THE SURVEY OF INCOME AND
PROGRAM PARTICIPATION

NONSAMPLING ERROR ISSUES IN THE SIPP

No. 10

G. Kalton
University of Michigan
D. B. McMillen, and D. Kasprzyk
Bureau of the Census

U. S. Department of Commerce BUREAU OF THE CENSUS

Acknowledgement

This paper was prepared for presentation at the U.S. Bureau of the Census 2nd Annual Research Conference, March 23-26, 1986 and is also available in the Proceedings of the Bureau of the Census 2nd Annual Research Conference. Hazel Beaton provided expert clerical assistance in the preparation of the Working Paper.

Suggested Citation

Kalton, Graham, David B. McMillen, and Daniel Kasprzyk, "Nonsampling Error Issues in the Survey of Income and Program Participation," SIPP Working Paper Series No. 8602. Washington, D.C.: U.S. Bureau of the Census, 1986.

TABLE OF CONTENT

PREFACE

This paper reviews some of the design decisions affecting the quality of various forms of estimates produced from SIPP data. Many of the SIPP design alternatives are also applicable in other surveys, especially other panel surveys. Nonsampling errors which arise from the major SIPP decision to employ a panel design, namely panel nonresponse and panel conditioning, are discussed. Other design decisions, such as the length of reference period, respondent rules, following rules, and data collection mode, are reviewed in light of their possible effects on nonsampling errors. Finally, the measurement of gross change in a panel survey is discussed. Examples using SIPP data are provided to clarify the discussion.

I. INTRODUCTION

The Survey of Income and Program Participation (SIPP) is a national survey program introduced by the Bureau of the Census in 1983 to collect detailed information from survey respondents on sources of cash and noncash income and on participation in the labor force and in various goverment transfer programs. The SIPP is a series of overlapping panel surveys. A new sample of households is selected each year and all persons aged 15 and over in the selected households become panel members who are followed for 2 1/2 years even if they change addresses or move out of their sampled households. Individuals not originally in the sample who subsequently reside in the same households as panel members are also included in the survey for the time they are living with panel members. Panel members and associated persons are interviewed every 4 months about their income and program participation in the preceding 4 months (Nelson, McMillen, and Kasprzyk 1985).

As with any sample survey, the design of the SIPP involves a multitude of 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 have to 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. The objective of survey design is to achieve an allocation of resources that minimizes the total survey error for a given budget; this is the concept of total survey design (Horvitz 1978).

As an illustration, consider the decision to use a panel design with 4-month reference periods for the SIPP. The main reason for choosing the panel design over a single-round survey that collects the data by retrospective questioning for a longer reference period was to reduce the reporting errors in income data that occur with the latter method. However, the sample size for the panel survey has to be much smaller than that for the single-round survey, so that its sampling errors are larger. The choice of a 4-month reference period for the SIPP is itself a compromise. In the surveys conducted during the Income Survey Development Program (ISDP--the development program for the SIPP), 3-month reference periods were used. They were rejected for the SIPP, however, because of the higher cost involved in making the more frequent contacts with the respondents. Another consideration involved in the choice of a 4-month panel design is that, while the design aims to reduce reporting errors, it may give rise to greater nonresponse biases and to panel-conditioning effects.

In assessing the significance of various sources of error, their effects on the survey estimates need to be considered. In practice, of course, surveys are multipurpose and give rise to a wide variety of forms for estimates. Each of these forms can be affected differently by different types of error. Thus, for instance, in a panel survey, random measurement errors will lower the precision of the measurement of net change but will not create a bias; on the other hand, random measurement errors will lead to an overstatement of the extent of gross change. The measurement of gross change is potentially a major benefit of a panel design: yet, in practice, this benefit is hard to achieve because of the problems created by nonsampling errors (see, for instance, U.S. Bureau of the Census and Bureau of Labor Statistics 1985).

As Deming (1944) catalogs, nonsampling errors can arise from a wide variety of sources. These sources may be divided, for convenience, into the two broad classes of nonresponse and noncoverage errors on on the one hand and measurement errors on the other (Hansen and Waksberg 1970). Nonresponse errors occur when information is not collected for some sampled units; noncoverage errors occur when some units have no chance of being selected for the sample. Measurement errors occur when incorrect data for some sample units are used in the survey analysis. Measurement errors may be further subdivided into response and processing errors; the former occur when incorrect data are recorded on the survey questionnaires and the latter when errors are introduced in converting the survey responses into machine-readable form for the analysis. Response errors may arise because of a faulty questionnaire, the collection of data from inappropriate informants, memory errors, deliberate distortion of responses (e.g., prestige bias), interviewer effects, and misrecording of responses. Processing errors can arise from incorrect editing, coding, and data transfer. It may be noted that, while the division into the broad classes of nonresponse and noncoverage errors vs. measurement errors provides a useful framework, the division is not a rigid one. When item nonresponse is handled by imputation, for instance, the imputation of an incorrect value can be viewed as a measurement error rather than a nonresponse error.

The purpose of this paper is to review some of the design decisions affecting the quality of various forms of estimates produced from SIPP data. Many of the SIPP design alternatives are also applicable in other surveys, especially other panel surveys; in some cases research has been conducted on their relative advantages and disadvantages. Section II of the paper reviews the

nonsampling errors that can arise from the major SIPP decision to employ a panel design, and section III briefly reviews nonsampling errors arising from some of the other SIPP design decisions. Section IV discusses the measurement of gross change in a panel survey. A few concluding remarks are given in section V.

II. NONSAMPLING ERRORS IN A PANEL DESIGN

As noted earlier, a panel design with a short reference period was adopted for the SIPP in order to improve the measurement of annual income by reducing reporting errors that occur with long-term retrospective questioning on income data. The March Income Supplement of the Current Population Survey collects annual income data which are known to have a number of deficiencies. For example, the long recall period (15 months) causes difficulties for the recall of minor flows of income and the timing of the occurrence of changes in receipt of income and changes in household composition, and distorts the relationship between household composition and income (Holden, Burkhouser, and Myers 1985). Other deficiencies can be observed in the high item nonresponse rates (U.S. Bureau of the Census 1985a) and the misclassification of program data. By collecting data at 4-month intervals, the SIPP tries to avoid some of these deficiencies.

Although many factors in addition to the survey design, such as training and questionnaire design, are involved in obtaining improved survey measures, indications of lower item nonresponse through a survey design which has repeated interviews at short intervals with the same persons can be seen in the ISDP where proportionally much less aggregate income for each income source results from imputation (Vaughan, Whiteman, and Lininger 1984). A

Table 1. -- Item Nonresponse Rates for the 1984 SIPP Panel and March 1985 CPS, for Selected Income Types

Income Type	SIPP 1st Quarter 1984 Monthly Average 1/	SIPP 2nd Quarter 1984 Monthly Average 2/	SIPP 3rd Quarter 1984 Monthly Average 3/	SIPP 4th Quarter 1984 Monthly Average 4/	CPS March 1985 4/
Wage and Salary	7.2	7.5	7.5	7.6	18.9
Self-Employment Income	16.8	16.2	16.0	16.1	26.5
Supplemental Security Income (Federal)	7.6	8.4	8.1	8.4	19.9
Social Security	10.8	11.6	11.7	12.3	21.9
Aid to Families with Dependent Children	6.1	6.9	6.5	5.5	16.0
Unemployment Compensation	10.1	13.6	10.4	12.7	21.8
Company or Union Pension	13.9	14.0	12.8	14.7	24.0
Food Stamp Allotment	5.2	6.3	6.7	6.6	13.7
Veterans' Compensation or Pension	11.3	11.2	11.9	13.5	18.3

1. U.S. Bureau of the Census (1985b)

2. U.S. Bureau of the Census (1985c)

3. U.S. Bureau of the Census (1985d)

4. U.S. Bureau of the Census (1986a)

similar experience may also be observed in table 1 where the SIPP monthly average item nonresponse rates for amounts of selected income types for the four calendar quarters in 1984 are compared to 1985 March Current Population Survey item nonresponse rates.

