Survey of Income and Program Participation

SIPP Quality Profile

1998

SIPP Working Paper
Number 230

3rd Edition

U.S. Department of Commerce
Bureau of the Census

PREFACE FOR 3RD EDITION

This third edition of the *Survey of Income and Program Participation Quality Profile* was prepared for the U.S. Bureau of the Census by Westat under contract number 50-YABC-2-66025, task order number 46-YABC-6-00017. This edition builds on the earlier editions to provide a cumulative review of the first 10 years of SIPP, covering the 1984 to 1993 SIPP panels. It also summarizes the new features and design for the 1996 panel.

Graham Kalton was the project director and principal editor for this edition. Revisions were made by Marianne Winglee, Louis Rizzo, and Thomas Jabine. Daniel Levine provided much support and useful comments through the entire process. Vicki Huggins and Karen King of the Bureau of the Census provided the relevant source materials and helped clarify many technical issues. Stephen Mack of the Bureau of the Census was also helpful in reviewing some technical material. Thomas Jabine reviewed the final draft and Pat Dean Brick conducted an editorial review. The preparation of this edition was ably supported by Janice Scullion, Chantell Adams, Mary Lou Pieranunzi, and Charlotte Lass.

This edition updates the previous editions by integrating new materials into the existing chapters and deleting obsolete materials. All chapters have been revised and brought up-to-date. Chapters 8 and 10 have undergone the most changes. Chapter 8 now includes information about the longitudinal weights, and the new weighting and imputation procedures that were implemented in 1990. Chapter 10 is rewritten to include more recent evaluations of estimates and their quality over time. A new chapter, Chapter 11, has been added to describe the major redesign of SIPP that was implemented with the 1996 Panel.

This edition of the *Quality Profile* relies heavily on the previous editions, and those who worked on those editions should be acknowledged. The second edition was prepared by Thomas Jabine with the assistance of Karen King and Rita Petroni, and that edition built on the first edition prepared by Karen King, Rita Petroni, and Rajendra Singh. Acknowledgments to others contributing to the earlier editions are given in the Preface to the Second Edition that follows.

PREFACE FOR 2ND EDITION

The *Survey of Income and Program Participation Quality Profile* was written for the U.S. Bureau of the Census by Thomas B. Jabine under Purchase Order Number 43-YABC-922193. Karen E. King and Rita J. Petroni of the Statistical Methods Division, U.S. Bureau of the Census, provided invaluable support to the author by providing relevant source material and clarifying many technical issues that arose in the preparation of this monograph.

This, the second edition of the *Survey of Income and Program Participation Quality Profile*, is an expansion and revision of the first edition *Quality Profile*, available as SIPP Working Paper No. 8708 (Karen E. King, Rita J. Petroni, and Rajendra P. Singh, authors). It is the result of the combined efforts of many people. In particular, the American Statistical Association Survey Research Methods Section Working Group on the Technical Aspects of the SIPP played an important role in providing ideas and comments for the first edition, and then articulating the need for periodically providing current information on the topics addressed in the *Quality Profile*. Their perceptive observations and suggestions helped in defining the structure, content, and scope of the *Quality Profile*. The encouragement and strong support of the Technical Working Group, in particular its chair, Constance Citro (National Academy of Sciences), and her predecessor as chair, Graham Kalton (University of Michigan), is appreciated.

Members of the American Statistical Association Survey Research Methods Section Working Group on the Technical Aspects of the SIPP during the period the *Quality Profile* was under preparation include: Constance Citro (National Research Council), Chair; Graham Kalton (University of Michigan); Michael Brick (Westat); Gordon Brackstone (Statistics Canada); Ralph Folsom (Research Triangle Institute); Martin Frankel (Baruch College, CUNY); Robert Groves (University of Michigan); Thomas B. Jabine (Statistical Consultant); Roderick Little (UCLA); and Nancy Mathiowetz (Agency for Health Care Policy and Research). Daniel Kasprzyk (U.S. Bureau of the Census) and Rajendra P. Singh (U.S. Bureau of the Census) served as technical liaisons for the working group. The working group's comments and suggestions helped to improve the quality and clarity of the *Quality Profile*.

Staff from a number of divisions within the Bureau of the Census provided many useful comments; in particular, Rajendra P. Singh (Statistical Methods Division), Chester E. Bowie (Demographic Surveys Division), and Daniel Kasprzyk (SIPP Research and Coordination Staff) were very helpful. Hazel V. Beaton (SIPP Research and Coordination Staff) worked accurately, diligently, and patiently to prepare the manuscript for publication; her dedication is exemplary and her efforts are appreciated.

SURVEY OF INCOME AND PROGRAM PARTICIPATION (SIPP): QUALITY PROFILE

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

List of Tables

TABLE OF CONTENTS (Continued)

List of Tables (continued)

TABLE OF CONTENTS (Continued)

List of Tables (continued)

TABLE OF CONTENTS (Continued)

List of Tables (continued)

TABLE OF CONTENTS (Continued)

List of Tables (continued)

List of Figures

TABLE OF CONTENTS (Continued)

List of Exhibits

1. INTRODUCTION

This is the third edition of the *Quality Profile* for the Survey of Income and Program Participation (SIPP). This edition is an update of the previous editions (King et al., 1987; Jabine et al., 1990) and it provides a cumulative review of the data quality of SIPP panels. Users' comments on the earlier versions also are reflected in the current edition. We value user input and continue to welcome comments for improvements.

1.1 Purpose and audience

The purpose of the *Quality Profile* is to summarize, in convenient form, what is known about the sources and magnitude of errors in estimates from SIPP. Without such a summary, anyone wanting information about the quality of SIPP estimates would have to search through a large body of literature, some of it not easily accessible. The *Quality Profile* draws on that literature and provides references for readers who want more detailed information. The report covers both sampling and nonsampling error; however, the primary emphasis is on the latter. SIPP data users requiring detailed guidance on how to take sampling errors into account in their analyses should refer to the chapter on Sampling Error in the *SIPP User's Guide* (U.S. Bureau of the Census, 1998).

The *Quality Profile* is addressed to two kinds of readers: users of SIPP data and those who are responsible for or have an interest in the SIPP design and methodology. The needs of the two groups are somewhat different. Data users are interested primarily in knowing about the levels of error associated with the specific categories or classes of estimates that they are using in their analyses. Census Bureau staff and others interested in improving the SIPP design and procedures will want to know the magnitude of errors associated with different features of the design, such as length of reference period, procedures for following persons who move, respondent rules, and quality control procedures for data collection and processing operations. As explained further in Section 1.2, we have tried to accommodate the needs of both groups.

Persons who are associated or concerned with the design and methodology of panel surveys other than SIPP should find much in this report that is relevant to their interests. As a multipurpose national survey with a relatively complex design, SIPP illustrates most of the design issues and methodological problems that are likely to arise in any household panel survey.

1.2 Sources of information on data quality

The quality of SIPP data is affected by both sampling and nonsampling errors. Sampling errors for estimates derived from the SIPP data are calculated using standard techniques. Also, estimates of sampling error for specific survey estimates are used to develop generalized sampling error functions or tables to which users can refer for their particular needs.

Information about nonsampling errors comes from many sources. Some sources, such as the record check studies, can provide direct estimates of the level of error associated with certain estimates. Other sources provide only indications of the likelihood of error. An illustration of this kind of information is data on item nonresponse rates. Imputation procedures may overcome some of the effects of item nonresponse; however, one would expect the level of bias by item to be correlated with the level of item nonresponse.

The main sources of information on nonsampling errors are:

- Performance data, such as interview completion rates, item nonresponse rates, and results of reinterviews.
- Methodological experiments designed to measure the effects on quality of changing one or more features of the survey design or procedures.
- Micro-evaluation studies, that is, studies which attempt to determine the size and direction of errors associated with individual survey responses.
- Macro-evaluation studies which review aggregate data for internal consistency and in comparison with similar data from other sources.

All of these sources provide a substantial amount of information on SIPP data quality.

1.3 Structure of the report

There are several possible methods of organizing data on errors in surveys. The method used in this report is to present information on errors associated with each phase of survey operations: frame design and maintenance, sample selection, data collection, data processing, estimation, and data dissemination. Information not related to specific survey operations, such as sampling errors and the results of macro-evaluation studies are presented in separate chapters.

Chapter 2 provides a brief overview of SIPP: its objectives and principal design features. It is recommended that readers who are not familiar with the SIPP design read this chapter before proceeding to the later chapters. Greater detail on procedures for different phases of survey operations is presented in subsequent chapters.

Chapters 3 to 10 review various aspects of the quality of SIPP panels between 1984 and 1993. Chapter 3 covers frame construction, frame maintenance, and sample selection procedures. Chapters 4 to 6 cover the data collection phase of the survey. Chapter 4 describes the data collection procedures. Chapters 5 and 6 cover nonresponse and measurement errors, respectively. Chapter 7 deals with the data preparation phase (data entry, coding, and editing) and Chapter 8 covers estimation and data dissemination. Chapter 9 briefly discusses sampling errors and their relation to sample size. Chapter 10 presents the results of macro-evaluation studies and reviews survey estimates topic by topic. This chapter should be of particular interest to data users.

Based on the experience gained in the first decade of SIPP, a major redesign was introduced with the 1996 Panel, with changes made to the sample design and survey operations. Chapter 11 outlines the new features of the 1996 Panel.

Chapter 12 is a concluding chapter that offers users of SIPP data some general guidelines for making effective use of SIPP data products, summarizes the main sources of error in SIPP and what is known about their effects, and provides an overview of current research. References are listed at the end of the report.

1.4 Sources of additional information

Several types of Census Bureau publications are cited in this report. The P-70 series presents tabulations and analyses of SIPP data. Each publication in the series includes an appendix on the sources and reliability of estimates. An extensive series of SIPP working papers (which includes the first version of the *SIPP Quality Profile*) provides information about methodological aspects of the survey as well as analyses of SIPP data. Such working papers are not cleared for formal publication, but are readily available upon request (some are available through the Internet). Since 1984, papers on SIPP results and methodology presented

at the annual meetings of the American Statistical Association have been reproduced in the working paper series. Several important papers on SIPP methodology and evaluation studies have been presented and published in the proceedings of the Census Bureau's research conferences.

This report relies wherever possible on published or readily available sources, but occasionally cites Census Bureau memoranda and other unpublished materials. Readers interested in obtaining such unpublished materials should write to the SIPP Branch, Demographic Statistical Methods Division, U. S. Bureau of the Census, Washington, DC 20233 or call (301) 457-4192.

The *SIPP User's Guide* contains a detailed account of the design and content of the survey. The SIPP Home Page at the Internet web site <http://www.sipp.census.gov/sipp> provides information on various aspects of SIPP. This web site will include an on-line tutorial that helps users learn how to analyze the SIPP database, and an extensive bibliography of references on SIPP-related research and documentation.

Readers with questions about specific aspects of SIPP findings and methodology may use the on-line service Surveys-on-Call at the SIPP Home Page by selecting Data Access, and then Data Extraction System (DES) or may call the following numbers:

Subject matter fields

Aging population	301-457-2738
Child care	301-457-2416
Education	301-457-2464
Health insurance	301-457-3215
Income	301-457-3243
Marriage and family	301-457-2416
Migration	301-457-2454
Pensions	301-457-3243
Poverty	301-457-3245
Wealth	301-457-3224
Women	301-457-2378

Methodology

Data collection procedures	301-457-3819
Questionnaire design	301-457-3819
Estimation and weighting	301-457-4192
Nonsampling and sampling errors	301-457-4192
Survey design	301-457-4192

2. OVERVIEW OF SIPP DESIGN

This chapter provides an overview of the SIPP design. Later chapters provide greater detail on design features and procedures for specific phases of the survey. Useful references about the goals and design for the first 10 years of SIPP include: Nelson et al. (1985), Kasprzyk (1988), and Petroni (1989). The issues considered for the 1996 redesign are summarized in Citro and Kalton (1993) and Huggins and Fischer (1994).

2.1 Evolution of the survey design

Early in the 1970s, it became evident that a new survey was required to provide information for detailed analyses of income transfer programs. At that time the Department of Health, Education and Welfare initiated the Income Survey Development Program (ISDP) to develop concepts and content and to test alternative procedures for collecting the data. Field experiments focused on general feasibility and design features such as questionnaire length, length of reference period, respondent rules, and possible linkages of survey and administrative record information (Olson, 1980; Ycas and Lininger, 1981; David, 1983).

The lessons learned in the ISDP were taken into account in the design of SIPP and interviewing began in October 1983. The design, although not regarded as final at that time, was used for the panels between 1984 and 1993. It was expected that adjustments would be necessary and lessons would be learned that would lead to significant changes. An ongoing program of research, experimentation, and evaluation was established to identify the strengths and weaknesses of the design and to test alternative design features, such as use of the telephone for some interviews, incentives to encourage participation in the survey, and feedback of prior-wave information to respondents.

In 1990, the Census Bureau asked the Committee on National Statistics (CNSTAT) of the National Research Council to convene a panel to undertake a wide-ranging review of SIPP. The panel's report, *The Future of the Survey of Income and Program Participation* (Citro and Kalton, 1993), provides a comprehensive summary of the first 9 years of SIPP along with a number of recommendations for the future of the survey. Some of the recommendations have been implemented with the 1996 SIPP Panel. The year 1996 marked the beginning of a new design for SIPP panels and the initiation of the Survey of Program Dynamics (SPD), authorized by The Personal Responsibility and Work Opportunity Reconciliation Act of 1996 (P.L. 104-193). In 1996, the 1992 and 1993 Panels of the SIPP were recontacted through the SPD and data will be collected annually from these panels through 2001. The objective is to collect 10 years of data (1992-2001) to support analyses of changes in income, work experience, and program participation resulting from the implementation of welfare reform initiatives.

2.2 Survey objectives

The main objective of SIPP is to provide policy makers with more accurate and comprehensive information than previously available about the income and program participation of persons and households in the United States and about the principal determinants of income and program participation. SIPP was designed with the expectation that it would provide key data to assist in the formulation and evaluation of initiatives in welfare reform, tax reform, and the improvement of entitlement programs such as Social Security and Medicare. Because the differences between the temporarily poor and the long-term poor were considered important, it was of interest to have detailed information on cash and noncash income on a subannual basis. SIPP is also of value when studying tax reform as it provides annual information on income, taxes, assets, and liabilities.

To provide the kinds of information described above, both cross-sectional and longitudinal estimates (the latter including gross flows, transition probabilities, and spell durations) are needed. Key cross-sectional estimates include the proportions of persons receiving various types of income (recipiency), the amounts received, and the person and family characteristics that may influence income and program participation. Longitudinal estimates require the collection of information about a person over a period of time. For example, it may be desired to measure the cumulative effect of patterns of irregular employment on income, savings, assets, and program benefits or the patterns of change in multiple receipt of program benefits.

SIPP is designed to produce reliable estimates at the national level for all persons and families in the United States and for subgroups defined in terms of characteristics related to program eligibility, such as age, family type, labor force status, income, and assets. Subgroups of special interest include: low and high-income households, households headed by females, and elderly persons. The ability to produce reliable estimates for small subgroups is limited by the sample size. Information on the reliability of estimates of level and change for subgroups of various sizes is presented in Section 9.2.

2.3 Survey design

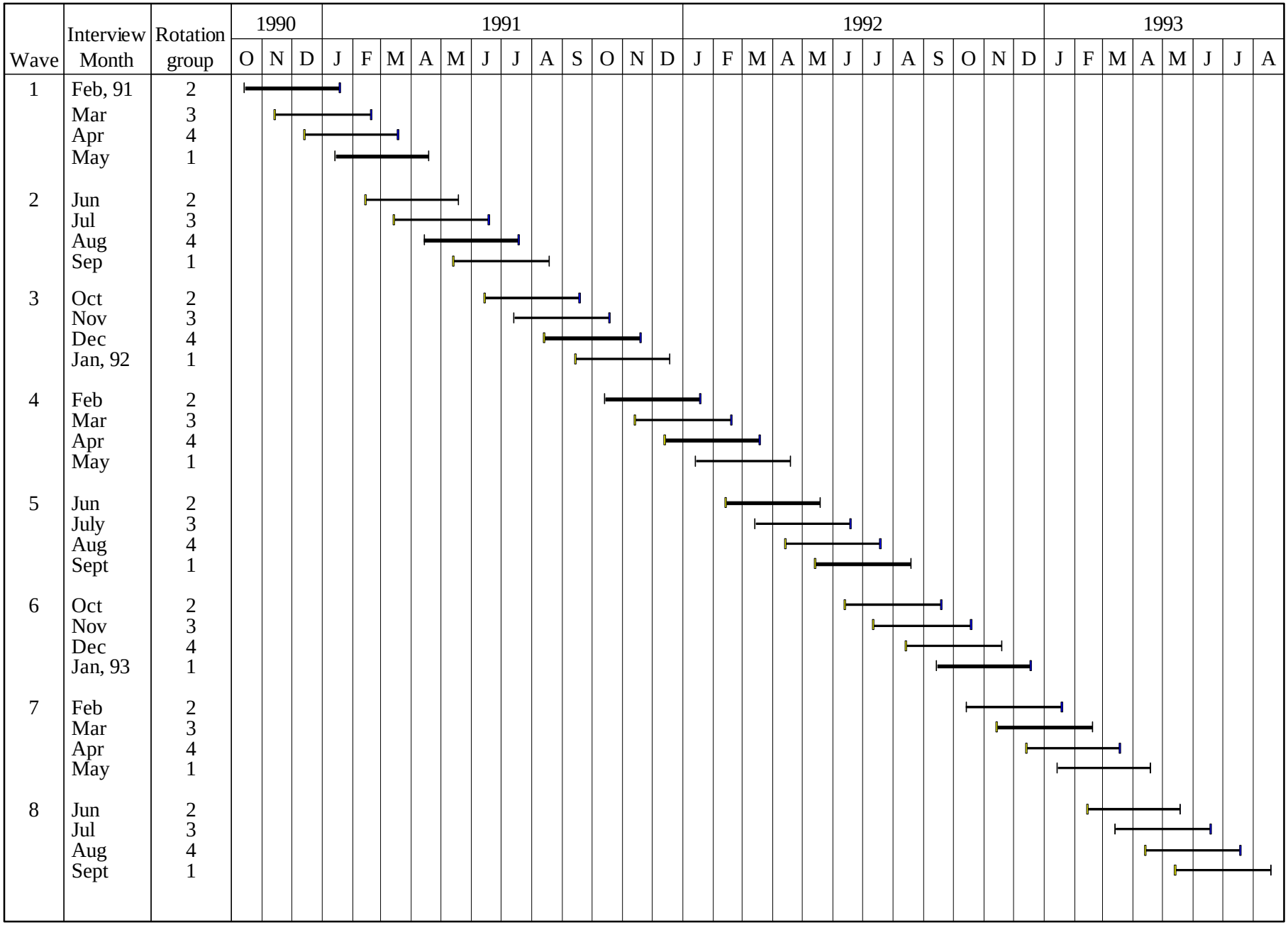
The survey population for SIPP consists of persons resident in United States households and persons living in group quarters, such as dormitories, rooming houses, religious group dwellings, and family-type housing on military bases. Persons living in military barracks and in institutions, such as prisons and nursing homes, are excluded.

SIPP is a panel survey. In the 10 years between 1984 and 1993, the basic design of SIPP involved the introduction of a new panel of sample households at the beginning of each calendar year, and interviewing members of the panel at 4-month intervals over a period of 2⅔ years for a total of eight interview waves. There were a number of variations from this basic design for individual panels as described later. With the 1996 Panel, the panel length has been extended to 4 years and a total of 12 waves.

A wave is defined as one round of interviews for a panel. To even out the interviewing workload, each panel is subdivided into four rotation groups, with one rotation group interviewed each month. Interviewing for each panel began in February, with the exception of the first panel in 1984 which began in October 1983. The reference period for most items covered in the interview consists of the 4 months immediately preceding the month of interview. Figure 2.1 shows the interview months and the reference months, by wave and rotation group, for the 1991 Panel. For example, rotation group 2 was first interviewed in February 1991, and asked about the reference months of October, November, and December 1990, and January 1991.

SIPP original sample members are adult members (ages 15 and older) of responding households at the first interview. Each original sample member is followed until the end of the panel or until the person becomes ineligible (by dying, entering an institution, moving to Armed Forces barracks, or moving abroad) or leaves the sample (for example, by refusing to continue to be interviewed, moving to an unknown or out-of-range address).

Figure 2.1. Reference months for each interview month: 1991 panel



SIPP is a true panel survey in that it follows original sample members, including those who change their addresses; this is in contrast to a quasi-panel survey such as the Current Population Survey (CPS) that returns to the same addresses and interviews the people who currently reside there. Starting in May 1985 for the 1984 Panel, and in October 1985 for the 1985 Panel, and at the beginning of all other panels, persons leaving the survey population to enter institutions have been followed and are interviewed again if they reenter the survey population. During a panel, persons who join the household of an original sample member are included in the survey as long as they remain with an original sample member. Those who joined after Wave 1 are not followed if they move from the households of the original sample members.

Table 2.1 shows the interviewing periods spanned by SIPP Panels 1984-1993, the number of eligible households sampled at Wave 1 (including nonresponding households) for each panel, and the number of original sample members in eligible households at Wave 1. The 1984, 1985, and 1986 Panels had one or two short waves where three rotation groups were used instead of four groups. Some panels had fewer waves of data collection and smaller sample sizes because of budget constraints. The 1992 and 1993 Panels had more waves because no new panels were introduced in 1994 and 1995, pending the debut of the SIPP redesign in 1996. The first SIPP panel in 1984 sampled about 21,000 households. The sample size for subsequent panels varied from about 12,500 to 23,500 households, and from about 33,000 to 62,000 original sample persons.

Between 1984 and 1993, SIPP did not oversample specific population groups, with the one exception that the 1990 Panel included about 3,800 extra households continued from the 1989 Panel. These extra households were selected because they were headed by blacks, Hispanics, or female single parents at the first wave of the 1989 Panel. Beginning in 1996, low-income households are oversampled to provide a larger sample of such households for analyses.

Table 2.1 A summary of the 1984-1993 SIPP Panels

Panel	First interview	Last interview	Number of waves	Wave 1	
				Eligible households[1]	Original sample members[2]
1984	Oct. 83	Jul. 86	9	20,897	55,400
1985	Feb. 85	Aug. 87	8	14,306	37,800
1986	Feb. 86	Apr. 88	7	12,425	32,800
1987	Feb. 87	May 89	7	12,527	33,100
1988	Feb. 88	Jan. 90	6	12,725	33,500
1989	Feb. 89	Jan. 90	3	12,867	33,800
1990	Feb. 90	Sept. 92	8	23,627	61,900
1991	Feb. 91	Sept. 93	8	15,626	40,800
1992	Feb. 92	May 95	10	21,577	56,300
1993	Feb. 93	Jan. 96	9	21,823	56,800

[1]Eligible households are households sampled for Wave 1, including both responding and nonresponding households.
[2]Number includes an estimate of the persons in Wave 1 noninterviewed households.

In the first decade of SIPP, a new panel was introduced each year and each panel stayed in the field for about 32 months. Therefore, there were two or three panels being interviewed at any given point of time. Data from different panels covering the same time period can be combined to produce estimates that have smaller sampling errors than those based on a single panel. For example, when the 1992 Panel was introduced in February 1992, two other panels were in the field: the 1991 Panel was being interviewed for its Wave 4, and the 1990 Panel was being interviewed for its Wave 7. By combining the 1990, 1991, and 1992 Panels for this period, users can obtain a total sample size of over 50,000 households, after taking into account attrition and nonresponse. The 1996 Panel has a sample size of 37,000 households and no overlap with other panels.

2.4 Content

For the 1984 to 1993 Panels, SIPP data were collected by means of paper and pencil instruments that consisted of a control card and a questionnaire.[1] Basic demographic characteristics and other classification variables associated with a household and its members were recorded on the control card in the initial interviews for a panel and updated in each subsequent wave. The survey questionnaire consisted of core questions, which were repeated at each wave, and topical modules, which included questions on selected topics. The topical modules varied from wave to wave. The main topics covered by the core questions were labor force participation and sources and amounts of income. Information for most items in these categories was obtained at every interview for each of the 4 months included in the interview reference period.

SIPP distinguishes between two kinds of topical modules: fixed and variable. Fixed topical modules are modules that are included in 1 or more waves during the life of each panel to augment the core data; they include, for example, modules on annual income, retirement accounts, income taxes, educational financing and enrollment, personal history, and wealth. Variable topical modules, which are designed to satisfy the special programmatic needs of other Federal agencies, are not necessarily repeated from one panel to the next. Some topics that have been covered are: child care arrangements, child support agreements, support for nonhousehold members, long-term care, pension plan coverage, housing costs, and energy usage. Variable modules are usually included in Waves 3 and 6 while fixed modules appear in other waves.

The data collection instruments for the 1996 Panel were revised to accommodate the introduction of computer-assisted interviewing (CAI). The content of the CAI instruments are comparable to the paper and pencil versions. The main changes are in the organization and wordings of the questions (see Section 11.4).

2.5 Data collection procedures

The preferred mode of data collection for the period 1984-1991 was face-to-face interviewing and most interviews conducted during this period used this method. In February 1992, SIPP switched to maximum telephone interviewing to reduce cost. The interviews for Waves 1 and 2 were conducted by face-to-face interviews as before, but interviews at subsequent waves were conducted by telephone to the extent possible. For the 1996 Panel, computer-assisted personal interviewing (CAPI) was used for Waves 1 and 2. For subsequent waves, one personal visit is planned each year; the remaining interviews will be conducted through computer-assisted telephone interviews (CATI).

In every SIPP Panel, persons ages 15 and older who are present at the time of the interview are asked to provide information for themselves (self-response). For those who are not available, another adult member of the household is asked to provide the information (proxy response). SIPP interviews are conducted by Census Bureau field representatives (FRs) under the supervision of the Census Bureau's 12 permanent regional offices. Telephone interviews and personal visits are carried out by the same FR using the same questionnaire.

[1] Copies of the control cards, questionnaires, and associated materials used in the data collection for the 1984 through 1993 Panels are available in SIPP informational booklets, which can be obtained from Customer Services, Data User Services Division, Bureau of the Census, Washington, DC 20233, telephone (301) 457-4100. The booklets for the different panels are: SIPP-4020 for the 1984 Panel; SIPP-5020 for the 1985 and 1986 Panels; SIPP-7020 for the 1987, 1988, and 1989 Panels; SIPP-9020 for the 1990 and 1991 Panels; and SIPP-9220 for the 1992 and 1993 Panels.

3. SAMPLE SELECTION: 1984-1993 PANELS

This chapter describes the sample selection procedures for SIPP panels between 1984 and 1993, discusses possible sources of error in carrying out these procedures, and presents data that provide indirect evidence of the levels of coverage error associated with the sample selection procedures. Section 11.2 describes the changes in sample design for the 1996 Panel.

Sample selection for SIPP has three stages: the selection of primary sampling units (PSUs), the selection of address units in sample PSUs, and the determination of persons and households to be included in the sample for the initial and subsequent waves of each panel. The first two stages are common to all household surveys, whether cross-sectional or longitudinal, that use multistage sample designs; the third stage is an additional requirement for panel surveys. Each stage requires the definition of a survey or target population. Each stage also requires the development and maintenance of one or more sampling frames, that is, a complete list of defined areas, addresses, or other units from which to select a sample.

The Census Bureau uses an integrated sample design, based primarily on the decennial Census of Population, to provide samples for all of its major household surveys, including the Current Population Survey (CPS), the National Crime and Victimization Survey (NCVS), the American Housing Survey, and SIPP. The sample PSUs selected for the different surveys are often the same. The sample addresses, however, are typically nonoverlapping. The design and sampling frames are updated after each Census. The old design is replaced by the new one about 5 years after the decennial Census. The 1984 SIPP Panel was based on the 1970 Census, the 1985-1993 Panels were based on the 1980 Census, and the 1996 Panel is based on the 1990 Census.

3.1 Selection of Sample PSUs

The frame for the selection of sample PSUs consists of a listing of U.S. counties and independent cities, along with population counts and other data for these units from the most recent Census of Population. Counties are either grouped with adjacent counties to form PSUs or constitute PSUs by themselves. Following the formation of the PSUs, the smaller ones, which are called non-self-representing (NSR), are then grouped with similar PSUs in the same region (South, Northeast, Midwest, West) to form strata. Census data for a variety of demographic and socioeconomic variables are used to determine the optimum groupings. A sample of NSR PSUs is selected in each stratum. All of the larger PSUs are included in the sample and are therefore called self-representing (SR).

Table 3.1 shows the number of sample PSUs in the 1984-1993 Panels. The sample of PSUs for the 1984 Panel consisted of 45 SR PSUs and 129 NSR PSUs, for a total of 174 PSUs. The NSR PSUs were grouped into 129 strata and one PSU was selected to represent each stratum. For the 1985-1993 Panels, which were based on the 1980 Census, all 86 PSUs with 184,000 or more housing units were treated as SR. The population in SR PSUs constitutes more than half of the 1980 Census population. The number of NSR PSUs for the 1985-1991 Panels was cut back from the planned 198 to 144 because of budget constraints. The planned sample of 198 NSR PSUs was restored for the 1992-1993 Panels. Since 1985, two sample PSUs were selected in most of the NSR strata.

Table 3.1 Characteristics of the SIPP first-stage sample

Type of stratum	Number of sample PSUs in panels		
	1984	1985-1991	1992-1993
Self-representing	45	86	86
Non-self-representing	129	144	198
Total	174	230	284

The introduction of a new set of PSUs for the 1985 and later panels affects the sampling errors of SIPP estimates in two ways. The use of more recent Census data to form the PSUs and strata and to determine selection probabilities almost certainly improved the efficiency of the first-stage sample, that is, reduced the contribution of between PSU variation to total sampling variability. In addition the selection, in most cases, of two PSUs per stratum improves the estimates of sampling error.

3.2 Selection of addresses in sample PSUs

The survey population for SIPP consists of adults (persons ages 15 and older) in the resident noninstitutional population of the United States. To represent the survey population, a sample of addresses is selected from each of the sample PSUs, using five nonoverlapping sampling frames. Three of the frames are based on information from the population Census, covering residential addresses and group quarters that were included in the Census. The other two frames represent address units that did not exist at the time of the Census (new construction) or were found to have been missed in the Census (coverage improvement).

A detailed description of the address sampling frames and selection procedures used for the 1984 Panel is presented in a technical paper by the U.S. Census Bureau (1978). This section gives a brief description of the frame development and sample selection procedures for each of the five frames used for the 1985-1993 Panels. A detailed technical description of sample selection procedures for the address and area frames based on the 1980 Census is given in Shapiro (1983a, b, c).

- *The Address Enumeration Districts (EDs) Frame* is a list of addresses in EDs which were located in permit issuing areas and for which at least 96 percent of the addresses were complete (street name and house number) when the frame was created. These addresses were on the 1980 Census files or were keyed into computer files created prior to sample selection. Addresses listed as special places in the 1980 Census EDs were excluded from this frame, but were included in the Special Place Frame discussed below. In each sample PSU, the addresses in the EDs were divided into clusters, each containing two neighboring housing units. Samples of these clusters were then selected and assigned to specific SIPP panels (Shapiro, 1983c, 1984). EDs are Census defined areas contained within the boundaries of a political unit (city, township, etc.), consisting of approximately 250-350 housing units.

- *The Area Enumeration Districts Frame* consists of all other Census EDs, that is, those in which more than 4 percent of the addresses in the ED were incomplete or which were located in areas where building permits were not issued or were not available. Most of these EDs are in rural areas. Area EDs in sample PSUs were subdivided into "blocks." Based on Census counts of housing units, an expected number of clusters of four housing units was assigned to each block. If a block contained fewer than three expected clusters, it was combined with another block. Sample blocks were then selected with probability proportional to their expected numbers of clusters. The selected blocks were visited several weeks prior

to scheduled interviewing and all residential addresses were listed. The listed addresses were divided into clusters of four housing units each. The required number of clusters were then selected from each sample block and assigned to a specific panel for interview (Shapiro, 1983b, 1984).

- *The Special Places Frame* consists of noninstitutional group quarters, which are a subset of all special places enumerated in the 1980 Census. Group quarters are housing units occupied by nine or more persons unrelated to the head and having shared kitchen facilities or a common entrance. For PSUs included in SIPP, samples of addresses for special places were assigned to each panel.

- *The New Construction Frame* contains addresses of structures for which building permits have been issued since 1979. Exact dates used as the starting point vary for different types of structures (Statt et al., 1981). This frame is updated continuously. In sampling from the New Construction Frame, permit issuing offices are treated like EDs in the first two frames. Based on the numbers of housing units authorized by the permits, clusters (measures) of four housing units are formed, always from the same permit office and with approximately the same date of issuance. For each new panel, a sample of clusters is selected shortly before interviewing is scheduled to start, so that coverage of new construction will be as up-to-date as possible.

- *The Coverage Improvement Frame* is a small frame consisting of addresses missed in address EDs in the 1980 Census. The frame for the 1985 and subsequent SIPP panels was obtained by comparing 1970 Census addresses sampled for the Current Population Survey with the 1980 Census address lists for the same areas (Parmer, 1985). For PSUs included in SIPP, samples of the 1970 addresses not found on the 1980 address lists were assigned to each panel. For the 1984 SIPP Panel the addresses for the coverage improvement frame were taken from canvassing a sample of areas in "list EDs" following the 1970 Census and matching the results against the Census address register.

Table 3.2 shows the proportion of sample addresses in the 1986 Panel that came from each of the five frames. Addresses enumerated in the 1980 Census (the first three frames) account for 91.3 percent of the total and addresses of units that existed on April 1, 1980, but were missed in the Census, account for only 0.1 percent. The proportion of addresses from the New Construction Frame was 8.6 percent in 1986, and has increased annually to provide adequate coverage of new construction.

Table 3.2 Distribution of sample addresses by frame: 1986 Panel

Frame	Percent of addresses
Address enumeration districts	61.7
Area enumeration districts	28.2
Special places	1.4
New construction	8.6
Coverage improvement	0.1

Source: King et al. (1987).

3.3 Evaluation of errors associated with address frames

This section discusses the sources of error associated with the address frames for the 1984 Panel based on the 1970 Census, and the 1985-1993 Panels based on the 1980 Census. Frame development and sample selection within sample PSUs require a complex system of automated and manual office operations. For the Area ED Frame, the listing of addresses in sample blocks is a field operation. All of these operations are subject to error.

One aspect of frame coverage that is specific to SIPP, as a panel survey, is that no additional new construction sample is added once a panel has been introduced. However, as discussed in Section 3.4, most persons moving into units built since the sample was selected have a chance of selection at their previous addresses and would be followed to the new address.

Coverage in the 1984 Panel

A comprehensive account of quality control procedures and information about errors associated with frame development and sample selection based on the 1970 Census is given by Brooks and Bailar (1978, Chapter II). Although their focus was on the Current Population Survey, their findings apply equally to the 1984 SIPP Panel, which was based on the same set of frames and sampling procedures.

Brooks and Bailar identify several potential coverage problems associated with the address frames used:

- Units converted from nonresidential to residential use may be missed unless a permit is issued in connection with the conversion.
- Units constructed without permits in permit-issuing areas may be missed.
- If a permit is issued for a new structure at an existing address, that address may receive a duplicate chance of selection.
- Adequate coverage of mobile homes presents a variety of coverage problems.

The magnitude of these coverage effects is not generally known, but is believed to be small in relation to the universe. Montfort (1988) estimated that undercoverage of mobile homes constructed after 1980 was close to 25 percent in the 1985 American Housing Survey. Because SIPP and the American Housing Survey use essentially the same sampling frames, one would expect undercoverage of these units to be of a similar magnitude in SIPP.

Sampling operations from the 1970 Census address frames were subjected to a variety of quality control procedures and evaluation studies. A careful evaluation of the quality of the Area and Address ED Frames detected errors of less than 1 percent in all of the categories checked (Fasteau, 1973). For the construction of clusters of addresses in the sample address EDs, a program error was detected part way through the process. The error combined addresses, within a block, that had the same street number but a different street name, so that one of these addresses would not have a chance of selection. Since the percentage of housing units affected was very small, the work was not redone in the affected EDs (Shapiro, 1972).

Coverage in the 1980 Census-based panels

The 1985-1993 SIPP Panels were based on frames developed from the 1980 Census. The procedures were quite similar to those used to develop frames from the 1970 Census. One difference, aimed at coverage improvement, was that the percentage of complete addresses required for an ED to be included in the Address ED Frame was increased from 90 to 96 percent. As explained in Section 3.2, the source of the Coverage Improvement Frame differed from that used after the 1970 Census. The coverage issues affecting the 1980 Census are discussed in Fay et al. (1988).

Relatively little direct information is available about the quality of the 1980 Census-based panels of sample addresses selected for SIPP interviewing. In the first wave of the 1985 Panel, interviews were conducted for about 50 housing units that should not have been included in the sample for that panel. In the second wave, the housing units interviewed by mistake were not revisited. There were no adverse effects on the quality of the data, since the units erroneously interviewed did not belong in the sample. For subsequent panels, procedures have been revised to make it easier to detect and eliminate incorrectly included addresses in the first month of interviewing.

Some coverage problems arise when permit-issuing offices change the boundaries of the areas within their jurisdiction or discontinue issuance of permits. If an area in the Area ED Frame is brought under the jurisdiction of an existing permit office, new units in that area will have a duplicate chance of selection, through the area ED frame and the New Construction Frame. Conversely, if an existing permit office stops issuing permits, new units in its areas will have no chance of selection. As of mid-1988, it was estimated that about 120 new housing units per month in such areas (equivalent to 0.08 percent of the total newly-constructed units authorized for the entire country) had no chance of selection (Loudermilk, 1989).

3.4 Association of persons and other units with sample addresses

The units of observation and analysis for SIPP are not addresses; they are persons ages 15 and older and units such as households and families. Therefore, once a sample of addresses has been selected for the first wave of a SIPP panel, rules must be established to determine which persons and units should be included in initial interviews at those sample addresses. During the life of a panel, persons move from one location to another and some persons enter or leave the survey population. A second set of rules, sometimes called "following rules" is needed to determine which persons and units to include in interviews after the first wave. To a considerable extent, the appropriate rules depend on the ways in which the data are to be analyzed.

For practical and economic reasons, the SIPP following rules represent a compromise with what would be done ideally to provide full longitudinal coverage of all conceivable survey populations. In this section, we summarize the effects on coverage of the rules adopted for SIPP. A fuller discussion is given by Kalton and Lepkowski (1985).

Effects of following rules

At the initial interview, all persons whose usual residence is the sample address are included in the panel. Most college students living in dormitories or other group quarters are considered to have their usual residences with their families. At the initial interview, the household is defined as all persons living at the sample address. The issue of defining households longitudinally has been and continues to be the subject of much discussion and research (McMillen and Herriot, 1985; Duncan and Hill, 1985; Citro, Hernandez, and Herriot, 1986; Citro, Hernandez, and Moorman, 1986). The present SIPP following rules permit

analyses using any one of several possible definitions. Under the present following rules, most original sample members (persons ages 15 and older interviewed at sample addresses in the initial wave) continue to be interviewed throughout the life of the panel, even when they move to other addresses. Exceptions include persons who leave the survey population through death, going abroad, or changing residence to an institution or military barracks. Table 3.3 shows the rate of sample loss as a result of these exceptions for Wave 5 of the 1984 Panel and Wave 7 of the 1992 Panel.

