sessions of interviewers and observers; (3) case-by-case reviews of completed questionnaires; (4) analysis of survey and item response rates; and (5) data analyses focussing on the quality of the data collected, and questionnaire edit failures, especially those associated with the inability of the interviewer to follow questionnaire skip patterns. The form adopted for further research and ultimately the SIPP was a variation of the long form. 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 with questionnaire formats was also included in the ISDP; this experiment contrasted a household screening format with a person-based format which had evolved from prior ISDP field tests. The household screening approach was based on a revised version of the questionnaire used in the April 1978 CPS Income Supplement Test and was intended to reduce burden by asking a single household respondent whether anyone in the household received a particular kind of income during the reference period. Each affirmative response was followed by a question to identify exactly which household member(s) received that type of income. Complete recipiency for all household members was recorded before asking about amounts of income received by specific individuals. This approach was expected to reduce interview time without reducing data quality.

The approach above 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 the quality of the data obtained with the two questionnaire formats and differences in the interview times appeared slight. Large differences were not observed between the two approaches in estimates of income recipiency rates, and in the incidence of "don't know" and "refusals." Interview time, expected to be significantly less under the household questionnaire approach, was about five minutes less per household and about three minutes less per person than the person approach. 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 part of the SIPP program. The general issue 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 at each interview, but the approach has not been systematically evaluated.

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

The rationale for this dependent or "feedback" approach was that respondents would provide more accurate estimates of change if they were first reminded of the amount they reported the previous year. If respondents know the amount of the change in asset values and were reminded of their beginning balance, then presumably their reporting of the current balance would be consistent with the true amount of change over the period. Lamas and McNeil (1987) analyze these data, but give no definite answer about the impact of the feedback approach since benchmark data are not available. They do, however, say that the dependent interview did not affect cross-sectional estimates and that the approach produced results consistent with expected differentials in net worth across subgroups. They also looked at micro-level changes in net worth

using only households with fully reported wealth data and found some evidence that the dependent interview reduced the estimates of the change in net worth.

The same questionnaire design issue, the dependent versus independent interview, has also occurred in the repeated measurement of industry and occupation. During the 1984 and 1985 SIPP panels these data were collected independently during each interview even though the individual had not changed employers. This procedure acknowledges the fact that an employee's duties may change from time to time and allows these changes to be recorded. Sufficient change in duties can result in a change in the person's occupation classification from interview to interview even though the employer has not changed.

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 of the number of times occupation and industry classifications change from interview to interview for persons with the *same* employer. Among individuals who reported the same employer during the first 12 months of the 1984 SIPP Panel, approximately 40 percent of these persons changed 3-digit occupation codes between two consecutive interviews and 20 percent changed 3-digit industry codes (Kalton, McMillen and Kasprzyk, 1986).

As a result, a modification was made in the 1986 SIPP 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 are then derived from responses given in the previous interview.

It is important to note that while this change was made for the 1986 SIPP Panel, industry and occupation data from the 1985 SIPP Panel, collected during the same time period, were still collected independently each wave, giving rise to a natural experiment embedded in the two panels. These data have not yet been analyzed.

3. DATA COLLECTION

Four topics affecting data collection in the SIPP are discussed below: (1) respondent rules; (2) data collection mode; (3) length of reference period; and (4) rules for following movers.

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 persons who are not present. The difficulty with this is that along some dimensions of the survey instrument, the proxy report may result in less accurate data than the self-report. Kalton, Kasprzyk and McMillen (1988) provide a discussion of this issue in the context of panel surveys.

A formal test of responent rules, conducted in the ISDP, compared the quality of reporting in a treatment group where proxy interviews are accepted from any household member who felt qualified to answer for a missing person with a treatment group where proxy interviews are not permitted except for extreme situations (respondent physically or mentally incapable,

unable to speak English, away from the household during the entire interviewing period, etc). About 85 percent of adults interviewed in the self-response rule households were self-respondents and about 65 percent were self-respondents in the usual or proxy response rule households. Thus, the implementation of the self-response rule resulted in approximately 20 percent more self-interviews than the other treatment (Coder 1980).