Panel surveys, however, give rise to their own potential sources of nonsampling errors. The two main concerns are panel nonresponse and panel conditioning. Panel nonresponse is a concern because, as a rule, the nonresponse rates in a panel survey increase as the panel ages. Panel conditioning occurs when panel membership has an effect on panel members that affects their survey responses.

Panel Nonresponse

Panel nonresponse can be separated into attrition and nonattrition nonresponse. With attrition nonresponse, a panel member drops out of the panel on one wave and remains out of the panel for all subsequent waves. With nonattrition nonresponse, a panel member fails to provide data on one or more waves but does return to the panel for at least one subsequent wave. When the interval between panel waves is long, such as a year or more, it is unlikely that a panel member lost on one wave will be recaptured at a later one; hence, attrition nonresponse is the dominant nonresponse pattern for this type of study. Indeed, in many such studies, once a panel member is lost on one wave, no attempt is made to contact the member on later waves.

When the interval between panel waves is shorter it is more common to try to recapture panel members lost on one wave; this is the practice for the SIPP. By this means, approximately 17 percent of SIPP nonrespondent households on one wave are converted to respondent households on the subsequent wave.

Table 2 provides the interviewing patterns observed during the first five waves of the SIPP for persons who were 15 years or over at the time of the first interview and who were interviewed (either through self or proxy interviews) during the first wave (U.S. Bureau of the Census 1986b). As can be seen from the table, most of the nonresponse is attrition nonresponse; in the first five waves, around 3 out of 4 nonrespondents were attrition nonrespondents.

The extent of panel nonresponse depends on a variety of factors, especially the efforts made to retain panel members and the burden on the respondents. The tracing of mobile panel members can present a particular challenge and can be time-consuming and very costly. The evidence from long-term panels that contact respondents infrequently (say, every year or so) shows that the level of panel loss from wave to wave can be kept surprisingly low provided that vigorous efforts are made to retain panel members. Thus, for instance, the Panel Study of Income Dynamics (PSID) which has interviewed respondents annually since 1968 has managed to keep around 97 to 98 percent of sample members from one wave to the next from the third wave onwards (Institute for Social Research 1985).

Nevertheless, panel losses accumulate over time, so that at later waves the overall nonresponse rate (i.e., the proportion of original sample members still in the survey population from whom data are not collected at a particular wave) may be large.

Panels with shorter intervals between waves usually suffer greater nonresponse losses from one wave to the next. As the interval between waves becomes shorter, the tracing problems become less severe, but the respondent burden

within a given time period is increased and less time is available to devote efforts to retaining panel members. With SIPP the burden on both the interviewer and respondent is large and the time between interviews is short, but, nevertheless, the average additional sample loss is only 2.7 percent of households from one wave to the next.

Table 2. -- Interview Patterns of SIPP Original Sample Persons for the First Five Interviews of the 1984 SIPP Panel 1/

		Percent
1. Response every interview (5 interviews)		
Pattern:	XXXXX	79.1
2. Apparent attrition cases		13.8
Patterns:	XXXXO	3.8
	XXXOO	3.1
	XXOOO	3.2
	XOOOO	3.7
3. First and fifth interviews conducted, but one intervening interview missing		3.4
Patterns:	XXXOX	1.6
	XOXXX	0.6
	XXOXX	1.2
4. First and fifth interviews conducted, two or more intervening interviews missing		0.7
Patterns:	XOOOX	0.1
	XOXOX	0.1
	XXOOX	0.3
	XOOXX	0.2
5. Fifth interview missing and one or more intervening interviews missing		0.7
Patterns:	XOXXO	0.1
	XOXOO	0.2
	XOOXO	0.1
	XXOXO	0.3
6. Left the universe (deceased, institutionalized, living in armed forces barracks, moved overseas)		2.3
	Total:	100.0
		(25,128)

1/ The universe for the table consists of all persons in rotation groups 1, 2, and 3 who were 15 years or over at the time of the first interview and for whom a personal interview was conducted (either self or proxy interviews) during the first wave of the 1984 SIPP Panel, and who were designated for interview for all five interviews. The symbol "X" represents a successful interview and the symbol "O" represents no interview (either no household interview or no personal interview).

Table 3 displays the sample loss for the 1984 SIPP Panel (the sample loss rate consists of cumulative noninterview rates adjusted for growth in the number of households as the panel ages). As can be seen from the table, the largest wave-to-wave sample losses occur during the first three interviews.

Table 3. -- Sample Loss: 1984 SIPP Panel

Wave	Sample Loss
1	4.9%
2	9.4%
3	12.3%
4	15.4%
5	17.4%
6	19.4%
7	21.0%

The prime concern about panel nonresponse is that nonrespondents may differ in systematic ways from respondents. If this is so, then a straightforward analysis of the respondents' data will produce biased estimates for at least some of the parameters being estimated. An investigation of possible biases resulting from the panel nonresponse in the PSID, in fact, found little evidence of bias across a wide range of estimates (Becketti et al. 1983). Also, a study of panel nonresponse in the ISDP 1979 Research Panel found that first wave responses were only slightly associated with second wave response status (Kalton, Lepkowski, and Lin 1985). Other studies, however, have found some differences in the characteristics of those who dropped out of a panel and those who continued to participate (e.g., Sobol 1959). The extent of nonresponse bias is likely to depend both on features of the panel design and on the nature of the subject matter of the survey. There must always be a risk that panel losses will lead to biases in survey estimates.

Table 4 presents distributions of characteristics of initially interviewed persons of the 1984 SIPP Panel by their status at the fifth interview. The table compares distributions of "stayers"--persons interviewed, either self or proxy, in each of the five waves--with "leavers"--persons who were nonrespondents in at least the fifth interview. The latter category excludes persons who left the survey universe; that is, those who died, were institutionalized, or moved overseas. The picture which emerges from the results in table 4 is that there appears to be some relationship between response status at the fifth wave and such characteristics as age, race, sex, marital status, education, whether living quarters are being bought or rented, and ownership of a savings account.

The assessment of nonresponse bias, however, should not be made in terms of the simple sample estimates based on the respondents alone, but rather in terms of the estimates obtained after nonresponse adjustments have been incorporated. Weighting adjustments are commonly employed in single-round surveys in an attempt to compensate for total nonresponses; similar adjustments may be made for those who fail to take part in any wave of a panel survey. The information available on which to base these weighting adjustments is generally very limited, so that they have to be made using what is available rather than what is most appropriate. In the case of panel wave nonresponse, however, the situation is different: for those who are nonrespondents on some but not all waves, a great deal of information is available from their responses for the waves in which they cooperated. This wealth of information holds the promise of enabling effective adjustments to be made to reduce the wave nonresponse bias in the survey estimates. In practice, however, it is difficult to take full advantage of all the available auxiliary information in making general multipurpose adjustments.