Table 3.3 Sample loss from original sample persons who leave the survey population: 1984 and 1992 Panels

Circumstance	Percent sample loss	
	Wave 5 1984 Panel	Wave 7 1992 Panel
Deceased	1.0	1.0
Institutionalized	0.5	0.6
Outside of country	0.5	0.8
In military barracks	0.3	0.2

Starting in May 1985 for the 1984 Panel and in October 1985 for the 1985 Panel, and at the beginning of all other panels, persons leaving the survey population to enter institutions are followed but not interviewed; if they reenter the survey population, they are again interviewed. Persons who live with original sample members after the initial wave (known as additional persons) are also interviewed as long as they reside with original sample members.

Following children

Coverage of persons who enter the survey population during the life of a panel, including those who reach age 15, must also be considered. For the early SIPP panels, children in original sample households were followed as long as they lived in households that contained original sample persons. Children under age 15 at the time of the attempted interview who moved unaccompanied by an original sample person were not followed, so that those who became age 15 during the panel period were not interviewed. Beginning at Wave 4 of the 1993 Panel, attempts were made to follow children and collect control card information for the households in which the children were living.

Other new entrants

Other new entrants, for example, those who have been abroad, in an institution, or living in a military barracks, can be interviewed if they moved to an address that was eligible for interviewing at Wave 1 and that has not been dropped because of noninterview or because all initial sample persons have left. However, at present, insufficient information is obtained from such persons the first time they are interviewed to follow the procedures suggested by Kalton and Lepkowski (1985) which entail treating such persons as panel members in their own right, rather than as additional persons.

Following movers

Original sample persons who move to a location more than 100 miles from any SIPP sample PSU are interviewed only if they can be reached by telephone. Approximately 96.7 percent of the U.S. population lived within 100 miles of the sample PSUs used in the 1984 Panel (Kalton and Lepkowski, 1985). The increase in the number of PSUs for subsequent panels means that this figure is now higher. Since most

movers go to another location in the same county, very few persons are lost as a result of this operational restriction.

Changes between waves

Another concern raised by the longitudinal character of the survey design is treatment of persons who leave the survey population between waves. Ideally, data should be obtained for such persons for that part of the 4-month reference period for which they were still members of the survey population. A similar problem exists for additional persons who move out of households occupied by sample persons. The procedure used for such cases is to find out how long the person remained in the survey population and to impute current wave data for that period based on demographic information collected previously.

A final issue is the coverage of persons who both enter and leave a SIPP household during the reference period for a single wave (Petroni, 1994). The household rostering questions for the initial and subsequent wave interviews do not identify such persons. It has been decided not to correct for this deficiency because the resulting number of omissions is small and because the application of the second-stage ratio adjustment (see Section 8.2.1) should at least partially compensate for them.

3.5 Evaluation of survey coverage

The current sample selection and following rules for SIPP, which are dictated by resource availability and other practical limitations, lead to some loss over time of members of the theoretical survey population. To some extent, such loss in coverage can be compensated by appropriate estimation and imputation procedures. The Census Bureau uses ratio-estimation procedures to adjust the SIPP sampling weights for population undercoverage. Weights are adjusted so that cross-sectional and longitudinal estimates derived from SIPP are aligned with current estimates of the survey population by age, sex, race, and Hispanic origin. The reciprocals of these weighting adjustments are known as "coverage ratios" because they provide an indication of the degree of undercoverage in the survey for various demographic subgroups of the population.

Coverage ratios

Tables 3.4 and 3.5 show the coverage ratios by age, sex, and race for SIPP and CPS estimates for March 1984 and March 1986. Most of the ratios are less than 1.00, some by large amounts. The level of undercoverage in SIPP is comparable to the CPS. Table 3.6 shows the coverage ratios for Wave 1 of the 1990 and 1991 SIPP Panels (Butler, 1993). The Wave 1 survey estimate is the average of the estimates for the 4 interview months (February, March, April, and May) for each of the rotation groups. All three tables show that the overall ratios for black males and females are smaller than the ratios for nonblack males and females. Ratios for blacks by age and sex in SIPP are quite variable because of the relatively small sample sizes in each cell. However, it is clear that the coverage of young black males, whose ages are in the late twenties and early thirties, is especially low. This trend is observed in most SIPP panels.

Table 3.4 Coverage ratios for SIPP and CPS samples: March 1984[1]

Age group	Male				Female			
	CPS		SIPP		CPS		SIPP	
	Black	Non-black	Black	Non-black	Black	Non-black	Black	Non-black
16-17	0.949	0.939	0.965	0.950	0.867	0.967	1.037	0.956
18-19	0.913	0.896	0.930	0.967	0.876	0.909	0.902	0.883
20-21	0.747	0.887	0.886	0.921	0.819	0.914	0.870	0.996
22-24	0.656	0.862	0.643	0.814	0.848	0.885	0.793	0.884
25-29	0.803	0.907	0.742	0.846	0.907	0.928	0.921	0.928
30-34	0.705	0.908	0.870	0.896	0.849	0.950	0.934	0.886
35-39	0.768	0.911	0.729	0.871	0.844	0.947	0.849	0.902
40-44	0.904	0.929	0.877	0.887	0.980	0.938	0.865	0.945
45-49	0.863	0.922	0.758	1.004	0.950	0.968	1.132	0.993
50-54	0.842	0.960	0.936	0.938	0.905	0.979	0.717	0.972
55-59	0.830	0.960	0.986	0.927	0.894	0.948	0.849	0.936
60-61	1.003	0.962	0.927	0.964	0.968	0.913	1.056	0.980
62-64	0.859	0.926	↓	0.935	0.933	0.950	1.036	0.935
65-69	1.099	0.934	0.959	0.940	1.070	0.951	↓	0.984
70-74	0.894	0.929	↓	0.946	1.019	0.946	1.049	0.918
75-79	↓	↓	↓	0.921	↓	↓	↓	0.918
80-84	1.014	0.927	0.973	1.036	0.980	0.938	0.928	0.952
85+	↓	↓	↓	0.783	↓	↓	↓	0.976
All	0.835	0.919	0.846	0.910	0.901	0.941	0.917	0.933

[1]Coverage ratios for other months are similar.

Source: Unpublished Census tabulations: SIPP 1984 Panel Weighting Output for Processing Cycle 3 and CPS Weighting Output for March 1984.

Table 3.5 Coverage ratios for SIPP and CPS samples: March 1986

Age group	Male				Female			
	CPS		SIPP		CPS		SIPP	
	Black	Non-black[1]	Black	Non-black	Black	Non-black[1]	Black	Non-black
16-17	0.922	0.974	0.897	0.971	0.889	0.954	1.021	0.988
18-19	0.744	0.892		0.796	0.862	0.881		1.134
20-21	0.798	0.869		0.742	0.781	0.880	1.001	0.873
22-24	0.723	0.889		0.926	0.791	0.914		0.981
25-29	0.728	0.895	0.579	0.787	0.866	0.939	0.696	0.879
30-34	0.756	0.909		0.827	0.904	0.945		0.950
35-39	0.850	0.939	0.867	0.963	0.856	0.978	0.834	0.997
40-44	0.861	0.935		0.938	0.951	0.948		1.034
45-49	0.948	0.919		0.983	0.859	0.930		0.970
50-54	0.921	0.946		0.899	0.897	0.940		0.923
55-59	0.800	0.964	0.874	0.990	0.865	0.949	0.897	0.952
60-61	0.710	0.916		1.060	0.953	0.917		1.185
62-64	0.857	0.883		1.108	0.867	0.970		1.069
65-69	0.860	0.937		0.981	0.951	0.966		1.125
70-74	1.007	0.979		1.171	1.001	0.997	1.269	1.036
75-79	0.868	0.960		0.995	1.093	1.003		0.991
80-84								0.850
85+								
All	0.817	0.926	0.797	0.925	0.893	0.949	0.889	0.979

[1]CPS coverage ratios for whites are used here. Spanish persons may be classified as either white or black; however, most probably fall in the white category.

Source: For SIPP, 1985 Panel weighting output for one rotation group from Wave 4 and Wave 5 was used to compute the coverage ratios. The CPS coverage ratios were provided by the CPS Branch in Demographic Surveys Division, U.S. Bureau of the Census. The coverage rates for other months are similar.

Table 3.6 Coverage ratios for SIPP Panels 1990-1991 (Wave 1 average)[1]

Age group	1990 Panel				1991 Panel			
	Male		Female		Male		Female	
	Black	Non-black	Black	Non-black	Black	Non-black	Black	Non-black
15	0.998	0.938	0.889	0.890	0.815	0.897	0.852	0.920
16-17	0.843	0.928	0.881	0.913	1.017	0.978	1.074	0.832
18-19	0.822	0.912	0.766	0.866	0.821	0.833	0.741	0.918
20-21	0.830	0.875	0.873	1.000	0.801	0.899	0.893	0.895
22-24	0.791	0.910	0.775	0.874	0.595	0.822	0.778	0.889
25-29	0.658	0.898	0.814	0.933	0.652	0.932	0.795	0.965
30-34	0.672	0.875	0.854	0.922	0.930	0.901	0.922	0.891
35-39	0.805	0.914	0.784	0.908	0.785	0.903	0.861	0.871
40-44	0.791	0.903	1.002	0.957	0.966	0.886	1.027	0.985
45-49	0.807	0.931	0.888	0.913	0.867	0.863	1.015	0.897
50-54	0.897	0.937	0.873	0.937	0.754	0.916	0.752	0.944
55-59	0.766	0.840	0.810	0.937	0.836	0.932	0.780	0.935
60-61	0.777	0.922	0.901	0.943	0.569	0.854	0.778	0.931
62-64	↓	0.981	↓	0.972	↓	0.956	↓	0.848
65-69	0.770	0.917	0.851	0.974	0.878	0.904	0.988	0.906
70-74	1.006	0.975	0.944	0.986	0.889	0.913	0.853	0.954
75-79	|	1.039	↓	0.979	|	0.914	1.067	0.928
80-84	|	0.942	0.991	0.970	|	0.900	|	1.000
85+	↓	0.947	↓	1.049	↓	0.737	↓	0.958
Average 15+	0.816	0.925	0.868	0.943	0.812	0.892	0.886	0.920

[1]Average of the interview month (February, March, April, and May) for the four rotation groups.

Source: Butler (1993).

The coverage ratios for SIPP and CPS in Tables 3.4, 3.5, and 3.6 understate total undercoverage in these surveys because the benchmark estimates, which are projections from the 1980 Census, do not include any adjustment for the undercoverage in the Census. There have been no analyses of how SIPP estimates might be affected by introducing an additional adjustment factor to account for Census undercoverage. A comparable analysis for CPS estimates prior to 1980 indicated that use of extended adjustment factors would lead to decreases in estimates of the proportion of persons in the labor force who were employed (Hirschberg et al., 1977). Undercoverage in the 1980 Census is believed to have been significantly smaller than it was for the 1970 Census (Fay et al., 1988). Therefore, for SIPP panels based on the 1980 Census, the effects of adjusting survey estimates for Census undercoverage would be smaller.

Ratio adjustments, regardless of what benchmarks are used, do not completely eliminate errors resulting from undercoverage. To the extent that persons in a population subgroup who are missed in the survey differ from those covered, the estimates for that subgroup will be biased. In general, the absolute size of the bias will be positively correlated with the proportion of the group not covered in the survey. (For a more detailed discussion of the effects of undercoverage on the quality of survey data in general see Shapiro and Kostanich, 1988.)

Reasons for and magnitude of undercoverage

Several studies have examined the reasons for and magnitude of undercoverage in household surveys (Fay, 1989a; Shapiro, 1992; Shapiro et al., 1993). Shapiro et al. (1993) examined the extent of survey undercoverage due to omissions of households, errors within households, and possible errors in the Census-based control totals. While this study was conducted on the CPS, the findings also apply to the SIPP and other household surveys conducted by the Census Bureau that use the Census listings as a sampling frame.

Shapiro et al. (1993) compared the March 1980 CPS estimates with the 1980 Census counts to estimate the proportion of undercoverage due to whole-household and within-household omissions. The results show that survey undercoverage occurs mostly because of omissions of persons within interviewed households rather than omissions of whole housing units. The survey coverage of young black males is about 14.5 percentage points below the decennial Census, and within-household omission accounts for 70 percent of the undercoverage. The survey coverage for the white population is about 6.8 percentage points below the decennial Census, with within-household omission accounting for 60 percent of the undercoverage.

Shapiro et al. (1993) note that Census figures are subject to errors of overcoverage (erroneous enumeration) as well as undercoverage. Some Census erroneous enumerations lead to two chances of selection for the housing units involved in surveys and thus are equally reflected in the numerator and the denominator of the coverage ratios. However, some Census erroneous enumerations are correctly treated in surveys and are overrepresented in only the denominator of the coverage ratios. Therefore, using Census-based estimates for population controls as a means to evaluate survey undercoverage has some limitations. Shapiro et al. used the results from the 1990 Post-Enumeration Survey (PES) to adjust the Census-based control totals. The resulting coverage ratios[1] are higher (closer to 1), reducing the apparent survey undercoverage by about 8 percentage points for young black males in the 18-29 age group, and by lesser amounts for other sex-race groups.

[1] At the time the coverage ratios did NOT include the effect of Census undercoverage. Inclusion of Census undercoverage would result in significantly lower coverage ratios.

Even taking account of erroneous enumerations, the general conclusion from this study is that substantial survey undercoverage exists. Survey undercoverage occurs because of omissions of whole housing units and omissions of persons within interviewed households, with the latter accounting for a large proportion of the undercoverage.

Roster research

The Census Bureau has sponsored several research studies to examine the roster questions and the rostering procedures used to enumerate members of sample households in surveys (Cantor and Edwards, 1992; Kearney et al., 1993; Shapiro et al., 1993; Tourangeau et al., 1993). These studies evaluated the effects of two types of within-household omissions. One type of omission occurs because the responding households deliberately conceal some persons (for example, for fear of deportation, or losing welfare benefits). Another type of omission results from difficulties in applying the residence rules to the living situation of persons in households with complex family structures. The studies compared the enumeration results based on the SIPP rostering questions and procedure against alternatives aimed at increasing the apparent confidentiality of the interviews and providing a clearer standard of residency. The results suggest that anonymous rostering shows promising improvements, while clarifying the living situation has only marginal gains, because relatively few persons have unconventional living arrangements.

The Living Situation Survey (LSS) is another research project conducted by the Census Bureau to assess the rostering techniques used in the decennial Census (Schwede, 1993). The LSS, conducted in 1993, was a one-time national survey that oversampled subpopulations at high risk of undercoverage in the Census. It was designed to test the efficiency of a battery of roster probes to identify all persons with some association to the sample housing unit within a specified reference period. Sweet (1994) analyzed the results of this survey and suggested that the Census Bureau consider new rostering procedures to improve coverage. Martin (1996) suggested the possibility of measuring residency by degrees rather than by an all-or-none selection.

4. DATA COLLECTION PROCEDURES: 1984-1993 PANELS

This chapter describes the data collection procedures for the 1984-1993 Panels, highlighting those features most likely to influence the quality of the data, for example, interview mode, respondent rules, the structure of the interview and the questionnaire, interviewer characteristics and training, and quality assurance procedures. Section 11.4 describes the changes in the procedures for the 1996 Panel due to the introduction of computer-assisted interviewing.

References that provide useful information about data collection strategies and procedures, as well as changes since the start of interviewing in 1983, include Kalton et al. (1986), Nelson et al. (1985), and Kasprzyk (1988). Persons interested in more detail may also wish to consult the *SIPP-5010 Interviewers' Manual* (U. S. Bureau of Census, 1988), which is periodically updated, and may be obtained by writing or calling Customer Services, Data User Services Division, U. S. Bureau of the Census, Washington, DC 20233, (301) 457-4100.

4.1 Basic data collection features

Interviewing every 4 months

Interviews of sample persons and households take place at 4-month intervals. Information on core items, such as labor force participation and income recipiency and amounts, is collected for each of the 4 months included in the reference period. Interviews are conducted in the month immediately following the 4-month reference period and the majority of interviews are completed in the first 2 weeks of that month.

Mode of interviewing

For SIPP panels between 1984 and 1991, face-to-face interviewing was the preferred mode of data collection and was used in the great majority of cases. Telephone interviewing was permitted to follow-up for information not obtained at a face-to-face interview, to interview persons who would not or could not participate otherwise, and to interview sample persons who had moved to locations more than 100 miles from a SIPP sample PSU. For the 1984 Panel, about 5.3 percent of the interviews were conducted by telephone because the households were inaccessible for face-to-face interviews. The corresponding rate for the 1985 Panel was 6.3 percent. For the most part, the proportion of interviews conducted by telephone increased from the second through the final wave of each panel.

The mode of data collection was switched to maximum telephone interviewing in 1992 commencing with Wave 7 of the 1990 Panel and Wave 4 of the 1991 Panel. For the 1992 and 1993 Panels, Waves 1, 2, and 6 were conducted primarily using face-to-face interviewing as before. The remaining waves were conducted through telephone interviewing, to the extent feasible. Section 4.5.2 discusses the studies that monitored effects of maximum telephone interviewing on data quality.

Interviewing time

The length of a household interview depends in part on the number of adults (persons ages 15 and older) in the household. Table 4.1a shows the median length of interview (in minutes) by wave and number of household members ages 15 and older for the 1985 Panel. The median for all households ranged from a low of 27 minutes to a high of 40 minutes for the eight waves. The high occurred in the first wave and the low in the second wave, which was the only wave after the first which had no topical modules. For the 1993 Panel (see Table 4.1b), the high again occurred in the first wave and the interview duration for the subsequent waves varied, depending on the topical modules included in each wave.

Table 4.1a Median household interview duration for face-to-face interviews by wave and number of persons ages 15 and older in the household, 1985 Panel

Number of persons ages 15 and older	Interview Duration (minutes) for Wave:							
	1	2	3	4	5	6	7	8
1	24	18	21	21	23	20	21	22
2	41	27	38	38	41	34	39	40
3	52	39	51	51	54	45	50	52
4	64	51	61	62	68	57	59	63
5	71	56	71	71	73	66	67	71
6	92	67	87	76	83	78	88	81
7 or more	116	90	99	98	96	101	98	99
All households	40	27	37	37	39	33	37	38

Source: Table prepared by A. Feldman-Harkins, Division of Housing and Household Economic Statistics, U.S. Bureau of Census.

Table 4.1b Median household interview duration for face-to-face and telephone interviews by wave and number of persons ages 15 and older in the household, 1993 Panel

Number of persons ages 15 and older	Interview Duration (minutes) for Wave:								
	1	2	3	4	5	6	7	8	9
1	27	20	20	19	17	22	19	17	20
2	45	35	35	32	30	40	33	30	37
3	59	49	50	45	42	55	45	42	50
4	73	60	65	55	53	68	53	50	61
5	90	77	77	71	68	89	73	68	77
6	94	96	103	82	78	104	78	71	92
7 or more	125	120	109	85	90	113	93	84	103
All households	40	32	32	27	21	35	30	27	32
Mode of interview[1]	Per	Per	Tel	Tel	Tel	Per	Tel	Tel	Tel

[1]Per = face-to-face interview, Tel = telephone interviews.

Self- and proxy respondents

Response to the survey is voluntary. The respondent rule for SIPP is that all persons ages 15 and older who are present at the time of the interview should report for themselves unless not physically or mentally able to do so. Proxy informants are accepted for persons absent or incapable of responding. Exhibit 4.1 shows the hierarchy of preferred proxy respondents established for SIPP. The procedure for identifying a proxy with telephone interviews is identical to that used for personal visits.

Exhibit 4.1 SIPP Respondent Rules

CARD D

RESPONDENT RULES

HOUSEHOLD RESPONDENT

Any household member 15 years old or older who is physically and mentally competent and knowledgeable may answer the control card questions and questions about the household as a unit.

HOUSEHOLD MEMBER 15 YEARS OF AGE OR OLDER

Each household member 15 years old or older, present at the time of interview, should respond for himself/herself. If a 15+ person is physically or mentally incompetent, select a proxy respondent. Also select a proxy respondent for a person absent at the time of interview. Any **knowledgeable** household member who is 15 years old or older may serve as proxy. Following is a chart for your use in determining who to interview. The choices are listed in order of priority.

INTERVIEW RESPONDENT PRIORITY RULES	
WAVE	FIRST FAMILY MEMBER INTERVIEW OR RETURN VISIT INTERVIEW
1	1. Self 2. Spouse (if any) 3. Other proxy
2	1. Self 2. Spouse (if any) 3. Proxy last visit 4. New proxy
3-8	1. Self 2. Spouse (if any) 3. Proxy last visit 4. Proxy at another visit 5. First time proxy

If a person wishes to act as a proxy but is not a household member, you **must** call your supervisor for permission before interviewing the proxy. Such cases should have an INTERCOMM attached to the questionnaire.

FORM SIPP-4004 (5-2-83)

During the first wave of the 1985 Panel, 67 percent of the sample persons for whom data were obtained reported for themselves. In subsequent waves this figure was quite stable at the level of 63-64 percent (Kasprzyk and McMillen, 1987). Similar results have been observed for other panels. For persons completing all interviews in the 1984 Panel, Table 4.2 shows that about 40 percent reported for themselves in every wave (that is, no proxy interviews). For persons completing the first seven waves of the 1992 Panel, the table shows that about 36 percent reported for themselves in all of those waves. The distributions of sample persons by patterns of self- and proxy response for the first three waves of the 1984 and 1992 SIPP Panels are shown in Table 4.3.

Table 4.2 Number of proxy interviews for persons ages[1] 15 and older completing all interviews in the 1984 Panel and completing the first seven interviews in the 1992 Panel

	1984 Panel		1992 Panel	
Number of proxy interviews	Number of persons	Percent	Number of persons	Percent
0	9,822	40.0	11,223	35.9
1	2,764	11.2	3,724	11.9
2	1,887	7.7	2,922	9.3
3	1,492	6.1	2,131	6.8
4	1,465	6.0	2,487	8.0
5	1,378	5.6	2,291	7.3
6	1,468	6.0	2,804	9.0
7	1,722	7.0	3,684	11.8
8	2,570	10.5	--	--
Total	24,568	100.0	31,266	100.0

[1]Age determined as of first interview.

Source: Data for the 1984 Panel were from Kasprzyk and McMillen (1987).

Table 4.3 Self (S) and proxy (P) response patterns for Waves 1 to 3: 1984 and 1992 Panels

Wave			1984	1992
1	2	3	Percent	Percent
S	S	S	49.4	50.5
S	S	P	5.9	7.7
S	P	S	6.0	4.6
S	P	P	6.2	7.2
P	S	S	5.8	4.5
P	S	P	3.5	3.6
P	P	S	5.1	3.1
P	P	P	18.1	18.8
All patterns			100.0	100.0

Source: Data for the 1984 Panel were from Kalton et al. (1986).

Following sample persons

The SIPP "following rules," which determine which persons should be interviewed after the initial wave, are considered to be a part of the sample selection procedure and are described in Section 3.4. The preferred procedure is for the same interviewer to interview sample persons and households at each wave, but because of interviewer turnover and moves of sample persons to other locations, this is not always possible.

4.2 Data collection instruments

This section describes the SIPP data collection instruments for the 1984-1993 Panels and how they were used in the interviews for successive waves of a panel. Copies of the data collection instruments for each completed panel are included in a SIPP Informational Booklet, which may be obtained by writing or calling Customer Services, Data User Services Division, Bureau of the Census, Washington, DC 20233, (301) 457-4100. The primary data collection instruments were the control card and the questionnaire. The control card was the basic field record for the sample unit and it remained in the regional office during the life of the panel. Questionnaires for each wave were sent to a central point for processing.

SIPP control card

For each sample address, a control card was completed at the first (Wave 1) interview and updated at each subsequent wave. At the start of the Wave 1 interview, the interviewer used the control card to record basic demographic characteristics for each person residing at the sample address, plus a few characteristics of the household and the housing unit. The telephone number of the household was recorded both for use in telephone interviewing in subsequent waves (introduced in February 1992) and for callbacks needed to obtain information not available at the time of the interview. To assist in following movers, each household was asked for the name and telephone number of a nonhousehold member who would be likely to know the new location of households or persons who had moved.

At the close of the first-wave interview, the interviewer transcribed a few key data items for each adult to the control card, such as names of employers, sources of income received, and assets, to aid in collecting data for these items in subsequent interviews. Social Security numbers were recorded for household members who had them, for possible use in connection with estimation and research projects requiring linkage of data from SIPP interviews with data from administrative records. A separate section of the control card was used to record information about household noninterviews when they occurred.

In interviews for the second and later waves, the interviewer began by updating the control card to identify changes in the household roster and in the demographic characteristics of persons, for example, a change in marital status. Selected data items from the control card were transcribed to the current-wave questionnaire to help the interviewer determine which specific questions needed to be asked for each person. At the close of each interview, except in the final wave, changes in the data items used for tracking employment, income, and assets were recorded on the control card.

SIPP questionnaire

A different version of the questionnaire was used for each wave and at each wave a separate questionnaire was completed for each adult. Core items, mainly about labor force participation and sources and amounts of income, were included on the questionnaire for each wave. Except in Wave 1 for all panels and Wave 2 for the 1984 and 1985 Panels, each questionnaire included one or more fixed and/or variable topical modules.

The interview proceeded on a person-by-person basis until a questionnaire had been completed for each adult in the household. The core items for a person were divided into four sections:

- Section 1, *Labor Force and Recipiency*, had two main goals—to obtain labor force status on a weekly basis for the 4-month reference period and to identify all sources of income during the 4-month period, including earnings from wages and self-employment, program benefits, and income-producing assets, such as stocks and savings accounts.

- Section 2, *Earnings and Employment*, obtained details, including monthly earnings, for jobs and separate types of self-employment reported by or for the person in Section 1. Data were obtained for employment with up to two employers and for self-employment in up to two businesses.

- Section 3, *Amounts*, obtained the amounts of income received from all sources other than earnings. In Part A, General Amounts, monthly income amounts were obtained for program benefits and pensions. A standard set of questions was completed for each income source in this category. Parts B to F covered all other income sources, such as interest, dividends, rents, and royalties. For each of these sources, only the total amount for the 4-month reference period was obtained.

- Section 4, *Program Questions*, covered participation in subsidized housing, school lunch, and school breakfast programs.

Section 4 was followed by the topical modules. After they were completed, the interview proceeded to the next adult in the household. Exhibit 4.2 shows the main steps in completing or updating the control card and completing the questionnaires for a single interview at a sample address.

Other SIPP instruments

Other instruments used in the data collection process were advance letters and respondent aids. Before each wave of interviewing, an advance letter signed by the Director of the Census Bureau was sent to each sample address. The letter informed the residents that an interviewer would be calling in a few days, explained the purpose of the survey, gave the authority for conducting the survey, and strongly encouraged the recipient(s) to participate. Exhibit 4.3 shows the version of the letter used for the 1993 Panel.

Respondent aids, sometimes called flash cards, played an important role in the interview. An important respondent aid was the Calendar of Reference Months, which was frequently kept in sight during the entire interview to clarify the specific months and weeks for which respondents were being asked to provide information. Other aids were lists of the categories used to respond to questions on topics such as race, origin or descent, types of income, and types of assets. These cards were shown to respondents at the appropriate stages in the interview. There were also aids for use primarily by interviewers, such as an Age Verification Chart, which the interviewer could use to determine whether reported ages and dates of birth were consistent. Checklists led interviewers through the correct work routines both before the interview was conducted and after its completion. Respondent aids were left with the respondents after the initial face-to-face interviews. If a respondent could not find the aids at the time of a telephone interview, the interviewer would read the pertinent information to the respondent.

Exhibit 4.2. The SIPP Interview

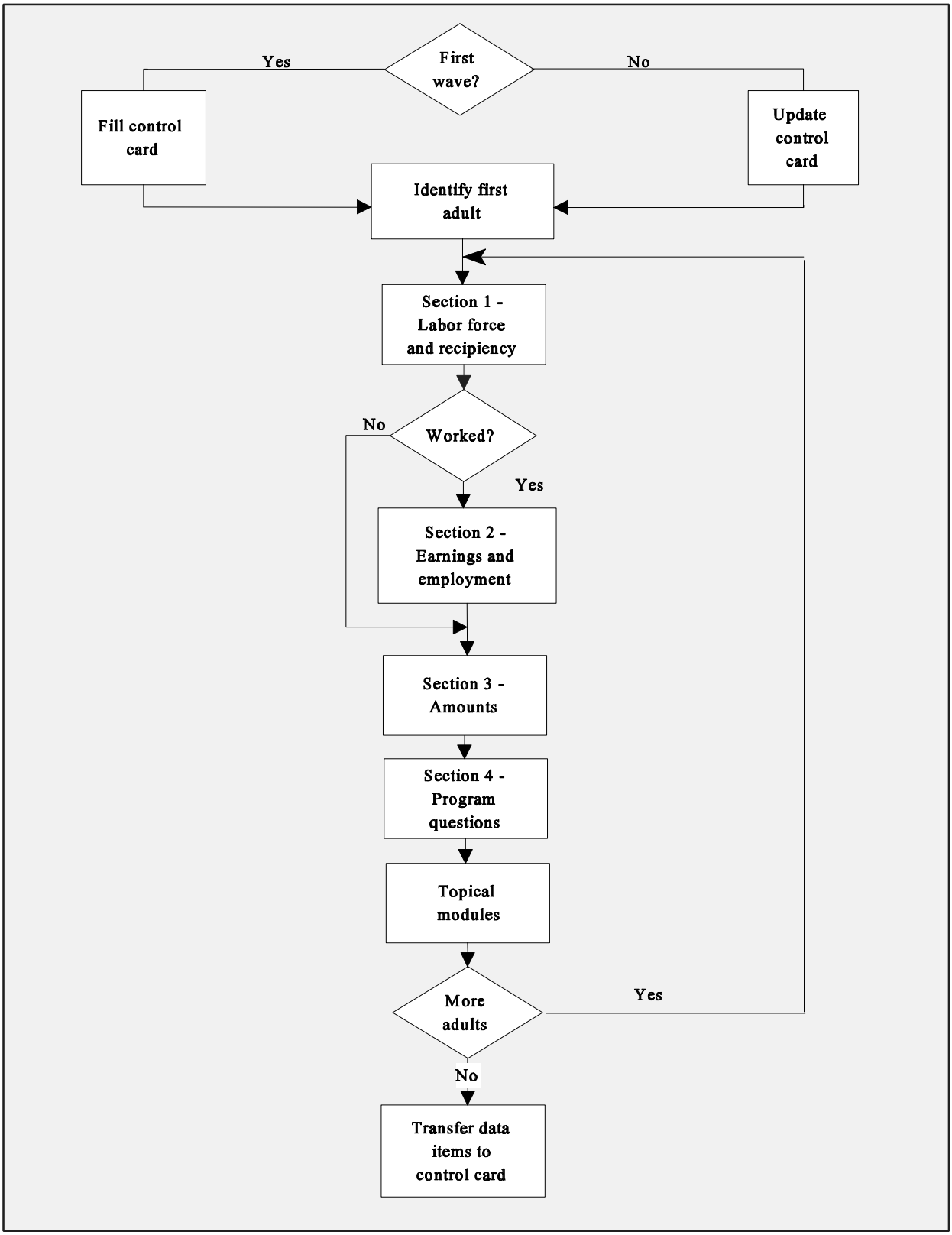


Exhibit 4.3 Advance Letter to Respondents: 1993 Panel

SIPP-13105 (L)
(10-88)

UNITED STATES DEPARTMENT OF COMMERCE
Bureau of the Census
Washington, D.C. 20233

OFFICE OF THE DIRECTOR

FROM THE DIRECTOR
BUREAU OF THE CENSUS

A Bureau of the Census field representative, who will show you an official identification card, will call on you within the next few days. This field representative is conducting a survey on the economic situation of persons and families in the United States. This survey will collect information on topics such as jobs, earnings, and participation in various government programs. We will use information obtained in this survey, together with data from government agencies and other sources, to learn more about how people are meeting their everyday expenses and how government programs are working.

This is a sample survey in which the Census Bureau selected your household to represent other households in the United States. We are conducting the survey under the authority of Title 13, United States Code, Section 182. Section 9 of Title 13 requires that we hold all information about you and your household in strictest confidence. By law, we may use this information only for statistical purposes and only in ways that no information about you as an individual can be identified.

Your participation is extremely important to ensure the completeness and accuracy of the survey data. Although there are no penalties for failure to answer any question in this voluntary survey, each unanswered question substantially lessens the accuracy of the final results.

On the reverse of this letter are the answers to questions most frequently asked about this survey.

Thank you for your cooperation. The Census Bureau is grateful to you for your help.

Sincerely,

BARBARA EVERITT BRYANT

Further information is available from:

Regional Director
Bureau of the Census
Jacob K. Javits Federal Building, Room 37-130
26 Federal Plaza
New York, NY 10278-0044

Exhibit 4.3 (continued) Reverse Side of Advance Letter

WHAT IS THE PURPOSE OF THIS SURVEY?

We are taking this survey to get a picture of the economic situation of people throughout the Nation. We are interested in the situation of people from all walks of life and from every part of the country.

To understand the economic situation of persons and families, we need information about jobs, income, unemployment, disabilities and so forth. We will combine your answers with the information we get from other households throughout the country. Then we will use the information to find out how the economy is affecting the young, the old, the workers, the retired, the disabled, the disadvantaged, and others. We also will use the data to find out if the government assistance programs are reaching the people who need help the most.

WHY DOES THE CENSUS BUREAU NEED THIS INFORMATION?

In a Nation as large and rapidly changing as ours, we need up-to-date facts to plan effective programs for the future. Although we take a complete Census (like the one we conducted in April 1990) every 10 years, we need to collect some kinds of information much more often. A Census Bureau field representative will contact you once every 4 months for the Survey of Income and Program Participation. Thus, we can keep abreast of changes people have in their jobs, the kind of work they do, the number of people looking for work, the situation of people on government programs, and other information.

WHY DOES THE CENSUS BUREAU WANT TO KNOW MY SOCIAL SECURITY NUMBER?

We would like to know your Social Security number so we can obtain information that you have provided to other government agencies. This will help us avoid asking questions for which information is already available and will help ensure the accuracy and completeness of the survey results. We will protect administrative records information that we obtain from these agencies from unauthorized use just as the survey responses are protected. Providing your Social Security number is voluntary. We collect all data in this survey under Section 182 of Title 13, United States Code, which gives us the authority to conduct surveys to produce demographic and economic data.

HOW WAS I SELECTED?

We selected households from a list of all the residential addresses in the Nation. We did not choose you personally but whoever was living at the address when our field representative arrived.

Because this is a scientific sample, we must interview all persons living at each address selected for the survey. We cannot substitute the household next door or down the street.

WHY ARE MY ANSWERS IMPORTANT

People in government and private organizations need the statistics from this survey to develop and evaluate policies and programs that meet the needs of Americans today. For this reason, this information must be as accurate and complete as possible. The only way we can get this information is through the cooperation of sample households such as yours. Your answers represent approximately 5,000 households. In all, we interview about 13,500 households each month across the Nation.

INFORMATION ABOUT THIS INTERVIEW

For each person 15 years old or older, we expect the interview to take about 30 minutes. Your interview may be somewhat shorter or longer than this depending on your circumstances. If you have any comments about this survey or have recommendations for reducing its length, send them to the Associate Director for Management Services, Paperwork Reduction Project 0607-____, Room 2027, FB 3, Bureau of the Census, Washington, DC 20233-0001; and to the Office of Management and Budget, Paperwork Reduction Project 0607-____, Washington, D.C. 20503.

WHAT GUARANTEE DO I HAVE THAT THE INFORMATION I GIVE TO THE CENSUS BUREAU ABOUT MY PERSONAL BUSINESS IS NOT REPORTED TO OTHER PERSONS OR ORGANIZATIONS?

All the information you give to the Census Bureau for this survey is confidential by law (Section 9 of Title 13, United States Code). Every Census Bureau employee takes an oath and is subject to a jail term, a fine, or both if he or she discloses any information that would identify an individual. We will release information only for statistical purposes, and we will never release identifiable information.

4.3 Data collection strategies

The quality of SIPP data is affected by certain broad strategies that were adopted in designing the data collection instrument and the interview process. These strategies are identified in this section. Some could apply to any survey; others apply only to panel surveys. How these strategies affect quality and some results of experiments to test alternative strategies are discussed in Chapters 5 and 6.

One data collection strategy entails *redundancy*, that is, collecting more than the minimum information needed to meet the survey's data requirements. The redundant information is then used either to check internal consistency or impute missing information. The most obvious example of this in the SIPP was the Annual Roundup, a topical module used in Waves 5 and 8 (Waves 6 and 9 in the 1984 Panel) to collect calendar year information on certain kinds of income that had already been covered on a month-by-month basis in prior interviews. The Annual Roundup was timed so that most respondents would already have completed their tax returns for the prior year. For some types of income, the availability of such records might lead to more accurate reporting.

Another important factor in data collection is *sequencing*. As previously discussed, all types of income were identified for each person before amounts and other details about each type were collected. An alternative would have been to collect detailed information on each type of income recipiency before proceeding to other types. Similarly, the decision to collect all data for each person before proceeding to the next person is another aspect of sequencing. An alternative would be to cover each topic for the entire household before proceeding to the next topic.

Use of records, such as pay vouchers, tax returns, and bank statements, during the interview may help respondents provide more accurate information. In particular, one would expect tax returns and wage and tax statements (Form W-2) to be especially helpful in completing the Annual Roundup modules. Field representatives were instructed to encourage the use of available records showing amounts received whenever possible.

An important strategy question that applies only to panel surveys is whether each interview should be *independent* or *dependent* with respect to preceding interviews. A mixed strategy is followed in SIPP. Questions on sources of income are asked in a dependent mode. Respondents are asked about sources of income reported in the previous interview, and then are asked about new sources of income during the 4 months covered by the current interview. Starting with the 1986 Panel, questions on occupation and industry of employed persons have also been asked in a dependent mode. Respondents are asked whether their activities or duties have changed during the past 8 months. A negative response would eliminate the detailed occupation and industry questions. On the other hand, questions on assets in Wave 7 have been asked without reference to information on assets reported a year earlier, in Wave 4. An exception was made for half of the 1984 Panel. This Asset Feedback Experiment is discussed in Section 6.3.1.

4.4 Field representatives: Characteristics, training, and supervision

Characteristics

The formal titles of Census Bureau interviewers and supervisors are field representatives (FRs) and supervisory field representatives (SFRs), respectively. The FRs locate sample addresses and collect most of the data. The SFRs conduct follow-up work to convert potential refusals, handle assignments that are sometimes partially completed by FRs, or act as team leaders for a group of FRs.