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. Person noninterview rates in households where at least one other adult was interviewed were higher under self-response rules than under usual response rules. 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 the self-response rules, records were used more often by persons when answering wages and salary questions, and response rates for 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 (4-6 percent higher than the proxy rule), and the implementation of a "call back" procedure to obtain certain 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 is to treat them as household members who are temporarily absent and obtain 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 an ISDP field test 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 by Roman and O'Brien (1984). The analysis presented is limited due to flaws in the administration and implementation of the test. The authors observed, however, that quite often a proxy cannot identify a particular source of student income and even if they can identify it, they are more likely to respond "don't know" to the particulars about that source. They also noted that the larger the income or expense, the better the proxy response becomes.

Data Collection Mode

The SIPP has conducted most interviews (approximately 95 percent) face to face (Kalton, McMillen, and Kasprzyk, 1986). Because of the rising costs of a face to face interviews, the Census Bureau is considering the possibility of conducting a substantially larger number of SIPP interviews by telephone.

As a result, a SIPP telephone interview pretest was conducted in June 1985 to assess the feasibility of "warm" telephone interviewing for SIPP — that is, telephone interviews for households which received a face to face interview at an earlier wave. The pretest was conducted in 2 of the Census Bureau's Regional Offices with a sample of 280 households. Refusal rates (about 2.5%) and noncontact rates (about 11%) were within staff's expectations. Item nonresponse rates showed no unexpectedly high nonresponse rates (U.S. Bureau of the Census 1986).

Following this, a SIPP National Telephone Test took place from August to November 1986 and February to April 1987; the purpose of the test was to study the large-scale use of warm telephoning in SIPP and to learn whether people are willing to furnish data by telephone for

2 interviews in a row. Households within 50 percent of the segments were designated as maximum telephone interview cases; the remaining 50 percent were maximum personal visit cases. Interviewers conducted almost all of the telephone interviews from their homes. Gbur and Durant (1987) report preliminary results from the first phase of the experiment. They indicate that household response rates did not seem to be seriously affected by the use of the telephone and person nonresponse rates were comparable by mode. Item nonresponse rates were only slightly affected by telephone interviewing. Additional results are forthcoming.

Length of Reference Period

The ISDP focussed 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 an important 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. Second, an experiment was conducted comparing reported property income amounts using a 3-month recall versus those with a 6-month recall period.

Olson (1980) describes some analyses conducted on the first experiment. Not surprisingly, using a 6 month recall period understates the proportion of income reported in earlier periods. This pattern held for a number of specific sources of income such as wages, Aid to Families with Dependent Children, and unemployment compensation. These findings though not definitive, support the presumption that longer recall periods increase chances of omission due to memory loss. Other analysis showed that the number of sources of income reported per household in the first three months of the six month reference period was lower than for the corresponding time using a three month reference period. Analyses of the second 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 for the SIPP; this decision maintains cost at the appropriate budget level while trying to maintain satisfactory data quality.

Rules for Following Movers

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 2-½ year period of the panel; this rule holds even if one or more persons should move to a new address. For cost and operational reasons, face to face interviews are conducted at new addresses that satisfy some geographic constraint — in the ISDP, the address had to lie within 50 miles of an ISDP primary sampling unit, while in SIPP, the address must lie within 100 miles of a SIPP primary sampling unit.

For each panel a sample of addresses is selected and individuals are identified at these addresses at the time of the first interview. After the first interview, the sample is no longer address-based but rather person-based, consisting of all individuals enumerated during the first interview. These people and anyone with whom they share living quarters are interviewed in subsequent interviews.

During the ISDP two issues concerning movers were important: (1) the production of cross-sectional point in time estimates at each interview; and (2) the costs associated with following movers. Huang (1984) presents several unbiased base weights for cross-sectional estimates of the noninstitutionalized population when the sample contains movers. He associates observations at any given point in time with the known inclusion probabilities of the original sample

households. Two approaches are described: (1) a multiplicity approach, which depends on the number of ways that a new household can be included in the sample; and (2) a "fair share" approach which assumes all household members contribute equally to their household. The SIPP as well as the ISDP adopted the "fair share" approach.