Table 4. -- Distributions of Characteristics of Initially Interviewed Persons by Status at Fifth Interview: 1984 SIPP Panel 1/

	Stayers (5 interviews)	Leavers (missing at least the 5th interview)
Total Number	19878	3655
LIVING QUARTERS:		
Owned/Being bought	71.6	63.5
Rented for cash	25.9	34.9
Occ'd w/o cash pmt	2.4	1.6
INTERVIEW LENGTH: (Minutes)		
Less than 15	26.8	29.3
15 to 29	44.1	43.7
30 to 44	21.0	20.1
45 to 59	6.1	5.3
60 or more	2.0	1.6
AGE:		
15-24	21.0	25.6
25-44	37.9	39.0
45-64	26.0	24.6
65 and over	15.2	10.8
SEX:		
Male	46.1	49.0
Female	53.9	51.0
RACE:		
White	87.6	84.2
Black	9.8	12.1
Am.Ind/Esk/AlNativ	0.4	0.8
Asian/Pac.Isl.	2.2	2.8
ETHNICITY:		
Spanish Origin	5.3	6.6
Not Spanish Origin	94.7	93.4
MARITAL STATUS:		
Mar'd, spouse present	59.9	53.1
Mar'd, spouse absent	0.5	0.8
Widowed	7.4	5.6
Divorced	6.4	7.3
Separated	2.0	3.2
Never Married	23.7	30.1
HIGHEST GRADE ATTENDED:		
Less than 9 years	11.4	8.5
9-11 years	16.5	18.5
12 years	35.9	36.6
More than 12 years	36.2	36.4
HOURS WORKED/WEEK:		
Not applicable	36.6	35.4
1 to 19	5.9	5.0
20 to 34	8.7	10.0
35 to 40	33.6	43.2
41 or more	15.2	15.3
PERSON MONTHLY INCOME:		
Less than 300	29.9	32.7
300 to 599	15.8	14.7
600 to 899	12.3	13.0
900 to 1199	9.8	10.5
1200 to 1599	10.4	10.2
1600 to 1999	7.0	5.3
2000 to 2999	9.0	8.1
3000 to 3999	3.2	3.0
4000 or more	2.7	2.7
SAVINGS ACCOUNT:		
Yes	58.1	50.2
No	41.9	49.8

1/ This table is restricted to interviewed persons in rotation groups 1, 2, and 3 who were 15 years old and over at the time of the first interview. Because of budget considerations, the SIPP sample was reduced by approximately 15 percent during the fifth wave of the 1984 Panel. As a result, some individuals were not eligible to be interviewed five times and consequently are not included in this table (U.S. Bureau of the Census 1986b).

There are two alternative ways of attempting to compensate for wave nonresponse-either by weighting adjustments or by imputation. With weighting adjustments, a number of different sets of weights are used--with the choice of the appropriate set of weights for a particular analysis depending on the waves from which the data for that analysis are obtained. It is possible to reduce the number of different sets of weights needed, but this is achieved at the cost of discarding some of the data (e.g., by discarding some waves of data for nonattrition cases to convert them into attrition cases). The alternative strategy of imputation can take advantage of high correlations of responses across waves and can produce a simple data set for analysis. However, while imputation may serve well for some estimates, it can lead to distortions in others. The choice between weighting adjustments and imputation for handling wave nonresponse is not clear cut. It may be that some combination of the two ways may be best, for instance using imputations for some patterns of response/nonresponse across waves and weighting adjustments for others (Kalton 1985). Whichever method is used, it is difficult to make effective use of all the data available for the wave nonrespondents in making the nonresponse adjustments.

For the same overall level, panel nonresponse is a less serious problem than total nonresponse because of the availability of data from other waves that can be used either for making nonresponse adjustments or for conducting checks on potential biases. Nevertheless, panel nonresponse is a major cause of concern, especially since its level increases with successive waves. When imputation is used to compensate for wave nonresponse (and also item nonresponse in a panel survey), the analyst needs to evaluate carefully the possible effects of the imputations on the analyses being conducted. In

particular, if, for a specific variable being imputed, the imputation scheme fails to employ the same variable measured on other waves as auxiliary data, then the amount of gross change in that variable across waves will be overstated (Kalton and Lepkowski 1983).

Panel Conditioning

The second concern in panel surveys is panel conditioning. Conditioning refers to a change in response that occurs because the respondent has had one or more prior interviews. The effect may come about because the prior interviews change respondents' behaviors or because they change the way respondents answer questions. The prior interviews may, for example, provide respondents with information or cause them to reflect on an issue, thus changing their subsequent behavior or attitude. Alternatively, they may simply cause a change in the quality of respondents' reports. On the one hand, respondents may become better reporters because the prior interviews have taught them what is expected of them or have sensitized them to the events to be reported so that their recall is improved; on the other hand, they may become worse reporters because they become unwilling to repeatedly make the effort required to report accurately.

Panel conditioning is thus the reactive effect of prior interviews on current responses. In practice, it is difficult to separate the effects of conditioning from those of other changes between waves, especially attrition. In experimental studies, Mooney (1962) found that older persons reported higher levels of illness on the first time in the panel than on subsequent times, and Neter and Waksberg (1964, 1965) found that resident owners reported fewer jobs and less expenditure on housing repairs and alterations on the second

than on the third time in the panel. In a survey of British social attitudes, Lievesley and Waterton (1985) found that panel members gave fewer "don't knows" and more socially undesirable answers on the second round of the survey (one year after the first round) than did a fresh cross-sectional sample. Moreover, they provide evidence to show that not all of the difference between the panel and cross-sectional results was due to panel attrition. In addition, there are results from two validation studies that compared panel respondents reports with record data. In one, Traugott and Katosh (1979) found that longer term members of a panel survey of election behavior gave more accurate responses on voting behavior than the newer members, and, moreover, had higher levels of voting. These effects may be due to panel conditioning or, perhaps, to attrition. In the other validation study, Ferber (1964) found that the quality of reporting on savings holdings improved with length of time in the panel; this improvement was partly due to panel attrition, with the poorer reporters dropping out of the panel, and partly to an increase in accuracy of reporting for those that remained.

In theory, panel conditioning may be examined by means of a rotating panel design in which fresh replicate samples are added to the panel at each wave. Provided that all other survey conditions are the same for all the rotation groups at a particular wave, the comparison of the results for the various rotation groups for that wave provides a means of examining conditioning effects. In practice, however, it is virtually impossible to keep all other survey conditions constant for the different rotation groups. Sometimes the questionnaire is somewhat different for the various rotation groups (e.g., the incoming rotation group may be given a longer questionnaire to collect basic sociodemographic data that need not be collected again at

reinterviews), the mode of data collection differs (e.g., all members of the incoming rotation group may be interviewed face-to-face, whereas many of the members of other rotation groups may be interviewed by telephone), or the type of informant differs (there may, for example, be more female household informants for the incoming rotation group than for the others). Even if factors such as these are controlled, there remains the problem that the rotation group samples are not exactly comparable because of the effects of panel nonresponse. It is consequently difficult to disentangle the effects of conditioning from the effects of other changes.

A number of studies have been conducted to compare the results obtained by different rotation groups, the variation between the results being termed the rotation group bias. The existence of rotation group bias has been established for various labor force characteristics in the Current Population Survey and several possible sources of the bias have been suggested (Bailar 1975, 1979; U.S. Bureau of the Census 1978; McCarthy 1980). For instance, the estimate of the percentage unemployed from the households in the incoming rotation group is 10 percent larger than the average for all eight rotation groups. Rotation group bias has also been found in the Canadian Labour Force Survey (Ghangurde 1982). An analysis of rotation group bias in the National Crime Survey found that victimization rates declined with length of time in panel, with the largest drop occurring between the second and third times in the panel (Woltman and Bushery 1975).

This range of studies indicates that rotation group bias is a pervasive effect in rotating panel surveys; it is plausible that often some of this effect is caused by conditioning. To date, no research on panel bias has

been conducted on the SIPP, but it is entirely possible, indeed likely, that some conditioning effects occur. It is, for instance, conceivable that repeated interviews about welfare programs may draw respondents' attention to the exis tence of the programs and stimulate applications to them. The overlapping panels in SIPP are similar to rotation groups in other panel surveys; hence it is possible to examine panel bias by comparing the estimates from different panels for the same period. In the early years of SIPP, however, the study of panel bias by this approach is problematic because of the confounding effects of changes in the sampling frame, the questionnaires, and the field procedures. When the program matures and the content and procedures are stabilized, comparisons of equivalent results from the overlapping SIPP panels will provide a means to examine the existence and nature of any panel bias.

III. THE EFFECTS OF OTHER SIPP DESIGN DECISIONS

This section briefly reviews a number of other major SIPP design decisions and discusses their possible effects on nonsampling errors. The following four design decisions are discussed in turn: length of reference period, respondent rules, following rules, and mode of data collection.