In September 1995, there were 335 FRs and about 50 SFRs working on the 1992 and 1993 Panels. The turnover rate for SIPP FRs in Fiscal Year 1995 was 19 percent. This rate was calculated by dividing the number of FRs trained in Fiscal Year 1995 (64) by the total number of FRs (335). The average monthly assignment size in Fiscal Year 1995 was about 14 households. FRs had the entire month to complete the work, although they were encouraged to finish about 90 percent during the first 2 weeks; the last 2 weeks were spent following up movers and potential refusal cases. Both FRs and SFRs were paid on an hourly basis. In September 1995, the starting salary ranged from $7.55 to $8.48 per hour, excluding any special increases for high cost-of-living geographic areas.

Training

Training for SIPP FRs included home study, classroom training, on-the-job training, and refresher training. Training was continuous and included:

- *Initial training*. FRs new to SIPP received intensive training, including 1 day of advance self-study and 3 days of classroom training. For the first wave of a panel, there was an additional half day of training on listing operations. Training sessions included lectures, audiovisual presentations, a mock interview exercise, and discussions. Trainees received detailed information on their jobs, the concepts and definitions used in the survey, and specific interviewing techniques, such as probing. As part of the initial training, a supervisor or senior field representative (SFR) observed each new FR during his or her first 2 or 3 days of interviewing.

- *Refresher training.* About twice a year, FRs received training on new topical modules and on special aspects of the survey, such as conversion of noninterviews, movers' rules, editing, and transcription.

- *Supplemental training*. FRs found to be weak in certain aspects of the survey, such as completion rates and accuracy, were given supplemental training to address and correct such weaknesses.

Supervision and quality assurance

The work of FRs was monitored and feedback provided in several ways.

- *Questionnaire checks*. Completed questionnaires were sent to the Census Bureau's regional offices, where they underwent clerical edits and simple computer edits that were incorporated into the data-entry programs. More complex edits were performed on the computerized records when they were received at Census headquarters. In some instances, FRs might be contacted to resolve problems identified in these edits.

- *Performance observation.* The work of each experienced FR was observed once a year by a SFR who checked the FR's performance in establishing rapport with respondents, asking questions in an appropriate manner, probing, and recording answers accurately. The results of the observations were discussed with the FRs. FRs whose performance was below standard in some respect had their work observed more often, as needed.

- *Reinterviews.* A systematic reinterview program served the dual purposes of checking a sample of the work of individual FR's and identifying aspects of the field procedures which may have needed improvement. The reinterview sample covered 5 to 8 of the sample addresses for one-sixth of the SIPP interviewers each month. The reinterviews, which were completed as soon as possible after the original interview, were conducted on the telephone by SFRs or other members of the supervisory staff. The reinterviews were used to determine whether the FRs visited the correct units, classified noninterviews correctly, and determined household composition correctly. Several questionnaire items were checked to verify that the FR asked these items during the original interview. The results of the reinterviews

were used to take corrective action, such as supplemental training and observation for FRs whose work was below standard.

- *Performance standards*. Specific performance standard guidelines have been established for the SIPP. These guidelines were conveyed to the regional offices as examples of how they might measure performance. However, the standards were only guidelines and the regional offices were empowered to change the standards as they deemed appropriate. If they chose to change the standards, they were required to apply the standards equally across all SIPP FRs for the same period of time. Performance standard guidelines have been established for response rates (proportion of eligible households interviewed) and productivity (average time per household, including travel), as follows (U.S. Bureau of the Census, 1988):

FR Response Rates:

Adjective rating	**Response rate**
Outstanding	100.0 - 97.5
Commendable	97.4 - 95.5
Fully successful	95.4 - 91.5
Marginal	91.4 - 88.0
Unsatisfactory	87.9 and under

Productivity:

Adjective rating	**Hours per household**
Outstanding	0.0 - 3.0
Commendable	3.1 - 3.5
Fully successful	3.6 - 4.0
Marginal	4.1 - 4.5
Unsatisfactory	4.6 and over

FRs received annual ratings based on these standards. SFRs could take account of extenuating circumstances when they assigned adjective ratings.

There was a concern that the method for calculating FR response rates was not sensitive to increases in the number of noninterviews for unlocated movers (Type D noninterview). A Type D noninterview reduced the eligible workload by 1 and the response rate was computed based on the reduced workload (in the denominator). With this method, increases in the number of Type D noninterviews have little effect on the response rate, and therefore offer little incentive for FRs to find unlocated movers. The Census Bureau has been reviewing alternative methods to calculate FR response rates and to provide incentives for locating movers (see Sections 4.5 and 11.4).

4.5 Evaluations of data collection procedures

This section discusses response errors that are associated with the SIPP data collection procedures. The level of response error is affected by virtually all design features including: length of reference period, interview mode, respondent rules, errors associated with interviewers, questionnaire length and structure, and use of records by respondents. Evaluation studies of each of these design features and improvements to the data collection procedures are discussed.

4.5.1 Length of reference period

SIPP interviews are conducted at 4-month intervals. In the core portion of the SIPP interview, respondents are asked to report on labor force activity, income recipiency, and amounts of income for the 4 months immediately preceding the month of the survey interview. For most items, they are asked to report separate data for each of the months included in the 4-month reference period. Although respondents sometimes refer to records when answering these questions, they often rely only on memory.

Memory error in surveys has been the subject of many evaluation studies and experiments (for example, Neter and Waksberg, 1965). Experience suggests that memory error can be reduced by the use of relatively short reference periods and, in panel surveys, by using some form of dependent interviewing such as bounded recall, where interviewers at each round of a survey have access to information reported by respondents on previous rounds.

3-month versus 6-month

The frequency of interviews and length of reference periods were key issues in the design of SIPP and were addressed in two studies conducted as part of the ISDP. In the first of these studies (Olson, 1980; Kasprzyk, 1988), a single interview using a 6-month recall period was compared with two consecutive interviews, both using 3-month reference periods. The same 6 calendar months were covered in both sets of interviews. Not surprisingly, the proportion of respondents who reported some income from sources such as wages, Aid to Families with Dependent Children (AFDC), and unemployment compensation for the first 3 months of this period was greater for the interviews that used a 3-month reference period.

Hill (1986) used an additive error model to analyze the joint effects of length of recall and time in sample (number of waves) on earnings data for the first 9 months of the 1984 Panel. The analysis was restricted to men ages 25 to 55 who reported for themselves in all three interviews, had at least 2 months with some employment, and had no self-employment and no imputed data - a total of 1,378 sample persons. He found evidence of significant differential reporting bias and response variance associated with both length of recall and time in sample.

4-month versus 6-month

The choice between 4-month and 6-month reference periods was considered for the 1996 SIPP redesign. Since a 6-month recall period involves only two interviews per year, SIPP could afford to increase sample size and reduce sampling error using a 6-month recall. However, a longer recall may reduce reports for small sources of income and incomes received for short durations, and exacerbate the seam effect observed at the transition between waves of data collection (see Section 6.1 for a discussion of the seam problem). Huang (1993) examined this choice. His report presents a literature review and the results of simulation studies of the effects of length of recall period on sampling error and a bias of estimates of poverty and program participation. The study concluded that the increase in the precision of estimates cannot compensate for the increase in bias. The recommendation was therefore to maintain the 4-month reference period. Other findings on recall bias from record-check studies are discussed in Chapter 6.

4.5.2 Interview mode

In the early SIPP panels, face-to-face interviewing was the preferred SIPP interview mode. Telephone interviews were allowed when a sample person had moved more than 100 miles from a SIPP primary sampling unit or when the interview could not otherwise be completed. However, in a continuing effort to improve the efficiency of the SIPP design, several tests of a mode that maximized telephone interviewing (that is, conducting interviews by telephone whenever possible) were conducted in the period 1985-87 to examine how the quality of results and the cost of data collection might be affected by shifting to greater

use of telephone interviews. The results of these tests were encouraging and a change to maximize telephone interviewing was introduced in 1992. Since that time, the effects of the change have been monitored.

Testing telephone interviewing

A feasibility test of maximum telephone interviewing was conducted in June 1985 with 279 households, in two regional offices. These households had previously been dropped from the 1984 Panel because of budget cuts. Subsequently, an experiment was carried out in two phases as part of regular data collection. From August to November 1986, half of the national sample for the 1986 Panel was designated for maximum telephone interviewing in either the second or third wave of the panel. The allocation to the two modes was by segment and most of the telephone interviews were done by regular SIPP field representatives, calling from their homes. In the second phase, from February to April 1987, some of the households assigned to the telephone mode in the first phase were assigned to that mode again for the Wave 4 interviews. As a result, there were three experimental groups and a control group, as follows:

Group	Proportion of national sample	Maximum telephone mode
Experimental A	One-fourth	Waves 3 and 4
Experimental B	One-eighth	Waves 2 and 4
Experimental C	One-eighth	Wave 2 only
Control	One-half	Not used

The experimental procedures and some results have been discussed by Durant and Gbur (1988), Carmody et al. (1988), Kasprzyk (1988), Petroni et al. (1989), Gbur and Petroni (1989), and Waite (1990). Household and person noninterview rates did not differ significantly for the experimental and control groups; item nonresponse rates were slightly but not significantly higher in the telephone interviews. The proportion of persons whose data were supplied by proxy respondents was higher for the maximum telephone group, 37.8 versus 35.3 percent. In interviewer debriefings, about three-fourths of the interviewers felt that telephone interviewing could be used successfully in SIPP, but others believed it would not work because of the sensitive nature of the questions and because it would be more difficult to establish rapport and trust with the respondent.

Implementing maximum telephone interviewing

This experiment yielded both encouraging results and cost savings, and subsequently led to greater use of telephone interviewing in SIPP. Maximum telephone interviewing began in February 1992 with Wave 7 of the 1990 Panel and Wave 4 of the 1991 Panel. For the 1992 and 1993 Panels, Waves 1, 2, and 6 were conducted primarily using face-to-face interviewing as before, but the remaining waves were conducted using telephone interviewing to the extent feasible. For waves in which telephone interviewing was the preferred mode, FRs were instructed to record the reason for obtaining a personal versus telephone interview (for example, respondent requests, no telephone or unlisted number, reluctant households, elderly persons, people with hearing problems, and language barriers).

Monitoring the effects

Since implementation, the Census Bureau has monitored the effects of telephone interviewing on cost savings and data quality. Analyses of cost showed savings in terms of interviewing hours and miles per household unit (Allen, 1993). Telephone interviewing showed no adverse effects on response rates (King, 1995) or on record usage by respondents during interviews (Huggins, 1994).

King (1995) found that although both Type A (refusal, no one home, etc.) and Type D (unlocated movers) nonresponse rates have risen during the 1990s, there was no consistent evidence that telephone interviewing was the cause of the increases. She compared the nonresponse rates by wave for the 1986-1993 Panels and concluded that the increase in Type A rate could not be attributed to telephone interviewing because most of the Type A nonresponse occurred in the first two waves which continued using face-to-face interviews. The Type D nonresponse rate showed an increase of approximately 12 percentage points in the unlocated mover rate per wave in panels where telephone interviewing was used heavily. However, there was no evidence that telephone interviewing caused the increase.

James (1994) reviewed studies by Ikeda (1993) and Sutterlin (1993) that evaluated SIPP estimates from the Education Financing module. Estimates of educational cost were derived based on data from the telephone interview wave (Wave 5 of the 1991 Panel and Wave 8 of the 1990 Panel) and from a personal visit wave (Wave 5 of the 1990 Panel). These estimates were compared with estimates from administrative data from the College Board and Department of Education. The results showed no evidence of an effect from telephone interviewing.

4.5.3 Respondent rules: Self- versus proxy respondents

As noted in Section 4.1, all sample persons ages 15 and older who are present at the time of an interview are asked to respond for themselves unless not physically or mentally competent to do so. Proxy responses are accepted for those who are absent or incapable of responding (see Exhibit 4.1 above). Under this rule, about 65 percent of sample persons report for themselves at each interview wave. It has been generally believed that the level of accuracy for self-response is higher than for proxy response (see Moore, 1988 for a contrary view); however, achieving a higher proportion of self-response would increase data collection costs and might lead to some increase in person nonresponse rates.

Evaluation studies

In the ISDP, the current SIPP respondent rules were experimentally compared with a maximum self-response rule, which required callbacks to interview sample persons not present at a household's first interview (Coder, 1980; Kasprzyk, 1988). Household and person noninterview rates were slightly higher for the maximum self-response group and the estimated costs of data collection were about 4 to 6 percent higher. Under the self-response rule, records were consulted more often by persons answering wage and salary questions and response rates for hourly wage rates were higher. Conclusive evidence favoring either set of respondent rules was not found. The less stringent requirement was adopted, with the addition of a callback procedure to obtain certain data items if not obtained at the initial interview.

Under the current respondent rules, most unmarried college students living away from home while attending college are considered to be members of their parents' households and information for them is usually obtained from proxy respondents. Roman and O'Brien (1984) describe the results of an ISDP study in which data for students obtained by proxy from members of their parents' households were compared with data collected from them by self-response at their school residences. The analysis was limited by flaws in the implementation of the test; however, it suggested that proxy response led to greater underreporting of recipiency and more nonresponse on amounts of income and expenses, especially for smaller amounts.

Hill (1987) used data from a single month of the 1984 Panel in a model-based analysis of the effects of proxy versus self-response on reported earnings for a sample of men ages 25 to 55. He estimated that self-reports of earnings were about 20 percent below proxy reports. The validity of Hill's results depended on the correctness of his assumptions concerning a common behavioral structure for proxy and self-respondents. Furthermore, without an external measure of validity, it could not be determined which type of report was of higher quality.

The above addresses the effects of self- and proxy responses on SIPP cross-sectional estimates. There is an additional concern about the effects of changing the informant between waves (for example, from self-respondent to proxy respondent, or from one proxy respondent to another) on longitudinal estimates, particularly on estimates of change (Kalton et al., 1989). No research has been conducted on this issue.

The evidence currently available on the effects of proxy reporting does not clearly indicate the advisability of changing SIPP respondent rules regarding proxy or self-respondents. There are, therefore, no plans at present to change these rules.

4.5.4 Errors associated with interviewers

Interviewer variance

Formal interviewer variance studies carried out in connection with U.S. Population Censuses have shown that when data are collected by interviewers (as opposed to the use of self-administered questionnaires), interviewer variance can contribute a substantial proportion of the total mean squared error for small area counts and sample estimates (see, for example, U.S. Bureau of the Census, 1968). There have been few studies of the contribution of interviewer variance to error in the Census Bureau's household surveys, but a study carried out in connection with the National Crime Survey in eight cities demonstrated that it can be an important source of error for some variables (Bailey et al., 1978). An analysis by Tepping and Boland (1972) of interviewer variance in the Current Population Survey gave estimates of 0.50 or greater for the ratio of interviewer (correlated response) variance to sampling variance for several items included in the survey.

There have been no formal interviewer variance studies in connection with SIPP. Nevertheless, the findings cited from other surveys and Censuses suggest that interviewer variance could be a significant source of error for some items. The contribution of interviewer variance to total error depends on, among other things, the size of the interviewer's assignment. The larger the assignment, the greater the effect on total error. In SIPP the average interviewer assignment is about 14 per month, or 56 households per wave (each interviewed three times during a 12-month period), as compared to approximately 21 households per month or 125 households per 6-month interviewing cycle in the Central Cities Sample for the National Crime and Victimization Survey. Monthly interviewer assignments in the Current Population Survey currently average between 25 and 30 households.

Reinterview checks

The continuing SIPP reinterview program (see Section 4.4) provides some information on measurement error attributable to interviewers. Selected questions from five content groups: job, income, benefits, health insurance, and assets are asked again in the reinterviews. The original and reinterview responses are compared and differences reconciled. When differences are confirmed, the reinterviewer judges whether or not the original interviewer is responsible for the difference in responses. Pennie (1988) summarizes results of this process for 1984 through 1987. The rates presented are based on the number of content groups in each reinterview with one or more confirmed differences. Overall difference rates for the five content groups ranged from a high of 3.1 percent in 1984 to 2.4 percent in 1987. The proportion of confirmed differences attributed to interviewers stayed within a fairly narrow range and averaged 43.3 percent for the 4-year period. Although the manner in which the rates were calculated and presented does not permit precise comparisons between content groups, it appears that the health insurance and asset groups have relatively high difference rates as compared to the other three groups. The health insurance questions were modified after the 1984 Panel, but the difference rates for that content group remained relatively high.

Reinterviewers also check the original interviewers' control card listing of persons present in the household and usual residents of the household who were temporarily absent at the time of the interview. The gross

error detected in the household composition check has been consistently small, averaging 0.23 percent of persons listed in the initial interview over the 4-year period (Pennie, 1988). The average net error rate was -0.05 percent, representing a very small undercount in the initial interviews. However, results of demographic analyses and other types of evidence suggest that the proportion of persons missed in interviewed households may be considerably larger than is indicated by the reinterview results (see, for example, Hainer et al., 1988; Shapiro et al., 1993). The low figure just cited may simply mean that reinterviewers are not much more successful than initial interviewers in obtaining complete listings of household members. The Census Bureau is examining methods to modify the roster questions to improve coverage (see Section 3.5).

Standards have been established for each section of the interview that is checked in the reinterviews. Failure rates have been low; generally, less than 2 percent of the interviewers for each section checked. Usually, the errors uncovered by reinterviewing are discussed with the interviewers. Occasionally interviewers are reassigned or are dismissed or suspended.

Research with reinterview data

Hill (1989) undertook a detailed analysis of reinterview data for Waves 2 and 3 of the 1984 Panel (1,559 cases), using a data file containing reinterview reports merged with data for the same respondents from the public use file. The highest overall discrepancy rates found were 5 to 7 percent for items relating to Medicare and health insurance coverage. For many items included in the reinterview, he found that a high proportion of the discrepancies were procedural, that is, the question was asked by the interviewer and not by the reinterviewer, or vice versa. He attributed this finding to unnecessarily complex skip patterns in the questionnaire.

Hill also examined the relationship between response inconsistencies and respondent characteristics, using a Poisson regression model. He concluded that "data quality appears to be significantly lower for low income, black, and either very young or very old respondents." His general conclusion was that the SIPP reinterview results can be valuable in understanding nonsampling error issues, and he made some recommendations to increase the utility of the reinterview program.

4.5.5 Questionnaire length and structure

The ISDP included two methodological experiments which had a direct influence on the length and structure of the SIPP core modules (Olson, 1980; Kasprzyk, 1988). The first of these experiments compared the use of short and long forms to collect income data for each household member. With the short form, respondents were asked a series of direct questions about receipt of selected types of income. For all "yes" answers, questions on amounts were asked before proceeding to the next income type. The long form made extensive use of cues and probes in an attempt to help respondents recall types of income received by putting them in a broader context. Questions on amounts were not asked until all sources of income had been determined, and a longer set of questions was used to determine the amounts.

Several kinds of data were used in the analysis. Except for Social Security income, which was reported at a higher level on the long form, the comparison of income reporting levels was inconclusive (Olson, 1980, pp. 6-46). In debriefing sessions, the interviewer consensus was that the short form was easier for respondents, but that respondents thought more carefully about their answers when the long form was used (Olson, 1980, pp. 2-22). Not surprisingly, the long form took longer to complete. The form ultimately adopted for SIPP was a variation of the long form used in the ISDP experiment.

The second experiment compared a household screening format with a person-by-person format for obtaining information on income recipiency and amounts. The household format was intended to reduce response burden by asking a single household respondent, for each income type, whether any household

members had received that type of income. For affirmative responses, the household members who had received that type of income were identified. Data on amounts were obtained only after completion of screening for all household members. The person-by-person format used a separate form for each adult household member. No significant differences in income recipiency rates and item nonresponse were observed for the two formats. Use of the person-by-person format added about 5 minutes per household to the length of the interview. The latter format, with moderate changes, was adopted for SIPP (Kasprzyk, 1988).

4.5.6 Use of records by respondents

SIPP respondents are encouraged to use financial records to aid response. However, the percentage of respondents who had used records during interviews were not well known for the earlier panels. In a debriefing of a sample of 1985 Panel respondents in connection with reinterviews following Wave 8 interviews, 56 percent of the respondents said they had referred to their W-2 forms and completed tax forms in the final (Wave 8) interview (Meier, 1988). In response to a question about the use of records such as pay stubs and bank statements in all interviews, 61 percent said they routinely referred to these kinds of records. About 80 percent of those who did not use records said that nothing could be done to encourage them to use records. About half of those who did not refer to bank statements or pay stubs in regular interviews said that they knew the information without referring to records.

For persons responding to questions in the tax and Annual Roundup module, interviewers indicated in check boxes on the questionnaire whether the persons referred to W-2 forms and tax forms during the interview. Table 4.4 shows the results for the 1984 and 1985 Panels (Altman, 1989) which indicate that records were used by around 35-45 percent of the respondents. Anecdotal information from interviewer debriefings suggests that the frequency of record use was significantly affected by the extent to which individual interviewers encouraged it (Chapman, 1988).

Table 4.4 Rate of record use reported by interviewers

Wave and panel	Percentage of respondents referring to	
	W-2 Form	Tax form worksheet
Wave 6, 1984 Panel	34.9	34.0
Wave 9, 1984 Panel	43.2	41.6
Wave 5, 1985 Panel	35.6	35.5
Wave 8, 1985 Panel	36.1	36.3

Source: Altman (1989).

About 20 percent of respondents used at least one type of record during an interview in the 1990 to 1993 Panels, and there was no difference in usage between face-to-face and telephone interviews (Huggins, 1994; James, 1992; and Jones, 1993). Guarino (1996) found that the record-usage tended to increase slightly by wave within a panel. She suggests that this finding may be due to better rapport between respondents and the field representatives, respondents having prior knowledge of what will be asked, and the possibility that respondents who are willing to refer to records are more likely to remain in the survey.

4.5.7 Other improvements to data collection procedures

The data collection procedures for the 1990 to 1993 Panels underwent several changes as a result of the ongoing evaluation and quality improvement processes. For example, the Bureau implemented a number of procedural changes designed to improve data quality, including providing better training for FRs on the importance of data quality, encouraging FRs to probe for more complete responses, providing FRs with expanded and improved information about SIPP, and increasing efforts to reassure respondents about confidentiality (Singh 1989a, 1989b, 1991; and Allen 1991). Further, the questionnaire was modified to clarify some questions in the Earnings, Employment, and Amounts sections. To improve the accuracy of responses on monthly earnings, the FRs first remind the respondents that certain months of the year contain 5 paydays for workers paid on the weekly basis, and 3 paydays for workers paid on a biweekly basis. The FRs also remind respondents that it is sometimes possible to receive more than one payment per month for some programs, such as unemployment compensation and Aid to Families with Dependent Children.

5. NONRESPONSE ERROR: 1984-1993 PANELS

Three kinds of nonresponse occur in SIPP: household (unit) nonresponse, person (within unit) nonresponse, and item nonresponse. Survey procedures are designed to minimize all three kinds of nonresponse, and weighting and imputation procedures minimize errors resulting from differences between responding and nonresponding units and persons. Section 5.1 describes the steps taken in SIPP to maximize response rates. Sections 5.2 through 5.4 present data on the level and trend of each kind of nonresponse in the 1984-1993 Panels and on the characteristics of respondents and nonrespondents. Nonresponse indicators for SIPP and other surveys are compared. Section 5.5 describes the results of experiments with procedures designed to improve response rates in SIPP. The weighting and imputation procedures designed to minimize the effects of nonresponse are covered in Chapter 8.

5.1 Steps to maximize response rates

Wave 1 response

Several steps are taken to encourage response in the first wave of interviewing:

- An advance letter from the Director of the Census Bureau explains the authority for and purposes of the survey and urges participation.
- Field representatives (FRs) carry cards and portfolios identifying them as Census Bureau employees.
- If no one is home at the time of the first visit, FRs try to determine the best time for a callback, either by asking neighbors or by telephoning.
- Potential respondents are assured that their answers will be held in confidence and used only for statistical purposes.
- Senior FRs and supervisors may be called in to convert refusals.
- The interviewing schedule of later panels has been made more flexible to allow for more extensive followup of refusals.
- FRs have talking points to explain SIPP and its value to policy makers. The talking points explain the importance of including different segments of the population: the elderly, the poor, the wealthy, and the middle-income groups.

Wave 2 and beyond

Maintaining high response rates in panel surveys presents particular difficulties since units and persons who initially respond can be lost in later waves. Steps taken in SIPP to minimize attrition after the initial wave include:

- Before each wave, the Census Director sends responding households a letter thanking them for their support and encouraging continued participation.
- Respondents are given a brochure containing some interesting results from the survey.

- FRs are instructed to call ahead and schedule interviews at times that are convenient for respondents.

- To assist in locating households that move, contact persons who would be likely to know the new locations are identified at the first interview. FRs are instructed in tracing movers by contacting the new occupants of the housing unit, the Post Office, and other likely sources.

Success in completing interviews is an important factor in evaluating FRs performance. The standards that have been established for household response rates, calculated on an annual basis, are shown in Section 4.4. The Census Bureau has continued to review these standards to ensure that they truly reflect the efforts of FRs in completing interviews with both movers and nonmovers.

5.2 Household nonresponse

Household or unit nonresponse occurs at the initial wave when all members of an eligible household refuse to participate or when no one can be contacted after repeated attempts. Except in the 1985 Panel, there have been no attempts to contact Wave 1 noninterview households in subsequent waves. When this was tried in the second wave of the 1985 Panel, relatively few households were brought back into the sample and the processing of these households turned out to be difficult and resulted in delays in releasing data products, so the practice was discontinued (Jean and McArthur, 1987).

After Wave 1, sampled households comprise those households that contain original sample persons. Household nonresponse for these households can occur for the same reasons as in Wave 1. In addition, these households may be temporarily unavailable because the sample person has moved, and the household cannot be located, or because the household is located more than 100 miles from a SIPP sample PSU and cannot be interviewed by telephone. Whenever possible, a household that is not interviewed in the second or a subsequent wave is visited once more in the next wave to attempt an interview. If that attempt fails, there are no further attempts to interview the household in later waves.

Nonresponse rates

The calculation of response (or nonresponse) rates is complicated by the fact that the exact number of eligible households after the first wave is not known. Households interviewed in Wave 1 may split up to form additional eligible households, or may leave the survey population entirely. When all members of an interviewed household move and cannot be located, they may account for 0, 1, or 2 or more eligible households in subsequent waves. If all members of the household leave the survey population, there will be no eligible households. If they split up and move to different housing units, there can be two or more households.

The accuracy of reported response rates can also be affected by interviewer misclassification of noninterviews. If a noninterview household is incorrectly classified as not eligible for interview, it will be incorrectly excluded from both the numerator and denominator of the calculated nonresponse rate, and vice versa. Reinterview results for the years 1984 to 1987 show that only 3.5 percent of noninterviews were incorrectly classified, so the effect of such errors on reported response rates is minimal (Pennie, 1988).

Type A and Type D nonresponse

Table 5.1 shows household noninterview rates by panel and wave for the 1984-1993 SIPP Panels for Type A and Type D noninterviews. Type A noninterviews occur when interviews are not obtained for known eligible households. Type D noninterviews occur when some or all members of a household leave and cannot be traced or move more than 100 miles from a SIPP PSU and cannot be interviewed by telephone. The rates shown are cumulative. The sample loss rates after Wave 1 include an adjustment for unobserved

Table 5.1 Cumulative household noninterview[1] and sample loss[2] rates: 1984-1993 Panels[3]

	1984 Panel[3]			1985 Panel			1986 Panel			1987 Panel			1988 Panel		
Wave	Type A	Type D	Loss	Type A	Type D	Loss	Type A	Type D	Loss	Type A	Type D	Loss	Type A	Type D	Loss
1	4.9	--	4.9	6.7	--	6.7	7.3	--	7.3	6.7	--	6.7	7.5	--	7.5
2	8.3	1.0	9.4	8.5	2.1	10.8	10.8	1.5	13.4	11.1	1.5	12.6	11.4	1.5	13.1
3	10.2	1.9	12.3	10.2	2.7	13.2	12.6	2.3	15.2	11.5	2.6	14.2	12.0	2.3	14.7
4	12.1	2.9	15.4	12.4	3.4	16.3	13.8	3.0	17.1	12.3	3.3	15.9	13.0	3.0	16.5
5	13.4	3.5	17.4	14.0	4.1	18.8	15.2	3.7	19.3	13.7	4.1	18.1	13.9	3.3	17.8
6	14.9	4.1	19.4	14.2	4.8	19.7	15.2	4.3	20.0	13.6	4.9	18.9	13.6	4.0	18.3
7	15.6	4.9	21.0	14.4	5.2	20.5	15.3	4.8	20.7	13.6	4.9	19.0			
8	15.8	5.7	22.0	14.4	5.5	20.8									
9	15.8	5.7	22.3												

	1989 Panel			1990 Panel			1991 Panel			1992 Panel			1993 Panel		
Wave	Type A	Type D	Loss	Type A	Type D	Loss	Type A	Type D	Loss	Type A	Type D	Loss	Type A	Type D	Loss
1	7.6	--	7.6	7.3	--	7.3	8.4	--	8.4	9.3	--	9.3	8.9	--	8.9
2	11.0	1.4	12.5	10.9	1.5	12.6	12.3	1.5	13.9	12.8	1.7	14.6	12.3	1.7	14.2
3	11.2	2.3	13.8	11.5	2.6	14.4	13.1	2.7	16.1	13.1	2.8	16.4	12.9	2.8	16.0
4				12.5	3.4	16.5	13.6	3.6	17.7	13.8	3.6	18.0	13.9	3.5	17.9
5				13.6	4.6	18.8	14.5	4.2	19.3	14.9	4.7	20.3	14.9	4.4	19.9
6				14.1	5.3	20.2	14.4	5.1	20.3	15.3	5.4	21.6	15.9	5.5	22.2
7				14.3	5.9	21.1	14.7	5.6	21.0	16.0	5.9	23.0	17.1	6.2	24.0
8				14.4	5.9	21.3	14.5	5.9	21.4	16.9	6.7	24.7	17.5	6.9	25.1
9										17.7	7.3	26.2	18.1	7.5	26.5
10										17.5	7.6	26.7			

[1]Type A noninterviews consist of households occupied by persons eligible for interview and for whom a questionnaire would have been completed had an interview been obtained. Reasons for Type A noninterview include: no one at home in spite of repeated visits, temporarily absent during the entire interview period, refusal, and unable to locate the sample unit. Type D noninterviews consist of households of original sample persons who are living at an unknown new address or at an address located more than 100 miles from a SIPP PSU, provided a telephone interview is not conducted.

[2]The sample loss rate consists of cumulative noninterview rates adjusted for unobserved growth in the Type A noninterview units (created by splits).

[3]Differences in rates for the 1984 Panel compared to subsequent panels may be due in part to differences in the panel designs.

Source: Data for the 1984-1988 Panels came from Kasprzyk (1988) and Bowie (1990).

growth resulting from splits of households that were Type A noninterviews in the first wave and later waves. Details of the calculation of the rates shown in Table 5.1 are given by Nelson et al. (1987). Apart from the 1984 Panel, most of the sample loss occurs in the first two waves, and most of the early losses are due to Type A nonresponse. The number of additional Type A noninterviews is quite small after the first four waves. Most households that have participated up to that point continue in their participation until the end of the panel. Losses due to moves (Type D noninterviews) continue at a fairly constant rate throughout the duration of each panel.

Table 5.1 shows a consistent pattern in the wave nonresponse rates for SIPP panels between 1984 and 1990. The rate of loss was about 5-8 percent of eligible households at Wave 1, an additional 4-6 percent of eligible households at Wave 2, another 2-3 percent in each of Waves 3-5, and less than 1 percent in each subsequent wave thereafter. For these panels, by Wave 6 (after 2 years of interviewing), cumulative sample loss for SIPP was 18-20 percent of eligible households; by Wave 8 it was 21-22 percent. The 1992 and 1993 Panels experienced higher household nonresponse. The loss rates were around 9 percent at Wave 1, 22 percent at Wave 6, and 25 percent at Wave 8.

The causes for the increase in household nonresponse in the 1992 and 1993 panels are unknown. One suggestion from the FRs is that maximum telephone interviewing had an adverse effect on Type A noninterview rates. Investigations by King (1995) and James (1993) found no evidence to support this suggestion. However, the Type D nonresponse rates were found to be higher when telephone interviewing was used heavily.

It is relevant to note here that several other major demographic household surveys conducted by the Census Bureau also experienced increases in Type A noninterview rates and refusal rates in the period 1990-1995 (Davey, 1996). The surveys investigated include: the CPS, the National Crime and Victimization Survey, the National Health Interview Survey (NHIS), the Consumer Expenditure (CE) Quarterly, and the Consumer Expenditure (CE) Diary. Comparisons of the annual average Type A noninterview rates between 1990 and 1995 showed increases ranging from 1.4 percent for the NHIS to almost 13 percent for the CE Diary. The increases in the annual average refusal rates ranged from 1.4 percent to 4.1 percent.

Rates relative to other surveys

Because of important differences in survey content and design features, it is difficult and possibly misleading to compare the SIPP response experience with that of other panel surveys, let alone cross-sectional surveys. Nelson et al. (1987) made a careful comparison of sample losses in SIPP, the Income Survey Development Program (ISDP), the National Longitudinal Surveys (NLS), the National Medical Care Utilization and Expenditure Survey, and the Panel Study of Income Dynamics (PSID). They concluded that: "In summary, SIPP sample losses appear to be comparable to those experienced by other longitudinal surveys. We see the same pattern repeated in all these surveys; that is, a heavy initial loss which increases at a much slower rate throughout the subsequent interviews until it levels off." (Nelson et al., 1987, p.636). Further information on nonresponse and other nonsampling errors in panel surveys can be found in Kasprzyk et al. (1989).

Refusals

Most Type A noninterviews are the result of refusals. Refusals accounted for between 70 and 76 percent of Wave 1 noninterviews in the first four panels (Bowie, 1988b, p.8). Some data on reasons for refusals in the first six waves of the 1984 Panel are given by Nelson et al. (1987). Only one reason for refusing was

coded for each household even though multiple reasons may have been given. For Wave 1, the most frequently cited reasons were:

Reason given	Percent of households that refused
Not interested in participating	18.7
No time, too busy	14.7
Invasion of privacy	9.9
Voluntary survey	9.3
Offended by income questions, too	9.1
Didn't believe information was confidential	3.4

Of those who refused in the second wave, 6.2 percent said they had not understood that the interviewer would be returning for another interview. In the third and later waves, about a quarter of the refusers said that they did not want to answer any more questions.

Characteristics of nonresponding households

For obvious reasons, relatively little information is available about the characteristics of sample households that are not interviewed in Wave 1 of a panel. It is possible, however, to compare the characteristics of households that are interviewed in the first wave and drop out later with the characteristics of households that continue to participate. The Census Bureau's SIPP Nonresponse Workgroup (1988) compared distributions based on first-wave characteristics of household reference persons for sample households interviewed in all of the first five waves of the 1984 Panel and sample households interviewed in the first wave but not in the fifth wave (households in the second group may or may not have been interviewed in Waves 2 to 4). Types of households more likely to be noninterviews in the fifth wave included those that were: occupied by renters, located in Standard Metropolitan Statistical Areas (SMSAs) with a population of 500,000 or more, or had reference persons in the age group 15 to 24. No statistically significant differences were observed for distributions of households by household monthly income, sex of the reference person, or receipt/nonreceipt of cash or noncash benefits.

5.3 Person nonresponse

Person nonresponse occurs when a member of an interviewed household refuses to participate or is unavailable for interview and a proxy interview is not obtained. Original sample persons can miss any number of waves and still reenter the sample as long as their households have not missed two or more consecutive interviews in the interim. Most analyses of person nonresponse subsequent to Wave 1 include all persons not interviewed in each wave, whether or not other members of their households were interviewed in that wave.

As discussed in Chapter 8, the presence of person nonresponse in interviewed households makes it difficult to decide what combinations of imputation and estimation procedures are best suited for undertaking various kinds of longitudinal analyses with SIPP data.

Nonresponse patterns

Table 5.2 (Lepkowski et al., 1989) shows the person response patterns for the eight interviews in the 1984 Panel. Panel members are original sample persons who responded for all eight interviews for which they were eligible (i.e., for which they were in the SIPP universe). Nonpanel members are original sample

Table 5.2 Person response (1) or nonresponse (0) patterns for the eight interviews in the 1984 SIPP Panel research file

Person response pattern	Percent
Panel members	**70.4**
Responded to all 8 interviews	67.0
Responded to all interviews prior to death	1.5
Responded to all interviews except when:	
in an institution	0.7
in the Armed Forces	0.4
abroad	0.6
Other panel members	0.2
Nonpanel members	**24.8**
Attrition nonresponse patterns:	
11111110	1.4
11111100	1.2
11111000	1.5
11110000	2.0
11100000	2.2
11000000	2.6
10000000	3.1
Other patterns with only one missing interview	4.3
Other patterns with two or more missing interviews	2.9
Nonpanel members with missing interviews who were known to have died, become institutionalized, entered the Armed Forces, or moved abroad	0.3
Nonpanel members with missing interviews who left for other reasons	3.3
Total nonrespondents[1]	**4.9**
Total	**100.0**
Total number of sample persons	35,027

[1]Estimated -- the number of total nonrespondents is not known; all that is available is the number of total nonresponding households, from which the number of persons has been crudely estimated by multiplying the average number of sample persons in responding households by the number of nonresponding households.

Source: Lepkowski et al. (1989).

persons who failed to respond to one or more interviews when they were eligible. Total nonrespondents are those who did not respond to any of the eight interviews. Table 5.2 shows that 70.4 percent of the SIPP sample were interviewed in all interviews in which they were eligible. Persons who were interviewed for one or more consecutive interviews and then dropped out for all of the remaining interviews, known as attrition cases, accounted for 14 percent. Persons with other response patterns, who might be called "in and outers," accounted for 7.2 percent.

Table 5.3 on the following page shows the nonresponse patterns for panel and nonpanel members from Waves 1 through 7 of the 1985, 1986, 1987, and the 1991 Panels. This table does not include an estimate for total nonrespondents, that is, persons in nonresponding households at Wave 1. Therefore, the percentages are not directly comparable to those in Table 5.2. Table 5.3 shows that about 76 to 80 percent of the SIPP sample who were in responding households at Wave 1 responded in all seven waves for which they were eligible. Persons who were interviewed for one or more consecutive interviews, and then dropped out for all remaining interviews accounted for about 13 to 16 percent. Persons with other nonresponse patterns accounted for about 6-7 percent. About half of the persons who were Type Z nonrespondents at Wave 1 failed to respond at any of the seven waves. Type Z nonrespondents are members of otherwise cooperating households for whom no information is obtained, either by self- or proxy response.

Movers

Not surprisingly, person nonresponse is higher for persons who move than for those who are residents of the same housing unit for the duration of the panel. Table 5.4 compares person nonresponse for movers and nonmovers in Waves 1 to 5 of the 1984 Panel, for rotation groups 1, 2, and 3. About one-fifth of the original sample persons moved at least once between Waves 1 and 5. One or more interviews were missed for about one-third of the movers, as compared with less than one-fifth of the nonmovers (Jean and McArthur, 1987). For the 1991 Panel, Mack et al. (1995) found that at Wave 8, the nonresponse rate for movers was 22.5 percent, about double the rate of 10.3 percent for nonmovers.