The issue of costs was addressed by a "Mover's Cost Study". This 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. They found that the number of eligible households for interview increased by 8.8 percent as a result of following movers during a one year time period; they also found that movers represented about 22 percent of the total sample after 15 months, and that during this period of time the number of interviewing hours increased by 7 percent and the number of miles charged by interviewers increased by 11.4 percent.

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. Kalton and Lepkowski (1985) also discuss the procedures for following movers in SIPP, and propose a research program aimed at measuring the extent of noncoverage from various sources and its concentration in particular subgroups. More recently, Jean and McArthur (1987), considering five waves of SIPP data, report that among persons who moved sometime after the first interview (that is, between Waves 2 and 5), 69 percent completed all 5 interviews, 23 percent did not complete the fifth interview, and 9 percent were interviewed in the fifth wave but were missing at least one intervening interview.

4. CONCEPTS, DESIGN AND ESTIMATION

During the ISDP and continuing with the SIPP program, significant research activity has taken place in the area of conceptualizing annual units of analysis using subannual data, and the statistical estimation of these concepts. The treatment of nonresponse in panel surveys has also been a topic of considerable research interest. Finally, estimation techniques to reduce sampling error and methods to sample subgroups have also been under study in the ISDP and SIPP programs.

Longitudinal Concepts

Annual family and household statistics are important indicators of the Nation's economic well-being. The SIPP collects subannual data, indeed monthly data, reflecting changes in the composition of households; these data allow the development of annual household statistics which reflect actual household composition experiences during the year, unlike current household statistics which simply ignore intrayear changes in household composition. The construction of annual units of analysis, whether they are households, families, or program units, raises methodological issues concerning longitudinal weights and imputation techniques. The main issue is, however, conceptual. Given intrayear composition change, when is it appropriate for annual measures to recognize change in household composition and when is it not? Put another way, how should households and families be defined 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.

Analysts at the Census Bureau have given considerable thought to the question of defining households and families over time (McMillen and Herriot 1985; Citro 1985). Empirical research to examine several 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). The empirical research emphasized four alternative concepts: (1) a household is the same over time if it has the same reference person; (2) a houshold is the same over time if it has the same principal person (this definition differs from the first in its treatment of married couple households for which the reference person may be either the husband or wife, but the principal person is always the wife); (3) a household is the same over time if it has the same reference person and is the same family type over time; and (4) a household continues over time if it has the same reference person, is the same family type, and has the same membership size.

This research has provided preliminary indications that the choice of definition does not appreciably 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 factors in the use of a longitudinal household definition.

Statistical Estimation for Longitudinal Concepts

Research on estimation for longitudinal concepts has proceeded along two paths — longitudinal person estimation and longitudinal household (family or program 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).

Kobilarcik and Singh define the longitudinal universe as the noninstitutional population (excluding military barracks) on December 1, 1983, the midpoint of the Wave 1 interview months. The sample from the longitudinal universe consists of eligible persons living in the selected living quarters at the time of the first interview. "Interviewed" persons for purposes of this estimation procedure are those who responded to each of the first three SIPP interviews, and who during the first interview lived in a household in which all eligible members responded to the interview, and those who resided in a Wave 1 interviewed household, but during the second or third interview died or moved outside the geographic boundaries of the survey.

Thus, noninterviewed persons in the estimation procedure are those who at the time of the first interview lived in a household in which at least one household member failed to respond to the first interview, and those who resided in a Wave 1 interviewed household but failed to respond at the second and/or third interview. All persons classifed as interviewed are assigned positive weights. Weights for this universe are derived in the usual way, using the reciprocal of the probability of selection, calculating an adjustment for noninterviews, and adjusting to demographic population controls. The nonresponse adjustment has two phases, an adjustment first for household nonresponse and then for person nonresponse, the latter using information collected during the first interview.

The topic of longitudinal household (family or program unit) estimation is also under study. Several approaches to this issue were reported by Ernst, Hubble and Judkins (1984) and more recently by Ernst (1988). The latter work describes why weighting by the reciprocal of the probability of selection does not, in general, work for longitudinal household and family estimates, and presents a class of weighting procedures which can accomplish this task. He, furthermore, describes the difficulties that can arise in applying these weighting procedures because the information necessary to create the weight may not be available. Ernst also presents conditions which,

if satisfied, by the longitudinal concept, are sufficient for there to exist a weighting procedure that avoids these problems. Finally, he discusses procedures for adjusting longitudinal concepts for nonresponse and for controlling demographic variables to independent estimates.