Length of Reference Period

Three factors are commonly considered in choice of length of reference period: recall loss, telescoping effects, and respondent burden. Studies such as those by Cannell, Fisher, and Bakker (1965) on hospitalizations and Turner (1972) on victimizations have shown that events that happened longer ago are less likely to be reported in survey interviews. Events that are less salient to the respondents appear to be the ones that are more likely to be missed. Therefore, as the reference period becomes longer, the more likely it is that some events will fail to be reported.

Telescoping refers to the frequently encountered phenomenon that respondents tend to draw into the reference period events that occurred before (or after) the period. The extent of telescoping can be examined in a panel survey by determining those events reported on a given wave that had already been reported on a previous wave. Using this approach, Neter and Waksberg (1964) found that with a onemonth reference period the reporting of the number of jobs of $100 or more for house alterations or repairs was about 55 percent higher with unbounded recall (i.e., before excluding those reported on a previous wave) than with bounded recall (i.e., after excluding those reported on the previous wave). With a 3-month reference period, the extent of telescoping was reduced to a 26 percent higher rate of reporting for the unbounded recall. Telescoping effects can be avoided after the first wave in a panel survey by using bounded recall, as is done in the National Crime Survey (Garofalo and Hindelang 1977).

The longer the reference period, the greater the effort the respondent may need to recall the information being collected. Apart from this factor, however, the total respondent burden in reporting specified events in a panel survey of given duration is independent of the length of the reference period. Thus, for instance, collecting twelve monthly incomes in two waves, with six monthly incomes collected at each wave, places the same total burden on the respondent as collecting the twelve monthly incomes in four waves, with three monthly incomes collected at each wave. However, the respondent burden at a particular wave does depend directly on the length of the reference period. In their examination of a respondent load effect, Neter and Waksberg (1965) concluded that it affected the number of jobs for house alterations and repairs under $10 reported; at least 20 percent fewer such jobs were

reported for the preceding month when the reference period was lengthened from 1 to 6 months, and at least 11 percent fewer such jobs were reported for the preceding 3 months when the period was lengthened from 3 to 6 months. They detected no respondent load effect for these small jobs when the reference period was lengthened from 1 to 3 months, and they suggest that the effect is also likely to decline with size of job.

As mentioned earlier, from the earliest days of the ISDP, a short (subannual) reference period for reporting difficult-to-remember information such as detailed labor force participation and monthly Federal program participation was assumed to be desirable. The reference period issue was addressed during the ISDP in a pilot test conducted in five cities in which an experiment that varied the reference period was conducted; the experiment compared a single interview using a 6-month recall period, with two consecutive interviews, both using 3-month reference periods. A summary of the analysis conducted in the pilot test is provided by Olson (1980). The proportion of respondents reporting some positive amount of income in the initial 3 month reference period was higher for the 3 month reference period group than for the 6 month reference period group; that is, a 6 month recall period may understate the proportion of income reported in the earlier part of the period. This pattern appeared to hold for a number of specific income sources such as wages, Aid to Families with Dependent Children, general assistance, and unemployment compensation. The findings, though not definitive, support the presumption that longer recall periods increase chances of omission due to memory loss. Other analyses also observed that the number of person-sources of income reported per household in the first 3 months of the 6 month reference period was lower than for the corresponding time using a 3 month reference period. The results

of this experiment, together with additional ISDP experience using a short reference period, have led to a 4-month recall period for SIPP. This decision maintains cost at an appropriate budget level while attempting to ensure satisfactory data quality.

Respondent Rules

A SIPP panel member is broadly defined to be anyone living in a household sampled for the first wave of the panel who is an adult (aged 15 and over), or who will become an adult during the life of the panel. For the first wave of the panel, the sample comprises all adult panel members in the selected households. For a subsequent wave, it comprises all adult panel members at the time of the interview together with all adults living in the households of adult panel members. The sample at any wave thus aims to collect data on all adults in the specified households.

There are many ways of designating informants in a household survey which collects data on all members of the sampled households. The "respondent rules" vary in the extent to which sample members are required to be self-reporters, and in the range of household members who are acceptable as proxy informants. For example, every adult in the household may be required to respond personally, one adult in the household may be the household informant providing data for all members of the household, or all adults present at the time of interview may be required to respond personally and a related adult may be interviewed for an absent adult. See Roshwalb (1982) for a review of this topic.

The main factors involved in the choice of appropriate respondent rules for a particular survey are the quality of the resultant data, the costs of data

collection, and the level of nonresponse. The use of the more restrictive respondent rules is based on the premise that self-reporters will provide higher quality responses than proxy informants, and that, if proxy informants are used, informants related to the sample members will generally provide higher quality responses than unrelated informants. While in general these premises seem reasonable, the amount of differences in data quality from using different respondent rules is likely to depend heavily on the data being collected; in some cases, proxy informants may be able to give almost as accurate responses as self-reporters, whereas in other cases their responses may be much less accurate. Sometimes, it may even happen that proxy informants provide more accurate responses; proxy informants may, for instance, give more honest answers on some sensitive matters than self-reporters. There is also another way in which the respondent rules may affect the quality of the data, namely through the burden on the informant. In particular, when one household informant provides data for all household members, the heavy reporting load may lead to lower quality responses.

The disadvantage of the more restrictive respondent rules is the greater difficulty they cause in data collection. They require more callbacks to collect the data--hence, the data collection costs are increased--and, in addition, they may lead to a higher nonresponse rate. This disadvantage has to be balanced against the potential advantage of better quality data from the application of the more restrictive rules. The appropriate choice of rule for a particular survey, thus, must depend on the survey's subject matter.

The respondent rules adopted for the SIPP are that adults present at the time of interview report for themselves while proxy informants are accepted for

absent adults. A hierarchy of proxy informants has been established for SIPP so that a spouse is always the first choice as a proxy; the second level proxy is the adult who was the proxy informant at the previous interview; the third level proxy is an individual who was proxy at any other interview; and, finally, a first time proxy is accepted. In the event that a knowledgeable proxy is not available, a personal visit is scheduled to interview the uninterviewed persons. Table 5 provides the distribution of self- and proxyreports across the first three waves of the 1984 Panel. As can be seen from the table, about two-thirds of respondents are self-respondents on each wave, while less than half are self-respondents for all three interviews.

Table 5. -- Distribution of Self (S) and Proxy (P) Reports for Panel Members Responding to the First 3 Waves of the 1984 SIPP Panel: Rotation Groups 1, 2, and 3

Waves			Percent
1	2	3	
S	S	S	49.4
S	S	P	5.9
S	P	S	6.0
P	S	S	5.8
S	P	P	6.2
P	S	P	3.5
P	P	S	5.1
P	P	P	18.1
Total			100.0 (27,002)

No experimental studies have been conducted on the SIPP respondent rules. There is some evidence that higher levels of item nonresponse occur in data collected from proxy informants than from self-reporters in the 1978 ISDP Research Panel (Kalton, Kasprzyk, and Santos 1981) and in the 1984 SIPP Panel (Coder and Feldman 1984), but this does not necessarily mean that self-reporting

is preferable. It may simply be due to the differences between sample members who are at home when the interview is conducted and those who are not. Mathiowetz and Groves (1985) provide an example of such a selection bias in selfreporters in a health survey. When self-reporters were taken to be those who answered the telephone call for the interview, they appeared to report more or about the same number of health events for themselves as they reported for others. However, when the self-reporters were selected at random within the household, they reported fewer health events for themselves than for others.

Although there have been no experiments in the SIPP on respondent rules, an experiment on this topic was conducted during the 1979 ISDP Research Panel. In this experiment, one sample of households was interviewed using rules that require self-reports except in special circumstances (such as, when the respondent was physically or mentally incapable, unable to speak English, or away from the household during the entire interview period), whereas another sample was interviewed using rules that allowed proxy reports for household members absent at the time of interview (i.e., using the rules adopted for the SIPP). An average of 85.7 percent of adults were self-respondents in the self-response rule households and an average of 65.0 percent were self-respondents in the usual or proxy response rule households; thus, in the 1979 ISDP Research Panel, the implementation of the self-response rule resulted in approximately 20 percent more self-interviews than the other treatment. The evidence from the experiment is, unfortunately, inconclusive. The self-response sample had a higher total nonresponse rate and was 4 to 6 percent more expensive than the usual response rule sample, but it generated somewhat higher quality data, in that respondents made greater use of records in answering questions,

there were fewer item nonresponses for key items, less rounding of answers, and less variance in nonzero amounts (Kulka 1984).