Table 5.4 Response status for movers and nonmovers in Waves 1 to 5: 1984 Panel[1]

Response status	Total		Movers		Nonmovers	
	Number	Percent	Number	Percent	Number	Percent
Five interviews	19,878	79.1	3,485	68.8	16,393	81.7
Missing at least interview 5	4,222	16.8	1,148	22.6	3,074	15.3
All other	1,038	4.1	436	8.6	602	3.0
Total	25,138	100.0	5,069	100.0	20,069	100.0

[1]Persons ages 15 and older in rotation groups 1, 2, and 3.

Source: Jean and McArthur (1987).

The higher nonresponse among movers is a concern because of the mobility of the SIPP sample. Waite (1995a) reports that over 25 percent of households with three or more persons would have at least one person moving out of their households in a 4-month period. On average, 41 percent of the new noninterview cases reported in each wave of the 1992 Panel were unlocated movers. In any given wave, the Type D noninterview rate (mainly unlocated movers) for persons in poverty was twice the rate for persons not in poverty. In contrast, nonresponse due to refusals was higher for persons not in poverty than for persons in poverty.

Table 5.3 Nonresponse patterns for panel and nonpanel members in Waves 1-7 of the 1985-87 and 1991 Panels

Pattern group	1985 Panel Percent	1986 Panel Percent	1987 Panel Percent	1991 Panel Percent
Panel members	**76.1**	**77.5**	**79.5**	**77.0**
A. Responded on all 7 waves	73.7	74.7	76.9	73.8
B. Ineligible on some waves:				
Died	1.0	1.1	1.2	1.4
Entered institution	0.5	0.6	0.5	0.8
Entered Armed Forces barracks	0.3	0.3	0.2	0.2
Moved abroad	0.6	0.7	0.7	0.8
Reason ineligible not listed	0.0	0.0	0.0	0.0
Nonpanel members	**23.9**	**22.5**	**20.5**	**23.0**
C. Eligibility for all 7 waves				
1. Attrition nonresponse only at:				
Wave 2	5.1	5.0	4.7	4.6
Wave 3	3.4	2.9	2.3	2.5
Wave 4	2.8	2.1	1.9	1.8
Wave 5	1.8	2.0	1.6	1.4
Wave 6	1.3	1.1	1.1	1.2
Wave 7	1.4	1.3	1.1	1.8
2. Nonattrition nonresponse only:	5.7	5.7	5.6	6.8
1 missing wave	4.5	4.6	4.3	5.2
2 missing waves	0.9	0.9	1.0	1.1
3-5 missing waves	0.3	0.2	0.3	0.5
3. Nonattrition and attrition nonresponse: At least one wave of each	1.3	1.3	1.2	1.4
4. Type Z noninterview at Wave 1:				
Noninterview all waves	0.4	0.4	0.4	0.6
Interviewed at least one wave	0.4	0.4	0.4	0.6
D. Some ineligibility	0.3	0.3	0.2	0.3
Total	**100.0**	**100.0**	**100.0**	**100.0**
Number of sample members	30,605	23,157	30,769	28,184

Characteristics of nonrespondents

A study by McArthur (1987), based on all 9 waves of the 1984 Panel, compared the characteristics of persons for whom all interviews were completed with the characteristics of persons for whom there were one or more noninterviews. Table 5.5 shows selected results from this study and the corresponding data from Waves 1-7 of the 1992 Panel. In both panels, those with higher than average person nonresponse have similar characteristics. These characteristics are: residence in rented housing units, members of racial minorities, children and other relatives of the reference person, ages 15 to 24, never married, and lack of a savings account or other assets. Residents of large SMSAs also have higher nonresponse rates.

Effects of Sample Loss on SIPP Estimates

There is evidence that sample attrition may affect SIPP estimates of income and poverty, and estimates of benefits for some means-tested assistance programs. Although weighting adjustments are used in an attempt to compensate for attrition, it appears that they are not entirely successful (see Chapter 8 for research on weighting).

From a review of patterns of SIPP estimates over time, Huggins and Winters (1995) observed the following trends:

- Within the same panel, estimates of the number of households with low income typically decrease over the life of the panel.

- For the same time period, estimates of households with low incomes for two overlapping panels differ markedly, with the later panel having a much higher estimate than the earlier one.

Possible explanations for these findings include reporting errors (such as telescoping), time-in-sample bias due to length in panel, seasonal effects, and attrition bias. While there are some seasonal differences in the estimates, these trends are likely to be mainly the result of attrition.

Sae-Ung and Winters (1998) further analyzed the effects of nonresponse on income and poverty estimates. They examined response propensities by various characteristics of sample members and found that low-income households are more likely than other households to drop out of the panel. This was true for all panels between 1984 and 1993. They suggest that improvements are needed for the weighting adjustment procedure to better compensate for the differential rate of nonresponse low-income households as the panel progresses.

Tin (1996) examined the relationship between attrition and estimates of benefits from some means-tested programs in the 1990 Panel. The result indicates that attrition affects estimates of benefits from Food Stamps and Supplemental Security Income, but not estimates of benefits from AFDC (Aid for Family with Dependent Children), General Assistance, and WIC (Food Program for Women, Infants, and Children).

5.4 Item nonresponse

Item nonresponse occurs when a respondent refuses or is unable to answer a question. It can also occur when a FR fails to ask a question that should have been asked or asks the question but fails to record the answer. Missing data for an item can also be the result of a data entry failure, that is, the response was not keyed in, or was deleted in editing because it is inconsistent with answers to other items.

Table 5.5 Selected characteristics of original sample persons by their interview experience for the full 1984 Panel and for the first 7 waves of the 1992 Panel

Characteristics of persons in Wave 1	1984 Interview experience (% distribution)				1992 Interview experience (% distribution)			
	Completed all	Missed last 2	Other[1]		Completed all	Missed last 2	Other[1]	
			Missed 1	Missed 2+			Missed 1	Missed 2+
Relationship:								
Ref. person	37.3	33.1	33.9	30.7	36.3	31.3	30.5	31.2
Primary ind.	12.7	14.7	13.4	12.0	13.9	18.0	15.4	17.9
Spouse	31.3	23.9	25.4	23.9	29.4	21.3	22.4	21.0
Child	13.9	17.9	18.4	24.1	14.2	14.4	17.8	14.8
All other	4.8	10.3	9.0	9.4	6.3	15.0	13.9	15.1
Age:								
15-24	18.2	24.2	25.3	29.5	16.3	23.4	25.5	23.8
25-34	22.1	21.4	22.6	22.9	21.1	23.3	25.3	23.4
35-44	17.3	14.4	17.0	14.9	21.1	16.2	16.4	16.1
45-64	27.0	23.2	21.5	23.2	26.0	19.9	18.0	19.8
65+	15.4	16.8	13.6	9.5	15.4	17.2	14.9	16.9
Race:								
White	88.5	83.8	82.9	78.2	87.0	80.9	78.9	80.8
Black	9.2	12.9	13.9	17.6	9.6	14.5	15.6	14.6
Other	2.3	3.3	3.2	4.2	3.5	4.6	5.6	4.6
Living quarters:								
Owned	72.7	62.9	65.6	59.8	71.9	55.6	57.8	55.2
Rented	24.8	35.1	32.1	37.3	25.9	41.6	39.1	42.1
Rent free	2.4	2.0	2.3	2.8	2.3	2.8	3.1	2.8
Marital status:								
Never married	20.8	28.8	28.3	33.6	23.4	33.0	32.8	33.5
Married	63.5	52.2	53.3	50.4	59.9	46.9	47.9	46.6
Other	15.8	18.9	18.5	15.9	16.7	20.0	19.3	19.9
Savings account:								
Yes	59.9	51.1	51.9	51.0	55.0	42.8	43.3	42.5
No	40.1	48.9	48.1	49.0	45.0	57.1	56.7	57.5

[1]Interview experience categories are mutually exclusive. At least one of the last two interviews was completed for persons in the "Other" category.

Source: Data for the 1984 Panel were adapted from McArthur (1987).

Item nonresponse can be analyzed either on a cross-sectional or a longitudinal basis. Cross-sectional item nonresponse refers to missing data items for persons interviewed in a single wave. Longitudinal item nonresponse refers to data items common to two or more waves of interviewing: There is said to be longitudinal item nonresponse for an item if the person was interviewed in all the relevant waves and data are missing for one or more of those waves. This section examines the extent of item nonresponse in SIPP and how it compares with other surveys. The methods used to compensate for cross-sectional and longitudinal item nonresponse are discussed in Chapter 8.

Core items

Table 5.6 shows the item nonresponse rates for five of the panels between 1984 and 1993 for selected core items in three categories: labor force activity; income recipiency or asset ownership; and income amounts. The rates for the first two categories are quite low and do not show much change over the first three panels. The results for the 1992 and 1993 Panels are lower for the labor force activity items, but they are not truly comparable with the results for earlier panels since they relate only to Wave 1. Nonresponse rates for income amounts, especially self-employment income, and interest and dividends, are substantially higher, indicating that respondents who report income recipiency or asset ownership are not always able or willing to report amounts. For interest, as shown by the rates in parentheses for the 1984-1986 Panels, the majority of respondents who did not report amounts of interest were willing to report the balances in their interest-bearing accounts, thus providing a reasonable basis for imputation of the amounts of interest. For dividends, separate questions are asked for dividends paid directly to shareholders and those credited to reinvestment or margin accounts (shown in parenthesis). Nonresponse rates are considerably higher for the latter category.

Table 5.7 shows longitudinal nonresponse rates for selected income amounts based on 12 months and 32 months from the 1984 Panels. For income types appearing in both Tables 5.6 and 5.7 the overall nonresponse rates are lower in Table 5.6. For example, Food Stamps had an average cross-sectional nonresponse rate of 3.6 percent in the 1984 Panel as compared with a longitudinal nonresponse rate of 8.1 percent for the first 12 months of that panel, and 7.2 percent for all 32 months of the panel. The 32-month estimate appears lower because it included only those panel members who completed all 8 waves of the survey. Table 5.7 also shows that only 1.9 percent of Food Stamp recipients had no Food Stamps amounts reported for any of the 12 months, and 3.0 percent had no amounts reported for any of 32 months. Thus, for many Food Stamp recipients with some missing values for Food Stamp benefit amounts, it would be possible to impute these values based on data for other months.

Items on asset amounts

The highest item nonresponse rates experienced in SIPP have been for questions on asset amounts. Table 5.8 shows nonresponse rates for selected asset amounts collected in Waves 3 and 4 of the 1984 SIPP Panel, Wave 4 of the 1986 Panel and, for comparative purposes, from the Income Survey Development Program (ISDP). SIPP nonresponse rates for these items range from 13 to 42 percent, lower than the corresponding rates in the ISDP, but still much higher than rates for most other types of questions. Curtin, et al. (1989) have compared item nonresponse for questions on asset ownership and amounts in SIPP (data for Waves 3 and 4 of the 1984 Panel), the Panel Study of Income Dynamics (1984 Wealth Supplement), and the 1983 Survey of Consumer Finances (SCF). Nonresponse on the ownership questions is low for all types of assets in all three surveys. For the asset categories included in all three surveys, item nonresponse rates for questions on amounts were, without exception, higher for SIPP than for either of the other two surveys.

Table 5.6 Nonresponse rates for selected SIPP core items by panel

Question	1984	1985	1986	1992[3]	1993[3]
Labor force activity:					
Identification of weeks absent without pay (item 4)	0.1	(Z)	0.1	0.3	0.2
Identification of weeks with a job or business (item 6a)	2.2	2.0	2.5	0.2	0.1
Presence of weeks looking or on layoff (item 7a)	1.0	1.3	2.0	0.1	0.1
Identification of weeks looking or on layoff (item 7b)	3.2	2.4	2.9	0.2	0.1
Income recipiency or asset ownership:					
Social Security	0.6	0.6	1.0	0.1	0.1
Unemployment compensation	0.1	0.1	0.2	0.2	0.1
Food Stamps	0.3	0.4	0.5	0.6	0.6
Savings accounts	1.0	0.9	0.9	0.8	1.0
Shares of stock	1.3	1.4	1.5	1.3	1.5
Income amounts:					
Hourly wage rate	9.5	10.4	10.8	7.2	7.7
Monthly wage and salary	6.2	7.2	8.4	4.0	4.2
Self-employment salary or draw	14.0	16.9	14.6	12.4	13.5
Social Security	8.8	9.5	10.0	14.0	14.7
Unemployment compensation	9.1	9.7	9.9	9.2	10.7
Food Stamps	3.6	4.1	4.4	6.4	5.4
Interest[1]	34.6(24.2)	29.8(28.9)	30.8(30.2)	10.3	10.2
Dividends[2]	9.4(30.7)	10.5(30.5)	9.4(29.1)	7.5	7.6

Z = Less than .05 percent.

[1]The figure in parentheses is the nonresponse rate on the balance in the account. This question was asked of the 34.6 percent that did not provide an estimate of the amount of interest received.

[2]The figure in parentheses is the nonresponse rate for dividends credited to accounts.

[3]The rates shown for labor force activity items in 1992 and 1993 refer to Wave 1 only.

Source: Rates for the 1984-86 Panels adapted from Bowie (1986).

Table 5.7 Longitudinal item nonresponse rates for amounts of selected income types: 1984 SIPP Panel 12-month summary[1] and 32-month summary[2]

Income type	All amounts reported	One or more amounts not reported	One or more but not all amounts not reported	No amounts reported
12 months				
Hourly wage rate	83.0	17.0	9.0	8.0
Social Security	82.8	17.2	13.1	4.1
Private pension	78.8	21.8	13.6	8.2
AFDC	91.0	9.0	5.6	3.4
Food Stamps	91.9	8.1	6.2	1.9
Unemployment compensation	87.9	12.1	4.0	8.0
Federal SSI	88.0	12.0	7.6	4.4
32 months				
Social Security	87.7	12.4	8.6	3.8
AFDC	92.1	7.9	4.8	3.1
Food Stamps	92.7	7.2	4.2	3.0
Unemployment compensation	86.6	13.4	4.5	8.9
Federal SSI	90.2	9.8	4.3	5.5

[1]These rates are based on the total number of persons with recipiency in one or more of the 12 months. Also these rates do not reflect imputations made to type Z person noninterviews.

[2]These are rates of missing data based on panel members included in the 1984 SIPP longitudinal file. The rates include imputations due to item nonresponse only. Type Z imputations are not included. Data are adapted from Pennell (1993, Table 4.3).

Source: Data for the 12-month summary were adapted from Kasprzyk and Herriot (1986) and data for the 32-month summary were adapted from Pennell (1993, Table 4-3).

Table 5.8 Item nonresponse rates for asset amounts in SIPP and the ISDP

Asset type	Percent nonresponse		
	SIPP		
	1984 Waves 3 and 4	1986 Wave 4	ISDP
Amount in savings accounts	16.8	22.8	24.9
Amount in checking accounts	13.3	21.2	23.1
Amount in bonds and government securities	25.9	23.1	32.2
Market value of stocks and mutual fund shares	41.5	36.9	65.8
Debt on stocks and mutual fund shares	41.1	38.6	87.3
Face value of U.S. Savings Bonds	24.9	24.9	35.8
Value of rental property	33.5	31.3	39.9
Value of own business	37.9	41.8	55.3
Debt on own business	28.8	31.5	50.4

Sources: Data from ISDP and SIPP 1984 Panel from U.S. Bureau of the Census (1986).
Data from SIPP 1986 Panel from unpublished tabulations.

Topical modules

Turning to item nonresponse for the topical modules, we note that there are some interviewed persons who do not provide any data, either directly or through a proxy respondent, for a topical module for which they are eligible. As shown in Table 5.9 for the 1984-1993 Panels, complete nonresponse to topical modules ranges from about 3 to 9 percent. Nonresponse to topical modules tends to increase moderately over the life of a panel. Topics covered in later waves of a panel are doubly disadvantaged: a smaller proportion of households and persons are interviewed, and of those who are interviewed, a slightly higher proportion fail to respond to the topical module.

Income nonresponse rates in SIPP and CPS

A question of considerable interest to users of SIPP income data is how the quality of those data compares with the quality of income data from the annual Current Population Survey (CPS) March Income Supplement. This is not an easy question to answer. A full analysis requires consideration of all sources of error that affect SIPP and CPS estimates, as well as differences in coverage and definitions between the two surveys.

Table 5.9 Person nonresponse rates for topical modules

Panel	Wave	Percent of persons with no topical module data	Type of module
1984	4	3.3	Fixed: Assets and liabilities Variable: Retirement and pensions, housing, energy
	5	3.1	Variable: Child care and support, welfare history, work expenses, reasons not working, reservation wage
	6	4.2	Fixed: Annual income, taxes, employee benefits, education Variable: Training (ETA)
	8	4.5	Fixed: Marital, fertility and migration history Variable: Household relationships, work expenses, support of persons not in household
	9	5.0	Fixed: Annual income, taxes, IRAs, education
1985	3	2.8	Fixed: Assets and liabilities
1990	2	5.0	Fixed: Education and work training history
	5	6.6	Fixed: Education financing
1991	2	5.0	Fixed: Education and training history
	5	6.0	Fixed: Education financing
	6	4.0	Variable: Extended measure of well-being
	8	6.5	Fixed: School enrollment and financing
1992	2	4.0	Fixed: Education and work
1993	8	9.3	Fixed: School enrollment and financing

Source: Bowie (1986, 1988a).

Looking only at item nonresponse rates for income amounts, SIPP has a clear advantage, as shown in Tables 5.10 and 5.11. Moreover, the SIPP sample persons who do not respond to particular income items are not necessarily the same ones in each quarter (see Table 5.7), so in many instances of item nonresponse, part-year data will be available for use in imputing missing values. In the March 1985 CPS, 5.2 percent of eligible households were not interviewed. In the interviewed households, data for the income supplement either were not obtained or could not be tabulated because of matching failures for 9.4 percent of the eligible persons. Combining these two rates indicates that the item nonresponse rates shown in Tables 5.10 and 5.11 for CPS apply to an interviewed population already subject to a 14.1 percent overall person nonresponse rate.

Table 5.10 Item nonresponse rates[1] for income amounts in the 1984 SIPP Panel and March 1985 CPS, for selected income types

Income type	SIPP 1st quarter 1984 monthly average	SIPP 2nd quarter 1984 monthly average	SIPP 3rd quarter 1984 monthly average	SIPP 4th quarter 1984 monthly average	CPS March 1985
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
AFDC	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
Veteran's compensation or pension	11.3	11.2	11.9	13.5	18.3

[1]Noninterviews or complete failure to obtain cooperation from any household member have not been considered in this examination of nonresponse rates.

Source: U.S. Bureau of the Census (1985).

Table 5.11 Item nonresponse rates[1] for income amounts in the 1985 SIPP Panel and March 1986 CPS, for selected income types

Income type	SIPP 1st quarter 1985 monthly average	SIPP 2nd quarter 1985 monthly average	SIPP 3rd quarter 1985 monthly average	SIPP 4th quarter 1985 monthly average	CPS March 1986
Wage and salary	7.8	6.7	8.5	8.3	16.5
Self-employment income	17.0	9.3	19.4	17.6	21.9
Supplemental Security Income (Federal)	8.5	8.7	8.0	8.5	16.5
Social Security	13.0	12.5	12.7	13.2	19.3
AFDC	8.9	8.1	8.5	8.5	14.4
Unemployment compensation	15.0	15.6	14.9	16.0	18.6
Company or union pension	17.0	16.1	16.2	16.5	21.0
Food Stamp allotment	7.1	5.9	6.3	6.6	11.1
Veteran's compensation or pension	13.1	13.4	13.2	16.2	18.8

[1]Noninterviews or complete failure to obtain cooperation from any household member have not been considered in this examination of nonresponse rates.

Comparable rates for SIPP are more difficult to determine. The SIPP rates increase somewhat each quarter because of additional sample loss at each successive wave. Table 5.1 that appears earlier, shows that 4.9 percent of eligible households were not interviewed in Wave 1 of the 1984 Panel. McArthur (1987) shows that of sample persons ages 15 and older who were interviewed in Wave 1, the percentages not interviewed in the next three waves were:

Wave 2	5.6 percent
Wave 3	10.3 percent
Wave 4	14.1 percent

Overall, the comparison favors SIPP, but the advantage of SIPP over the CPS March Income Supplement in terms of total nonresponse would be considerably smaller in the second calendar year covered by a panel.

5.5 Methodological research and experimentation

This section describes Census Bureau research and experimentation aimed at reducing nonresponse in SIPP and using supplemental data collection procedures to enhance survey content and to compensate for missing data in regular interviews.

5.5.1 The gift experiment

The purpose of the gift experiment was to determine whether giving an inexpensive gift to households at their first interview would reduce household nonresponse. All households in rotation group 4 of the 1987 Panel were given an inexpensive solar-powered calculator immediately following the interviewer's introductory statement in the first interview, conducted in April 1987. No obligation to participate either in the initial interview or subsequent ones was associated with the gift.

The other three rotation groups of the 1987 Panel did not receive gifts and served as the control group for the experiment. In the analysis it was necessary to consider possible seasonal effects on the outcome measures; the experimental rotation group started interviewing in April and the three rotation groups used as the control started interviewing in February, March, and May.

The main analysis of the experiment compared cumulative Type A noninterview rates for the experimental and control groups. The results for the first three waves of the 1987 Panel were inconclusive. Cumulative noninterview rates were significantly lower (significance level = .10) in the experimental group at each wave, but comparisons with the 1985 and 1986 Panels indicated that the differences might have been at least partly due to seasonal trends in response rates.

Because the effectiveness of a procedure can depend in part on interviewers' reactions to it, interviewers in the experimental group were asked to complete a self-administered evaluation form after the interview in which the gift was given. Their reactions were mixed: although about one-third thought that giving a gift to participating households was a good idea, only 12 percent believed that giving it helped to gain cooperation. Some interviewers felt that they were "buying the interviews" and that the survey was becoming "too commercial."

Further details on the experimental procedures and results of the gift experiment are given by the Gift Experiment Workgroup (1988), Petroni et al. (1989), and Butler (1991). Since the gift of a calculator resulted in only a 2 percent increase in final response rates compared to the control (no gift group), the increase was deemed too small to justify adoption of the gift procedures.

Mack and Waite (1995) reviewed a number of incentive studies which found that incentives increased response rates and reduced the number of callbacks to contact sample persons. Based on this review, an experiment with a cash incentive was conducted in Wave 1 of the 1996 Panel. This experiment is discussed in Section 11.4.

5.5.2 A missing wave module

In a panel survey like SIPP, if a sample person is not interviewed in one or more waves, it is possible to try to collect the missing information in subsequent waves. This possibility was tested in the 1984, 1985, and 1986 Panels. Starting with Wave 4 of the 1984 Panel, a missing wave module was added to the questionnaire and administered, following the core and topical module sections, to persons who had missed the preceding wave but had been interviewed in the wave prior to one missed. The missing wave module contained an abbreviated set of questions on labor force status, receipt of income and assets, and program participation for the reference period covered by the missed wave.

A study of the missing wave data collected in Wave 9 of the 1984 Panel evaluated the potential utility of the data for imputation and noninterview adjustment. The main criteria selected for evaluation were item nonresponse rates, completion rates, and reporting of transitions. A transition was defined as a change in recipiency status that occurred during the reference period covered by the missing wave questions.

It was found that interviewers succeeded in collecting at least some missing wave information for 94 percent of eligible persons. However, the number of transitions reported was much smaller than predicted. The additional information did not justify the respondent burden and cost of collection, because other methods of compensating for the missing data, such as imputation by direct substitution, are likely to be equally effective. Consequently, the missing wave module was dropped from the SIPP questionnaire midway through the 1986 Panel. Additional information on the evaluation study is provided by Huggins (1987), Kasprzyk (1988), and Petroni et al. (1989).

5.5.3 The employer-provided benefits study

For some kinds of data about sample persons, the person may not be the best source of information. In several household surveys, information collected directly from respondents has been supplemented by data obtained, with respondents' permission, from record sources such as health care providers, utility companies, and government agencies. In SIPP, it was felt that information on the economic well-being of employed sample persons might be enhanced by contacting their employers to obtain information on employee and employer contributions to health insurance, pension, and life insurance plans.

Procedures for obtaining such information from employers were tested for a probability sample of 500 households taken from one rotation group of the 1985 Panel. At the time of the final interview for this rotation group, in August 1987, employed persons in the sample households were shown the questionnaire that would be sent to their employers and, if they gave permission, were asked to provide employer addresses and names of contact persons at their places of employment. When permission was given, the authorized questionnaires were sent to the employers.

For SIPP interviews conducted by telephone or with a proxy respondent, the interviewer mailed or left the employer questionnaire/authorization form, along with a letter of explanation and an envelope to mail the form to the Census Bureau. There was no follow-up when these forms were not returned. Follow-up on the forms that were sent to employers was conducted first by letter and subsequently by telephone.

This procedure carried the possibility of two new kinds of nonresponse, beyond those normally encountered in SIPP. First, sample persons might decline to authorize the Census Bureau to obtain information from their employers. Second, employers might fail to complete and return the questionnaires. Somewhat unexpectedly, the first of these sources accounted for most of the total nonresponse. Authorizations were obtained from only 596 (44.1 percent) of the 1,352 employed persons in the study sample. Of the 596 forms mailed to employers, 549 (92.1 percent) were mailed back. Thus, the overall response rate for eligible persons was 40.6 percent. Overall response rates for self-respondents to face-to-face interviews were moderately higher than they were for sample persons interviewed by telephone or through a proxy respondent.

The high response rate by employers was encouraging and it is believed that with some additional effort the rate of authorization by eligible sample persons to contact their employers can be increased. Additional information about the employer-provided benefits study is given by Kasprzyk (1988), Singh (1988b), Petroni et al. (1989), and Haber (1989). This research, however, was discontinued in the 1990s.

6. MEASUREMENT ERROR

This chapter treats measurement or response errors that are associated with the data collection phase of SIPP. Measurement errors can contribute to the total mean square error of survey estimates by introducing nonsampling variance, bias, or both. The level of measurement error is affected by virtually all design features including: interview mode, respondent rules, questionnaire content and structure, and interviewer training and supervision (see discussion in Section 4.5). The longitudinal nature of SIPP brings opportunities for additional types of measurement error not found in cross-sectional surveys. It also provides opportunities to detect errors that exist in cross-sectional surveys that might otherwise go undetected.

This section focuses on measurement errors that are mainly related to the longitudinal nature of SIPP. Also included are the results of several methodological experiments and studies that have been undertaken, either as part of the Income Survey Development Program (ISDP) or SIPP, to evaluate data quality and alternative data collection strategies. Macro-evaluation studies, which examine differences between aggregate data from SIPP and comparable data from other sources, without distinguishing the sources of observed differences, are covered in Chapter 10.

Section 6.1 deals with what may be the single most perplexing problem that has emerged in SIPP. This is popularly referred to as the "seam phenomenon." It reflects the fact that many more month-to-month changes in survey variables occur at the seam between two waves of interviews than occur between months within a wave. Section 6.2 covers other types of error associated with panel surveys, such as conditioning or time-in-sample bias. Section 6.3 presents the results of evaluation studies designed to provide direct or indirect information about measurement errors and the results of some tests of procedures for reducing various kinds of measurement error.

6.1 The seam phenomenon

Use of a panel survey design makes it possible to collect information on changes for individual members of the survey population. At the individual level, SIPP provides information on *transitions*, for example, changes in labor force activity, income recipiency, and income amounts during the period covered by the survey. At the aggregate level, such information provides data on *gross flows*, for example, the number of persons receiving initial or final benefits from a specified program (that is, moving on or off the program) during a given period. Subject to the limitation of the 2⅔ year period for which data are collected for the members of each panel, SIPP also provides information on spells, that is, the duration of a particular state or characteristic for the individual.

Use of a panel design introduces new kinds of measurement error. Response variance, which does not lead to bias in the cross-sectional estimate for a single variable, can cause substantial bias in the kinds of longitudinal estimates described in the preceding paragraph. Changes from proxy to self-response, or vice versa, from one interview to the next are a potential source of measurement error; others are interviewer changes, which are sometimes unavoidable, and data processing errors. This section describes the seam phenomenon, which is a singularly perplexing manifestation of longitudinal measurement error. Other kinds of longitudinal errors are covered in Section 6.2.

Definition

The seam is the boundary between the 4-month reference periods for interviews in successive waves of a panel. The seam phenomenon is the tendency to overreport changes in status (such as going on or off the Food Stamp program) and in amounts received (such as value of Food Stamps) between adjacent calendar

months included in the reference periods for different interviews (such as the months 4-5, 8-9, 12-13, etc., in a panel), and to underreport changes between adjacent months covered by the reference period for a single interview.

The seam phenomenon affects most variables for which monthly data are collected in SIPP. It was observed in the ISDP (Moore and Kasprzyk, 1984) and has been extensively documented in SIPP (e.g., Burkhead and Coder, 1985; Hill, 1987; Kalton et al., 1990; Kalton and Miller, 1991). A similar phenomenon has been observed in the Panel Study of Income Dynamics (Hill, 1987). For some variables the size of the seam effect is substantial. Tables 6.1 to 6.3 show the size of the seam effect for recipiency of earnings, Social Security benefits and Food Stamps, respectively, in the first three waves of the 1984 Panel. All three tables show a clear clustering of changes in recipiency status at the seam, that is, the number of transitions at the on-seam months (that is, 4th to 5th month, and 8th to 9th month) are much higher than the number at the off-seam months. Other tables and charts (see Burkhead and Coder, 1985; Coder et al., 1987a, Appendix J) provide clear evidence of this phenomenon for many labor force and income items.

Causes are unknown

The main design features associated with the seam phemomenon have not been identified with certainty, but it has been suggested that questionnaire wording and design, length of recall, and the interaction between them, have an important role (Kasprzyk, 1988). In general, research on this phemomenon with recipiency items found no association with the characteristics of the respondents, edits and imputations, proxy versus self- response, and changes in interviewer assignments (Lepkowski et al., 1990; Moore and Marquis, 1989; Marquis and Moore, 1990). In one study, Kalton and Miller (1991) took advantage of a 3.5 percent increase in Social Security payments that began in January 1984 to examine the reporting of the December 1983 to January 1984 change in payment amounts. Their analysis included only three rotation groups that collected data for December and January in the same wave. Using a logistic regression analysis, they found that the significant predictors of reporting this change in January were: self- versus proxy reporting, race, the amount of the January payment, and rotation group (change is reported most often by the group interviewed in February and least often by the group interviewed in April).

Some cognitive research on the processes of responding in SIPP suggests a form of inertia in reporting. This research found that SIPP respondents often adopt simple rules for determining their responses instead of engaging in more strenuous mental processes to recall the information (Cantor et al., 1991; Marquis et al., 1991). As a result, some respondents who experience a transition in the middle of a reference period may simply report their current status for all 4 months of the reference period. Thus, a respondent interviewed in May, who entered the AFDC program in March, may report receipt of AFDC for January and February as well as for March and April. Comparison of these reports with reports of no AFDC participation in the previous January's interview would (erroneously) date the transition at the seam between the two interviews.

Table 6.1 SIPP month-to-month transitions for persons by interview status[1]: earnings

Interview status	Month-to-month											Mean
	1st to 2nd	2nd to 3rd	3rd to 4th	4th to 5th	5th to 6th	6th to 7th	7th to 8th	8th to 9th	9th to 10th	10th to 11th	11th to 12th	
NUMBER												
Receiving both months												
Total	18,854	18,934	19,089	18,455	19,172	19,135	19,166	18,536	19,377	19,388	19,345	19,041
Self-interview	8,583	8,662	8,783	8,470	8,776	8,790	8,798	8,519	8,804	8,857	8,699	8,695
Other	10,271	10,272	10,306	9,985	10,396	10,345	10,368	10,017	10,573	10,631	10,646	10,346
Not receiving to receiving												
Total	639	726	662	1,118	432	515	499	1,136	500	533	568	666
Self-interview	334	369	308	499	228	241	244	455	211	239	255	308
Other	305	357	354	619	204	274	255	681	289	294	313	359
Receiving to not receiving												
Total	539	559	571	1,296	401	469	484	1,129	295	489	576	619
Self-interview	263	255	248	621	193	214	233	523	170	258	297	298
Other	276	304	323	675	208	255	251	606	125	231	279	321
INDEX												
Receiving both months												
Total	100.0	100.4	101.2	97.9	101.7	101.5	101.7	98.3	102.8	102.8	102.6	(X)
Self-interview	100.0	100.9	102.3	98.7	102.2	102.4	102.5	99.3	102.6	102.0	101.4	(X)
Other	100.0	100.0	100.3	97.2	101.1	100.7	100.9	97.5	102.9	103.5	103.7	(X)
Not receiving to receiving												
Total	100.0	113.6	103.6	175.0	67.6	80.6	78.1	177.8	78.2	83.4	88.9	(X)
Self-interview	100.0	110.5	92.2	149.4	68.3	72.2	73.1	136.2	63.2	71.6	76.3	(X)
Other	100.0	117.0	116.1	203.0	66.9	89.8	83.6	223.3	94.8	96.4	102.6	(X)
Receiving to not receiving												
Total	100.0	103.7	105.9	240.4	74.4	87.0	89.8	209.5	54.7	90.7	106.9	(X)
Self-interview	100.0	97.0	94.3	236.1	73.4	81.4	88.6	198.9	64.6	98.1	112.9	(X)
Other	100.0	110.1	117.0	244.6	75.4	92.4	90.9	219.6	45.3	83.7	101.1	(X)

(X) Not applicable.

[1]Persons interviewed first 12 months in sample.
Source: Burkhead and Coder (1985).

Table 6.2 SIPP month-to-month transitions for persons by interview status[1]: Social Security

Interview status	Month-to-month											Mean
	1st to 2nd	2nd to 3rd	3rd to 4th	4th to 5th	5th to 6th	6th to 7th	7th to 8th	8th to 9th	9th to 10th	10th to 11th	11th to 12th	
NUMBER												
Receiving both months												
Total	6,447	6,484	6,520	6,473	6,620	6,647	6,686	6,650	6,766	6,783	6,795	6,625
Self-interview	4,484	4,511	4,530	4,502	4,584	4,604	4,629	4,616	4,678	4,689	4,705	4,594
Other	1,963	1,973	1,990	1,971	2,036	2,043	2,057	2,034	2.088	2,094	2,090	2,031
Not receiving to receiving												
Total	58	51	58	157	42	52	37	134	33	28	32	62
Self-interview	40	31	36	91	29	32	21	68	20	20	17	37
Other	18	20	22	66	13	20	16	66	13	8	15	25
Receiving to not receiving												
Total	22	21	15	105	10	15	13	73	18	16	16	29
Self-interview	11	13	12	64	9	9	7	34	6	9	4	16
Other	11	8	3	41	1	6	6	39	12	7	12	13
INDEX												
Receiving both months												
Total	100.0	100.6	101.1	100.4	102.7	103.1	103.7	103.1	104.9	105.2	105.4	(X)
Self-interview	100.0	100.6	101.0	100.4	102.2	102.7	103.2	102.9	104.3	104.6	104.9	(X)
Other	100.0	100.5	101.4	100.4	103.7	104.1	104.8	103.6	106.4	106.7	106.5	(X)
Not receiving to receiving												
Total	100.0	87.9	100.0	270.7	72.4	89.7	63.8	231.0	56.9	48.3	55.2	(X)
Self-interview	100.0	77.5	90.0	227.5	72.5	80.0	52.5	170.0	50.0	50.0	42.5	(X)
Other	100.0	111.1	122.2	366.7	72.2	111.1	88.9	366.7	72.2	44.4	83.3	(X)
Receiving to not receiving												
Total	100.0	95.5	68.2	477.3	45.5	68.2	59.1	331.8	81.8	72.7	72.7	(X)
Self-interview	100.0	118.2	109.1	581.8	81.8	81.8	63.6	309.1	54.5	81.8	36.4	(X)
Other	100.0	72.7	27.3	372.7	9.1	54.5	54.5	354.5	109.1	63.6	109.1	(X)

(X) Not applicable.

[1]Persons interviewed first 12 months in sample.

Source: Burkhead and Coder (1985).

Table 6.3 SIPP month-to-month transitions for persons by interview status[1]: Food Stamps

Interview status	Month-to-month											Mean
	1st to 2nd	2nd to 3rd	3rd to 4th	4th to 5th	5th to 6th	6th to 7th	7th to 8th	8th to 9th	9th to 10th	10th to 11th	11th to 12th	
NUMBER												
Receiving both months												
Total	1,323	1,335	1,359	1,224	1,357	1,363	1,376	1,243	1,347	1,343	1,341	1,328
Self-interview	975	984	999	930	1,004	1,008	1,012	942	1,002	1,000	995	986
Other	348	351	360	294	353	355	364	301	345	343	346	342
Not receiving to receiving												
Total	69	75	72	162	49	60	57	140	39	32	48	73
Self-interview	42	52	48	95	26	37	35	83	31	21	27	45
Other	27	23	24	67	23	23	22	57	8	11	21	28
Receiving to not receiving												
Total	52	57	51	207	29	43	47	190	36	43	34	72
Self-interview	37	33	37	117	21	22	33	105	23	33	26	44
Other	15	24	14	90	8	21	14	85	13	10	8	27
INDEX												
Receiving both months												
Total	100.0	100.9	102.7	92.5	102.6	103.0	104.0	94.0	101.8	101.5	101.4	(X)
Self-interview	100.0	100.9	102.5	95.4	103.0	103.4	103.8	96.6	102.8	102.6	102.1	(X)
Other	100.0	100.9	103.4	84.5	101.4	102.0	104.6	86.5	99.1	98.6	99.4	(X)
Not receiving to receiving												
Total	100.0	108.7	104.3	234.8	71.0	87.0	82.6	202.9	56.5	46.4	69.6	(X)
Self-interview	100.0	123.8	114.3	226.2	61.9	88.1	83.3	197.6	73.8	50.0	64.3	(X)
Other	100.0	85.2	88.9	248.1	85.2	85.2	81.5	211.1	29.6	40.7	77.8	(X)
Receiving to not receiving												
Total	100.0	109.6	98.1	398.1	55.8	82.7	90.4	365.4	69.2	82.7	65.4	(X)
Self-interview	100.0	89.2	100.0	316.2	56.8	59.5	89.2	283.8	62.2	89.2	70.3	(X)
Other	100.0	160.0	93.3	600.0	53.3	140.0	93.3	566.7	86.7	66.7	53.3	(X)

(X) Not applicable.

[1]Persons interviewed first 12 months in sample.
Source: Burkhead and Coder (1985).

Effects on estimates

The effect of the seam phenomenon on uses of SIPP data depends on the kinds of analyses for which the data are being used. Because of SIPP's rotation pattern, only one-fourth of the sample is at the seam for any given pair of calendar months. If it were assumed that transitions are overreported at the seam and underreported elsewhere, one might expect the net effect on cross-sectional estimates to be small. To test this expectation, transition rates based on estimates from the SIPP 1984 Panel have been compared with rates derived from administrative records for the Food Stamp, AFDC, and Supplemental Security Income programs (Singh, 1987). Results of the comparisons for the three programs are shown in Table 6.4. A summary of the findings is given by Singh et al. (1989). They suggest that effects for each program should be evaluated separately. Estimates for Food Stamp transitions from the two sources agreed fairly well. For AFDC, the results were inconclusive because of small sample sizes. For Supplemental Security Income, entrance rates (transitions to recipiency) from SIPP were significantly higher than those shown by program records.