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 has been the study of methods for compensating for nonresponse. To that end, Kalton (1983) reviewed procedures used in survey research. Imputation procedures were also discussed by Kalton and Kasprzyk (1982, 1986), 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 the SIPP as well as other Census Bureau surveys are described in Chapman, Bailey and Kasprzyk (1986). Complications arising in the treatment of unit nonresponse in a multi-interview survey are described. In a panel survey, however, nonresponse may also occur, as item nonresponse, where a unit takes part in the survey but does not provide answers to all items, and as wave nonresponse where a unit provides data for some, but not all of the interviews.

Heeringa and Lepkowski (1986) describe general classes of longitudinal imputation methods which might be considered as an alternative to a cross-sectional hot deck imputation approach. They also empirically compare a simple longitudinal imputation method, longitudinal direct substitution, where the value of a nonmissing item is substituted from one time period to another when the same item is missing, with a cross-sectional hot deck scheme. Not surprisingly, they demonstrate that the direct substitution method for longitudinal imputation understates change. They concluded, however, that this may be preferable to the gross overstatement of change resulting from the use of the cross-sectional hot deck method.

Panel surveys have an additional type of missing data problem called wave nonresponse. The amount of missing data for an individual with wave nonresponse is typically greater than that encountered for records with item nonresponse. Data available from completed waves of interviewing provide more detailed information about the nonresponding unit than is available for total nonrespondents. Thus, nonresponse compensation strategies may include weighting, imputation, or a combination of both. Kalton, Lepkowski and Lin (1985) discuss this issue and empirical findings in the context of the ISDP. This work made it clear that the choice between weighting and imputation for missing data of this type is far from obvious. Kalton (1986) and Kalton and Miller (1986) further refine the understanding of this problem and conclude that imputation can distort some forms of estimates and that weighting may be the preferred solution for large subclasses when the reduction in effective sample size is tolerable. They caution, however, that imputation may be better for estimates based on small subclasses when the loss of sample is important. In the case of a three interview longitudinal SIPP file the difference in sample size between weighting and imputation is not substantial, and consequently the weighting approach is the safer general purpose solution. Finally, Lepkowski (1988) after further empirical research concludes that a specific strategy for wave nonresponse can only be developed after consideration of such factors as the major survey design objectives, the panel design, and the distribution of wave nonresponse patterns. He provides criteria to be considered in developing missing data strategies and concludes that weighting strategies appear to be preferable for compensating for wave nonresponse.

Sampling Error Reduction through Estimation Techniques

Two methods for reducing sampling error through estimation techniques are under study: composite estimation and the use of administrative records in SIPP 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 the SIPP. Chakrabarty (1986) has conducted a preliminary review of the types of composite estimates appropriate for the SIPP data structure. The content of the survey has not been sufficiently stable during the first few years of the SIPP to seriously consider adoption of a composite estimator.

Another approach to variance reduction is through the use of administrative records for post-stratification. Currently, cross-section estimation procedures for SIPP make use of a second-stage adjustment to increase the precision of estimates by ratio adjusting collection month and reference month estimates to population estimates. However, the Census Bureau has access to some Internal Revenue Service and Social Security Administration files which can be used to produce detailed age, race, and sex distributions by adjusted gross income. The issue, which we have just begun to explore, is how these administrative data can be used for post-stratification to improve estimates of mean and median personal and household income as well as the estimates of the deciles of the personal and household income distribution. The basic question under study is the magnitude of the reduction in variances of these estimates achieved through such a procedure. Fay and Huggins (1988) will provide some indications.

Sampling for Special Subpopulations

Subgroups of the population are often cited as being more affected by governmental policy than others — the population of persons in poverty, the aged, the Blacks, Hispanics, and participants of Federal income security programs. 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 (Kasprzyk 1983; Bowie and Kasprzyk 1987).