Following Rules

A panel survey like the SIPP aims to provide data for a wide variety of analytic objectives. Some analyses are cross-sectional, producing estimates from individual waves of data, whereas others are longitudinal, focusing on temporal characteristics of individuals or other analytic units. Since populations are not static over time, these analyses often have different populations of inference. An important consideration in the design of a panel survey is then to ensure that the survey collects the data needed to provide valid estimates for the various cross-sectional and longitudinal populations of analytic interest. The sample design for the first wave of a panel survey is chosen to provide coverage of the defined survey population at the time the sample is selected. After that, "following rules" are needed to determine how the samples for subsequent waves are to be generated. These following rules specify which units of the sample remain and which drop out from one wave to the next, and, if necessary, how additional sample members are to be added to the panel, in order that the panel will provide valid samples of the desired populations of inference.

The SIPP is designed to cover the U.S. civilian noninstitutionalized population aged 15 and over at each wave of the survey. The initial sample of households gives a sample of this population at the first wave. The following rules then need to update this sample for the subsequent waves. The ideal following rules for the SIPP would specify that all members of the original sample who remain in the U.S. civilian noninstitutionalized population should be followed, and that supplementary samples should be added at each wave to represent the

new entrants to that population. In practice these rules have to be somewhat restricted because of the difficulties and high costs of implementing them. We will comment briefly on some of these restrictions; Kalton and Lepkowski (1985) provide a more detailed discussion of the SIPP following rules.

As noted, the SIPP rules aim to follow original sample persons throughout the life of the panel or, if they leave the U.S. civilian noninstitutionalized population (through death, emigration, joining the Armed Forces and living in barracks, or becoming an inmate of an institution), until the time they leave. For practical reasons two restrictions have been imposed on these rules. First, for processing reasons, persons in the 1984 SIPP Panel who were in the 4.8 percent of nonresponding households at the first wave were not followed in subsequent waves. Second, sample persons who moved to an address that was more than 100 miles from a SIPP primary sampling area within the United States were not followed for face-to-face interviews (but interviewers were instructed to conduct telephone interviews with them where possible). Since about 96.7 percent of the U.S. population lives within 100 miles of the SIPP primary sampling areas, this restriction is not a severe one. In fact, of the original sample persons who were interviewed in the first three SIPP rotation groups, only 30 persons can be identified as having moved beyond the SIPP geographic limits by the conclusion of the fifth interview. This is an underestimate since by this same time 564 persons moved to unknown addresses, and some of these persons are likely to have moved outside the SIPP geographic limits.

In constructing following rules for panel surveys, it is necessary to consider not only the theoretical issues related to their coverage of the populations

of inference, but also the practical feasibility of carrying them out. The SIPP rules specify that panel members are to be followed even when they move, and those who move can be difficult to trace. Jean and McArthur (1984) report that only 80 percent of movers between the first and second waves of the 1984 SIPP Panel were traced. Since then, attempts have been made to improve the success rate in tracing by asking sample persons to return a postcard with their new address if they move, but this has not proved effective. In fact, the postcard approach has been used so infrequently that its use has been discontinued.

Another issue relating to movers concerns those who temporarily move out of the U.S. civilian noninstitutionalized population, and return during the life of the panel. The particular group of concern contains those who move into institutions, such as nursing homes, for a period of time. If such persons return to a household containing another panel member, they are recaptured for the panel. Using this procedure 3 percent of persons entering institutions returned to the panel. If institutionalized persons do not return to a panel household, however, they are lost. An attempted remedy for this loss is to have interviewers check on whether a sample person is still institutionalized; if not, he or she is followed as any other mover. This rule was not used in the 1984 Panel but has been implemented in the 1985 and 1986 SIPP Panels.

One part of the following rules deals with original panel members. The other part deals with the supplementary samples that may be needed to represent new entrants to the population. There are no effective economical ways of updating a panel sample for most types of new entrants to the population and, therefore,

the deficiencies in coverage have largely to be accepted. There are, however, remedies that can be used in certain cases.

The one type of entrant that can in general be readily handled comprises the "births"--in this case, persons who reach the age of 15 during the life of the panel. Children under 15 in original sampled households are followed in a SIPP panel provided that they stay with one or more panel members. Those who reach 15 during the life of the panel become full panel members to be interviewed and followed for the rest of the panel. This procedure thus captures all the births from the original sample, except for the few that leave panel members before reaching their 15th birthday. By the fourth interview, approximately 49 children who were less than 15 years old were not followed. Other entrants to the population are persons coming from abroad, persons returning from institutions, and persons returning from living in Armed Forces barracks. It would be a major undertaking to fully represent these entrants in the sample since no straightforward sampling frames are available for them. The SIPP does not attempt to incorporate these entrants into a panel.

Decisions concerning following rules must be balanced by analytic and cost concerns. During the ISDP the issue of costs was addressed by a study conducted as part of the 1979 Research Panel. The purpose of the study was to shed some light on the data collection costs resulting from following movers to their new addresses. The study showed that, as of the final interview of the 1979 ISDP Panel (a period of 15 months), movers represented about 22 percent of the total sample and that over the entire Panel there was approximately a 7-percent increase in the number of hours and an 11.4-percent increase in

the number of miles charged by interviewers due to the following movers rules adopted for that Panel (White and Huang 1982).

Mode of Data Collection

The main modes of survey data collection are face-to-face interviews, telephone interviews, and self-completion questionnaires. For a particular survey, the choice between modes involves an assessment for each mode of such factors as the feasibility of collecting the data required, the quality of the data that would be collected, the response rate that would be achieved, the availability of an adequate sampling frame, and the costs of data collection.

For a single round survey of the general population, the likelihood of a very low response rate and the lack of a sampling frame nearly always rule out reliance on self-completion questionnaires. The choice between face-to-face and telephone interviewing is more open to debate. Face-to-face interviewing usually produces a higher response rate, and the noncoverage of the area sampling frame generally employed with face-to-face interviewing is probably less biased than the noncoverage resulting from households without telephones when telephone interviewing is used. On the other hand, telephone interviewing is less expensive than face-to-face interviewing. As a rule, face-to-face interviewing tends to be preferred when a long and complex questionnaire is employed.

A number of studies have been conducted to compare the quality of responses to face-to-face and telephone interviews. In practice it is difficult to disentangle the effect of mode on quality from its effects on other features of the survey, including nonresponse, different questionnaire forms, and different interviewers. Nevertheless, the finding that emerges from these

studies is that responses to face-to-face and to telephone interviewing are generally quite comparable. A detailed study by Groves and Kahn (1979), for example, found few significant differences in univariate and bivariate response distributions by mode. The differences that did appear suggest that telephone respondents tend to give shorter responses to open questions and to give financial data in somewhat more rounded form than respondents to face-to-face interviewing. Cannell, Groves, and Miller (1981), on the other hand, found that the telephone respondents in their study reported somewhat more health events than the respondents to face-to-face interviewing, and they interpreted this as indicating higher quality reporting of this information on the telephone. In general, the differences that have been found are not large, and thus do not indicate that one mode is appreciably superior to the other in terms of the quality of the data produced.

In panel surveys, the use of a combination of modes of data collection often proves effective. The Current Population Survey (CPS), for instance, requires interviewers to collect data by face-to-face interviewing for households entering the panel but on certain other waves they may collect the data by telephone. At the end of 1976, about three out of five CPS interviews were completed by telephone (U.S. Bureau of the Census 1978). A similar approach is used in the National Crime Survey, in which, nowadays, about four out of five interviews are conducted by telephone (Paez and Dodge 1982). This mixed mode strategy capitalizes on the advantages of both modes, while also reducing their disadvantages. The use of face-to-face interviewing at the first wave avoids the noncoverage of nontelephone households and establishes a relationship with the sampled household. This relationship, together with the collection of the households' telephone numbers, helps with the response rate and the conduct of subsequent telephone interviews.