Table 6.4 Comparison of survey and administrative data on entrance and exit rates for selected benefit programs: 1984 Panel[1]

Program	Type of rate	Year and transition months							
		1983				1984			
		S/O	O/N	N/D	Qtr. Av.	D/J	J/F	F/M	Qtr.
Food Stamps	Entrance rate								
	Admin. records	6.2	6.7	5.0	6.0				
		(0.5)	(0.6)	(0.5)	(0.4)				
	SIPP	6.9	5.8	5.0	5.9				
		(1.1)	(1.0)	(0.9)	(0.6)				
	Exit rate								
	Admin. records	7.0	6.1	5.1	6.1				
		(0.6)	(0.5)	(0.5)	(0.4)				
	SIPP	5.3	5.0	4.9	5.1				
		(1.0)	(1.0)	(0.9)	(0.6)				
AFDC	Entrance rate								
	Admin. records	4.8	4.8	4.8	4.8	4.6	4.5	4.5	4.5
	SIPP	4.1	4.2	4.5	4.3	4.0	3.9	4.4	4.1
		(1.3)	(1.3)	(1.3)	(0.8)	(1.3)	(1.2)	(1.3)	(0.7)
	Exit rate								
	Admin. records	4.6	4.6	4.5	4.6	4.2	4.2	4.2	4.2
	SIPP	3.7	2.8	2.3	2.9	3.0	3.4	3.4	3.3
		(1.2)	(1.1)	(1.0)	(0.6)	(1.1)	(1.2)	(1.2)	(0.7)
SSI	Entrance rate								
	Admin. records	1.0	0.9	0.9	0.9	0.9	1.0	1.0	1.0
	SIPP	2.8	2.1	1.4	2.1	3.1	1.9	2.9	2.6
		(1.0)	(0.9)	(0.7)	(0.5)	(1.1)	(0.8)	(1.0)	(0.6)

[1]Standard errors of estimates are shown in parentheses.

Source: Singh (1987).

Less is known about how the seam phenomenon affects analyses based on SIPP longitudinal data. Since correlations between variables underlie most kinds of causal modeling and multivariate analysis, it is reasonable to examine the effects of the seam phenomenon on correlations. Young (1989), using the 1984 Full Panel Longitudinal Research File, has studied the effects of the seam phenomenon on estimates of correlation between transitions reported for pairs of variables. This study compared the estimated correlations for the first (seam transition) month of each wave with the average of those based on the second, third, and fourth months. Some results are shown in Table 6.5. Of the 21 pairs formed from the 7 variables studied, 13 exhibited statistically significant differences between the correlation for month 1 of a wave and the average correlation for months 2, 3, and 4. The largest observed differences were for the correlations of employment status with personal earnings and with total family income.

Table 6.5 Correlation between occurrence of monthly transitions, by wave month

Variables	Month	Marital status	Employ-ment status	Personal earnings	Total family income	Social Security	Family AFDC receipt	Family Food Stamp receipt
Marital status	1	1.000	0.023	0.027*	0.027	-0.014*	-0.004	0.012
	2, 3, 4	1.000	0.014	0.009	0.033	-0.001	0.004	0.011
Employment status	1		1.000	0.380*	0.172*	-0.058*	0.048*	0.073*
	2, 3, 4		1.000	0.486	0.114	-0.004	0.029	0.039
Personal earnings	1			1.000	0.463	-0.097*	0.018	0.032
	2, 3, 4			1.000	0.497	-0.006	0.022	0.034
Total family	1				1.000	-0.055*	0.054	0.050
income	2, 3, 4				1.000	0.008	0.063	0.052
Social Security	1					1.000	-0.015*	-0.009*
	2, 3, 4					1.000	-0.001	-0.004
Family AFDC	1						1.000	0.162*
Receipt	2, 3, 4						1.000	0.200

*Indicates significant difference for a .01 two-tailed probability level.
Source: Young (1989).

Kalton et al. (1992) analyzed spell durations for various welfare programs and for being without health insurance in the 1987 Panel. Their results showed that across all the analyses, spell durations of multiples of 4 months (e.g., 4 months, 8 months, 12 months) were particularly common, a feature that can be explained by the seam effect. Spells of veteran's compensation exhibited the greatest effect, with about four out of five spells lasting for exactly 4, 8, or 12 months.

Efforts to reduce seam effects

- *Changes to questionnaire wording*. In an attempt to reduce the magnitude of seam effects, several operational and experimental changes have been introduced in SIPP instruments and procedures (Petroni et al., 1989). Two changes were made to the questionnaire at the start of the 1988 Panel. The first change applies to persons who reported receipt of an income type during the prior wave interview and do not report it for the current wave. Such persons are asked to report the last month in which that type of income was received in order to verify that it was not received during the reference period for the current interview. Second, persons who report receipt of an income type for the current reference period but did not report it at the previous interview are asked to state the month in which the income was first received. An analysis comparing transitions reported for Wave 5 of the 1987 Panel and Wave 2 of the 1988 Panel did not detect any improvement in the accuracy of reporting transitions.

- *Training of FRs*. In January 1989 a description of the seam problem and its effects was included in the interviewer training package and interviewer comments and suggestions were solicited. Interviewer training is now placing greater emphasis on data quality and the advance letters to respondents also emphasize quality issues.

- *Comparison with independent sources*. An evaluation study that matched individual data from SIPP interviews with program records for nine Federal and State programs in four States has produced direct information on the nature of measurement errors in SIPP, including those that affect reports of transitions (Marquis and Moore, 1989a, 1989b). This study and some of its results are discussed in Section 6.3.

- *Event history calendar*. The Census Bureau studied the effectiveness of an event history calendar in an effort to reduce the seam effect. An event history calendar of program participation and basic demographic information was developed and used in the Chicago region for the 1989-1991 Panels. In this study, respondents received a copy of the calendar at the beginning of the panel to record month-by-month participation and income from program sources. They were asked to refer to their calendars during the interview to facilitate recall. The FRs also kept a copy of the calendar which was updated after each interview. Kominski (1990) reported the results from the 1989 Panel (which ended after three waves). The basic conclusion was that the overall level of transitions at the seam was reduced by using the calendar, while the within-wave level was only marginally affected. As a result, the overall level of transitions dropped slightly. For the 1990 and 1991 Panels, the calendar completed by the FR was presented to the respondent at the following interview to serve as a reminder of previous responses. Since the calendar contained reported amounts of income, it was not shown to proxy respondents who had not also been the proxy in the previous interview. Due to budget constraints, the data from the 1990 and 1991 Panels have not been fully analyzed nor has the idea of using a calendar been pursued for later panels.

6.2 Other longitudinal measurement errors

6.2.1 Consistency

Measurement errors can have a particularly serious effect on longitudinal analysis of data from panel surveys, as seen in the estimation of wave-to-wave changes. Differential measurement error in successive waves can give the appearance of change, when no actual change occurred. Kalton et al. (1986) give several examples of errors in measuring change in SIPP. One example concerned items such as sex, race, and age that should remain constant between panels, or change only in predetermined ways. For each of the first four pairs of waves of the 1984 Panel, between-wave changes were observed in race for about 0.10 percent of all persons and in sex for about 0.08 percent. Changes of 2 or more years in age were somewhat more frequent,

occurring for about 0.4 percent of persons between Waves 2 and 3. The reasons for these changes are not known, but they are believed to result from some combination of interviewer, data entry, and processing errors.

Another example cited by Kalton et al. (1986) concerns the consistency of occupation and industry codes between interview waves. During the 1984 and 1985 panels, occupation and industry data were collected and coded independently at each wave, even when the individual's employer and duties had not changed. Examination of data for persons who reported the same employer for all of the first 12 months of the 1984 panel showed that only about 50 percent had the same 3-digit occupation code throughout and only about 70 percent had the same 3-digit industry code in all three waves. Although changes in occupation and industry (especially the former) are possible without a change in employer, on a priori grounds one would expect greater stability. It was believed that a high proportion of these changes were spurious, and resulted from data collection or processing errors. As described in Section 4.3, a dependent procedure for collecting occupation and industry data was introduced at the start of the 1986 Panel.

6.2.2 Time-in-sample bias

Some of the measurement errors that occur in panel surveys are associated with the duration and number of interviews for a sample panel. One possibility is that panel members change their behavior as a result of participating in the survey. This would cause a bias in that the expected values of estimates based on the second and subsequent waves differ from the values for the entire survey population. In debriefing interviews of 462 SIPP respondents from the 1985 Panel who had completed their final interviews, 2.2 percent reported that they had learned of a government program through SIPP and then applied for it (Petroni et al., 1989).

In addition to changing their behavior with respect to topics covered by the survey, sample persons and households may change their behavior with respect to participation and reporting in the survey. Participation in interviews is a learning process which can lead to a lesser or greater desire to participate in subsequent interviews and, for those who continue to participate, improvement or deterioration in the accuracy of their reporting.

Research on time-in-sample bias

The SIPP design makes it possible to study effects of time-in-sample (or panel conditioning) bias in various ways. Because of the pattern of waves and rotation groups within a panel, it is possible, for much of the length of a panel, to compare estimates from different waves that refer to the same calendar month. For example, referring to Figure 2.1 in Section 2.3, it can be seen that separate estimates for March 1991 will be available from rotation groups 1 and 4 in Wave 1 and from rotation groups 2 and 3 in Wave 2 of the 1991 Panel. Chakrabarty (1988, 1989b) has performed this type of comparison with data from the 1984 Panel. He detected moderate effects of time-in-sample bias for some labor force activity items, but little if any effect on income recipiency and amount items. On the other hand, Hill (1986), as described in Section 4.5.1, found a statistically significant time-in-sample effect on earnings reported for employed men ages 25 to 55. The scope of both analyses was limited to biasing effects between successive waves. They were not capable of measuring time-in-sample effects over the life of a panel.

The SIPP design also allows comparison of estimates for the same calendar month from different panels. For example, estimates for March 1992 are available from Waves 4 and 5 of the 1991 Panel and from Waves 1 and 2 of the 1992 Panel. Comparisons between panels permit evaluation of the effects of time-in-sample bias over longer periods, subject to the limitation that some observed differences may result from

questionnaire and procedural changes from one panel to the next. Studies that have compared estimates from different panels include Lepkowski et al. (1990), Pennell and Lepkowski (1992), McNeil (1990, 1991), and McCormick et al. (1992). Lepkowski et al. (1990) compared estimates of program participation, income from programs, and employment using Wave 4 of the 1984 Panel and Wave 1 of the 1985 Panel. They found significant differences in reported levels of AFDC income and in reports of unemployment, but no significant differences for estimates for other programs, such as Social Security, and Food Stamps.

Pennell and Lepkowski (1992) compared calendar year estimates from different panels on income recipiency from various sources, income amounts, health insurance coverage, and labor force participation. They found no evidence that the 1985 and the 1986 Panels produced different calendar year estimates for 1986. When monthly estimates from the 1985 and the 1986 Panels, and the 1986 and 1987 Panels were compared for January, May, and September, some significant differences were found. For example, state unemployment compensation recipiency was higher for the 1987 Panel than the 1986 Panel in all three reference months in 1987. However, there was no significant difference between the recipiency data for the 1985 and 1986 Panels. The comparisons were repeated using only persons who responded to all seven waves of the panel, in an effort to remove confounding effects due to sample attrition. These comparisons yielded no evidence of significant panel effects.

McNeil (1989, 1990, and 1991) and McCormick et al. (1992) compared SIPP quarterly estimates for the 1984-1987 Panels and found no evidence that time-in-sample affects estimates of income, poverty, labor force participation, health insurance coverage, or program participation status. McCormick et al. note significant differences were often observed when comparing quarter 1 estimates across panels. They attributed this to a possible seasonal effect.

Citro and Kalton (1993:104-107) provide a review of time-in-sample effects, examining evidence from the SIPP and from other surveys. They note that, because of the difficulty of distinguishing between time-in-sample effects and other changes across waves, such as attrition, it is difficult to draw firm conclusions about the extent or level of time-in-sample biases. The available evidence, however, suggests that the effects are limited.

6.3 Evaluation studies and experiments

This section describes several special studies and experiments that were designed to obtain a better understanding of measurement errors in SIPP and to test methods for reducing their effects. Sections 6.3.1 to 6.3.3 discuss experiments undertaken with ongoing or expired panels. Section 6.3.4 describes a record check study that estimated measurement error by matching survey reports from the first two waves of the 1984 Panel against administrative record information for the same individuals. Section 6.3.5 describes several tests of experimental survey data collection procedures, based on cognitive theory and research, aimed at reduction of reporting errors. These tests were conducted "off line," that is, the households participating were not part of ongoing SIPP panels.

6.3.1 The Asset Feedback Experiment

In the 1984 Panel, an experiment was undertaken to determine whether reporting of asset amounts could be improved by reminding respondents of the amounts they had reported in a prior interview. In that panel, information on ownership of financial assets was collected in every wave, but information on asset and liability amounts was collected only in fixed topical modules in Waves 4 and 7. (This pattern was followed for the first four panels, but starting with the 1988 panel, the data on amounts were collected in only one wave for each panel). Since the interviews in Waves 4 and 7 were separated by 12 months, it was possible

to use the reported data to make both cross-sectional and longitudinal estimates of net changes in assets over this period. However, both kinds of estimates, especially the latter, may be significantly affected by the kinds of measurement error discussed earlier in this section.

The Asset Feedback Experiment was conducted in Wave 7, using a split-panel design. Sample segments were assigned either to the experimental or control group depending on whether the ending digits of their identification numbers were odd or even. For the experimental half of the sample households, a computer-generated form was prepared for each sample person showing amounts reported 12 months earlier for key asset types. To avoid violating confidentiality rules, these feedback forms were used only when the person responding in Wave 7 was the one who had responded in Wave 4. During the administration of the asset module, respondents were given their feedback forms and asked to refer to their earlier responses before providing current information for each asset type.

Analysis of the results was complicated by several factors: changes in household structure between Waves 4 and 7, a relatively high level of imputation in both waves, and high levels of sampling variability resulting from the effects of lump-sum (as opposed to incremental) changes in assets. A more fundamental problem was the absence of any external measure of the validity of individual reports, so that even when statistically significant differences between the experimental and control groups were observed, it could not be established with certainty which estimates were of better quality.

Detailed descriptions of the experimental procedures and analyses are given by Lamas and McNeil (1987) and Weidman et al. (1988). The former summarized their conclusions in the following:

> "When we examined estimates of change based on cross-sectional estimates of mean or median net worth, we found few changes which were statistically significant for the feedback or control group. We also examined microlevel changes in net worth using only households with fully reported wealth data. We found some evidence that the feedback approach reduced the estimates of the change ... It is possible that the technique of providing previously reported data to respondents during the interview may lead respondents to give more careful consideration in their answers. However, the results also suggest that a one year time period between the point estimates may be too close to measure changes in net worth."

Weidman et al. (1988) concluded that "... the data ... do not give any statistical evidence of consistent differences in the measure of annual changes in asset and liability amounts due to the use of the feedback procedure." Summaries of the experimental procedures and findings are also given by Carmody et al. (1988) and Petroni et al. (1989). The feedback procedure for asset amounts has not been used subsequent to the 1984 Panel.

6.3.2 Respondent debriefing interviews

In 1987, respondent debriefing interviews were conducted with 462 respondents in rotation groups 2, 3, and 4 in conjunction with scheduled reinterviews following the final wave interviews of the 1985 Panel. The topics covered in the debriefing interviews included: respondents' use of records during interviews, learning about benefit programs through participation in SIPP, and reasons for continued participation in SIPP. Findings on record use were discussed in Section 4.5.6 and those on effects of learning through participation in Section 6.2.2.

The two main reasons given by respondents for continued participation were that they liked the interviewer(s) (27.8 percent) and that they felt it was their patriotic duty to participate (22.6 percent). In response to a request for other comments and suggestions at the end of the debriefing interview, the most frequent comment was that the interviewers were of good caliber. Additional results and discussion are given by Carmody et al. (1988) and Petroni et al. (1989).

6.3.3 The Eighth Interview Study

In the late summer of 1988, a pilot study known as the Eighth Interview Study was conducted to test the feasibility of interview procedures designed to improve the accuracy of reporting on changes in income recipiency and amounts and, in particular, to reduce the effect of the seam phenomenon. The study included two experimental groups and a control group. About 100 Wave 8 interviews were conducted for each of the three groups. The regular Wave 8 interviews for the 1986 Panel had been eliminated due to budget constraints, so the households for the pilot study were selected purposively from one rotation group, using criteria designed to increase the likelihood of transitions in the sample households.

One of the experimental treatments (procedure W) attempted to improve the quality of reporting within the 4-month reference period, without regard to prior responses. The procedure emphasized attempts to encourage greater use of records by respondents, the use of probes when reported amounts did not change over the 4-month reference period, and positive feedback statements by interviewers whenever changes were reported. The other experimental treatment (procedure B) attempted to reduce the number of reported changes between adjacent interview periods. This treatment included attempting to interview the same respondent who was interviewed in Wave 7, giving the respondent a feedback form with a summary of previous responses, and using special probes when the respondent's answers for the current reference period indicated one or more changes at the seam.

To obtain information on the feasibility and perceived utility of the experimental procedures, four interviewer debriefing sessions were held in September 1988. Detailed notes from these sessions are provided by Chapman (1988). Most of the interviewers felt that procedure W was very similar to standard interview procedures. One finding of interest was that interviewers from the Seattle region reported substantially higher respondent use of records than was reported by interviewers from other regions.

With respect to procedure B, at least half of the interviewers felt that the feedback form was useful. However, some interviewers felt that its use might affect some respondents' perceptions of the confidential treatment given their information, with potentially unfavorable effects on cooperation rates. The use of special probes when seam changes were reported was viewed favorably by most interviewers, and there were a few instances of incorrectly reported seam changes being identified as a result of the probes. When asked for general comments on SIPP, some interviewers recommended that self-response be required for Wave 1, and interviewers from the Seattle region said that this was already their practice. Most participants thought the debriefings useful and expressed interest in participating in future sessions.

6.3.4 The SIPP Record Check Study

In the SIPP Record Check Study, survey data on program participation and benefit amounts were matched against administrative records for the same programs. By treating data from the latter source as true values, estimates of the levels of response bias and response error variance were obtained for each variable. These estimates have been used to test hypotheses about sources of nonsampling error and to guide efforts to reduce the levels of error.

The study was based on a full record check design, that is, it could detect both false positive and false negative survey reports of program participation. To make this possible, the Census Bureau obtained the complete program data files for the States and time periods covered by the study and performed all of the linkages, using an elaborate combination of computer and manual matching procedures. The key matching variable was the Social Security number (SSN), which had been obtained and verified for about 95 percent of SIPP sample persons. The SSN was supplemented by other identifiers, including name, address, sex, and date of birth, so the overall quality of the matching was considered to be high.

The study covered four States--Florida, New York, Pennsylvania, and Wisconsin--and eight Federal and State benefit programs, as follows:

> **State-administered programs**: Aid to Families with Dependent Children (AFDC), Food Stamps, Unemployment Insurance, and Workers' Compensation.
>
> **Federal programs**: Federal Civil Service Retirement, Supplemental Security Income, Social Security Retirement, and Veterans' Pensions and Compensation.

Since there were some unresolved questions about the quality of selected data fields in the administrative files for the State-administered programs for the State of New York, only the Federal programs were included in the analysis for New York.

The SIPP data came from Waves 1 and 2 of the 1984 Panel, covering 11 months from June 1983 through April 1984 (for each rotation group there were 8 months of data). Most of the comparisons of SIPP and program data were made at the individual level, except for AFDC and Food Stamps, for which both individual and family benefit units could be compared. The SIPP sample for the four States combined was approximately 5,000 households. Not all of the approximately 11,000 persons in these households were included in the Record Check Study; children less than 15, adults who did not provide their Social Security numbers in the survey, and persons for whom data were not available for the full 8 months were excluded. Overall, the number of sample persons included in the analyses was about 7,550 for Federal programs and 5,200 for State programs.

The design and results of the Record Check Study have been described in several papers, for example, Marquis and Moore (1989a, 1989b, 1990) and Marquis et al. (1990). Table 6.6 shows Record Check estimates of response bias (in percent) in SIPP person-based estimates of the *level of participation* in the eight Federal and State benefit programs. All of the estimated biases are the net result of both under and overreporting. For all but one of the programs, Social Security retirement, the relative net bias is negative, denoting that the estimated proportion of participation from the survey is below the true proportion derived from program records. Biases for some programs were substantial, the highest being for the AFDC program, with a 39 percent net underestimate. Biases for some other programs were small, such as the 1 percent overestimate for Social Security retirement and the 3 percent underestimate for veterans' benefits.

Table 6.7 shows Record Check estimates of response bias (in percent) in SIPP person-based estimates of *month-to-month change* for six of the eight benefit programs. Two of the programs were omitted because, according to the program records for persons in the sample, no changes occurred for any of the pairs of months included in the study. Estimates of bias are shown separately for changes occurring within waves or "off-seam" and those occurring between successive waves or "on-seam." Estimates of all but one of the off-seam biases are negative, signifying underreporting of change in the survey. Estimates of bias for changes occurring at the seam are all positive, signifying overreporting, ranging from 20 to more than 200 percent.

Table 6.6 SIPP Record Check: Percent net bias in estimates of program participation

Program	Percent bias[1]
Social Security retirement	1
Veterans' benefits	-3
Civil Service retirement	-8
Supplemental Security Income	-12
Food Stamps	-13
Workers' compensation	-18
Unemployment insurance	-20
AFDC	-39

[1]Negative bias indicates net underreporting.
Source: Adapted from Marquis et al. (1990).

Table 6.7 SIPP Record Check: Percent net bias in estimates of change in program participation

Program	Percent bias	
	Off-seam	On-seam
Workers' compensation	3	80
Social Security retirement	-6	132
Food Stamps	-32	135
Unemployment insurance	-32	28
Supplemental Security Income	-34	>200[1]
AFDC	-64	20

[1]Value uncertain, see reference cited as source of table.
Source: Adapted from Marquis et al. (1990).

Analysts of SIPP data on program participation are interested in the relationships of levels of participation and changes in program participation (going "on" and "off" programs) with other variables, such as age, sex, and education. The record check study found that SIPP response errors have substantial downward biasing effects on estimates of correlation between program participation and other variables. For correlations involving SIPP measures of level, the estimated biases ranged from -10 to -50 percent for six of the eight programs. For measures of change, they ranged from -50 to -100 percent. (This result assumed that the other variables had been measured without error. For further details about the estimation model see Marquis and Moore, 1990, Appendix.)

Several kinds of analyses of the Record Check data were undertaken to seek a better understanding of the causes of these often substantial biases in the kinds of estimates that are of primary interest to many users of SIPP data (Marquis and Moore, 1990). These analyses did not identify or suggest any single cause accounting for most, or even a large proportion, of the errors. Unemployment Insurance was the only program for which there was statistically significant evidence of memory decay during the 4-month reference period. A small number of errors were caused by respondents' confusion about program names and, for Food Stamps, the name of the official recipient within the family. There was some evidence that experience in

the initial interview may have led respondents to underreport Unemployment Insurance and Workers' Compensation in the second interview in order to avoid detailed questioning about these types of benefits.

Bollinger and David (1993), using results from the Record Check Study, explored the effect of errors in reporting Food Stamp program participation on estimation of a model that related participation to variables based on individual and family earnings and assets, number of children, disability, and other survey items. Their analysis proceeded in two steps. First, a model was developed to predict the probability of response error in reporting participation on the basis of variables such as gender, marital status, and family income. Then, the estimates of error probability developed from this model were used to correct the basic participation model for response error. Comparisons of estimated model coefficients from the uncorrected participation model with corresponding coefficients for several alternative corrected models suggest that Food Stamp participation is significantly more sensitive to changes in wages and assets than is suggested by the uncorrected model.

In a continuation of this study, Bollinger and David (1995) added a new variable representing the response pattern of the family unit in Waves 3 through 8 to the probit model described above. They found a strong and stable relationship between reporting error and nonresponse. Those responding units that continue to give interviews are less error prone; and poor reporters are more likely to miss interviews and to drop out of the survey (attritors). Bollinger and David (1996) extended the model of the effects of response error on estimates of Food Stamp participation to a trivariate probit model of errors of omission in reporting participation, errors of commission in reporting participation, and interview nonresponse. The results of their research suggest that estimates of the models of Food Stamp participation should be adjusted for response error controlling for interview nonresponse.

6.3.5 The SIPP Cognitive Research Evaluation Experiment

Findings from the SIPP Record Check and other studies demonstrated a need to reduce the levels of response error in the reporting of program participation and related variables, but it was not at all obvious how this could be accomplished. Since traditional methods of reducing response error did not appear promising, Census Bureau research staff turned to recently developed models of the cognitive aspects of survey interviews and to some research tools borrowed from cognitive psychologists. It was clear that SIPP interviews place heavy recall and information processing demands on respondents. A better understanding of the cognitive features of the interviews might suggest how interviewing procedures could be redesigned to achieve desired improvements in the quality of response. This research effort lasted about 5 years, starting in 1990, and began with a series of simulated SIPP interviews that included follow-up questions by trained observers. The results of these exploratory interviews led to a proposal for a substantial redesign of the interview structure and goals. The proposed procedures were then tested in a controlled experiment that included a record check to evaluate the levels of response error for the experimental and control treatments.

In the spring and summer of 1990, approximately 50 exploratory interviews were conducted by teams consisting of a regular SIPP interviewer and one of 13 observers from the Census Bureau staff who had received training in cognitive "think-aloud" and hypothesis testing methods. The staff observers were free to pursue any of the survey topics, but for the most part focused on questions about program participation and amounts received. They sought, through their observations and follow-up questions, to gain a better understanding of respondents' problems of comprehension, recall, and response formulation for questions on these topics. All interviews were conducted in respondents' homes and were audiotaped. Each participating household received an incentive payment of $15.

Detailed findings from these interviews are given by Marquis (1990). A primary finding was that respondents, although generally well-intentioned, could seldom recall participation dates and income flows for the 4-month reference period directly from memory, generally substituting simple and sometimes inappropriate estimation procedures in place of detailed recall or record use. Amounts received were often reconstructed based on application of a current or typical amount to prior weeks or other time periods. Records were used infrequently and even when they were, they were sometimes used only as a starting point for completing the task by other means. Normal interviewer strategies were designed to achieve high response rates and high levels of efficiency as measured by the length of time taken to complete their assigned interviews.

Further support to these findings was provided in a series of exploratory SIPP Wave 1 and Wave 2 interviews conducted a few months later by Westat, a contractor to the Census Bureau (Cantor et al., 1991; Cantor et al., 1992). In the first wave, 125 interviews were conducted in the Washington, DC area with respondents recruited from several sources. Half of the interviews used the standard format for SIPP interviews; the other half included cognitive think-aloud procedures during the interview and a detailed debriefing of each respondent after the interview. Wave 2 interviews were conducted with 76 of the 125 respondents from the first wave. The procedures for the Wave 2 interviews were similar to those used for the experimental group in the Wave 1 interviews, with the addition of an attempt to reconcile reported changes occurring at the seam, that is, between the end of the reference period for the first wave and the beginning of the reference period for the second wave. In each wave, some of the interviews were conducted at Westat's focus group facility; others were conducted in the respondents' homes or at a job-training site where some of the respondents were receiving training. All interviews were audiotaped and those conducted at Westat's facility were also videotaped. All respondents received a $15 incentive payment.

Based on the findings from these exploratory interviews and on the goal of reducing underreporting of participation in selected programs by 25 percent, an experimental interviewing procedure was developed. It placed high priority on achieving more accurate response, even though doing so might increase costs or decrease response rates. The experimental treatment involved changes to interviewer training, the questionnaire, interviewing procedures, supervision, and data processing. Key features were:

- Strong emphasis on the use of records for reporting income details.
- Asking for reports of each income payment rather than monthly totals.
- Maximizing the use of self-response for all persons ages 15 and older in the first interview and conducting interviews with all eligible persons in a family.
- Using a less structured questionnaire, with more open-ended items, to collect information about sources and amounts of income.
- Using a special procedure in the Wave 2 interviews to determine the validity of reported changes in participation occurring "at the seam," that is, between interviews.
- Conducting interviews in distraction-free settings.
- Making callbacks as needed to achieve these objectives.
- Providing interviewers with monthly feedback on their use of the experimental procedures. This feedback was based on tape recordings of all interviews and a sample of these interviews were coded for each interviewer.

Prior to the main experiment, there were three tests of the experimental data collection instruments and procedures. The first test consisted of 70 Wave 1 and Wave 2 interviews conducted by three Westat interviewers in the early part of 1991 with a convenience sample of 39 households in the Washington, DC metropolitan area. The content of the interviews was limited to household roster information and income sources and amounts. Particular emphasis was placed on the feasibility of persuading respondents to use their financial records and to agree to audiotape recording of interviews. The main analyses of the test results were based on periodic debriefing sessions with the interviewers and review of a subset of the audiotapes. In general the proposed procedures for encouraging the use of records to improve the accuracy of response and for recording interviews were found to be feasible, and the revised survey instruments worked reasonably well. The results were sufficiently encouraging to proceed with the experiment (Cantor, 1991; Moore et al., 1996).

The next two tests, which were more formal and controlled, were conducted in the medium-sized Midwestern city that was to be the site of the main experiment. The first of these tests focused again on operational details of the interviews and the second tested procedures for the record check component of the main study. The five interviewers who worked in both of these tests were hired, trained, and supervised by the Census Bureau's Kansas City Regional Office. For each test, a sample of 130 addresses was used; however, they differed in that the sample for the first test was selected at random from the test site, with oversampling from poor areas. The sample for the second test was drawn from record systems associated with five income sources: AFDC, Food Stamps, Unemployment Insurance, Supplemental Security Income, or earnings from a specific area employer. Both tests included Wave 1 and Wave 2 interviews, with an abbreviated 2-month reference period for the Wave 2 interviews. Interviewing for the two tests covered the period from August 1991 through March 1992.

The results of these two tests were encouraging in that they confirmed the feasibility of achieving a high rate of record use and of taping most interviews. Indicators of data quality, especially in the second test, were also favorable. However, the feedback system, based on taped interviews, for monitoring interviewer performance did not work as well as hoped for. More significantly, the household response rates were substantially lower than those normally achieved in standard SIPP interviews and the cost per case was about 50 percent higher (Moore et al., 1993; Moore et al., 1996).

The formal SIPP Cognitive Research Evaluation Experiment was designed to provide clear statistical evidence of the effects of the new interviewing procedures on data quality. To focus on reports of participation in selected programs and to facilitate the record check component of the study, the sample addresses were those associated with a sample of persons receiving income from the same five sources that had been used in the final pretest. To achieve a full record check design, capable of detecting both over and underreporting of program participation, survey reports for "extra persons"--persons in interviewed households who were not in the initial sample of persons from the five income sources--were matched against the program records for these five programs.

The sample addresses were randomly assigned to the experimental and control (standard SIPP) interviewing procedures. For Wave 1 interviewing assignments, about 750 addresses were assigned to each group. To spread out the interviewing workload, the assignments were divided into five "rotation groups," one starting in each successive month. Separate interviewing staffs were used for the experimental and control groups. The experimental interviewing staff was larger, 15 interviewers for Wave 1 and 10 for Wave 2, compared to 9 and 6 for the control group. The interviewing period lasted from May 1992 through April 1993.

Detailed accounts of the findings are given by Marquis (1995) and Moore et al. (1996). Operationally speaking, the experimental procedure achieved many of its goals including: fewer proxy respondents, more

group interviews, a very substantial increase in the use of records, and successful audiotaping of most interviews. Except for Food Stamps, most of the more common sources of income were generally reported during the initial free recall section, prior to inquiries about specific sources not reported in that section. For most of the less common income sources, the proportion reported in the free recall section increased in the Wave 2 interviews.

Operational results on the debit side were similar to those observed in the pretests. Response rates were substantially lower for the experimental sample. This may have been due in part to differences between the interviewing staffs for the experimental and control procedures, which were not designated by random assignment of interviewers. The experimental interview staff itself was less experienced and more racially diverse; moreover, they worked under a less experienced and less productive crew leader. The cost per case for the experimental procedure was about double that of the standard SIPP interviewing procedure. The higher cost was due in part to longer interviews and in part to the need for more callbacks in order to achieve greater self-response and to conduct the group interviews. Although some of the interviewers saw benefits from the audiotape monitoring system, others felt that it was burdensome and gave rise to too many critical comments about their performance.

The effects of the experimental procedures on data quality were mixed and sometimes difficult to interpret. Table 6.8 shows estimates for five income sources of underreporting percentages for the experimental and control procedures used in the Cognitive Research Evaluation Experiment and for the 1984 SIPP Record Check. These estimates reflect only failures to report income that was received according to the program files for each income source; they do not reflect incorrect reports of recipiency by the "extra persons" at the sample addresses. Consequently, the 1984 data in this table differ from the estimates of *net* underreporting (that is, the amount of underreporting less the amount of overreporting) shown in Table 6.6 above.

Table 6.8 Percent underreporting of program participation and income from wages: Cognitive Research Evaluation Experiment and 1984 Record Check

Program/Income source	Evaluation experiment		1984 Record Check
	Experimental	Control	
AFDC	12	10	25
Food Stamps	17	12	24
Supplemental Security Income	13	8	23
Unemployment insurance	41	44	39
Job	11	4	N/A

N/A: No comparable data available.
Source: Adapted from Moore et al. (1996).

Table 6.8, shows that there was little difference between experimental and control procedures in the underreporting percentages for the five income sources included in the record check. Most of the underreporting occurred in cases where a respondent failed to mention a source at all, as opposed to cases where the source was reported for some months but missed for others. However, the table also shows that for three of the four income sources covered in both this experiment and the earlier SIPP Record Check, the underreporting rates for the standard SIPP interviews in the control group for the experiment were substantially below those observed in the Record Check. The populations sampled in the two studies were different in several respects; nevertheless, the lower rates in 1992-93 compared to 1984 suggest the possibility that the level of underreporting may have declined in the interim as a result of the development

of a more experienced interviewing staff or for other reasons. The much higher underreporting rate for unemployment insurance may be associated with the more transitory nature of income from that source, which is often paid on a weekly basis for a relatively short period.

As was the case for underreporting, there was relatively little difference between the experimental and the control procedures for overreporting, as measured by matching the survey reports of the "extra persons" against the program records. The observed rates for the experimental procedures were slightly lower for each of the four programs included in the record check but the difference was statistically significant only for Food Stamps.

As shown in Table 6.9, the observed effects of the experimental procedures on reporting amounts of income were more encouraging. The basic result is that the experiment treatment usually produced better reporting of income amounts by Wave 2 (Marquis, 1995). For the first three sources (AFDC, Food Stamps, and Supplemental Security Income), the experimental group attained the same percentage of correct responses as the control group in Wave 1, but surpassed the control group by Wave 2. For each of these income sources, the treatment-by-wave interaction was significant ($p<=0.05$) in a repeated measures analysis of variance using people who correctly reported their participation in both waves. For unemployment insurance income, even though the number of cases was small, the trends were similar to the long-term welfare programs. For this income source, however, the main effects of wave and treatment were statistically significant, but the interaction effect was not. Finally for earned income from a job, the main effect of the wave was significant, but the main effect of treatment and the interaction effect of treatment-by-wave were not statistically significant.

Table 6.9 Percent of matched cases with correct (±5 percent) income amounts: Cognitive Research Evaluation Experiment

Program/Income source	Wave	Number of matched cases		Percent of correct reports	
		Experimental	Control	Experimental	Control
AFDC	1	114	115	83	80
	2			87	72
Food Stamps	1	123	130	67	66
	2			83	63
Supplemental Security Income	1	67	77	78	78
	2			84	68
Unemployment insurance	1	9	8	29	20
	2			61	19
Job	1	46	46	67	52
	2			77	76

Source: Moore et al. (1996).

The experimental procedures for avoiding incorrect reports of changes in program participation at the seam did not produce the hoped for results. The number of off-seam changes reported in the survey was very close to the number found in the program records, but the number of changes at the seam was about twice that supported by the records. The net result was that the experimental procedures overreported the number of changes by about 8 percent, whereas the standard SIPP interviews underreported changes by about 26 percent.

Overall, although much was learned from the SIPP Cognitive Research Evaluation Experiment, it was not an unqualified success in meeting its objectives. The gains in quality were not sufficiently great to compensate for the lower response rates and substantially higher cost per case for the experimental procedures. The achievement of more complete reporting of program participation and more accurate measurement of changes in participation continues to present a difficult challenge.

7. DATA PREPARATION: 1984-1993 PANELS

This chapter describes the data preparation operations for SIPP Panels between 1984-1993. The operations involved two main phases. The first phase took place in the Census Bureau's 12 regional offices, where a receipt and control operation ensured that all assigned sample cases had been accounted for. Subsequently, clerical personnel edited the completed questionnaires and assigned geographic codes for sample persons who had moved since the preceding interview. Data entry clerks keyed the information from the questionnaires and the resulting data files were transmitted electronically to a central location for processing.

The second phase of data preparation began with an additional receipt and control operation for all questionnaires received from the regional offices to ensure that all assigned sample cases had been accounted for, followed by coding the written descriptions of occupation and industry, computer runs to edit for consistency, imputation of missing values, recode assignments, estimation, and preparation of the final tabulations. Except for the occupation and industry coding, which was performed at the Bureau's facility in Jeffersonville, Indiana, all of these activities were carried out at headquarters.

In order to provide a full overview of the data preparation operations for SIPP, the steps involving imputation and development of weights are mentioned briefly in this chapter. A more detailed discussion of imputation and weighting procedures, including relevant research and evaluation studies, appears in Chapter 8. Section 11.4 describes changes in data preparation operations for the 1996 Panel.

7.1 Regional office operations

Following the receipt and control operation of the workload, the first operation in the regional office was a clerical edit of a sample of questionnaires sent by each interviewer. This check, which covered both the core and topical modules, detected omissions and other errors in the completion of the questionnaires.

For new FRs, the first 10 questionnaires were fully edited. If their work was satisfactory, they became qualified for reduced editing. For qualified FRs, each questionnaire was subjected to a reduced clerical edit, in which only a few of the core items, such as control and person numbers, were checked. Topical modules were fully checked on every questionnaire.

The next step was the assignment of geographic (GRIN) codes to sample households that had moved. These codes identified where the sample household was located and permitted linkage to a file that contained a full set of geographic codes for different areas, such as States, counties, metropolitan statistical areas, and central cities.

The next step was data entry: keying the information from control cards and questionnaires. Edits were built into the data entry program to ensure that:

- The data were keyed in the proper sequence.
- Certain key identifiers, such as control number, name and relationship to householder, were present.
- Selected numeric items, mostly on the control card, were present.
- Data failing these edits were rekeyed after investigation and correction.