A Census Bureau Working Group analyzed subsampling (screening) proposals for oversampling special populations. The issue studied concerned the reliability of estimates when different subsampling schemes are introduced. Subsampling characteristics based on income and demographic variables were identified and estimates of reliability for different subsampling rates and characteristics were calculated (U.S. Bureau of the Census 1985).

This group concluded that subsampling proposals, for a general purpose income survey like the SIPP, provided only modest gains in precision for low-income items and did not outweigh the disadvantages, which included an increase in the complexity of the operation, the loss of a self-weighting design, and large decreases in precision for the middle income items.

5. RESPONSE ERROR

Response error is one aspect of a more general problem, nonsampling error, discussed by Kalton, Kasprzyk and McMillen (1988). Response error occurs when incorrect data are recorded on the questionnaire. This can occur for a variety of reasons, such as a faulty questionnaire, memory errors, inappropriate respondents, etc. In this section we briefly describe a response error issue with the SIPP gross flow data and a record check study aimed at providing insight into a better understanding of response errors in general.

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). Analysis using the SIPP data (Burkhead and Coder 1985) covering month-to-month changes in receipt of income benefit amounts for a 12 month period focussed on changes occurring between the last month of one reference period and the first months of the succeeding reference period. The results using SIPP and ISDP data are similar, where an uneven pattern of change is observed and this pattern is clearly associated with the interviewing scheme. Gross changes are significantly higher between the last month of one reference period and the first month of the next. Hill (1987) used monthly data from the 1984 and 1985 waves of the Panel Study of Income Dynamics (PSID) to investigate the extent and determinants of excessive change between waves relative to measured change within waves of a panel survey. He found that in spite of different question sequences, and recall periods, between wave transitions dominate the within wave transitions in the PSID just as they do in the SIPP. The main causes for the problem are not known, but questionnaire wording/design, respondent recall error, and the interaction between these two factors seem likely.

Weidman (1986) did an empirical analysis to look for obvious relationships between respondent characteristics and changes in receipt status of a number of income types. He did not detect any relationship between gross change distributions, self/proxy status and nine demographic variables (age, race, sex, education, marital status, household size, tenure, relationship to reference person, and size of metropolitan area) for consecutive months, but did note that more transitions occur when some of the data are imputed. The absence of any notable relationships indicates a need for exploring other ways to understand this problem.

Interest in gross flow estimates remains high. Hubble and Judkins (1986) developed a model to estimate biases in gross flows estimates resulting from response errors, the parameters of which are estimated using SIPP response error rates and the ratios of within-wave and between-wave gross flow estimates. Several strong assumptions, as well as a reinterview program which produces accurate reinterview data on gross flows within the period, are necessary. Weidman (1987) presents linear models that try to represent the relationships between observed and actual transitions. The models are admitedly oversimplified using only survey reported data, but nevertheless, illustrate the need to obtain more information about the SIPP error structure in reporting receipt of benefits from government transfer programs.

SIPP Record Check Study

One way to study the SIPP error structure in reporting receipt of program benefits and amounts is to develop validation studies of items common to both survey records and administrative records. The SIPP program has initiated such a study to investigate response quality issues.

The goal is the improved understanding of the quality of the SIPP data and, ultimately, the development of quantitative estimates of response and nonresponse errors in order to adjust the survey data or modify survey procedures to obtain better quality data. The research questions addressed in this study include: (1) the quality of the respondent reports of receipt of program benefits for a variety of state and Federally administered transfer programs; (2) the quality of benefit dollar amount reporting for these programs; (3) demographic correlates of report quality; (4) extent of misclassification errors; (5) the effects of self-proxy respondent status on report quality; and (6) between wave recipiency turnover effects. Four state administered programs and six Federally administered programs are included in the study. Moore and Marquis (1987) provide very preliminary results, suggesting that reporting problems

are different for the Aid to Families with Dependent Children (AFDC) and the Food Stamp Programs, the former having a net under-reporting as well as a time placement problem for reporting a transition in program status while the latter having only a time placement problem.

6. CONCLUSION

As in all large-scale continuing survey programs, research is needed to improve understanding of the effects of survey methods on the data collected. A survey, like the SIPP, which is complex in its implementation requires a commitment to understanding the measurement process. The wide range of topics discussed above — collection, longitudinal concepts and estimation, and response error — illustrate where the interest and emphasis was placed during the development program and the first few years of the SIPP program.