Another mixed mode strategy that can be useful with panel surveys is to employ both face-to-face interviewing and self-completion questionnaires. Again, the face-to-face interviews establish a relationship with the sampled households that can help with the response rate for the self-completion questionnaires. If face-to-face interviewing is used for the main data collection, self-completion questionnaires mailed to respondents a short time before the next wave of the survey can be collected by the interviewers. Self-completion questionnaires have a particular use in panel surveys in which one household member is interviewed in each sampled household, providing data for all household members. When individual responses are wanted from each member of the household--for instance, responses to attitude items--they may be collected by means of a self-completion questionnaire. Self-administered questionnaires of this type are being employed with the National Medical Expenditure Survey to obtain information on health behavior, health opinions, and attitudes.

At present, no mixed mode strategy of data collection is used with the SIPP, although, on the average, over the first six interviews of the 1984 Panel about 4.5 percent of interviews were conducted by telephone because the households were inaccessible for face-to-face interviews. An experiment is planned for August to November 1986 to investigate the large scale use of telephone interviewing in SIPP. The experiment is designed to provide estimates of data quality and cost per case for telephone interviewing versus personal visit interviewing. The SIPP National Telephone Test will be conducted with the last two rotations of Wave 2 and the first two rotations of Wave 3 of the 1986 Panel. Fifty percent of the households 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. The analysis will focus on comparing estimates of various characteristics, item nonresponse rates, and the cost data for the two interview modes.

IV. MEASURING GROSS CHANGE

A major analytic advantage of panel surveys is that they provide the data needed to measure gross change. The SIPP, for instance, collects data from which one can estimate not only the levels of use of different poverty programs, but also the flows into and out of those programs. While information on gross change is of great value, its measurement is bedevilled by serious response error and other problems. (See, for instance, Hogue and Flaim (1986) on the problems with the labor force gross flow data from the CPS.)

The critical difficulty in measuring gross change is that it is highly sensitive to changes in the measurement errors for sampled individuals between waves of the panel. Many aspects of a panel survey operation can give rise to changes in individual measurement errors, including:

1. Simple response variability. A respondent's answer may depend on his or her mood at the time of interview, the time and place of the interview, etc., irrespective of other changes.

2. Changing respondents between waves. In particular, information on a sampled person may be provided by a proxy informant on one wave, while the person may be a self-reporter on another.

3. Changing the mode of data collection. Although the discussion in section III suggests that responses to face-to-face and telephone interviews are quite comparable, any differences that do exist may have a serious effect on measures of gross change.

4. Changing interviewers. If the interviewers affect the responses obtained, changes in the interviewers for a respondent from one wave to the next will affect changes in measurement error. This issue is particularly relevant when the data are collected on one wave by face-to-face interviews while they are collected on the next

wave by centralized computer assisted telephone interviewing. In this case, all those interviewed by the latter procedure will have different interviewers on the two waves.

5. Changing questionnaires. Several studies have shown that responses to a survey question can be sensitive to the context in which the question is asked, that is, to the other questions on the questionnaire (see, for instance, Kalton and Schuman 1982). A change in the questionnaire from one wave to the next can then give rise to a change in the measurement error for a given (unchanged) question.

6. Changing interpretation of the meaning of a question. During the life of a long-term panel, the meaning of words may change in subtle ways, so that over time respondents interpret the identical question differently.

7. Panel conditioning. As discussed in section II, membership in a panel may cause a change in responses. Panel members may, for instance, report fewer events on later waves because of a loss of motivation, or alternatively they may report more because they better understand the task. Panel membership may also alter their behavior.

8. Changing coders. If coders vary in the ways in which they code answers, a change in coders for a particular respondent will cause a change in measurement errors.

9. Imputation. As noted at the end of the section on panel nonresponse, unless care is taken, imputation for item nonresponses can cause the amount of gross change to be overstated.

10. Matching files. If the sample is not properly accounted for, and the individual identification number assigned to a sample individual is not unique, mismatches of respondents across waves may generate a measure of change that is greater than is actually occurring.

11. Keying errors. Keying errors are a very small part of nonsampling errors in surveys. Nevertheless, they can give rise to incorrect changes in individual measurements.

Several examples will be given to illustrate the problems of measuring gross change using SIPP data. It must be emphasized that these examples are illustrative and serve to point out that, like other panel surveys, SIPP must continue to study and identify problems of the type discussed below, as well as continue to work toward solutions.

The first example illustrates that basic survey items which should either remain constant across time (race and sex) or change in predetermined ways

(age) can and do appear to change in inexplicable ways. The demographic changes observed in SIPP are few because of a control system designed to ensure that individual records can be matched successfully and accurately at successive survey waves. When inconsistencies in demographic data occur, the case is returned to the field for verification and correction. Thus, the demographic data in earlier waves may differ from demographic data in later waves because corrections are only made to the most recent wave of data collection. In this way, the most recent response is interpreted as the most accurate response.

Tables 6, 7, and 8 show the wave-to-wave inconsistencies in the race, sex, and age variables in the SIPP and are illustrative of a nonsampling error problem which occurs in all panel surveys. A change in the race of a respondent occurs less than 50 times between each wave of interviewing. This is approximately 0.1 percent of the entire sample. Similarly in table 8, a change in sex is observed no more than 34 times between waves of interviewing, or approximately 0.08 percent of the sample. Age changes are observed more frequently between waves of interviewing, the largest number occurring between waves 2 and 3, where 189 age changes occurred, or approximately 0.4 percent of the sample. The reasons for these changes are not known, but it is likely to be the combination of interviewer, keying or processing error.

The second example concerns the consistency of occupation and industry codes between interview waves. During the 1984 and 1985 SIPP Panels industry and occupation data were collected and coded independently during each interview, even though the individual's employer and job may not have changed. Thus, variation in occupation classification can result when the respondent's

description of his or her duties varies slightly or when the interpretation of the duties varies between the clerical staff members who assign the classification codes. For cases in which a characteristic remains stable between waves, a panel survey in effect provides a test-retest reliability check of the data. The issue of reliability is often ignored in single-round surveys; a panel survey, designed to measure change must, however, directly confront differences in reporting or coding the same phenomenon at two or more points in time.

Some evidence on the consistency of occupation and industry coding can be obtained by considering persons who reported the same employer for all of the first 12 months of the 1984 SIPP Panel (16,886 persons). Among these individuals 37.4 percent changed three-digit occupation codes between waves 1 and 2, 37.8 percent changed between waves 2 and 3, 19.8 percent changed three-digit industry codes between waves 1 and 2, and 20.1 percent changed between waves 2 and 3 (U.S. Bureau of the Census 1985e). Only about 50 percent of persons with the same employer in all 12 months had the same occupation and only about 70 percent had the same industry code in all three waves. Overall, only 39.9 percent of persons with the same employer in all 12 months had the same three-digit occupation and industry codes during this same time period.

In this example, the symptoms of nonsampling error are exhibited in what appear to be unduly large numbers of changes in industry and occupation codes between waves for persons who seem to have the same job. At first sight, this might appear to be a coding problem, but this is not necessarily the case. Even though the job is the same, the written description as provided by the respondent may vary greatly from one wave to the next due to response

variability, different interviewers, or a different respondent. Another possible source of error occurs when a person has more than one job and the interviewer incorrectly assigns a given job number to different jobs at different points in time. The problem in this case is that different jobs are matched across waves, thus generating industry and occupation changes.

A change in data collection procedures for occupation and industry was implemented in the 1986 Panel to reduce the number of changes in industry and occupation codes resulting from response error and clerical interpretation, as well as 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 and the classifications coded at the conclusion of the previous interview are then brought forward to the current interview during data processing. While this change is being made for the 1986 SIPP Panel, industry and occupation data from the 1985 Panel (collected during the same time period as the 1986 Panel) will be collected independently each wave, giving rise to a natural experiment embedded in the two panels.