Data entry was subject to formal quality control procedures. After training, a keyer's initial work was 100 percent verified, using a quasi-independent key verification system with correction of errors detected. Keyers whose error rate (based on fields or source codes with errors) for a batch of 10 units (control cards and associated questionnaires) was 0.43 percent or less, qualified for sample process control. One-sixth of the work of these qualified keyers was verified in each batch keyed thereafter. Each batch keyed and sample verified was either accepted or rejected and rekeyed, based on the number of fields verified and the number of errors detected. At the end of each month, error rates for all keyers were reviewed to determine which keyers would remain qualified in the following month. Subsequent work of those who failed to qualify was verified 100 percent until such time as they requalified (Jones, 1983a). Data files for the accepted batches were transmitted electronically to a central location for further processing.

7.2 Central operations

The initial step with files received from the regional offices was a further receipt and control run to ensure that all expected cases, both interviews and noninterviews, were received. Errors identified in this step were described in reject listings for the regional offices. Regional office personnel resolved the problems by reviewing the hard copy documents or contacting the interviewers. Corrections and additional data were keyed and transmitted to headquarters.

Subsequent steps in data preparation were:

- Keyed verbal descriptions of occupation and industry were transmitted electronically to the Census Bureau's processing facility in Jeffersonville, Indiana, where appropriate codes were assigned and keyed. All coding and keying of occupation and industry items was verified 100 percent. Verification of coding was dependent (for example, the verifiers had access to the coders' entries); verification of keying used a quasi-independent system.

- Data were imputed for noninterviewed persons in interviewed households (see Chapter 8).

- An edit was performed to ensure consistency of responses recorded for persons, families, and households. Consistency was examined within and between sections of the questionnaire and between the control card and the questionnaire.

- Each section of the questionnaire, including topical modules, was edited to ensure that responses appeared when they should and to impute missing values (see Chapter 8).

- Recodes based on combinations of data items were added to the records and the codes that identified geographical areas were corrected, if necessary.

- For households that had moved and could not be located, three items were imputed: number of persons, number of adults, and number of additions to the household since the previous wave. The donor universe for this imputation was the interviewed mover households. These three imputed items were used in the calculation of weights for movers (see Chapter 8 and Riccini, 1984).

At this point the data were ready for cross-sectional weighting and estimation, as described in Section 8.2.

In order to prepare for the development of longitudinal data files, additional edits were performed to make data consistent over time for selected data items. Changes made in these additional edits were not carried over to the cross-sectional data files.

Inconsistencies in longitudinal data may result both from response errors and from errors in various data preparation procedures. They may also result from cross-sectional imputations made without reference to corresponding values for adjacent waves. The longitudinal edits covered labor force activity, earnings, income sources, program participation, health insurance coverage, demographic characteristics, and household composition. Details of the edits for the first 12-month longitudinal file, which was created from data for Waves 1 to 4 of the 1984 Panel, are given by Coder et al. (1987, Appendix A). A description of the edits implemented on the 1984 SIPP Full Panel Research File (32 months of data) is presented in Appendix B of the Technical Documentation of the 1984 SIPP Full Panel Research File, available from the Bureau of the Census, Customer Services, Data Users Services Division, Washington, DC 20233, (301) 457-4100. Pennell (1993) discusses the processes of both the cross-sectional and longitudinal edits and imputation after 1984.

7.3 Information about quality

Except for data entry, little quantitative data on errors in the different phases of data preparation are readily available. For data entry, a quarterly report on verification results shows national and regional office estimates of keying error rates and average outgoing quality levels. Keying error rates were low for all panels. For example, in the third quarter of 1988, outgoing error levels nationally were estimated to be 0.14 percent for July and August and 0.10 percent for September. For the quarter, the estimated outgoing error level for each regional office was below the 0.4 percent target average outgoing quality level (Waite, 1988).

A review of early data from the 1984 Panel led to detection of a significant level of error in the assignment of geographic area (GRIN) codes. Estimates of the U.S. population in metropolitan areas appeared reasonable, but the central city estimates were 10 million in excess of projections based on the 1980 Census. SIPP estimates showed that growth in central cities was larger than growth in the suburbs, a result not supported by information from other sources. Coding problems were suspected. An investigation showed that instructions for interviewers were not clear as to the information needed to assign accurate codes for Census geographic areas and that inadequate coding instructions for regional office clerks were compounding the problem. Better coding instructions reduced the level of error dramatically (Bowie, 1984; Jones, 1985b). For the 1985-1993 Panels, the problem was minimal due to the improved coding instructions and the fact that coding was required only for households that moved after the first wave interview.

Another procedural error was detected in the imputation of characteristics for noninterviewed mover households for the 1984 Panel. As explained in Section 7.2, these imputed characteristics were used in the development of weights for households that had moved. Two of the variables used to select households from the donor population were incorrectly defined, resulting in the imputation of inappropriate values. This error did not affect data for later panels. The size of the error and its effect on estimates for the 1984 Panel is not known, but is believed to be small (Singh, 1986; Riccini 1984).

There have been concerns about delays in SIPP data preparation. Citro and Kalton (1993) report that the Census Bureau released the wave files for Waves 1-9 of the 1984 Panel on average about 13 months after the last month of data collection. The topical modules files took longer, with an average release date of about 22 months after data collection, and the 1984 Longitudinal Panel file was released in April 1988, 20 months after the end of the panel. The lag was longer for the next two panels. The wave files from the 1985 and 1986 Panel were released on average about 31 months and 26 months, respectively, after the last month of data collection. However, after that, timelines improved: By the summer of 1987, the Census Bureau had begun to achieve delivery times in the range of a year after data collection.

The Census Bureau made adjustments to provide data products that met user needs in a timely manner. The data processing staff at the Bureau, working with an advisory group from the Association of Public Data Users, redesigned the core data files in a person-month format making these files easier to use for many analyses. The Census Bureau has continued to seek improvement in the timeliness of SIPP files and, along with the use of CAPI for the 1996 Panel, has moved to an automated system of processing to speed up data preparation.

8. ESTIMATION: 1984-1993 PANELS

This chapter describes the estimation procedures for the 1984-1993 Panels, and research efforts to improve them. Estimation is used in this chapter in a broad sense to cover the weighting and imputation procedures that are applied in converting raw data from SIPP into statistics that can be used for descriptive and analytical purposes. After some introductory remarks in Section 8.1, Sections 8.2 and 8.3 describe the estimation procedures for the cross-sectional files that contain data from a single wave and for the longitudinal file with data from the full panel, respectively. Section 8.4 summarizes research efforts to evaluate and improve these procedures. The changes adopted for the 1996 Panel are discussed in Section 11.5.

8.1 Introduction

The three goals of the estimation procedures are: to minimize the biases that may result from item, unit, and wave nonresponse; to take account of the selection probabilities used at every stage of sample selection; and to make use of data from external sources such as the Decennial Census of Population and the Current Population Survey to improve the precision of SIPP estimates. The Census Bureau has developed complex weighting and imputation estimation procedures for the SIPP. The resultant imputed values and sets of weights are used in computing the estimates that are presented in SIPP publications, such as the P-70 series. They are also included with records in SIPP public-use microdata files.

Users of the microdata files have the flexibility to choose whether or not to use data items that are identified as imputed and whether to use one of the sets of weights provided in the files or to develop their own. Guidance on estimation procedures for users of the microdata files is provided in the chapter on Use of Weights in the *SIPP User's Guide* (U.S. Bureau of the Census, 1998). A review of this document will make it clear that effective use of SIPP microdata files calls for a substantial investment of time and effort by users.

The Census Bureau provides a number of public-use files from the SIPP database. For each panel the following types of cross-sectional and longitudinal files are produced:

- Wave files (prepared in person-month format since 1990),
- Topical module files, and
- A longitudinal file (also known as the full panel file).

The wave files provide data separately for each wave of each panel. The longitudinal file contains data collected during the life of the panel. The following sections describe the weighting and imputation procedures applied to these files.

8.2 Weighting and imputation for the cross-sectional files

This section describes the weighting and imputation procedures used for the wave files. Weighting procedures are described in Section 8.2.1 and imputation procedures are described in Section 8.2.2.

The weights for the topical module file for a particular wave are the weights for the interview month for that wave, as described below for the wave file. Missing data items in the topical modules are generally imputed by hot-deck methods that are similar to those used with the wave files.

The person-month files are reformatted versions of the wave files. The appropriate monthly weight on the wave file is associated with each person-month record in the person-month file.

8.2.1 Cross-sectional weighting

The cross-sectional data file for each wave of a SIPP panel includes five weights for each sample person: one weight for each of the 4 months in the reference period and one for the month of interview.[1] We will first describe the procedures used to derive the monthly weights for Wave 1, followed by the procedures for subsequent waves. Each weight is the product of several components, some of which are the same for all members of a household and some of which differ for members of the same household.

Wave 1 weights

The monthly weight for each person in Wave 1 has four components:

- *The base weight.* This component of the final weight is the reciprocal of the overall selection probability of the household. If there were no nonresponse and the survey frame were complete, application of this weight would provide unbiased estimates for the survey population.

- *The Wave 1 noninterview adjustment factor*. This component is applied to interviewed households to compensate for household noninterviews. Noninterview adjustment cells are defined using the following variables:

  1. Census region - Northeast, Midwest, South, West.

  2. Residence - Metropolitan Statistical Area (MSA), non-MSA. For the 1984 Panel only, Standard Metropolitan Statistical Areas (SMSAs) were used in place of MSAs for variables 2 and 3.

  3. Place/not place - for units not in an MSA. Central city/balance - for units in MSAs.

  4. Race of reference person - black, nonblack.

  5. Tenure - owner, renter.

  6. Household size - 1, 2, 3, 4 or more.

  7. Rotation group - 1, 2, 3, 4. (Used only for the 1984 Panel.)

  Each cell must contain at least 30 unweighted units and the adjustment factor, which is equal to the ratio of total eligible sample households to interviewed households, must not exceed 2.0. If either of these requirements is not met, cells are combined according to established rules until both requirements are met (Jones, 1985a; Chapman et al., 1986; Singh and Petroni, 1988). For the 1985 Panel only, a special new construction noninterview adjustment factor was applied to interviewed households in new construction segments to account for the fact that operational difficulties had precluded interviewing in some of these segments.

- *The first-stage ratio estimate factor*. This factor is intended to reduce the between primary sampling unit (PSU) component of sampling error and is applied to sample households in non-self-representing PSUs. Estimation cells are established based on the following variables: Census region, residence (Metropolitan Statistical Area/other), central city/balance (only for Metropolitan Statistical Areas), and race. Within each cell, the adjustment factor is the ratio of the 1980 Census count of population (1970

[1]The files also contain weights for households and families. See the *SIPP User's Guide* (U.S. Bureau of the Census, 1998) for details.

Census data were used for the 1984 Panel) to an estimate of that count based on census data for sample PSUs (Jones, 1983b).

- *The second-stage ratio adjustment factor*. The second-stage factors are applied at the level of persons, with a separate set of factors calculated for each month. Their purpose is to reduce the mean squared error of estimates by partially correcting for survey undercoverage. Independent estimates of persons in cells defined by age, race, Hispanic origin, and sex are obtained from Census counts updated to account for subsequent births, deaths, immigration, and emigration. These independent estimates serve as the numerators for the ratios: the denominator for each cell is the sample estimate of persons in that cell, reflecting the base weights, noninterview adjustment factors, and first-stage ratio adjustment factors. The Hispanic origin controls were not used for the 1984 Panel but were introduced starting with the 1985 Panel, after a review of the 1984 Panel estimates showed that estimates of the population of Hispanic origin were low initially and declined further in later waves (Hubble, 1986).

  Monthly Current Population Survey estimates of the number of persons in households by race (black/nonblack), sex, marital and family status of householders are also used as control totals. For males, the marital and family status categories are:

  Persons in households with a primary family or subfamily:

  Husband of primary family,
  Male householder, no spouse present,
  Husband of subfamily, and
  Other.

  Persons not in households with a primary family or subfamily:

  Householder, and
  Not a householder or person in group quarters.

  For females, the categories are essentially the same, with wife and female substituted for husband and male. In order for the weights to sum to these independently derived totals while keeping the number of husbands equal to the number of wives and the number of householders equal to the number of households, an iterative raking procedure is used to derive the weighting factors (Nelson et al., 1985).

The Wave 1 weight for each person is the product of these four components. Further details of these procedures are given in the Weighting Appendix in the *SIPP User's Guide* (U.S. Bureau of the Census, 1998).

Weights for later waves

For later waves, each Wave 1 sample person interviewed in the current wave receives an initial weight that is the product of the first three components of the Wave 1 weight: the base weight, the noninterview adjustment factor, and the first-stage ratio estimate factor. These initial weights are then adjusted by the following steps:

- *Mover's weight.* Weights are decreased for persons in housing units containing adult members who were not part of the original sample (excluding persons returning from institutions, overseas, or the armed forces). For example, if a third adult moves into a household occupied by two original sample persons, all three would receive the initial weight multiplied by a factor of two-thirds. This weighting adjustment is described in Kalton and Brick (1994).

- *Later wave noninterview adjustment*. This adjustment factor compensates for household nonresponse after the first interview. The variables used to establish weighting cells differ from those used in the Wave 1 household noninterview adjustment factor because more data are available for households lost after the first interview. Noninterview adjustment cells are based on control card information from the most recent wave for the following variables: non-Hispanic white reference person or other; reference person a female householder without a spouse and with children, or 65 or older, or other; education of reference person; whether household income includes welfare payments; possession of financial assets; household tenure; and household size. (Information for original sample reference persons in noninterviewed households comes from Wave 1. For other reference persons in noninterviewed households, the information comes from their first interview in SIPP.) Each cell must include at least 30 nonweighted sample households and have an adjustment factor of 2.00 or less (Jones, 1983b).

- The adjusted weights are subjected to the second-stage ratio adjustment procedure that was described for Wave 1.

The application of the above weighting procedures produces weights that vary across sample persons. Table 8.1 presents some summary statistics for the distributions of weights for a number of representative waves and panels. The weights need to be used in estimation to produce estimates that apply to the survey population, taking account of unequal selection probabilities, nonresponse, and noncoverage. However, there is an increase in sampling variance incurred from the weights. One measure of this increase is $L=1 + CV^2$ where CV is the coefficient of variation of the weights, and L denotes the factor by which the variance is increased from the use of variable weights (Kish, 1992).[2] As can be seen from Table 8.1, the coefficients of variation in the cross-sectional weights are fairly sizable.

In addition to the noninterview and noncoverage weighting adjustments, several different features of the sample design contribute to the variation in weights reflected in these coefficients of variation. After Wave 1, the weights of initial sample persons are reduced if their household contains additional sample persons. Certain classes of persons, such as those in group quarters or entirely military households, are exempt from some phases of the weighting procedures. Segments are occasionally subsampled in the field when they turn out to be much larger than expected. For the 1985 and subsequent panels, a cut in sample following the initial selection of PSUs was carried out in a way that led to some variation in the base weights associated with different PSUs.

[2]Note that this factor does not apply for poststratification adjustments for estimates related to those adjustments. Nevertheless, it is a useful general guide to the effects of weighting on the precision of estimates.

Table 8.1 Statistical properties of final person weights of reference month 4 for selected waves and panels

Panel and wave	Final person weight							
	Number nonzero	Mean	Standard deviation	Coefficient of variation (%)	25th percentile	Median	75th percentile	Maximum
1985 Panel								
Wave 1	36,113	6,481	1,328	20.5	5,687	6,377	7,098	38,478
Wave 2[1]	26,807	6,565	1,576	24.0	5,478	6,510	7,330	40,287
Wave 5	28,813	8,236	2,401	29.1	7,047	8,222	9,411	42,203
Wave 8	28,484	8,404	2,659	31.6	7,013	8,372	9,695	50,575
1990 Panel								
Wave 1	58,149	4,236	2,401	56.7	3,246	3,752	4,484	43,843
Wave 2	56,223	4,396	2,525	57.4	3,311	3,911	4,671	43,612
Wave 5	54,792	4,558	2,674	58.7	3,373	4,144	4,891	49,506
Wave 8	54,246	4,651	2,717	58.4	3,417	4,241	5,086	51,047
1993 Panel								
Wave 1	51,895	4,943	1,058	21.4	4,400	4,794	5,248	28,695
Wave 2	50,192	5,137	1,222	23.8	4,580	4,976	5,494	29,661
Wave 5	48,899	5,329	1,496	28.1	4,666	5,182	5,791	31,401
Wave 8	46,807	8,615	1,700	30.3	4,820	5,433	6,184	32,653

[1]Only three rotation groups were included in this wave.

8.2.2 Cross-sectional imputation

Imputation is used in SIPP to assign values for missing items for interviewed persons and to replace reported values that fail consistency edits. Imputation is also used to assign all data to noninterviewed persons in interviewed households. In some cases, separate imputations are made for the same missing item for the cross-sectional and longitudinal files because there is more information that can be used for imputation for the longitudinal files than is available when a cross-sectional file is created (in particular, the information for following waves). This section describes cross-sectional imputation procedures, with Section 8.3.2 providing a description of longitudinal imputation procedures. Research on the properties of both the cross-sectional and longitudinal imputation procedures used in the 1984-1993 SIPP Panels, and into proposed alternatives for both, is presented in Section 8.4.

The Census Bureau's traditional sequential hot-deck procedure is used in each wave of SIPP to impute missing or rejected values for selected items for interviewed persons. The variables used to define imputation matrices vary widely, depending on the item being imputed. They include age, race, sex, income, occupation, and education. For each missing value, the procedure assigns a value reported for a person with similar characteristics. For each item subject to imputation, an indicator variable (or flag) is added to the SIPP data file to show which values have been imputed. The procedure is described in more detail by Nelson et al. (1985), Heeringa and Lepkowski (1986), and Pennell (1993).

Imputation is also used to assign all the values in a wave for Type Z nonrespondents, that is, individuals in otherwise cooperating households for whom no information is collected either from the individual or from a proxy respondent. Complete records are imputed for Type Z nonrespondents using a sequential hot-deck procedure. Those for whom data were not obtained in the previous wave are matched based on the following characteristics: age, race, sex, marital status, householder status, education, veteran status, and parent or guardian status. Those who were interviewed in the previous wave are matched based on the above characteristics as well as income and asset sources. The imputation cells for Type Z imputations were selected because they are either related to incidence of Type Z nonresponse, are survey characteristics of interest, or are related to survey characteristics of interest.

Not all SIPP survey items are imputed. For those that are imputed, the imputations are performed in the following order (derived from Nelson et al., 1985):

- Control card items: sample unit characteristics such as household tenure,
- Control card items: personal demographic items such as race/ethnicity, sex, age, and marital status,
- Type Z imputations,
- Labor force and information about recipiency of types of income and assets,
- Other cash income items,
- Wage, salary, and self-employment information,
- Asset income, and
- Program participation information.

8.3 Weighting and imputation for the longitudinal file

The longitudinal file consists of longitudinal records of all persons who have participated at any time during the panel. Several sample cohorts are defined: calendar year cohorts and full panel cohorts. Essentially, each cohort is defined as those persons who are in sample at the beginning of a designated time period. Consider the 1990 Panel as an example. Those eligible for the first calendar year cohort (that is, the file for January-December, 1990) are original sample persons who were present in the sampled households on January 1, 1990. Such persons should have been interviewed in the first wave of interviews, which were conducted in February, March, April, or May, 1990, depending on the rotation group. Those eligible for the second calendar year cohort are those in the sample on January 1, 1991. This cohort includes associated persons joining original sample persons in 1990 and still residing with them on January 1, 1991. It excludes original sample persons who died in 1990 or otherwise left the survey universe. Like the cohort for the first calendar period, the cohort for the longitudinal file is based on original sample persons. However, the cohort for the panel file includes a few original sample persons who were interviewed at Wave 1 but were not in the sampled households on January 1, 1990.

The next step is to divide the eligible sample persons in a longitudinal cohort into two groups: interviewed persons and noninterviewed persons. Broadly speaking, interviewed persons are eligible sample persons for whom data were obtained by self- or proxy response for each month of the period under study (panel or calendar year) or for every month before dying or moving to an ineligible address, such as an institution, Armed Forces barracks, or foreign location. All interviewed persons receive positive weights for that

longitudinal file. Before the 1990 Panel, all original sample persons who were nonrespondents for any wave contributing data for a given time period were classified as noninterviewed persons for the associated longitudinal file and received weights of zero for that file. Starting with the 1990 Panel, a new procedure was put into place to allow more of the records to have positive weights in the longitudinal files. All noninterviewed original sample persons with one or more missing waves, each of which is surrounded on either side by interviewed waves, have all their data imputed for the missing waves. Under this procedure, such persons are treated as interviewed and their records receive positive weights in the longitudinal file. As shown in Table 5.3, about a quarter of the panel nonrespondents are interim-only panel nonrespondents (about 5 percent of the full panel sample), and most of these have only a single wave of missing data. As a result, this missing wave imputation procedure retains a significant portion of the panel nonrespondent records as usable records for longitudinal analyses.

Section 8.3.1 discusses the assignment of weights for each longitudinal cohort. Section 8.3.2 describes the imputation methods for both missing wave imputations (when the surrounding waves are not missing) and item imputations.

8.3.1 Longitudinal weighting

The steps in developing the weights for the three longitudinal cohorts are summarized in Exhibit 8.1. The first two steps of defining the sample cohort and determining the set of sample persons to receive positive weights have already been described. Here we describe the remaining four steps of the operation.

Eligible persons with incomplete records in a longitudinal period (that is, with missing data for at least 1 month of the longitudinal file in which they were in the survey universe, and which have not been filled in with a missing wave imputation) receive zero longitudinal weights. The longitudinal weights of the remaining panel members (those with reported data for all months for which they were eligible or reported data for most months with missing wave imputations for the remaining months) are adjusted upwards to compensate for the zero-weighted eligible sample persons. This longitudinal weight has the following three components:

- *The initial weight.* This component is constructed from all components of the cross-sectional weights for the start of the weighting period (see Section 8.2.1) except the second-stage ratio adjustment.

- *Noninterview adjustment factors*. These factors account for eligible sample persons classified as noninterviewed in the appropriate longitudinal period and not already accounted for in the noninterview adjustment component of the initial weight. These factors are computed separately for each of a set of noninterview adjustment cells. Noninterviewed eligible sample persons are assigned to these cells based on information from the first interview wave for the reference period for the following variables: non-Hispanic white person or other, whether person was self-employed or not, household income, education of reference person, whether household income includes welfare payments, possession of financial assets, household tenure, and whether person was in labor force or not.

- *Second-stage ratio adjustment factors*. These factors are determined by a process similar to that used for cross-sectional weighting, but without equalization of husbands' and wives' weights. The benchmark dates chosen for the three longitudinal weights are a designated month of the first wave of the panel for the panel file weight (for example, this month was March 1990 for the 1990 Panel), and January of the appropriate calendar year for the two annual file weights.

- *Longitudinal weight.* The longitudinal weight for the calendar year or full panel is then computed as the product of the above three terms.

Exhibit 8.1 Steps in the development of longitudinal weights

Weight[1]	Step 1 Sample cohort definition	Step 2 Persons with nonzero longitudinal weights	Step 3 Initial weight (IW)	Step 4 Noninterview adjustment factor (FC)	Step 5 Control date for second-stage adjustment (SS)	Step 6 Final longitudinal person weight
Calendar year 1 weights	In sample as of January 1 of the first calendar year for the panel	Persons with complete data (either reported or imputed) for every month of calendar year 1 of a panel (excluding months of ineligibility)	All components of cross-sectional weighting except second-stage adjustment from wave containing January of calendar year	Adjustment for persons assigned zero calendar year 1 weight not adjusted for already in initial weight	January 1 of first calendar year	CY1 weight: IWxFCxSS
Calendar year 2 weights	In sample as of January 1 of the second calendar year for the panel	Persons with complete data (either reported or imputed) for every month of calendar year 2 of a panel (excluding months of ineligibility)	All components of cross-sectional weighting except second-stage adjustment from wave containing January of calendar year	Adjustment for persons assigned zero calendar year 2 weight not adjusted for already in initial weight	January 1 of second calendar year	CY2 weight: IWxFCxSS
Panel weights	In sample in Wave 1 of the panel	Persons with complete data (either reported or imputed) for every month of a panel (excluding months of ineligibility)	All components of Wave 1 cross-sectional weighting except second-stage adjustment	Adjustment for persons assigned zero panel weight not adjusted for already in initial weight	Designated date in the first wave (March 1 for recent panels)	Panel weight: IWxFCxSS

[1]The 1992 Panel has an additional calendar year 3 weight.

The weighting for the 1985 Panel file also includes a sample cut adjustment factor, to account for persons lost as the result of the sample cut in March 1985 (see Chapter 2).

Further details on the longitudinal weighting procedures are given in the Weighting Appendix in the *SIPP User's Guide* (U.S. Bureau of the Census, 1998). Additional background on the development of longitudinal estimation procedures for SIPP is given by Ernst et al. (1984) and Kobilarcik and Singh (1986).

8.3.2 Longitudinal imputation

The cross-sectional imputation procedures used to assign values for item nonresponses in a wave file employ only auxiliary information that is available on that file. The cross-sectional imputations are replaced in the longitudinal file by longitudinal imputations that take advantage of the availability of information from other waves on those files. In particular, the use of reported values for the same item on adjacent waves in longitudinal imputation can markedly improve the quality of the imputations and can avoid inconsistencies in these values across waves. Longitudinal imputation is used to compensate both for item nonresponse at a given wave and for some forms of wave nonresponse, as described below.

For example, if a value is missing in a given wave for a person, then the cross-sectional imputation for that value on the wave file (which is drawn in a sequential hot-deck procedure from another sample person's record) is replaced with the nearest reported value of the given person themselves (that is, the reported value from the closest month with a reported value, giving priority to the earlier month if there are two equidistant months). The longitudinal imputation procedure applied for missing items in SIPP takes advantage of reported data for the same person on other waves in the following manner. Consider an item which is reported on a monthly basis. If the monthly values are missing for a person in a given wave, but the values for the item are available for that person in a previous and/or succeeding wave, then the missing monthly values are assigned values from the closest months with reported values (giving priority to the earlier month if there are two equidistant months). If no nonimputed values are available for the item from previous or succeeding waves, then the longitudinal imputations for each month are equal to the mean value of the cross-sectional imputations for the item (over all months). In general, the longitudinal imputation is designed to reduce the spurious variability in longitudinal records which could result from the use of cross-sectional imputation. See Pennell (1993) for a detailed discussion of longitudinal imputation. Table 5.7 in Chapter 5 (also from Pennell, 1993) indicates the percentage of months on average that longitudinal imputation is used to fill in missing values for panel respondents, for the 1984 Panel and a set of selected characteristics.

Beginning with the 1991 Panel, longitudinal imputation is also employed to assign values for whole waves of missing data, provided that data are available for both the preceeding and succeeding waves. A carry-over imputation method is used. When the last month of the preceding wave and the first month of the succeeding wave have the same value for an item, then that value is imputed for every month in the missing wave. If the last month of the preceding wave and the first month of the succeeding wave have different values for an item, then a randomized procedure is used. One of the months in the missing wave period is selected with equal probability for each household (the "change" month for the household). For all months in the missing wave, up to the randomly selected change month, the value from the last month of the preceding wave is imputed. For all months from the change month onwards, the value from the first month of the succeeding wave is imputed.

Sections 8.4.3 and 8.4.5 describe the extensive research into longitudinal imputation that led to these imputation methodologies.

8.4 Evaluation and research

8.4.1 Introduction

There are no direct sources of data on the magnitude of biases in SIPP estimates due to nonresponse or on how these biases are affected by the imputation and weighting procedures used to compensate for missing data. Indirect indicators are available from several sources:

- Information on levels and correlates of household, person, and item nonresponse is presented in Chapter 5. One would expect the size of nonresponse biases to be positively correlated with the levels of each type of nonresponse.

- The population estimates used to calculate the second-stage ratio adjustment factors are postcensal estimates based on the latest population census (the 1980 Census for months up to and including December, 1991, and the 1990 Census for January, 1992 and beyond). The 1980 Census postcensal counts did not take account of undercoverage in the census. As a result, population totals will be underestimated in SIPP prior to 1992 and certain groups, such as young adult black males, will be underrepresented in distributions by age, sex, and race (Fay et al., 1988). However, the postcensal estimates used with SIPP from 1992 onwards do incorporate undercoverage adjustments (see, for example, Rawlings and Saluter, 1995, Appendix D).

- In the 1984 Panel, there were a number of errors in the construction of cross-sectional and longitudinal weights which were corrected in that panel. Quality control improvements in the system have eliminated these errors from later panels.

There has been a substantial amount of research designed to inform choices between alternative imputation and weighting procedures. Section 8.4.2 describes research into the effects of panel attrition and the weighting adjustments used in the 1984 Panel on later wave cross-sectional files. Section 8.4.3 describes research on alternative procedures for the treatment of interim wave nonresponse for longitudinal files. Sections 8.4.4 and 8.4.5 describe research into specific approaches for weighting and imputation for longitudinal files. Section 8.4.6 describes recent research on alternative estimation procedures designed to reduce the sampling variances of SIPP estimates.

8.4.2 Handling wave nonresponse in later wave cross-sectional files

Petroni and King (1988) report on an evaluation of the cross-sectional adjustments for noninterview households subsequent to Wave 2 of the 1984 Panel. The purpose of the study was to evaluate the ability of the classification variables used in the adjustment procedure to account for attrition within the context of the overall weighting procedure. The full cross-sectional weighting procedures (described in Section 8.2.1) were applied twice to data collected for sample households in Wave 2: once including all households interviewed in Wave 2 and once excluding households that were eligible for interview in Wave 6 but were not interviewed at that time. In the 1984 Panel, the cumulative sample loss of households was 9.4 percent through Wave 2 and 19.4 percent through Wave 6. About 1,300 sample households were lost between Waves 2 and 6 and were therefore not included in the second set of estimates.

Under the assumption that a household's Wave 2 characteristics are similar to its characteristics at Wave 6, the researchers believed that a comparison of the two sets of estimates would approximate the actual situation at Wave 6. The estimates, referred to as the W2/W2 and W2/W6 estimates, were compared using *t*-tests, with sampling variances based on SIPP generalized variance parameters. Some of the results, with statistically significant differences identified by appropriate symbols, are shown in Tables 8.2a and 8.2b.

Comparisons at the household level show higher estimates of median income and fewer households with low monthly income for the W2/W6 estimates, that is, the ones based on interview status at Wave 6. These results are consistent with other findings that attrition is greater among low-income households. The largest numbers of significant differences were for households in the metropolitan/nonmetropolitan and race/ethnicity demographic categories.

Table 8.2a Results of noninterview adjustment study (I): Number of households (in thousands) receiving benefits or with low monthly income, first quarter 1984

	Unemployment compensation		Cash benefits		Food Stamps		Low monthly household income[6]	
	W2/W2	W2/W 6	W2/W2	W2/W6	W2/W2	W2/W6	W2/W2	W2/W6
Race/ethnicity								
All races	2,707	2,712	7,246	7,350	6,582	6,582	11,819	11,504 [2]
White	2,231	2,217	4,879	4,986 [1]	4,238	4,244	8,659	8,374
Black	385	399	2,155	2,142	2,133	2,119	2,890	2,832 [1]
Hispanic[3]	208	186 [1]	779	767	728	682 [1]	1,132	1,096
Metro/nonmetro								
Metro	1,861	1,852	5,355	5,360	4,671	4,556 [1]	8,194	7,790 [2]
≥1,000,000	917	897	2,844	2,752 [1]	2,444	2,370 [1]	4,278	3,978 [2]
<1,000,000	944	955	2,510	2,608 [2]	2,227	2,186	3,916	3,812 [1]
Nonmetro	846	860	1,892	1,989 [2]	1,911	2,026 [2]	3,625	3,714
Family households	2,270	2,284	5,348	5,401	5,001	4,982	7,363	7,127 [2]
MC households[4]	1,814	1,799	2,366	2,463 [2]	1,859	1,877	3,838	3,677 [2]
Other family	310	331	932	887	795	742 [2]	884	834 [1]
FHH NSP WC[5]	147	153	2,051	2,052	2,347	2,363	2,640	2,616
Nonfamily households								
Male householder	272	271	582	590	466	463	1,536	1,471 [1]
Female householder	165	158	1,316	1,359	1,115	1,138	2,921	2,907

[1]W2/W2 and W2/W6 are significantly different at the 10 percent significance level.

[2]W2/W2 and W2/W6 are significantly different at the 5 percent significance level.

[3]Persons of Hispanic origin are also included in white or black.

[4]MC = Married Couple.

[5]Female householder, no spouse present, with own children under 18 years of age.

[6]Households with low monthly income are households below the poverty threshold for that month.

Source: Petroni and King (1988).

Table 8.2b Results of noninterview adjustment study (II): Monthly cash income for households, first quarter 1984

	Number of HHs (in thousands)		Mean income		Median income	
	W2/W2	W2/W6	W2/W2	W2/W6	W2/W2	W2/W6
Race/ethnicity						
All races	83,845	83,871	2,210	2,203	1,707	1,717 [1]
White	72,681	72,718	2,299	2,295	1,791	1,803 [1]
Black	9,314	9,347 [1]	1,454	1,427 [2]	1,165	1,165
Hispanic[3]	4,118	4,091	1,661	1,702 [1]	1,391	1,434
Metro/nonmetro						
Metro	63,763	63,206 [2]	2,301	2,291	1,797	1,813 [2]
≥1,000,000	34,348	33,489 [2]	2,448	2,463	1,893	1,924 [2]
<1,000,000	29,415	29,717 [2]	2,129	2,097 [2]	1,697	1,704
Nonmetro	20,083	20,665 [2]	1,920	1,933	1,485	1,491
Age groups						
<25	5,633	5,654	1,459	1,476	1,278	1,289
25-34	19,618	19,557	2,104	2,120	1,812	1,845 [2]
35-44	16,420	16,360	2,660	2,654	2,244	2,255
45-54	12,127	12,197	2,934	2,959	2,386	2,419
55-64	12,635	12,656	2,489	2,406 [2]	1,787	1,767
=65	17,412	17,447	1,439	1,432	1,000	986 [1]
Married couple HHs						
All races	48,847	48,857	2,772	2,762	2,257	2,265
White	44,229	44,229	2,816	2,812	2,298	2,312
Black	3,454	3,483	2,096	2,014 [2]	1,807	1,807
Hispanic[3]	2,483	2,490	2,034	2,096 [2]	1,723	1,775 [1]

[1]W2/W2 and W2/W6 are significantly different at the 10 percent significance level.

[2]W2/W2 and W2/W6 are significantly different at the 5 percent significance level.

[3]Persons of Hispanic origin are also included in white or black.

Source: Petroni and King (1988).

8.4.3 Handling interim wave nonresponse in the longitudinal files

The development of cross-sectional imputation and weighting procedures is relatively straightforward; deciding how to apply these tools in a longitudinal context is much more difficult. Several theoretical and empirical analyses have addressed the selection of imputation and weighting procedures for use in the creation of SIPP longitudinal files. No single set of procedures is optimum for all kinds of analyses; thus some users take the position that the files should contain only reported data and weights based on the sample selection probabilities, leaving users to decide how to deal with missing data in their analyses. It is evident, however, that most users of SIPP longitudinal files prefer the data files to contain imputed values and adjusted weights, together with sufficient information included in the data files and in the file documentation to allow for any adaptation that they may wish to implement.

Early empirical work on the evaluation of longitudinal imputation strategies was based on data files from the Income Survey Development Program (Huggins et al., 1985; Kalton et al., 1985; Kalton, 1986; Lepkowski et al., 1989). An extensive research study on the treatment of person-wave nonresponse in panel surveys was undertaken by the Census Bureau and the Survey Research Center, University of Michigan. The

research procedures and findings, which include two empirical analyses based on data from the 1984 SIPP Panel, are reported in Kalton et al. (1987).

One of the studies in this project was a simulation study in which the results of a simple carry-over method of longitudinal imputation were compared with the results of a weighting adjustment for person-wave nonresponse for a subset of persons interviewed in all of Waves 1 to 3 of the 1984. Nonresponse was created in the data set by random deletion of Wave 2 or Wave 3 data for selected persons, based on patterns of nonresponse actually observed for the full set of sample persons interviewed in Wave 1. The weighting and imputation procedures were applied and the results compared.

As illustrated in Table 8.3, the researchers found that the carry-over imputation procedure "...fails to track net changes in means or proportions when these vary over time." They also point out that the procedure should be used carefully when there is the possibility of extreme outliers. There were a few persons in the database who received unusually large Social Security benefits in a single month and at least some of these instances were believed to represent correctly reported information. The investigators concluded that weighting was the safer general-purpose solution for the three-wave file, but pointed out that a different conclusion might be reached for an eight-wave file, where a greater loss in effective sample size would be associated with the weighting procedure.

Ernst and Gillman (1988) used the longitudinal research file based on the first three waves of the 1984 Panel for another empirical study. The researchers divided the sample persons in the file into four groups:

- *Group 1*. Original sample persons interviewed for all 32 months during the full panel reference period (or until leaving the survey universe).

- *Group 2*. Original sample persons who (a) missed the eighth interview and were in the SIPP universe at the time of their first missed interview, or (b) missed any three consecutive interviews.

- *Group 3*. All other original sample persons.

- *Group 4*. All associated sample persons, that is, those persons living with an original sample person who were part of the SIPP universe in Wave 1 but were not selected at that time.

Distributions for several demographic and economic characteristics as of the initial reference month of the panel were compared for the four groups. As shown in Tables 8.4 and 8.5, there were significant differences between groups 1 and 2 and between groups 1 and 3 for distributions by age, marital status, and relationship to reference person. No significant differences were observed for the three income and program participation variables.

Other comparisons were made to examine the differences among the four groups with respect to the frequency of gross changes and the effects of the longitudinal weighting system used for the three-wave longitudinal research panel. The researchers concluded there was some evidence that, at least for certain characteristics, groups 2 and 3 differed significantly from group 1. Their results indicated that longitudinal weighting only partially compensates for this problem. They suggested that further consideration be given to the inclusion of more than three sets of weights on the longitudinal file for the full 1984 Panel, in order to make greater use of cases in groups 2, 3, and 4 for which there are no missing interviews during specified time periods.

Table 8.3 Distributions of responses across waves for three items for the wave nonrespondents (a) with the actual responses (b) with imputed responses for missing waves and (c) with weighting adjustment for wave nonrespondents

Response pattern across waves	Percent		
	Actual	Imputed	Weighted
Having a job			
YYY	58.1	63.3	57.4
YYN	2.4	0.4	2.4
YNY	2.5	-	2.5
YNN	3.2	2.6	3.1
NYY	2.5	1.5	2.6
NYN	0.7	-	0.7
NNY	2.7	0.4	2.7
NNN	27.8	31.8	28.6
Total	100.0	100.0	100.0
Receiving Social Security income			
YYY	14.4	14.8	14.7
YYN	0.3	-	0.1
YNY	0.1	-	0.1
YNN	0.3	0.2	0.2
NYY	0.3	0.3	0.6
NYN	0.2	-	0.1
NNY	0.6	0.1	0.6
NNN	83.9	84.6	83.7
Total	100.0	100.0	100.0
Having savings accounts			
YYY	45.1	49.9	48.9
YYN	2.4	0.7	2.7
YNY	1.2	-	1.2
YNN	4.4	2.4	3.3
NYY	2.8	1.3	2.7
NYN	0.2	-	0.4
NNY	2.4	0.8	2.3
NNN	41.5	44.9	38.5
Total	100.0	100.0	100.0
Number of persons (sum of weights)	1,846	1,846	(1,846)

Source: Kalton, et al. (1987).