ACKNOWLEDGEMENTS

The author gratefully acknowledges the helpful comments of the referees and the secretarial and clerical help of Ms. Hazel Beaton in the preparation of this manuscript.

REFERENCES

BOWIE, C., and KASPRZYK, D. (1987). A review of the use of administrative records in the Survey of Income and Program Participation. *SIPP Working Paper Series No. 8721*, Washington, D.C.: U.S. Bureau of the Census.

BURKHEAD, D., and CODER, J. (1985). Gross changes in income recipiency from the Survey of Income and Program Participation. *Proceedings of the Social Statistics Section, American Statistical Association*, Washington, D.C., 351-356.

CHAKRABARTY, R. (1986). Composite estimation for SIPP: A preliminary report. *SIPP Working Paper Series No. 8610*. Washington, D.C.: U.S. Bureau of the Census.

CHAPMAN, D., BAILEY, L., and KASPRZYK, D. (1986). Nonresponse adjustment procedures at the U.S. Bureau of the Census; *Survey Methodology*, 12, 161-180.

CITRO, C. (1985). Alternative definitions of longitudinal households in the Income Survey Development Program: Implications for annual statistics. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, Washington, D.C., 381-386.

CITRO, C., HERNANDEZ, D., and HERRIOT, R. (1986). Longitudinal household concepts in SIPP: Preliminary results. *Proceedings of the U.S. Bureau of the Census Second Annual Research Conference*, 598-619.

CITRO, C., HERNANDEZ, D., and MOORMAN, J. (1986). Longitudinal household concepts in SIPP. *Proceedings of the Social Statistics Section, American Statistical Association*, Washington, D.C., 361-366.

CODER, J. (1980). Some results from the 1979 Income Survey Development Program research panel. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, Washington, D.C., 540-545.

ERNST, L. (1988). Weighting issues for longitudinal household and family estimates. In *Panel Surveys* (Eds. D. Kasprzyk, G. Duncan, G. Kalton, and M.P. Singh), New York: John Wiley, forthcoming.

ERNST, L., HUBBLE, D., and JUDKINS, D. (1984). Longitudinal family and household estimation in SIPP. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, Washington, D.C., 682-687.

FAY, R.E., and HUGGINS, V.J. (1988). Use of administrative data in SIPP longitudinal estimation. Paper to be presented to the Section on Survey Research Methods, American Statistical Association, August 1988.

GBUR, P., and DURANT, S. (1987). Testing telephone interviewing in the Survey of Income and Program Participation and some early results. Paper presented at the International Symposium on Telephone Survey Methodology, 1987, Charlotte, North Carolina.

HEERINGA, S., and LEPKOWSKI, J. (1986). Longitudinal imputation for the SIPP. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, Washington, D.C., 206-210.

HILL, D. (1987). Response errors around the seam: Analysis of change in a panel with overlapping reference records. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, Washington, D. C., 210-215.

HUANG, H. (1984). Obtaining a cross-sectional estimate from a longitudinal survey: Experience of the ISDP. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, Washington, D.C., 670-675.

HUBBLE, D., and JUDKINS, D. (1986). Measuring the bias in gross flow estimates in the presence of auto-correlated response errors. *Proceedings of the Section on Survey Research Methods, American American Statistical Association*, Washington, D.C., 237-242.

JEAN, A., and McARTHUR, E. (1987.) Tracking Persons Over Time. *SIPP Working Paper Series No. 8701*, Washington, D.C.: U.S. Bureau of the Census.

JEAN, A., and McARTHUR, E. (1984). Some data collection issues for panel surveys with application to the Survey of Income and Program Participation. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, Washington, D.C., 745-750.

JUDKINS, D., HUBBLE, D., DORSCH, J.R., McMILLEN, D., and ERNST, L. (1984). Weighting of persons for SIPP longitudinal tabulations. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, Washington, D.C., 676-681.

KALTON, G. (1986). Handling wave nonresponse in panel surveys. *Journal of Official Statistics*, 2, 303-314.