Another symptom of nonsampling error in SIPP, which is troublesome for the measurement of gross flow statistics, was first noted in the Income Survey Development Program (Moore and Kasprzyk 1984). The problem occurs in the monthly observations of transitions between recipiency and nonrecipiency of some social programs. Program turnover occurred more often between waves of interviewing than within the wave. Consider 12 monthly SIPP observations; these monthly data are collected in three interviews with each interview obtaining data for the previous 4 months. That is, interview 1 collects data

on months 1 to 4; interview 2 collects data on months 5 to 8; and interview 3 on months 9 to 12. Burkhead and Coder (1985) observed that the rate of transitions is higher between months 4 and 5 and between months 8 and 9 than between the remaining pairs of adjacent months for a number of income sources.

Table 6. -- Inconsistencies in 1984 Panel Demographic Data: Wave-to-Wave Race Change 1/

From Wave	To Wave	White to Black	White to Other	Black to White/ Other	Other/ refused to White	TOTAL
1	2	31	1	2	12	46
2	3	16	7	2	2	27
3	4	31	7	4	2	44
4	5	13	15	4	7	39
TOTAL		91	30	12	23	156

Table 7. -- Inconsistencies in 1984 Panel Demographic Data: Wave-to-Wave Sex Change 1/

From Wave	To Wave	Male to Female	Female to Male	Total
1	2	22	12	34
2	3	16	18	34
3	4	6	16	22
4	5	20	13	33
TOTAL		64	59	123

1/ These tables were clerically developed from records maintained in the Demographic Surveys Division and used to monitor and control the SIPP sample. The total number of persons (adults and children) in the Wave 1 SIPP sample is approximately 53,700, in the Wave 2 sample approximately 42,000, in the Wave 3 sample approximately 56,200, and in the Wave 4 sample approximately 56,000.

Table 8. -- Inconsistencies in 1984 Panel Demographic Data: Wave-to-Wave Age Change 1/

From Wave	To Wave	AGE INCREASED 2 - 5 years	AGE INCREASED 6 -10 years	AGE INCREASED 11 or more years	AGE DECREASED 2 - 5 years	AGE DECREASED 6 -10 years	AGE DECREASED 11 or more years	Total
1	2	33	8	3	24	5	9	82
2	3	50	13	19	88	10	9	189
3	4	51	8	3	65	5	1	133
4	5	14	8	2	51	3	1	79
TOTAL		148	37	27	228	23	20	483

Table 9 drawn from Burkhead and Coder (1985) provides some examples of this phenomenon. Clearly an uneven pattern of change is observed and the pattern seems to be associated with the interviewing scheme. The inconsistency in the amount of change within and between interviewing waves is clearly a symptom of nonsampling error; the causes for the problem are not known, but seem to be related to questionnaire wording/design and respondent recall error. Kalton, Lepkowski, and Lin (1985) noticed similar findings relating to the correlation between monthly income amounts from various sources in 1979 ISDP Research Panel: the correlation of amounts one month apart are higher when the months are within the same wave than when they are obtained from two adjacent waves.

1/ This table was clerically developed from records maintained in the Demographic Surveys Division and used to monitor and control the SIPP sample. The total number of persons (adults and children) in the Wave 1 SIPP sample is approximately 53,700, in the Wave 2 sample approximately 42,000, in the Wave 3 sample approximately 56,200, and in the Wave 4 sample approximately 56,000.

Table 9. -- Monthly Transitions for Persons: Earnings, Unemployment Compensation, Social Security, and Food Stamps

(Persons interviewed first 12 months in sample)

Interview status	Mean Number of Monthly Transitions within Interview 1	4th month to 5th month	Mean Number of Monthly Transitions within Interview 2	8th month to 9th month	Mean Number of Monthly Transitions within Interview 3
EARNINGS					
Receiving both months	18,959	18,455	19,158	18,536	19,360
Not receiving to receiving	676	1,118	482	1,136	534
Receiving to not receiving	557	1,296	451	1,129	453
UNEMPLOYMENT COMPENSATION					
Receiving both months	435	313	462	294	343
Not receiving to receiving	129	234	141	165	90
Receiving to not receiving	155	241	128	294	105
SOCIAL SECURITY					
Receiving both months	6,484	6,473	6,651	6,650	6,781
Not receiving to receiving	56	157	44	134	31
Receiving to not receiving	19	105	13	73	17
FOOD STAMPS					
Receiving both months	1,339	1,224	1,365	1,243	1,344
Not receiving to receiving	72	162	55	140	40
Receiving to not receiving	53	207	40	190	38

Adapted from Burkhead and Coder (1985)

V. CONCLUDING REMARKS

This paper has discussed the effects of some of the major design decisions in the SIPP on nonsampling error. In many cases, no research has yet been conducted on the effects of these design decisions on SIPP data, and reliance has had to be placed on the findings obtained in other studies. This approach is typical of what must be done in the early years of an ongoing data collection program like the SIPP. There are, however, risks in trying to generalize to SIPP from other surveys with different subject matter and methods. We hope that a program of methodological research will provide direct evidence on the effects of the major SIPP design decisions on the quality of SIPP data.

The design decisions we have chosen to discuss are only a few of the multitude of decisions that have to be made in a complex survey like the SIPP. Research on the effects of other design decisions on the quality of SIPP data is also needed.

Finally, it should be noted that, although we have considered the various design decisions independently, they are in fact interrelated. Thus, for instance, a change from face-to-face to telephone interviewing would have implications for the respondent rules to be used and also affect the following rules. Survey design is a complex web of a very large number of interwoven design decisions.

REFERENCES

BAILAR, BARBARA A. (1975), "The Effects of Rotation Group Bias on Estimates from Panel Surveys," Journal of the American Statistical Association, 70, 23-30.

BAILAR, BARBARA A. (1979), "Rotation Sample Biases and their Effects on Estimates of Change," Bulletin of the International Statistical Institute, 48(2), 385-407.

BECKETTI, SEAN, GOULD, WILLIAM, LILLARD, LEE, and WELCH, FINIS (1983), Attrition from the PSID, Santa Monica: Unicorn Research Corporation.

BURKHEAD, DAN and CODER, JOHN (1985), "Gross Changes in Income Recipiency from the Survey of Income and Program Participation," Proceedings of the Section on Survey Research Methods, American Statistical Association, forthcoming.

CANNELL, CHARLES F., FISHER, GORDON, and BAKKER, THOMAS (1965), Reporting of Hospitalization in the Health Interview Survey, Vital and Health Statistics, Series 2, No. 6. Washington, D.C.: U.S. Government Printing Office.

CANNELL, CHARLES F., GROVES, ROBERT, and MILLER, PETER V. (1981), "The Effects of Mode of Data Collection on Health Survey Data," Proceedings of the Social Statistics Section, American Statistical Association, 1-6.

CODER, JOHN F. and FELDMAN, ANGELA M. (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, 693-697.

DEMING, WILLIAM E. (1944), "On Errors in Surveys," American Sociological Review, 9, 359-369.

FERBER, ROBERT (1964), "Does a Panel Operation Increase the Reliability of Survey Data: The Case of Consumer Savings," Proceedings of the Social Statistics Section, American Statistical Association, 210-216.

GAROFALO, JAMES and HINDELANG, MICHAEL J. (1977), An Introduction to the National Crime Survey, National Criminal Justice Information and Statistics Service, Law Enforcement Assistance Administration, U.S. Department of Justice, Washington, D.C.: U.S. Government Printing Office.

GHANGURDE, P.D. (1982), "Rotation Group Bias in the LFS Estimates," Survey Methodology, 8, 86-101.

GROVES, ROBERT M. and KAHN, ROBERT R. (1979), Surveys by Telephone: A National Comparison with Personal Interviews, New York: Academic Press.