Table 8.4 Distributions of characteristics of 1984 SIPP Panel at initial reference month by group[1]

Characteristics	Groups				
	1	2	3	4	1+2+3
Age					
15-17	5.8	6.6	6.9	5.5	6.0
18-24	14.5	20.7	21.4	39.2	16.2
25-34	21.8	23.1	23.6	29.0	22.2
35-44	16.5	16.2	16.1	9.6	16.4
45-59	19.0	18.1	17.3	9.7	18.7
60 and over	22.4	15.3	14.7	7.1	20.5
Sex					
Male	46.8	50.1	49.8	54.7	47.7
Female	53.2	49.9	50.2	45.3	52.3
Marital status					
Married, spouse present	59.7	50.4	50.6	31.3	57.2
Other	40.3	49.6	49.4	68.7	42.8
Relationship to reference person					
Reference person	48.5	43.6	44.1	18.1	47.2
Spouse of reference person	29.4	24.2	24.7	17.1	28.0
Child/relative of reference person	19.6	26.8	25.9	35.2	21.5
Nonrelative of reference person with household relative	0.3	0.9	0.8	2.7	0.5
Nonrelative of reference person without household relatives	2.2	4.5	4.6	26.8	2.8
Income as percentage of Food Stamp cutoff					
100 percent and less	11.8	13.8	12.9	10.9	12.3
101 - 130 percent	5.6	6.1	6.1	6.0	5.7
131 - 185 percent	10.5	10.1	9.7	11.0	10.2
Over 185 percent	72.1	70.0	71.3	72.1	71.6
Household receives Food Stamps					
Yes	6.5	6.6	9.6	10.2	6.7
No	93.5	93.4	90.4	89.8	93.3
Household receives means-tested cash benefits					
Yes	8.2	7.8	8.1	11.9	8.1
No	91.8	92.2	91.9	88.1	91.9

[1]See text for definition of groups 1, 2, 3, and 4.

Source: Ernst and Gillman (1988).

Table 8.5 Chi-square statistics for pairs of columns in Table 8.4

Characteristics	Degrees of freedom	Pairs[1]						
		1,2	1,3	1,4	2,3	2,4	3,4	1,1+2+3
Age	5	38.7[1]	18.8[1]	213.0[1]	0.3	88.6[1]	57.9[1]	50.6[1]
Sex	1	3.2	1.0	9.1[1]	0.0	2.7	2.2	3.8
Marital status	1	26.8[1]	10.0[1]	122.2[1]	0.0	47.4[1]	36.2[1]	32.7[1]
Relationship to reference person	4	52.0[1]	17.9	660.1[1]	0.2	226.2[1]	139.0[1]	59.9[1]
Income as percentage of Food Stamp cutoff	3	1.5	0.3	0.2	0.1	1.8	0.5	1.6
Household receives Food Stamps	1	0.0	2.0	3.6	1.5	2.8	0.0	0.4
Household receives means-tested cash benefits	1	0.1	0.0	2.9	0.0	3.0	1.7	0.0

[1]Significant at the 5 percent level.

Source: Ernst and Gillman (1988).

8.4.4 Alternative methods for computing panel weights

This section describes three independent research projects that were carried out to study alternative methods for adjusting the weights of panel respondents to account for panel nonrespondents (with panel nonrespondents defined as those missing any wave interview for which they were eligible). Rizzo et al. (1994, 1996) carried out an extensive analysis of the correlation of panel nonresponse to Wave 1 survey responses, and found generally that the current set of variables used to define nonresponse cells for longitudinal weights (see Section 8.3.1) captures most of the relationships of Wave 1 questionnaire data to later panel nonresponse, but with some exceptions. The additional variables they included in their weighting schemes were age, relationship to the reference person, class of work, food stamp recipiency, and the number of imputed items at Wave 1. The relationship of Wave 1 item nonresponse and subsequent panel nonresponse was quite strong; respondents who had high levels of item nonresponse for the Wave 1 questionnaire tended to have higher levels of wave nonresponse in later waves, even after adjusting for all of the other person and household characteristics known to be associated with panel nonresponse.

This relationship between item nonresponse and panel nonresponse is in line with a "cooperator/ noncooperator hypothesis" studied by Bollinger and David (1996). Under this hypothesis, persons with a propensity to be noncooperative express this propensity in a variety of ways: unit nonresponse, item nonresponse, and measurement error (that is, the incidences of these types of problems are positively correlated). In a study matching SIPP Food Stamp questionnaire answers to Food Stamp program administrative records, Bollinger and David (1995, 1996) found a correlation between panel nonresponse and measurement error for Food Stamp reporting. Panel nonrespondents tend to fail to report Food Stamp benefits received when they do respond.

Rizzo et al. (1994, 1996), Folsom and Witt (1994), and An et al. (1994) studied alternative methodologies for adjusting for panel nonresponse using weighting adjustments, and compared them with the current methodology of adjustments based on the reciprocals of weighted response rates within designated nonresponse cells (see Section 8.3.1). Rizzo et al. (1994) examined the following alternatives:

- Adjustments based on predicted values from logistic regression models. These are smoothed versions of the weighting adjustments computed for single cells.

- Adjustments computed for nonresponse cells selected by a tree search algorithm. This methodology finds the optimal set of nonresponse cells (based on heterogeneity of nonresponse rates) according to a designated criterion.

- Adjustments generated through raking methods, using overall totals as margins for the sample (raking the weights for responding panel members to totals computed for all panel members).

Folsom and Witt (1994) studied nonresponse adjustments based on a generalized raking method (adjusting weights for responding panel members to totals computed for all sample members). This raking method is related to the Logit (L,U) method of Deville et al. (1993), and is designed to smooth large calibration ratios which sometimes occur under iterative proportional fitting.

An et al. (1994) examined two alternatives to the current panel nonresponse weighting methodology. As has been described in Section 8.2.1, the current methodology is applied in two stages: first, carry out nonresponse adjustments within adjustment cells to make the weights of panel respondents sum to those of the total panel sample; second, calibrate the nonresponse adjusted weights to independent control totals. One of two alternatives studied by An et al. (1994) was to reverse the order of these two stages. The other alternative carried out the two stages of the adjustment simultaneously using generalized regression methods. This "full information" procedure used the two sets of auxiliary data—the data from the total sample and from the independent controls—more efficiently than the current methodology.

Rizzo et al. (1994, 1996), Folsom and Witt (1994), and An et al. (1994) compared estimates from the 1987 SIPP Panel calculated with the different weighting schemes and with results from benchmark data. Although the research suggested some improvement to the current methodology (for example, the addition of Census region, household income, and extent of item imputation as auxiliary variables in the panel nonresponse adjustments), it did not produce adjustments that corrected the large discrepancies that exist between SIPP estimates and external benchmarks. However, these discrepancies may well not be a result of panel attrition, but may rather arise from response error, in which case the adjustments would not be expected to be effective.

Allen and Petroni (1994) present research into another potential modification of the methodology for adjusting for panel nonrespondents through weighting adjustments. This research focused on households which could not be found because they moved; movers comprise a significant portion of panel nonrespondents (see Table 5.1). Two weighting adjustment approaches were studied.

The first method was the inclusion of mover status among the variables used to generate panel nonresponse cells. Under this approach, the weights of movers who are panel respondents are adjusted upwards to represent movers who are not panel respondents. If movers have characteristics that are different from nonmovers after adjustments have been made for other nonresponse cell characteristics (for example, race/ethnicity, income, employment status, etc.), then this approach should reduce the bias incurred from the higher panel nonresponse rate of movers. The weakness, however, found in this approach was that mover

status was unknown for many panel nonrespondents (since they may become nonrespondents before they move). As a result it was difficult to do this adjustment accurately, which, in turn, reduced its value in bias reduction.

The second method was the inclusion of control totals for movers and nonmovers (derived from CPS data) at the calibration stage. This approach avoids the difficulties of tracking movers in the sample itself, and appeared to have better bias reducing properties than the first method but the effects were not significant. Allen and Petroni (1994) concluded that research into movers' nonresponse adjustments merits further development, but their results did not as of themselves justify modifying the SIPP panel nonresponse weights along these lines.

8.4.5 Alternative longitudinal imputation methods

There has been considerable research on the effects of imputation procedures used for cross-sectional estimates when those estimates are used in longitudinal analyses. Alternative imputation procedures designed to better preserve longitudinal relationships have been proposed and evaluated. There has been some feedback from users of SIPP data files on how the imputation and weighting procedures have created problems in their analyses. For example, Doyle and Dalrymple (1987), in a study of the low-income population receiving benefits under the Food Stamp Program, found that " ... the imputation process does not preserve the known relationships between benefits and the determinants of benefit levels." Holden (1988), in a study of the effects of widowhood on income, reported that "because the Census Bureau's cross-sectional imputations exaggerate income changes and their magnitude ... this study reimputed all missing income data for husbands and wives."

An early study was conducted by Heeringa and Lepkowski (1986), used a data set consisting of all persons interviewed for all of the first three waves of the 1984 Panel. For three categorical and two amount items in this data set, they compared the effects of using a cross-sectional and a longitudinal imputation procedure for persons with item nonresponse in one or more waves. The cross-sectional procedure was the standard hot-deck procedure that had been used to produce the imputed values in the data set. The longitudinal procedure was a direct substitution procedure, in which the reported value for the previous wave was substituted for the missing value in the current wave. If no value had been reported in the previous wave, the cross-sectionally imputed value for the current wave was accepted. As might be expected, Heeringa and Lepkowski found that the direct substitution method for longitudinal imputation appeared to understate change. However, they concluded that this might be preferable to the greater overstatement of change resulting from use of the cross-sectional hot-deck method (Kasprzyk, 1988).

Little and Su (1989) proposed a model-based longitudinal imputation procedure based on row and column fits for use in imputing missing values for a single variable for which repeated observations are obtained. They applied the method to monthly SIPP wages and salary amounts, using a multiplicative model in which the column effects corresponded to time periods and the row effects to individuals. The model also included a residual component, taken from a complete case with characteristics similar to those of the incomplete case for the variable imputed. They found that the method had several attractive features and was comparatively easy to implement.

There have been a number of recent empirical studies of the properties of different imputation methods for cross-sectional and longitudinal purposes. Four of these studies are described in the following paragraphs (Lepkowski et al., 1993; Tremblay, 1994; Williams, 1997; and Williams and Bailey, 1996). These papers studied imputation procedures using artificial data sets created from the complete longitudinal records of SIPP Panels. "Artificial nonresponse" was generated by randomly selecting records from the complete data

set to have values set as missing. The random selection procedure used the distribution of observed missing data patterns within the panel as its probability distribution to create an artificial data set with roughly the same distribution of missing data patterns. Since the missing values are in this case known (but deleted) values, estimates for the various imputation procedures can be compared to known benchmark estimates. The papers overlap with regard to the imputation procedures studied; the discussion below is organized by imputation procedure.

The simplest and easiest to implement imputation procedure is the simple carry-over. In this procedure, the last nonmissing value is imputed for all missing values in a record. Lepkowski et al.[3] (1993) and Williams[4] (1997) studied this imputation procedure, and both papers provide evidence that the procedure is generally effective for cross-sectional monthly imputation (but Lepkowski et al. (1993) found that this was not the case for seasonally variable quantities). Both studies found that the simple carry-over is ineffective for longitudinal records of categorical variables (such as program participation), because the method tends to artificially extend spell lengths. For continuous variables such as program amounts the simple carry-over still is adequate for calendar year totals (when all months need to be imputed), according to Williams (1997), but Lepkowski et al. (1993) found a tendency for the carry-over to overestimate quantities such as calendar year earned income, and to show considerable instability for income quantities which have significant month-to-month differences.

Two imputation methodologies intended to correct some of the deficiencies of the simple carry-over are the random carry-over and the population-based random carry-over methods (studied by Tremblay[5], 1994; Williams, 1997; and Williams and Bailey[6], 1996). For the random carry-over (which is generally done for a single missing wave), a month (called the "change month" in this discussion) is randomly selected with equal probability from the last nonmissing month and the months in which the item is missing (that is, if there are 4 missing months, then 1 of the 4 months or the last nonmissing month is selected as the change month with probability 1/5). The previous nonmissing value is then imputed to all months up to and including the change month, and the remaining missing months after the change month are imputed using the first nonmissing value following the missing period. The population-based random carry-over is a modification of the random carry-over. The vector containing the probabilities assigned to each month for the random selection of the change month is computed based on observed data (for example, if among the observed values it is found that a change takes place in the second month 15 percent of the time and in the third month 10 percent of the time, then the second month will be selected as the change month 15 percent of the time for imputations and the third month 10 percent of the time). The equal probability random carry-over has been in production for the 1990 Panel onwards for longitudinal imputations for single missing waves (see Section 8.3.2).

Tremblay (1994) found that the two carry-over methods performed quite adequately as compared with the more complex procedures of flexible matching (see below) and the Little and Su method for Food Stamp longitudinal benefit totals. Williams (1997) showed that the random carry-over had a similar performance to the simple carry-over for calendar year program benefit amounts. Williams and Bailey (1996) also found

---

[3]Lepkowski et al. (1993) used the 1987 Panel and studied categorical variables on program participation (AFDC, Food Stamps, Social Security, and other Federal programs), and income variables (earned income, unearned income, and assets).

[4]Williams (1997) used the 1992 Panel (calendar year 1993) with nine variables on program participation and benefits.

[5]Tremblay (1994) used the 1990 Panel with Food Stamp benefits as the study variable.

[6]Williams and Bailey (1996) used the 1992 Panel with program benefit amounts (Food Stamps, AFDC, Social Security), and wage and salary amounts.

that the random carry-over and population-based random carry-over had a similar performance. Both were better in terms of average absolute deviation (of imputed from real value) than the Little and Su and flexible matching methods over the course of a wave. However, both performed poorer than the two more complex procedures in maintaining the true month-to-month correlation patterns (the carry-over methods produce imputations with correlations that are too high compared to the real values). The population-based random carry-over did somewhat better in this regard than the equal probability random carry-over.

The second set of imputation procedures examined in these four papers were hot-deck procedures. The cross-sectional hot-deck is the current procedure for generating imputations for cross-sectional purposes (see Section 8.2.2), and generally serves reasonably well for this purpose. Several types of hot-deck procedures intended to improve the longitudinal properties of imputations were examined. Lepkowski et al. (1993) modified the cross-sectional hot-deck by matching imputation recipients to donors using imputation cells based not only on within-wave characteristics, but also on earlier wave values. The best results occurred when earlier wave values of the imputed variable (for example, Food Stamp benefits) were used. Recipients were matched to donors using the cross-sectional information and also the values of the imputed characteristic for the first month and the last nonmissing month for the recipient. In other words, recipients in general were matched to donors who had the same values of the variables in the first month and the last nonmissing month. For continuous variables this matching was done after the variable was grouped into two or four categories (for example, low Food Stamp benefits, high Food Stamp benefits). Matching was also tried using a related characteristic for the previous wave match (for example, AFDC benefits for Food Stamp benefits). This generally worked less well.

The research found that this modified hot-deck avoids the bias of carry-over approaches, but has high variability. This high variability can be mitigated by imputing changes (absolute and relative) rather than values, but these methods also produce outlying imputations in certain situations, and therefore high variability. Lepkowski et al. (1993) recommended further research into these "longitudinal hot-deck" imputation procedures using more sophisticated matching and imputation approaches than used in their research.

Williams (1997) studied a spell length hot-deck method. This method matches recipients to donors using a full longitudinal string (this method is only appropriate for dichotomous variables or categorical variables with a small number of longitudinal patterns). This method performed very well for program participation values over a calendar year, compared to carry-over methods.

Tremblay (1994) and Williams and Bailey (1996) studied the flexible matching imputation approach, derived from an American Housing Survey methodology, and the Little and Su approach. The flexible matching approach is a modified hot-deck procedure which allows recipients to be matched to donors based on cross-sectional information as well as the values in the previous and following waves. For example, if Food Stamp benefits were imputed for a recipient, then Food Stamp benefit values for the recipient in the previous wave and in the following wave could be used to match to donors, if those values were nonmissing (as well as other values potentially). The matching is not done by imputation cells as in standard hot-deck methods, but by a matching algorithm chosen empirically using stepwise regression methods (using the complete data records to choose the model).

The flexible matching imputation approach and the Little and Su approach are both complex and computationally intensive. Tremblay (1994) and Williams and Bailey (1996) did not find any decisive advantages to either approach as compared to carry-over approaches with 1 month or longitudinal totals. Both methods did show an advantage over the random carry-over methods (equal probability and population-

based) in preserving the correct between-month correlation coefficients (the random carry-over methods tended to make these correlations too high).

8.4.6 Research on estimation procedures to reduce variance

As stated at the beginning of this chapter, one of the goals of estimation is to reduce the sampling variance of survey estimates by making use of data from other sources as part of the estimation process. To this end, the second-stage ratio adjustments factors used in developing both cross-sectional and longitudinal weights make use of current population estimates based on the most recent population Census and estimates of the number of households from the Current Population Survey, which had a considerably larger sample than any of the 1984 through 1993 SIPP Panels and, therefore, provides estimates with smaller sampling errors.

Huggins and Fay (1988), Dorinski and Huang (1994), and Dorinski (1995) report research carried out to determine whether use of data on income distributions from administrative record systems as a part of the second-stage ratio estimate procedure was feasible and would be likely to achieve further variance reductions. The Census Bureau has access, for statistical uses only, to extracts from Internal Revenue Service (IRS) files of individual income tax data and Social Security Administration (SSA) files containing data on age, race, and sex for all persons issued Social Security numbers. Some of these extracts are for all persons in the administrative record files and some are for samples. The samples are so large that estimates based on them have much smaller sampling errors than estimates based on the smaller samples used in SIPP. Linkages of SIPP sample persons to these files are possible because both administrative record systems contain Social Security numbers as identifiers and because most sample persons provide their Social Security numbers in SIPP (Dorinski, 1995, reports that the refusal rate among SIPP respondents for this item was 5.1 percent in the 1990 Panel).

The experimental use of administrative data in SIPP estimation followed procedures initially proposed by Herriot (1983) and Scheuren (1983). New weighting factors were developed for persons in the three-wave longitudinal research file from the 1984 and 1990 Panels who could be identified as primary or secondary tax filers on the basis of linkage to an IRS file for the reference year that was complete except for late filers.[7] For the 1984 and 1990 Panels, close to 80 percent of SIPP adults could be matched to an IRS record. Control totals of persons by age, race/ethnic category, sex, filing status as joint or nonjoint and adjusted gross income category were developed for this research from an IRS extract file with linked SSA information on age, sex, and race. These control totals were used in an iterative raking procedure to determine the weighting factor to assign to each of the matched persons in the SIPP sample.

Sample estimates were prepared with and without the weighting factors based on the administrative record data and their variances were compared. The main results from the Huggins and Fay (1988) research on the 1984 Panel are shown in Table 8.6. These results suggest substantial potential improvements for some characteristics, especially with respect to statistics on income. The largest gains were for statistics such as mean and median income, which were influenced strongly by the middle and upper end of the income distribution, but results for a poverty measure were also encouraging. The procedure also reduced variances for estimates of Food Stamp recipiency for most groups, but yielded a mixed outcome for AFDC recipiency. The large improvements shown for Social Security recipiency were believed to be largely an artifact resulting from the form of the variance calculations used in the study.

[7]A three-wave panel from the 1984 Panel was used in the research reported in Huggins and Fay (1988) and Dorinski and Huang (1994). A panel from the 1990 Panel was used in the research reported in Dorinski (1995).

Table 8.6 Ratios of estimated variances after and before adjustments to administrative totals, Waves 1 to 3 of the 1984 SIPP Panel

Category	Percentages of income distribution				
	Loss - $10K	$10K-$20K	$20K-$30K	$30K+	$20K+
Total age 25+	0.49	0.80	0.58	0.41	0.38
Males	0.53	0.93	0.70	0.38	0.35
Females	0.48	0.58	0.61	0.78	0.54
Black age 25+	0.74	0.91	0.87	0.80	0.75
Males	0.68	0.93	0.87	0.74	0.65
Females	0.81	0.97	1.15	1.12	1.15
Hispanic age 25+	1.03	0.83	0.82	1.01	0.69
Males	1.23	0.86	0.77	0.91	0.68
Females	0.79	0.81	0.81	1.07	0.83

Category	Mean income	Months in poverty	Food Stamps	AFDC recipient	Social Security[1]
Total age 25+	0.46	0.74	0.89	1.00	0.27
Males	0.46	0.71	1.01	1.13	0.27
Females	0.49	0.80	0.81	0.99	0.40
Black age 25+	0.69	0.71	0.76	0.89	1.14
Males	0.61	0.65	0.81	1.42	0.94
Females	0.74	0.78	0.77	0.87	1.21
Hispanic age 25+	0.83	0.89	1.21	1.15	0.85
Males	0.86	0.99	1.13	1.11	0.91
Females	0.94	1.00	1.23	1.17	0.98

[1]The researchers believe that the large reductions shown for estimates of Social Security recipiency were primarily an artifact resulting from the variance calculations used in this study.

Source: Huggins and Fay (1988).

The Dorinski and Huang (1994) research duplicated that of Huggins and Fay (1988) on the 1984 Panel with some changes, including a ratio adjustment of the weights of unmatched SIPP respondents (unmatched to the IRS file) to CPS control totals (subtracting out the weights of matched SIPP respondents). The results were similar to the Huggins and Fay (1988) results summarized in Table 8.6. The Dorinski (1995) research used the 1990 Panel and the IRS extract file based on a 20 percent sample rather than the extract file based on a 1 percent sample (as used in the earlier research). In addition, SIPP presecond-stage weights were used as inputs rather than SIPP final weights (used in the earlier research). However, no calibration of unmatched SIPP respondents to adjusted CPS control totals was carried out, as in the Dorinski and Huang (1994) research. Table 8.7 summarizes the results on sampling variance.

The results for the total population were not improved from the 1984 Panel results (but still indicated gains from using the IRS control totals), but results for blacks and especially Hispanics were markedly improved. Dorinski (1995) notes, however, considerable additional effort would be required to develop an operating prototype that integrates this new procedure with the present system and resolves problems related to coverage by IRS of persons "excluded from the SIPP universe, the omission of late filers from the IRS files and the nonreporting of Social Security number by about 5 percent of SIPP respondents."

Another technique that has been considered in a preliminary way is the use of composite estimation procedures, which take advantage of the correlations between responses for the same units of analysis at different time periods. Composite estimators are used effectively in the Current Population Survey, and can lead to significant variance reductions when the correlations over time are high, which is likely to be the case for several important data items in SIPP. Chakrabarty (1987, 1989a) reviewed the types of composite estimators appropriate for the SIPP sample design and data structure and recommended that a particular type, the Ernst-Breau estimator, be used, starting with data from the 1986 Panel. This recommendation, however, has not been implemented. Composite estimation is not applicable with the 1996 redesign since that design does not have overlapping panels.

Table 8.7 Ratios of estimated variances after and before adjustments to administrative data, first calendar year of the 1990 SIPP Panel

Category	Percentages of income distribution				
	Loss - $10K	$10K-$20K	$20K-$30K	$30K+	$20K+
Total age 25+	0.75	0.67	0.79	0.54	0.56
Males	0.75	0.76	0.81	0.63	0.62
Females	0.65	0.76	0.82	0.61	0.62
Black age 25+	0.79	0.73	0.90	0.78	0.58
Males	0.87	0.76	0.93	0.81	0.73
Females	0.62	0.73	0.95	0.90	0.61
Hispanic age 25+	0.73	0.88	0.66	0.83	0.66
Males	0.76	0.99	0.78	0.83	0.69
Females	0.82	0.91	1.12	1.11	1.03

Category	Mean income	Months in poverty	Food Stamps	AFDC recipient	Social Security
Total age 25+	0.62	0.83	0.91	0.93	0.42
Males	0.65	0.86	1.02	0.96	0.44
Females	0.71	0.87	0.90	0.96	0.49
Black age 25+	0.87	0.76	0.79	0.81	0.75
Males	0.90	0.80	0.91	0.82	0.76
Females	0.72	0.79	0.79	0.89	0.76
Hispanic age 25+	0.72	0.73	0.85	0.81	0.73
Males	0.73	0.69	0.74	0.86	0.70
Females	0.85	0.80	0.86	0.79	0.82

Source: Dorinski (1995).

9. SAMPLING ERRORS

Estimates based on SIPP, like those from any sample survey, are subject to sampling error. Because SIPP uses a probability sample, data for sample persons and households can be used to estimate the sampling variance of any sample estimate. The Census Bureau estimates sampling errors for selected variables and uses these estimates to develop values of parameters for use in generalized variance functions (GVFs). The GVFs can be used by the Census Bureau and other users of SIPP data to estimate the sampling variance associated with any statistic of interest. Users of SIPP microdata files also have the option of estimating sampling errors directly from those files.

This chapter is divided into two sections. Section 9.1 describes and evaluates the procedures used to estimate sampling errors and provides references to more detailed information for data users who wish to use the GVFs or to estimate sampling errors directly, using the public use microdata files. Section 9.2 looks at the relationship between sample size and sampling error, the effects of sampling error on uses of SIPP data, and the prospects for reduction of the levels of sampling error associated with SIPP estimates.

This section is not intended to provide detailed information on how to estimate sampling error and construct confidence intervals for specific items. For that purpose, users of SIPP data should refer to the *SIPP User's Guide* (U.S. Census Bureau, 1998) and to the source and accuracy statements that appear in SIPP publications that are provided to purchasers of public use files.

9.1 Estimation of sampling errors

9.1.1 Current procedures

The Census Bureau calculates direct estimates of sampling error for a large set of selected items using a replication procedure that employs random complementary half-samples (Fay, 1989b). Because the first-stage sample for the 1984 Panel contained only one sample PSU per stratum, a collapsed stratum variance estimator had to be incorporated in the procedure for the non-self-representing strata in that panel. The procedure (described in detail by Roebuck, 1985 and 1986) used 50 replicate half-samples, which were not balanced. Subsequent panels have used two sampled PSUs per stratum design with a few exceptions, thus generally obviating the need for the collapsed stratum technique. For the 1990 Panel, a partially balanced half-sample replication variance estimation procedure with 100 replicates was used (Siegel and Mack, 1995; Williams, 1993).

When replication is used for variance estimation, the replicate weights should ideally be recomputed entirely for each replicate, starting with adjusted inverse selection probabilities, and continuing with replicate-specific nonresponse and ratio estimation adjustments. For the 1984 Panel, the Census Bureau employed a simpler approach of constructing the replicate weights using the nonresponse and ratio estimation adjustments for the total sample without modification for the particular replicate (termed "unreweighted weights" by Williams, 1993). For the 1990 Panel, the replicate weights were adjusted separately for each replicate for the ratio estimation adjustments, but not for the nonresponse adjustments (termed "reweighted weights" by Williams, 1993). The Bureau is examining the development of fully replicated weights for future variance estimation work with SIPP.

For each of several classes of estimates, the calculated variance estimates are used to estimate values of the parameters a and b for a generalized variance function (GVF) of the form: $s = (ax^2 + bx)^{1/2}$ where x represents the value of the estimate and s represents its standard error (the square root of its variance).

As explained in the *SIPP User's Guide* (U.S. Bureau of the Census, 1998), by using the appropriate parameter values in the GVF, it is possible to derive an approximate estimate of the standard error for many estimates based on SIPP, including not only estimates of totals, but also estimates of proportions or percents, differences between totals or percents, means, medians, ratios, and differences between means or medians. For statistics that involve differences or ratios, it is necessary to have an estimate of, or make some assumptions about, covariances for the variables involved.

The Census Bureau uses the GVFs and the estimated parameter values in two ways. When analytical statements based on SIPP estimates are published, all actual or implied comparisons are tested for statistical significance. For example, a statement that two estimates are different will not be published unless the estimated difference is at least 1.6 times its standard error, as determined by using the GVFs (using a 10 percent significance level).

In addition to making this internal use of the GVFs and parameter values, the Census Bureau includes GVFs in various publications, along with explanations of how to use them. With this information, users may calculate, using relatively simple methods, approximate estimates of sampling error for many statistics of interest to them, whether the statistics are based on published data or on their own tabulations of the public-use microdata files. A detailed presentation of information on sampling error appears in the Sampling Errors chapter of the *SIPP User's Guide* (U.S. Bureau of the Census, 1998). Publications in the P-70 series contain similar information, including GVF parameter values appropriate for the data presented in each publication.

Users of SIPP public-use microdata files can also make direct estimates of sampling error, using estimation procedures that they consider suitable for the SIPP sample design and the kinds of statistics of interest to them. The *SIPP User's Guide* (U.S. Bureau of the Census, 1998) provides instructions on how to do this. A detailed description of procedures for applying a balanced half-sample replication method to data in SIPP public-use files for direct estimation of sampling errors is given by Bye and Gallicchio (1988, 1989), who used this method to calculate sampling errors for estimates relating to participants in various Social Security programs. The Census Bureau is currently developing systems to allow users access to direct variance calculations for a wide variety of tables and other estimates (Waite, 1995b).

9.1.2 Evaluation

Estimates of sampling error are themselves subject to sampling error and, often, to biases as well. The direct estimates of sampling errors have two possible sources of bias:

- For a sample design like that used for the 1984 SIPP Panel, with only one sample PSU per stratum, it is not possible to obtain an unbiased estimate of sampling errors. Strata were combined before the half-sample replicates were created. This "collapsed stratum" procedure results in positively-biased estimates of variance. This source of bias does not apply with subsequent panels since they have employed a sample design with two sampled PSUs per stratum.

- As noted above, to simplify the construction of the replicate weights, these weights were not fully reweighted in the 1984 or 1990 Panels. As a result, the replicate variance estimators do not correctly reflect the effects of the nonresponse adjustments (both panels) and ratio estimation adjustments (1984 Panel).

Estimates of sampling errors based on the GVFs and associated parameter values are subject to some sources of error that are not associated with direct estimates. The parameters used are based on curves fitted to direct estimates of sampling variance for a sample of items in each given category and do not reproduce each of the individual estimates. A single adjustment factor is used to adjust the parameter values for all categories to reflect changes in sample size within and between waves or the use of less than four rotation groups, but it is not likely that such changes will affect all variables equally. Assumptions about correlations for variables used in estimates of differences or ratios may also introduce error. Special GVFs are being considered for the future for variances of differences, to include any covariance terms (Dajani, 1996).

Direct estimation of sampling errors by users is subject to some loss of precision in comparison with the Census Bureau's own estimates. The PSU and stratum codes in the public-use files, which are needed for variance estimation, are artificial ones, whereas the Census Bureau can use the actual PSU and stratum codes for its estimates. The actual codes are withheld from the public-use files in order to minimize the risk of disclosing confidential information by identifying individual records as belonging to geographic areas with small populations. A comparison of sampling errors estimated with the actual and pseudo-design PSU and stratum codes showed very small differences for most items and no evidence of bias (Bye and Gallicchio, 1989).

Using the GVFs to estimate sampling errors is less work for users, but the calculation of direct estimates should, in general, provide them with better estimates of sampling error for the specific statistics in which they are interested. In addition to making their own direct estimates, Bye and Gallicchio (1988) used these estimates to develop GVF parameter values for the class of variables they were studying (receipt and median benefit amounts from the Old-Age, Survivors, and Disability Insurance and Supplemental Security Income programs) and compared their GVF estimates with those derived using the Census Bureau's parameter values for the broad category (program participation and benefits, poverty) that appeared to include their estimates. They found that the Census Bureau GVF estimates of standard errors for small estimates (250,000 or fewer persons) were 50 percent or more above the estimates based on their parameter values. They attributed the differences largely to different strategies in the selection of the variables used to estimate the GVFs. Their selection gave considerably more weight to estimates at the smaller end of the range of estimates by size.

Research has also been carried out on the GVF parameter estimation methodology (Dajani, 1996). Eight alternatives to the $s = (ax^2 + bx)^{1/2}$ functional form were evaluated in this research, with the current functional form proving best among the alternatives. In addition, new domains were developed and new ways of selecting final GVFs for a domain (that is, new goodness of fit criteria) were developed.

9.2 Sampling errors and sample size

Because of reductions in the funds allocated for SIPP, sample sizes in the 1985 through 1989 Panels were decreased to a little more than half of the original 1984 Panel. About 21,000 households were interviewed in Wave 1 of the 1984 Panel. The Wave 1 sample size was reduced to 14,300 for the 1985 Panel and to about 12,400 for the 1986-1989 Panels. The sample size for the 1990 Panel increased to about 24,000 households, but that for the 1991 Panel was only about 16,000 households. The 1992 and 1993 Panels each started with about 22,000 households (see Table 2.1).

For some panels, sample reductions had been made after the first wave. An across-the-board cut of about 18 percent was made after Wave 4 of the 1984 Panel. In the 1984 Panel, interviews for two waves (Waves 2 and 8) were conducted with only three of the four rotation groups.

The sample for the 1985 Panel was cut twice. In February 1985 the sample was reduced by 5,280 interviewed households. For non-self-representing primary sampling units (PSUs), this cut was accomplished by eliminating 54 of the 198 sample PSUs. In February 1986 a further cut, in all of the remaining sample PSUs, eliminated an additional 1,960 interviewed households. At that time, the 1986 Panel was cut by the same amount. The eighth waves of the 1986 and 1987 Panels were eliminated entirely. Detailed information on the earlier sample cuts is given by Kasprzyk and Herriot (1986).

In addition to these cuts in sample size, which have been dictated largely by reductions in the funds available for the survey, the effective sample size in each panel is gradually reduced in successive waves due to attrition of sample households and persons.

Table 9.1 provides a rough indication of how sample sizes for various subpopulations of interest have been affected by the successive cuts in the SIPP sample of households. The numbers in the first column of the table are based on the initial sample for the 1984 Panel, about 20,000 households, and those in the second column are based on the initial sample for the 1989 Panel, about 12,000 households. As the numbers in this table indicate, the sample sizes for many population subgroups of interest were quite small in the 1985 through 1989 Panels, precluding detailed analyses of these groups.

Table 9.2 provides an indication of the level of sampling errors for some key items and the effect of reductions in initial panel sample size. The table shows estimated coefficients of variation (the standard error divided by the estimate) for selected estimates of average monthly mean income for the fourth quarters of 1984 and 1989. The 1984 values are taken from Series P-70, Number 6 (U.S. Bureau of the Census, 1985) and the 1989 estimates were obtained by multiplying the 1984 coefficients of variation by a factor of (20,000/11,600) that reflects the reduced sample size of the 1989 Panel. The 1989 estimates are rough approximations, because they do not take account of changes in the between primary sampling unit component of variance. The 1990, 1992, and 1993 Panels will have coefficients of variation similar to those for the 1984 Panel.

Studies have been undertaken to compare the ability of the Current Population Survey (CPS) and SIPP to detect changes in key income variables (Gbur, 1987). The CPS monthly sample of eligible households was about five times as large as the initial sample of households for the 1989 SIPP Panel. Not surprisingly, the studies indicate that a larger number of statistically significant differences could be detected using the CPS data. However, as discussed in Chapter 10, there is some evidence that nonsampling biases for most of the important SIPP variables that are covered in both surveys are smaller for SIPP than for CPS. In addition, SIPP provides data for many variables not included in the CPS. These considerations have suggested the possible use of an integrated SIPP/CPS estimation procedure which would minimize mean squared errors by taking full advantage of the larger CPS sample and the smaller nonsampling errors associated with SIPP data.

Table 9.1 Estimated number of SIPP sample persons for selected subpopulations

Subpopulation	Sample size	
	20,000 households	12,000 households
All persons	53,700	32,200
Adults	41,400	24,850
Persons 65 and over	5,965	3,580
Persons 75 and over	2,600	1,560
Persons in households with income less than poverty (monthly)	7,400	4,440
Recipients of:		
Social Security (aged and disabled)	7,475	4,485
Railroad retirement	175	105
AFDC	705	420
General assistance	245	150
SSI (Federal and State)	750	450
Medicare	6,510	3,905
Medicaid	4,125	2,475
WIC	570	340
Multiple recipients of:		
Food Stamps and AFDC	675	405
Food Stamps and SSI	285	170
Social Security and Food Stamps	385	230
Social Security and Housing Assistance	335	200
Medicaid and SSI	795	480
Food Stamps and Housing Assistance	315	190

Source: Kasprzyk (1989).

Table 9.2 Coefficients of variation for mean monthly household income for persons and households, fourth quarter 1984 and 1989

Characteristics	Estimates for 1984 (in dollars)	Coefficient of variation[1] 1984	Coefficient of variation[1] 1989
Persons			
All persons	2,615	0.011	0.015
Male	2,730	0.016	0.021
Female	2,507	0.016	0.020
White	2,743	0.012	0.015
Black	1,660	0.013	0.017
Hispanic	1,939	0.058	0.077
Spouse in married-couple family	2,920	0.016	0.021
Householder, no spouse present	1,635	0.043	0.057
Other family members	2,722	0.016	0.021
Not a family member	1,685	0.046	0.061
Households			
All households	2,327	0.012	0.016
Family households	2,668	0.012	0.016
Married couple families	2,920	0.013	0.018
Female householder, no husband present, own children <18	1,205	0.039	0.051
All other	2,078	0.031	0.040
Nonfamily households	1,446	0.037	0.048
Single-person household	1,306	0.044	0.058
Male	1,644	0.040	0.053
Female	1,080	0.080	0.105
Multiple-person household	2,509	0.041	0.054

[1]Coefficient of variation equals the standard error divided by the estimate.

Source: 1984 values from U.S. Census Bureau (1986). The 1989 values were estimated by: cv(1989) = cv(1984) x [20,000/11,600].

There are various possibilities for reducing the sampling errors of SIPP estimates. The most obvious is to combine data for adjacent panels. For any given calendar month from 1985 to 1993 (except for 1990),[1] there are at least two panels providing estimates for that month. The main question to be resolved is how to weight the observations from the different panels to minimize mean squared errors, taking into account differences in sample sizes and differential effects of nonresponse and time-in-sample biases. A procedure for obtaining combined estimates is described in the *SIPP User's Guide* (U.S. Bureau of the Census, 1998). Research on improved estimation procedures aimed at reduction of sampling errors for single-panel estimates was described in Chapter 8, Section 8.4.6 (see also Huggins and Fay, 1988). Investigations of the use of administrative income data in the second-stage ratio adjustments has produced encouraging results.

Oversampling

Another possible approach for increasing the precision of certain estimates is to oversample selected subgroups, such as persons in poverty, the aged, blacks, Hispanics, and participants in selected Federal income security programs. One way to do this would be to supplement the SIPP sample with samples drawn from current lists of program participants. This procedure was tested on various occasions in the ISDP; however, multiple-frame estimates were never produced and operational considerations led to a decision not to use a multiple-frame sampling procedure in SIPP (Kasprzyk, 1983; Bowie and Kasprzyk, 1987).

A Census Bureau working group investigated another method of oversampling selected groups: increasing the overall sample size for the initial wave of a panel and then reducing the sample for nontargeted groups in subsequent waves. The working group concluded that the procedure would provide only modest gains in precision for the targeted subgroups and that these gains would not outweigh the disadvantages, which included an increase in the complexity of survey operations, loss of a self-weighting design and substantial decreases in the precision of estimates for subgroups not oversampled (Singh, 1985).