KALTON, G. (1983). *Compensating for missing survey data*. Survey Research Center, University of Michigan.

KALTON, G., and KASPRZYK, D. (1986). The treatment of missing survey data. *Survey Methodology*, 12, 1-16.

KALTON, G., and KASPRZYK, D. (1982). Imputing for missing survey responses. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, Washington, D.C., 22-31.

KALTON, G., KASPRZYK, D., and McMILLEN, D.B. (1988). Nonsampling errors in panel surveys. In *Panel Surveys* (Eds. D. Kasprzyk, G. Duncan, G. Kalton, and M.P. Singh), New York: John Wiley, forthcoming.

KALTON, G., and LEPKOWSKI, J. (1985). Following rules in SIPP. *Journal of Economic and Social Measurement*, 13, 319-329.

KALTON, G., LEPKOWSKI, J., and LIN, T. (1985). Compensating for wave nonresponse in the 1979 ISDP research panel. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, Washington, D.C., 372-377.

KALTON, G., McMILLEN, D., and KASPRZYK, D. (1986). A review of nonsampling error issues in the Survey of Income and Program Participation. *Proceedings of the U.S. Bureau of the Census Second Annual Research Conference*, 147-165.

KALTON, G., and MILLER, M. (1986). Effects of adjustments for wave nonresponse on panel survey estimates. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, Washington, D.C., 194-199.

KASPRZYK, D. (1983). Social Security number reporting, the use of administrative records, and the multiple frame design in the Income Survey Development Program. In *Technical, Conceptual, and Administrative Lessons of the Income Survey Development Program* (ISDP), (Ed. M. David), Washington, D.C.: Social Science Research Council, 171-198.

KOBILARCIK, E., and SINGH, R. (1986). SIPP longitudinal estimation for persons' characteristics. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, Washington, D.C., 214-219.

LAMAS, E.J., and McNEIL, J.M. (1987). An analysis of the SIPP asset and liability feedback experiment. *Proceedings of the Social Statistics Section, American Statistical Association*, 194-199.

LEPKOWSKI, J. (1988). The treatment of wave nonresponse in panel surveys. In *Panel Surveys*, (Eds. D. Kasprzyk, G. Duncan, G. Kalton, and M.P. Singh). New York: John Wiley, forthcoming.

McMILLEN, D.B., and HERRIOT, R. (1985). Toward a longitudinal definition of households. *Journal of Economic and Social Measurement*, 13, 349-360.

MOORE, J.C., and KASPRZYK, D. (1984). Month-to-month recipiency turnover in the ISDP. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, Washington, D.C., 726-731.

MOORE, J.C., and MARQUIS, K. (1987). Using administrative data to evaluate the quality of survey estimates. Paper presented at the International Symposium on the Statistical Uses of Administrative Data, 1987, Ottawa, Canada.

NELSON, D., McMILLEN, D., and KASPRZYK, D. (1985). An overview of the Survey of Income and Program Participation: Update 1. *SIPP Working Paper Series No. 8401*. Washington, D.C.: U.S. Bureau of the Census.

OLSON, J. (1980). *Reports from the Site Research Test*, (Ed. J. Olson). Office of the Assistant Secretary for Planning and Evaluation, Department of Health and Human Services, United States.

ROMAN, A.M., and O'BRIEN, D.V. (1984). The student follow-up investigation of the 1979 ISDP. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, Washington, D.C., 732-737.

U.S. Bureau of the Census (1986). Additional results from the SIPP telephone test. Memorandum from J. Coder to A. Norton, April 9, 1986.

U.S. Bureau of the Census (1985). SIPP research on SIPP oversampling/subsampling. Memorandum from R. Singh to G. Shapiro and D. Kasprzyk, August 12, 1985.

WEIDMAN, L. (1987). Examination of relationships between actual and reported changes in the SIPP. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 216-220.

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*, Washington, D.C., 231-236.

WHITE, G., and HUANG, H. (1982). Mover follow-up costs for the Income Survey Development Program. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, Washington, D.C., 376-381.

YCAS, M., and LININGER, C. (1981). The Income Survey Development Program: Design features and initial findings. *Social Security Bulletin*. Baltimore, Md.: Social Security Administration, 44.