HANSEN, MORRIS H. and WAKSBERG, JOSEPH (1970), "Research on Nonsampling Errors in Censuses and Surveys," Review of the International Statistical Institute, 38, 317-332.

HOGUE, CARMA R. and FLAIM, PAUL O. (1986), "Measuring Gross Flows in the Labor Force: An Overview of a Special Conference," Journal of Business and Economic Statistics, 4, 111-121.

HOLDEN, KAREN, BURKHOUSER, RICHARD, and MYERS, DANIEL (1985), "Marital Disruption and Poverty: the Role of Survey Procedures in Artificially Creating Poverty," Institute for Research on Poverty Discussion Paper #781-85, Madison, Wisconsin.

HORVITZ, DANIEL G. (1978), "Some Design Issues in Sample Surveys," in Survey Sampling and Measurement, N. Krishnan Namboodiri ed., pp. 1-11. New York: Academic Press.

INSTITUTE FOR SOCIAL RESEARCH (1985), A Panel Study of Income Dynamics: Procedures and Tape Codes 1983 Interview Year, Ann Arbor, Michigan: Institute for Social Research.

JEAN, ANNE C. and McARTHUR, EDITH K. (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, 745-750.

KALTON, GRAHAM (1985), "Handling Wave Nonresponse in Longitudinal Surveys," Proceedings of the Bureau of the Census First Annual Research Conference, 453-461.

KALTON, GRAHAM, KASPRZYK, DANIEL, and SANTOS, ROBERT (1981), "Issues of Nonresponse and Imputation in the Survey of Income and Program Participation," in Current Topics In Survey Sampling, Krewski, Daniel, Platek, Richard, Rao, Jon N.K. eds., pp. 455-480. New York: Academic Press.

KALTON, GRAHAM and LEPKOWSKI, JAMES (1983), "Cross-wave Item Imputation," in Technical, Conceptual, and Administrative Lessons of the Income Survey Development Program (ISDP), David, Martin H. ed., pp. 171-198, Washington, D.C.: Social Science Research Council.

KALTON, GRAHAM and LEPKOWSKI, JAMES (1985), "Following Rules in SIPP," Journal of Economic and Social Measurement 13, 319-329.

KALTON, GRAHAM, LEPKOWSKI, JAMES, and LIN, TING-KWONG (1985), "Compensating for Wave Nonresponse in the 1979 ISDP Research Panel," Proceedings of the Section on Survey Research Methods, American Statistical Association, forthcoming.

KALTON, GRAHAM and SCHUMAN, HOWARD (1982), "The Effect of the Question On Survey Responses: A Review," Journal of the Royal Statistical Society, A, 145, 42-73.

KULKA, RICHARD A. (1984), "The ISDP 1979 Research Panel as a Methodological Survey: Implications for Substantive Analysis," Proceedings of the Section on Survey Research Methods, American Statistical Association, 738-743.

LIEVESLEY, DENISE and WATERTON, JENNIFER (1985), "Measuring Individual Attitude Change," in British Social Attitudes: The 1985 Report, Jowell, Roger and Witherspoon, Sharon eds., pp. 177-194, Aldershot, Hants: Gower Publishing Company.

MATHIOWETZ, NANCY A. and GROVES, ROBERT M. (1985), "The Effects of Respondent Rules on Health Survey Reports," American Journal of Public Health, 75, 639-644.

MCCARTHY, PHILIP J. (1980), "Some Sources of Error in Labor Force Estimates from the Current Population Survey," in Data Collection, Processing and Presentation: National and Local, National Commission on Employment and Unemployment Statistics, pp. 2-39, Washington, D.C.: U.S. Government Printing Office.

MOONEY, H. WILLIAM (1962), Methodology in Two California Health Surveys, Public Health Monograph No. 70, Washington, D.C.: U.S. Government Printing Office.

MOORE, JEFFREY C. and KASPRZYK, DANIEL (1984), "Month-to-month Recipiency Turnover in the ISDP," Proceedings of the Section on Survey Research Methods, American Statistical Association, 726-731.

NELSON, DAWN, MCMILLEN, DAVID B., and KASPRZYK, DANIEL (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.

NETER, JOHN and WAKSBERG, JOSEPH (1964), "A Study of Response Errors in Expenditures Data from Household Surveys," Journal of the American Statistical Association, 59, 18-55.

NETER, JOHN and WAKSBERG, JOSEPH (1965), Response Errors in Collection of Expenditures Data by Household Interviews: An Experimental Study, Technical Paper No. 11, Washington, D.C.: U.S. Government Printing Office.

OLSON, JANICE, ed. (1980), Reports from the Site Research Test, Office of the Assistant Secretary for Planning and Evaluation, Department of Health and Human Services.

PAEZ, A.D., and R.N. DODGE (1982), Criminal Victimization in the U.S., 1979-80 Changes, 1973-80 Trends, Bureau of Justice Statistics Technical Report, NCJ-80838, Washington, D.C.: U.S. Department of Justice.

ROSHWALB, ALAN (1982), "Respondent Selection Procedures Within Households," Proceedings of the Section on Survey Research Methods, American Statistical Association, 93-98.

SOBOL, MARION G. (1959), "Panel Mortality and Panel Bias," Journal of the American Statistical Association, 54, 52-68.

TRAUGOTT, MICHAEL W. and KATOSH, JOHN P. (1979), "Response Validity in Surveys of Voting Behavior," Public Opinion Quarterly, 43, 359-377.

TURNER, ANTHONY G. (1972), The San Jose Methods Test of Known Crime Victims, Washington, D.C.: National Criminal Justice Information and Statistics Service, Law Enforcement Assistance Administration, U.S. Department of Justice.

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.

U.S. BUREAU OF THE CENSUS (1985a), Income Nonresponses: March 1983 CPS, Technical Paper 54, Washington, D.C.: U.S. Government Printing Office.

U.S. BUREAU OF THE CENSUS (1985b), Current Population Reports, Series P-70, No. 3, Economic Characteristics of Households in the United States: First Quarter 1984, U.S. Government Printing Office, Washington, D.C.

U.S. BUREAU OF THE CENSUS (1985c), Current Population Reports, Series P-70, No. 4, Economic Characteristics of Households in the United States: Second Quarter 1984, U.S. Government Printing Office, Washington, D.C.

U.S. BUREAU OF THE CENSUS (1985d), Current Population Reports, Series P-70, No. 5, Economic Characteristics of Households in the United States: Third Quarter 1984, U.S. Government Printing Office, Washington, D.C.

U.S. BUREAU OF THE CENSUS (1985e), "SIPP Waves 1-3: Changes in Occupation and Industry Codes Between Waves," memorandum from Dan Burkhead to John Coder, May 7, 1985.

U.S. BUREAU OF THE CENSUS (1986a), Current Population Reports, Series P-70, No. 6, Economic Characteristics of Households in the United States: Fourth Quarter 1984, U.S. Government Printing Office, Washington, D.C.

U.S. BUREAU OF THE CENSUS (1986b), "Measurement of Attrition from SIPP through the Fifth Wave of the 1984 Panel," memorandum from E. McArthur and K. Short to distribution list, April 10, 1986.

U.S. BUREAU OF THE CENSUS and BUREAU OF LABOR STATISTICS (1985), Proceedings of the Conference on Gross Flows in Labor Force Statistics, Washington, D.C.: U.S. Bureau of the Census and Bureau of Labor Statistics.

VAUGHAN, DENTON, WHITEMAN, CAMERON, and LININGER, CHARLES (1984), "The Quality of Income and Program Data in the 1979 ISDP Research Panel: Some Preliminary Findings," Review of Public Data Use, 12, 127-132.

WHITE, GLEN G. and HUANG, HERTZ (1982), "Mover Follow-up Costs for the Income Survey Development Program," Proceedings of the Section on Survey Research Methods, American Statistical Association, 376-381.

WOLTMAN, HENRY and BUSHERY, JOHN (1975), "A Panel Bias Study in the National Crime Survey," Proceedings of the Social Statistics Section, American Statistical Association, 159-167.