Oversampling was implemented in the 1990 Panel by taking advantage of the discontinuation of the 1989 Panel after the third wave. A subsample was selected from the households interviewed at the first wave of the 1989 Panel to include all households that were headed by blacks, Hispanics, and females with no spouse present and living with relatives, and a small fraction of other households. This subsample was included as an auxiliary sample in the 1990 Panel, thereby leading to an oversample in that panel of households headed by blacks, Hispanics, and females with no spouse present and living with relatives (King, 1990).

The oversampling approach that has received most attention in Census Bureau research for SIPP is the use of 1990 Census data for oversampling in the 1996 Panel. The general approach is to use Census data for stratification purposes and to oversample those strata that contain greater proportions of subpopulations of special interest. The procedures are described briefly below.

The SIPP sample is selected mainly from four separate frames: (1) the Unit Frame that comprises lists of housing units in Census blocks that contain a high proportion of complete Census addresses and that are covered by building permit offices; (2) the Area Frame that comprises lists of housing units and group quarters in Census blocks that contain a high proportion of incomplete Census addresses or are not covered by building permit offices; (3) a frame of New Construction Permits from building permit offices to capture construction since the time of the Census; and (4) a Group Quarters frame for blocks covered by the unit frame. The oversampling research has addressed only the unit and area frames, which cover over 90 percent of the 1996 SIPP sample. No oversampling is considered for the new construction and group quarters frames.

[1] The 1988 and 1989 Panels were shortened and did not overlap with the 1990 Panel.

The research has focused on oversampling the poor and near poor, with a particular concern for blacks, Hispanics, and female-headed households in or near poverty. The methodology examined in the research by Weller et al. (1991) and Weller (1991) partitions the unit and area frames within each PSU into a poverty stratum and a nonpoverty stratum (where poverty is defined as under 150 percent of the poverty threshold). The procedures for constructing the two strata differ between the two frames because of the different sampling procedures adopted.

With the unit frame, housing units are sampled for SIPP directly from the frame. As a result, the poverty and nonpoverty strata can be defined at the household level. However, since income is collected in the Census on the Census long form which is administered to a sample of on average one-sixth of all households, only Census sample households could be assigned directly to the poverty or the nonpoverty stratum. All other households were assigned to one of the strata based on auxiliary variables collected on the Census short form that were predictive of poverty status. Thus, for example, households with a female householder, no spouse present and children under 18 and households in a central city of an MSA and renter with rent under $150 were assigned to the poverty stratum.

For the area frame, the housing units in the blocks sampled for SIPP are listed and sampled in the field prior to the first wave of the panel. For this frame, Weller et al. (1991) proposed a procedure of stratification at the block rather than the household level. They assumed that this procedure would reduce variances by only about one-half of the reduction achieved with the unit frame.

Weller et al. (1991) conducted research on the above procedure for 27 PSU equivalents from 1980 Census data. They also used data from the Annual Housing Survey to investigate the effect of changes over time in the effectiveness of the procedures. They examined the effect of the procedures on a wide range of variables, both those related and those not related to poverty. A serious concern about oversampling on poverty is that it will lower the precision of other estimates, for example, estimates for persons ages 55 and over and estimates for persons with high incomes.

Table 9.3 presents a selection of results from the research of Weller et al. (1991) on the effect of the oversampling procedure for the unit frame, not taking into account effects of changes over time. The table presents the ratio of the variance of a given estimate under the oversampling design to that under a design in which there is no oversampling. There are sizable gains for estimates related to poverty (that is, ratios less than 1), but loses for some other estimates. The gains in this table are overestimates of the actual gains because they relate only to the unit frame and because they do not reflect losses in efficiency caused by changes over time. Taking both these factors into account, the corresponding ratios from the first three rows in the table (that is, for blacks, Hispanics, and female householders, under 150 percent of the poverty level) are estimated to be 0.69, 0.80, and 0.85 at a time 8 years after the Census (assuming that similar changes occur in the 1990s as occurred in the period 1974-1985).

Given the importance of poverty households in SIPP, the results of this and other research led the Census Bureau to implement this oversampling design for the 1996 Panel (Huggins and Fischer, 1994; Siegel and Mack, 1995). In view of the inherent uncertainties involved, a fallback plan was included in the 1996 Panel design to return to a design without oversampling if the variance results of the oversampling design were not favorable. The effectiveness of the oversampling in the 1996 Panel is discussed in Section 11.2.

Table 9.3 Ratios of variances with the proposed oversampling design to those with a design without oversampling, unit frame only

Subgroup	Ratio of variance with the oversampling design to the variance without oversampling
Under 150% of the poverty level:	
Blacks	0.62
Hispanics	0.78
Female householders	0.84
Total persons	0.76
Number of adult blacks	0.71
Number of persons in urban areas	1.06
Number of persons ages 55+	1.05
Number of persons ages 65+	1.04
Households with yearly income:	
Under $5,000	0.82
$5,000 - $9,999	0.97
$10,000 - $14,999	1.04
$15,000 - $24,999	1.05
$25,000 - $34,999	1.08
$35,000 - $49,999	1.10
$50,000 - $74,999	1.12
$75,000 and over	1.13

Source: Weller et al. (1991).

10. EVALUATION OF ESTIMATES: 1984-1993 PANELS

Previous chapters have described specific aspects of SIPP design and operations and have presented information about errors associated with each of those aspects of the survey. The information on errors was derived mostly from three sources: performance statistics, such as noninterview rates; methodological experiments to test alternative design features; and direct measurement of errors associated with individual responses, as in the SIPP Record Check Study. Another way to evaluate the quality of survey data is to compare the survey estimates with data from other sources that are at least roughly comparable with respect to population coverage and definition of variables. These external sources are generally of two kinds: other surveys (or censuses), and record systems maintained primarily for administrative rather than statistical purposes.

This chapter discusses evaluation studies that compare key estimates from SIPP with estimates from other data sources. Section 10.1 discusses the available data sources, the methods for constructing estimates for comparisons, and the interpretation of the results. Section 10.2 presents the results of evaluations of SIPP estimates of income and poverty. Section 10.3 discusses the evaluation of other SIPP estimates including: labor force participation, assets and liabilities, wealth, vital events, migration, and other personal information.

SIPP data users and others are encouraged to consult some of the references cited in this chapter. Vaughan (1989a, b, c) provides an especially comprehensive evaluation of the cross-sectional estimates of income recipiency and amounts in the 1984 Panel. Coder et al. (1987) present several comparisons of estimates of income based on the SIPP 1984 Panel three-interview, 12-month longitudinal research file with CPS estimates. Coder and Scoon-Rogers (1996) compare the SIPP and CPS estimates of income for 1984 and 1990, and examine changes in relationships over time. For other estimates, the data evaluation memoranda that are routinely prepared by Census Bureau subject-matter specialists and made available to purchasers of data tapes based on topical modules are a useful source of information. In addition to comparisons between SIPP estimates and external data, these evaluation memoranda frequently contain other kinds of information related to the quality of the data, such as sample loss rates for the wave in which the topical module was included, module noninterview rates, item nonresponse rates, and the frequency of edit changes for specific items.

10.1 Data sources

Administrative sources of comparative data for SIPP income estimates include the National Income and Product Accounts (NIPA), individual tax return data, and the Federal Reserve Board's Flow of Funds Accounts. Often, estimates from such external sources are not directly comparable with SIPP estimates, in which case the estimates from SIPP or the external source, or both, may need to be adjusted to make them more comparable. Observed differences between SIPP and external estimates are subject to both sampling and nonsampling errors and, in most instances, both sources of data contain these errors. Although administrative or program data are often thought to be closer to the truth, it is not always obvious that any statistically significant differences observed between administrative and survey estimates are entirely due to errors in the survey data. Even when it is fairly clear that the survey data are in error, both the size and the sources of the errors may be far from obvious.

Nevertheless, such macro-evaluation studies can be valuable. Many studies of this type, of varying scope and intensity, have been undertaken. Careful interpretation of the results of these studies has led to some specific procedural improvements, especially in the areas of question design, editing, and imputation procedures. In a broader context, these studies are an important part of the Census Bureau's continuing effort to inform users of SIPP data about the quality of the data and to ascertain how well the main goals of the survey (see Section 2.2) are met.

Other surveys

Besides comparisons with administrative data, SIPP estimates are often compared with estimates from other major surveys, including the Current Population Survey (CPS), the Panel Study of Income Dynamics (PSID), and the Survey of Consumer Finances (SCF). Most comparisons are with the CPS because there are many similarities between the two surveys. However, there are also many differences, including survey design, conceptual differences, and nonsampling errors. The most fundamental difference, noted by Coder and Scoon-Rogers (1996), is that the SIPP and the CPS are designed for different purposes. The focus for SIPP is income and program participation, whereas for CPS, it is employment statistics, and income questions are secondary. The different focus affects the content and the mode of data collection for the surveys. The SIPP uses multiple interviews at 4-month intervals and a reference period of 4 months between interviews to facilitate recall. Income is collected month-by-month, using person-based questions. The CPS, in contrast, is constrained by the need to produce monthly employment statistics within 3 weeks after the interviewing week. It collects income in a single annual interview, the reference period is the preceding calendar year, and questions are household-based.

There are several alternative ways in which aggregate income amounts for various sources can be computed from SIPP for comparison with the corresponding CPS aggregates. One method, employed by Coder and Scoon-Rogers (1996), was to create March-based estimates from SIPP, using a special purpose March-based file which fixed the SIPP panel household composition as of March following the reference year. For example, to derive the estimates for 1984, a data file was created based on the SIPP panel household composition fixed as of March of 1985. Then calendar year estimates were constructed by summing monthly values for the reference year from this March-based file. This method was developed specifically for the purpose of comparisons with the CPS; the CPS collects income data for the previous calendar year for persons who are members of the household in March.

Another method of constructing SIPP annual aggregate estimates, say income aggregates, is to sum the products of monthly estimates of income recipients multiplied by mean amounts derived from the "wave" files (used by Vaughan, 1989a, b, c). The number of income recipients is estimated using monthly cross-sectional weights controlled to monthly population estimates. This method is limited in that it is based on the sum of monthly cross-sectional estimates, and it does not provide an estimate of the number of income recipients "ever receiving income" during the calendar year. Relative to this sum-of-waves method, the March-based estimates are expected to be smaller because the March-based universe excludes the income received by persons who died or entered the institutional population or the Armed Forces between January 1 of the reference year and the time of interview in March of the succeeding year.

A third method for constructing SIPP annual aggregate estimates is to use the longitudinal files. These files contain income and recipiency data by month over the length of the SIPP panel. Annual income amounts for each sample person can be created by summing the amounts received during each month of the calendar year. Then the longitudinal-based aggregate income estimates can be derived by multiplying the annual amounts by the appropriate person weight and summing the products across the entire population. The calendar year weights on this file sum to the estimated noninstitutional population as of January 1 of the calendar year, thus excluding those persons entering the institutional population after January 1 (see Section 8.3). In theory, the longitudinal-based method is expected to produce estimates lower than the sum-of-waves monthly estimates but higher than the March-based estimates. In practice, the differences between the three estimates are relatively small. Table 10.1 shows the SIPP estimates of the 1990 income aggregates based on the three estimation methods from SIPP, the March CPS, and the independent estimate based on NIPA.

Table 10.1 Alternative estimates of aggregate income for 1990: SIPP (March-based, longitudinal-based, and sum-of-waves), March CPS, and independent source

Source of income	Aggregate Income ($ billions)				
	SIPP				
	March-based	Longitudinal-based	Sum-of-waves	March CPS	Independent Source[10]
Wage and salaries[1]	$2,475.7	$2,472.5	$2,337.5	$2,614.0	$2,695.6
Self-employment[2]	267.7	275.0	283.0	228.1	341.4
Interest income	150.8	157.9	152.5	172.8	282.8
Dividends income	58.2	62.4	61.6	39.5	126.3
Rents and royalties[3]	45.4	48.7	43.0	38.7	44.1
Estates and trusts[4]	6.9	7.1	7.0	7.1	N/A
Social Security income	221.6	229.2	227.5	209.8	225.5
Railroad retirement	6.6	6.8	6.8	4.6	6.9
Supplemental Security income	12.9	13.2	12.9	12.1	13.6
AFDC[5]	13.6	13.8	14.3	14.1	19.7
Other cash welfare[6]	2.5	2.4	2.4	2.5	2.9
Unemployment compensation	14.9	14.3	14.7	14.2	17.7
Worker's compensation[7]	12.6	12.4	12.8	13.8	14.6
Veterans' payments	11.6	11.9	12.1	10.7	13.8
Private pensions[8]	75.2	77.0	75.6	77.8	70.2
Federal employee pensions	22.3	23.6	23.6	25.1	30.4
Military retirement	18.8	19.3	18.9	18.2	20.4
State and local employee pensions	27.1	28.9	28.8	28.9	36.1
Alimony	2.9	2.9	2.9	3.1	2.5
Child support	13.7	13.8	13.5	11.8	N/A
Financial assistance	8.2	7.7	8.6	8.5	N/A
Other cash income[9]	10.2	10.3	9.3	4.6	N/A

[1]SIPP and CPS wage and salary estimates include earnings from incorporated businesses operated by self-employed persons.
[2]SIPP and CPS self-employment estimates exclude earnings from incorporated business operated by self-employed persons.
[3]The 1990 SIPP estimate includes "other financial investments," with "rents and royalties." The 1990 CPS rents and royalties estimate corrects the estimate printed in the P60, No. 174 report.
[4]The CPS estimate of estates and trusts for 1990 refers solely to survivor estates and trusts. The CPS estimate for 1984 includes all estates and trusts (that is, both survivor and other sources of estates and trusts).
[5]The CPS estimate of AFDC for 1990 includes the entire sum of public assistance when respondent indicated receipt of both types of public assistance (0.6 billion dollars in AFDC and other cash welfare).
[6]The CPS estimate of "other cash welfare" for 1990 excludes "other cash welfare" when respondent indicated receiving both other cash welfare and AFDC.
[7]The CPS estimate of worker's compensation includes (beyond P-60 series published, line item definition) black lung payments, accident insurance, and temporary insurance.
[8]The SIPP estimate of private pensions includes pensions not classified anywhere else. It excludes lump-sum payments. CPS private pensions include (beyond P-60 series published, components) survivor's annuities, other unspecified survivor benefits and unspecified disability benefits.
[9]Some cash income sources are excluded from "other cash income" and from all income sources identified above. "Other cash income" as shown is definitionally incomparable over time and across surveys.
[10]Independent estimates derived from National Income and Product Accounts (NIPA).
Source: Coder and Scoon-Rogers (1996: Tables 1, 2, and 8).
N/A = not available.

10.2 Estimates of income and poverty

This section summarizes evaluations of SIPP estimates of regular money incomes from various sources and percentages of persons in poverty. The sources of income include: earnings, property income, transfer payments, pension income, alimony, and other sources. Most of the results and their interpretations are taken from Coder and Scoon-Rogers (1996) who compared the SIPP estimates of income for 1984 and 1990 with estimates from other independent sources and estimates from the March CPS. The independent sources include the NIPA, Federal individual income tax returns, and caseload data from various transfer programs. Most of the comparisons were made after detailed adjustments had been carried out to compensate for differences in universe coverage and concepts of measurement. Tables 10.2 through 10.4 present annual estimates of aggregate income amounts, numbers of recipients, and mean income amounts. The SIPP estimates in these tables are the March-based estimates; the exception is for AFDC, the estimates of which are longitudinal-based.

10.2.1 Overall results

Evaluation studies have shown that survey estimates of income are consistently lower than estimates from independent sources. This is true for SIPP and for estimates from other surveys, such as the CPS (see Table 10.2). The SIPP estimates of aggregate income for many income types are comparable to those from CPS, and surpass the CPS estimates for some income types (for example, Social Security, Railroad Retirement, and Supplemental Security Income). As shown in Table 10.3, the SIPP estimates of the numbers of recipients also surpass the CPS estimates for most income sources. When estimates of the annual mean amounts are computed based on estimates of aggregate amount and number of recipients, SIPP estimates of mean amounts are lower than the CPS estimates in some cases and higher in others. The number of SIPP estimates which are lower is more pronounced in 1990 than in 1984.

Coder and Scoon-Rogers (1996) suggest several factors that may explain SIPP-CPS differences shown in Tables 10.2-10.4. These include: changes in the March CPS processing system, subtle differences in SIPP questionnaire wording, the special characteristics of the SIPP 1990 Panel (which included the potentially "low-income" households carried over from the 1989 Panel), and changes in the way income is received (especially for pension income where the incidence of lump-sum disbursements has been increasing). They argue that the larger number of recipients in SIPP, relative to CPS, suggests that the multiple interview scheme in the SIPP captured more short-term recipients who may have been missed in the CPS. The shorter interviewing intervals in SIPP are designed to improve the reporting of incomes that are typically received on an irregular or part-year basis. Coder and Scoon-Rogers, however, found no obvious explanation of why "the SIPP provides lower counts of higher income recipients" than does the CPS.

Table 10.2 Comparisons of SIPP and CPS estimates of aggregate income as percentages of independent estimates, 1984 and 1990

Source of income	1984			1990			Standard error of percent			
							1984		1990	
	Independent estimate ($ billions)	SIPP (percent)	MARCH CPS (percent)	Independent estimate ($ billions)	SIPP (percent)	MARCH CPS (percent)	SIPP	CPS	SIPP	CPS
Wage and salaries[1]	1,820.1	91.4	97.3	2,695.6	91.8	97.0	0.7	0.5	0.8	0.5
Self-employment[2]	192.6	103.1	70.2	341.4	78.4	66.8	1.9	1.7	2.8	1.9
Interest	244.8	48.3	56.7	282.8	53.3	61.1	2.2	0.8	2.1	0.8
Dividends	59.3	65.9	51.8	126.3	46.1	31.3	3.8	0.9	5.9	1.5
Rents and royalties[3]	19.4	211.3	95.4	44.1	102.9	87.8	6.8	3.4	5.0	3.6
Estates and trusts[4]	N/A	N/A	N/A	N/A	N/A	N/A				
Social Security	160.5	96.2	91.9	225.5	98.3	93.0	1.2	0.8	1.4	0.8
Railroad retirement	5.6	96.4	71.4	6.9	95.7	66.7	8.7	5.8	0.7	5.3
Supplemental Security Income	9.9	88.9	84.8	13.6	94.9	89.0	3.7	2.2	4.0	2.0
AFDC[5]	13.9	83.5	78.4	19.7	70.1	71.6	3.1	2.0	4.3	2.2
Other cash welfare[6]	2.0	135.0	120.0	2.9	86.2	86.2	1.2	0.8	1.4	0.8
Unemployment compensation	16.3	76.1	74.8	17.7	84.2	80.2	2.8	2.3	3.1	1.9
Worker's compensation[7]	14.1	56.7	48.2	14.6	86.3	94.5	5.5	4.8	4.3	2.9
Veterans' payments	13.9	82.0	59.7	13.8	84.1	77.5	5.8	3.6	5.8	2.8
Private pensions[8]	65.2	63.8	57.2	70.2	107.1	110.8	3.4	2.4	2.5	1.4
Federal employee pensions	20.3	98.0	84.7	30.4	73.4	82.6	4.9	3.3	6.9	3.5
Military retirement	15.6	105.1	98.1	20.4	92.2	89.2	6.9	4.4	9.0	4.5
State and local employee pensions	21.9	88.1	71.7	36.1	75.1	80.1	4.2	2.8	5.5	2.7
Alimony	2.7	100.0	81.5	2.5	116.0	124.0	6.0	2.0	8.5	1.1
Child support	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Financial assistance	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Other cash income[9]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Note: See footnotes to Table 10.1.
Source: Coder and Scoon-Rogers (1996: Table 2).
N/A = not available.

Table 10.3 Comparisons of SIPP and March CPS income recipient estimates (millions of persons), 1984 and 1990

Source of income	1984			1990			Standard error of percent			
							1984		1990	
	SIPP	March CPS	SIPP as a percent of CPS	SIPP	March CPS	SIPP as a percent of CPS	SIPP	March CPS	SIPP	March CPS
Wage and salaries[1]	116.0	114.4	101.4	127.3	124.6	102.2	0.5	0.3	0.4	0.3
Self-employment[2]	11.6	12.2	95.1	12.9	18.1	71.3	0.3	0.2	0.2	0.2
Interest	124.3	99.0	125.6	132.9	108.5	122.5	0.5	0.3	0.4	0.3
Dividends	26.9	20.6	130.6	31.8	23.3	136.5	0.4	0.2	0.4	0.2
Rents and royalties[3]	17.6	12.5	140.8	17.4	13.5	128.9	0.3	0.2	0.3	0.2
Estates and trusts[4]	0.7	1.5	46.7	0.5	0.5	100.0	0.1	0.1	0.0	0.0
Social Security	33.4	32.2	103.7	36.9	35.4	104.2	0.4	0.2	0.4	0.3
Railroad retirement	0.8	0.6	133.3	0.9	0.6	150.0	0.1	0.0	0.1	0.0
Supplemental Security Income	3.9	3.6	108.3	4.6	4.0	115.0	0.2	0.1	0.1	0.1
AFDC[5]	4.0	3.6	111.1	4.0	3.9	102.6	0.2	0.1	0.1	0.1
Other cash welfare[6]	2.0	1.4	142.9	1.5	1.3	115.4	0.1	0.1	0.1	0.1
Unemployment compensation	8.6	7.7	111.7	9.2	7.6	121.1	0.2	0.1	0.2	0.1
Worker's compensation[7]	3.2	2.5	128.0	3.7	2.8	132.1	0.1	0.1	0.1	0.1
Veterans' payments	4.4	2.9	151.7	3.4	2.6	130.8	0.2	0.1	0.1	0.1
Private pensions[8]	10.5	8.0	131.3	13.7	11.5	119.1	0.2	0.1	0.2	0.1
Federal employee pensions	2.0	1.6	125.0	1.8	1.9	94.7	0.1	0.1	0.1	0.1
Military retirement	1.6	1.5	106.7	1.6	1.4	114.3	0.1	0.1	0.1	0.1
State and local employee pensions	3.3	2.6	126.9	3.3	3.2	103.1	0.1	0.1	0.1	0.1
Alimony	0.6	0.3	200.0	0.6	0.4	150.0	0.1	0.0	0.1	0.0
Child support	4.0	3.2	125.0	5.2	4.1	126.8	0.2	0.1	0.2	0.1
Financial assistance	2.6	1.4	185.7	2.7	1.8	150.0	0.1	0.1	0.1	0.1
Other cash income[9]	2.0	7.8	25.6	2.3	2.0	115.0	0.1	0.1	0.1	0.1

Note: See footnotes to Table 10.1.
Source: Coder and Scoon-Rogers (1996: Table 3).

Table 10.4 Comparisons of SIPP and March CPS mean income estimates, 1984 and 1990

Source of income	1984			1990			Standard error of the mean			
							1984		1990	
	SIPP	March CPS	SIPP as a percent of CPS	SIPP	March CPS	SIPP as a percent of CPS	SIPP	March CPS	SIPP	March CPS
Wage and salaries[1]	$14,339	$15,478	93	$19,448	$20,979	93	$114	$67	$141	$85
Self-employment[2]	17,112	11,090	154	20,752	15,951	130	268	258	296	348
Interest	952	1,401	68	1,135	1,593	71	42	20	46	21
Dividends	1,454	1,490	98	1,830	1,695	108	130	44	149	47
Rents and royalties[3]	2,330	1,480	157	2,609	2,856	91	162	56	166	102
Estates and trusts[4]	8,143	5,200	157	13,800	13,170	105	1,413	557	2,181	1,318
Social Security	4,623	4,581	101	6,005	5,927	101	33	18	39	22
Railroad retirement	6,720	6,667	101	7,333	8,069	91	439	318	434	361
Supplemental Security Income	2,256	2,333	97	2,804	2,981	94	69	38	78	46
AFDC[5]	2,950	3,028	97	3,450	3,590	96	93	52	98	57
Other cash welfare[6]	1,350	1,714	79	1,667	2,095	80	103	66	116	76
Unemployment compensation	1,442	1,584	91	1,620	1,868	87	46	30	47	34
Worker's compensation[7]	2,500	2,720	92	3,405	4,759	72	154	147	191	204
Veterans' payments	2,591	2,862	91	3,412	4,082	84	159	102	191	137
Private pensions[8]	3,962	4,663	85	5,489	6,771	81	115	88	146	120
Federal employee pensions	9,950	10,750	93	12,389	12,966	96	460	265	522	299
Military retirement	10,250	10,200	100	11,750	12,522	94	546	319	570	368
State and local employee pensions	5,848	6,038	97	8,212	9,068	91	264	146	337	205
Alimony	4,500	7,333	61	4,833	6,566	74	601	650	588	546
Child support	2,250	2,250	100	2,635	2,850	92	95	59	102	71
Financial assistance	2,231	3,429	65	3,037	4,546	67	169	190	209	237
Other cash income[9]	2,650	2,090	127	4,435	2,342	189	384	130	586	137

Note: See footnotes to Table 10.1.
Source: Coder and Scoon-Rogers (1996: Table 4).

10.2.2 Earnings

Wage and salary

Wage and salary earnings are the main component of income. Independent estimates of the annual aggregate amount, based on the NIPA, were $1,820 billion for 1984 and $2,695 billion for 1990 (Table 10.2). The SIPP estimate amounted to 91 percent of the independent estimate in 1984 and to about 92 percent in 1990. The CPS estimates amounted to about 97 percent of the independent estimates for both years. The numbers of wage and salary earners estimated from SIPP were higher than those from CPS in both years (Table 10.3). As a result, the SIPP estimates of mean earnings were lower than those from the CPS by about 7 percent in both years (Table 10.4).

When compared to income reported on Federal individual tax returns, both 1990 SIPP and CPS survey estimates of wages and salaries, as shown in Table 10.1, are closer to the alternative benchmark, with the CPS estimate surpassing the tax aggregate (the 1990 tax return amount was $2,576 billion). The NIPA-based independent estimate (shown in Tables 10.1 and 10.2) includes deferred amounts of wage and salary income, the wages of nonfilers, and estimates of unreported wages, while the amounts reported on tax returns exclude these components. In 1984, the CPS estimate was below the aggregate tax return wage and salary amount. The SIPP estimates were lower than the tax return amount for both 1984 and 1990.

Coder and Scoon-Rogers speculate that the relative advantage of CPS over SIPP in estimating earnings may be attributed to differences in the data collection environments. The SIPP environment is conducive to reporting monthly "take-home" pay, and therefore, may have a tendency to miss the amounts from "extra" paychecks received by persons paid weekly or biweekly. To address this possibility, changes have been made to the SIPP data collection process to include instructions to respondents about the "extra" paycheck (see Section 4.5.7). The CPS environment, in contrast, is conducive to reporting the current earnings annualized, rather than the total amount received during the preceding calendar year.

Self-employment

Self-employment earnings are difficult to evaluate because comparable independent estimates are not readily available, and different independent sources provide conflicting results. Based on the NIPA, a survey-consistent estimate of the aggregate amount was $341.4 billion in 1990. Both the SIPP and the CPS aggregate estimates fall well short of that figure (see Table 10.2). However, the SIPP and CPS estimates are far greater than the estimate obtained from individual tax returns (approximately $136.3 billion). One explanation is that the NIPA estimate included self-employment income that was not, but should have been, reported on tax returns.

The SIPP and CPS comparison is also difficult because of conceptual differences in measurement. In SIPP, self-employment earnings are measured using the concept of "salary or draw." For persons with self-employment, the SIPP core questions ask for two different kinds of information related to earnings. For each month of the 4-month reference period, a self-employed person is asked to give the total amount of income received from the business each month, that is, the "draw," or income that the person took out of the business to meet personal or family expenses. In addition, the person is asked for an estimate of the net profit or loss (gross receipts minus expenses) for the entire 4-month period. The CPS estimates are based on the concept of annual net income.

Coder and Scoon-Rogers' compared the SIPP estimates based on the concept of "draw" with the CPS estimates and found the SIPP estimates to be higher. Estimates based on the "net income" concept were not investigated in this study. However, an earlier study by Vaughan (1989b) found that the SIPP estimates of net income came to less than half of the CPS estimates for 1984.

Since the concept of self-employment as a labor force activity is the same for both surveys, it is possible to make direct comparisons of estimates for the numbers of recipients. These comparisons identify a rare situation in which the SIPP estimates of recipients were lower than the CPS estimates in both years. In 1990, the CPS yielded an estimate of 18.1 million recipients compared to 12.9 million from the SIPP. For mean income, the SIPP estimate amounted to 154 percent of the CPS estimate in 1984, and to 130 percent in 1990.

10.2.3 Property income

The main components of property income are interest, dividends, net rent and royalties, estates and trusts. Evaluation studies have shown, consistently, that survey estimates of property income are generally poor. Although there is relatively high reporting of rents and royalties, there is significant underreporting of interest and dividend incomes, and confusion about these two sources. The methods used to compensate for missing data also add further variation.

The confusion over interest income and dividend income is partly due to respondents' inability to correctly distinguish between them, and partly due to the different practices used to classify these income sources. Income from mutual funds (money market funds and other mutual funds) is classified as interest in the NIPA. However, in SIPP, earnings from mutual funds are explicitly divided into interest and dividends. Earnings from "money market funds" are classified as interest and earnings from "other mutual funds" are classified as dividends. For tax purposes, returns on "mutual funds" are reported as dividends.

Comparisons between SIPP and CPS estimates of interest income are complicated by the use of different methods to handle missing data. The CPS estimates include a post-imputation adjustment that attempts to correct for known biases in the imputation of missing amounts, whereas this adjustment is not employed in SIPP. The application of this adjustment inflated the 1984 CPS estimate of interest income by 26 percent, a rate that contributes to most of the observed differences between the SIPP and CPS estimates.

The SIPP estimate of dividend income was $39.1 billion for 1984 and $58.2 billion for 1990. These estimates exceeded the CPS estimates of $30.7 billion and $39.5 billion for 1984 and 1990 respectively. The estimates from both surveys are short of independent estimates in both years. For 1990, the dividend income based on NIPA was $126.3 billion; whereas, the estimate based on tax return was $78.6 billion. The two sources provided different estimates because (1) dividend incomes received by IRA and KEOGH plans, and by nonprofit and other fiduciaries were included in the NIPA estimate but not in the tax return estimate; (2) the tax return estimate was lower because of non-filing and omissions in filed tax returns.

The SIPP estimate of income from rents and royalties was $41 billion for 1984, amounting to 211 percent of the independent estimate from NIPA at $19.4 billions. In contrast, the CPS estimate for 1984 amounted to 95.4 percent of the independent estimate. For 1990, the estimates from the three sources were $45.4 billion from SIPP, $44.1 billion from NIPA, and $38.7 billion from CPS. The combined rent and royalty, and estates and trusts income was $52.3 billion from SIPP and $45.8 billion from CPS. Coder and Scoon-Rogers (1996) discuss the difficulties with making comparisons on these incomes.

10.2.4 Transfer payment programs

This section discusses estimates of income from the following transfer payment programs: Social Security, Railroad Retirement, Supplement Security income (SSI), Aid to Families with Dependent Children (AFDC), other cash welfare programs, unemployment compensation, worker's compensation and Veterans' payments. The Food Stamps Program is not discussed because participants in this program do not receive direct income.

Social Security, Railroad Retirement, Supplemental Security Income

SIPP produces higher estimates of income from the Social Security, Railroad Retirement, and Supplemental Security Income programs, than the CPS. SIPP's estimates are also closer to information derived from program sources. Social Security is the largest of the transfer payment programs and is also the one for which SIPP does the best. This was true for both 1984 and 1990 and the same finding emerged from other evaluation studies. An independent estimate of the aggregate amount for this income type in 1990 was $225.5 billion. The SIPP estimate amounted to 98 percent of this estimate, exceeding the CPS estimate by about 5 percent. The situation was similar in 1984.

For the Railroad Retirement program, the SIPP aggregate estimate is over 96 percent of the estimate from program data. The CPS estimate of the aggregate amount is about 67 percent of the program estimate.

The SIPP estimate of aggregate income from Supplemental Security Income (SSI) amounted to 95 percent of the estimate from program sources in 1990. The SIPP estimate used in this comparison was adjusted to reflect a change in the SSI payments for children (see Coder and Scoon-Rogers, 1996).

AFDC and other public assistance

Public assistance income includes Aid to Families with Dependent Children (AFDC), and a small amount of other cash payments. Studies conducted with ISDP data showed that there were significant misclassification errors in reporting the sources of cash welfare benefits (Klein and Vaughan, 1980; Goudreau et al., 1984). In particular, AFDC payments were frequently misreported as general assistance or general welfare. Results of the SIPP Record Check Study also confirmed this finding (see Section 6.3.4). The independent estimate of total aggregate income from public assistance (that is, AFDC and other cash welfare programs combined) was $22.6 billion in 1990. Both the SIPP and the CPS estimates were significantly lower, amounting to about 72 percent of this estimate.

There has been concern that the SIPP estimates of the numbers of recipients for AFDC (and also for Food Stamps) are much lower than the numbers based on administrative data, and the extent of underestimation may have increased over time. Shea (1995a) examined this issue using point-in-time estimates of underreporting as of January of each year, for 1985 through 1993. The results show that estimates for a given point-in-time vary, depending on which panel was used. SIPP estimates based on the first wave of interviewing compared more favorably with administrative estimates than SIPP estimates based on later waves. Focusing on the Wave 1 estimates alone, there is no obvious increase in underreporting over the years.

Comparisons of SIPP estimates and program data on transitions in and out of AFDC (also Food Stamp and SSI programs) have been reported by Singh (1987). The results of these studies are summarized in connection with the discussion of the seam phenomenon in Section 6.1.

Unemployment and workers' compensation

An independent estimate of the 1990 aggregate amount of unemployment compensation derived from NIPA was $17.7 billion. The independent-estimate of the amount for workers' compensation was $14.6 billion for that year (see Coder and Scoon-Rogers, 1996, Appendix A). SIPP estimates amounted to 84 percent and 86 percent of these independent estimates.

Table 10.5 shows comparisons, for 10 quarters covered by the 1984 Panel, of SIPP estimates of recipiency and amounts of unemployment compensation with estimates derived from independent sources. The trend by quarter was erratic, with SIPP performing well for the last two quarters of 1983, the last quarter of 1984 and the first quarter of 1985, but considerably less well in other quarters. No clear cut explanation has emerged, but Vaughan (1989a) suggests that the apparent underreporting in SIPP may have been due in part to the wording of the relevant questions and the skip sequences used in interviews after the initial wave.

Table 10.5 Comparisons of SIPP state unemployment compensation estimates with estimates derived from independent sources, 1984 Panel

Period	SIPP[2]		Independent estimate		SIPP as a percent of independent estimate	
	Recipients (thousands)	Aggregate dollars (millions)	Recipients[1] (thousands)	Aggregate dollars[1] (millions)	Recipients	Aggregate dollars
1983						
3rd Quarter	3,084	$1,287	3,056	$1,259	100.9	102.2
4th Quarter	2,878	1,193	2,784	1,117	103.4	106.8
1984						
1st Quarter	2,982	1,206	3,608	1,415	82.6	85.2
2nd Quarter	2,212	897	2,682	1,079	82.5	83.1
3rd Quarter	1,927	762	2,456	949	78.5	80.3
4th Quarter	2,462	978	2,590	969	95.1	100.9
1985						
1st Quarter	3,225	1,393	3,771	1,470	85.5	94.8
2nd Quarter	2,220	927	2,872	1,193	77.3	77.7
3rd Quarter	1,917	783	2,633	1,078	72.8	72.6
4th Quarter	1,981	854	2,506	1,103	79.1	77.4

[1]Excludes Federal Supplemental Compensation.
[2]SIPP estimates are based on monthly averages for specified quarter.
Source: Coder (1987a).

Veterans' payments

The independent estimate of the aggregate amount of veterans' payments for 1990 was $13.8 billion. The SIPP estimate amounted to 84 percent of this figure, exceeding the CPS estimate by 6 percent. SIPP performance relative to the CPS for this income type shows some deterioration over time. Coder and Scoon-Rogers attribute this change to a substantial drop in the SIPP estimate of the number of recipients from 4.4 million in 1984 to 3.4 million in 1990.

10.2.5 Pension income

Pension income includes income from private pensions, Federal employee pensions, military retirement, and State and local pensions. An independent estimate of the combined total pension income for 1990 was about $157 billion. The SIPP estimate amounted to 91 percent of the independent estimate, and is comparable to the CPS estimate. An evaluation by Vaughan (1989a) found that the SIPP estimate appeared better than the CPS estimate in 1984, but that finding was not apparent in 1990. Coder and Scoon-Rogers (1996) provide a more detailed account of the comparison by type of pension income and the derivation of the independent estimates.

10.2.6 Child support and other sources of income

Other sources of income reported on SIPP include alimony, child support, financial assistance, and other cash income. Independent estimates for these income types are not readily available. For alimony, Coder and Scoon-Rogers (1996) used estimates from individual tax returns. The SIPP estimates of the aggregate alimony income compare well with this independent source in both 1984 and 1990.

For child support, the SIPP estimates in 1990 exceed the CPS in aggregate amounts by 16 percent (see Coder and Scoon-Rogers, 1996, Table 1). Scoon-Rogers and Lester (1995) suggest that this finding is probably due to SIPP's ability to capture more child support recipients through monthly data collection compared with the annual data collection in the CPS. The SIPP estimates of child support recipients exceed the CPS estimates by 25-27 percent (see Table 10.3). However, the CPS estimates of the mean amount are higher than the SIPP estimates.

Scoon-Rogers and Lester (1995) also examined child support, comparing SIPP estimates from the topical modules with CPS estimates for April 1992. Their results, are shown in Table 10.6. The SIPP topical module estimates 12.7 million custodial parents with children under 21 in the fourth quarter of 1991, compared with the 11.5 million estimated in the CPS in April 1992 for the period October 1991 to January 1992. Scoon-Rogers and Lester also found that the number of custodial parents with child support awards is greater in the SIPP (6.9 million) than in the CPS (6.2 million). However, the percentage of custodial parents with child support awards, 54 percent, is not different across the surveys. The percentage of custodial fathers among custodial parents, approximately 13 percent, is not different across the two surveys. The SIPP estimate of the average custodial parents' child support income, from the fourth quarter of 1990 to the fourth quarter of 1991, was $3,010 per year. This amount is not significantly different from the $2,961 reported in the April 1992 child support supplement of the CPS.

Table 10.6 Comparison of child support data: SIPP estimates based on the child support topical module in 1991 and CPS estimates based on the April 1992 Child Support Supplement

Characteristic	CPS	SIPP
Custodial parents		
Total (millions)	11.5	12.7
Female	9.9	11.1
Male	1.6	1.6
Custodial parents with awards		
Total (millions)	6.2	6.9
Female	5.5	6.4
Male	0.6	0.4
Mean child support in 1991[1]		
Total (dollars)	2,961	3,010
Female	3,011	2,263
Male	2,292	3,040

[1]CPS child support income reported for calendar year 1991. SIPP child support income reported for 12 consecutive months between September 1990 and December 1991.
Source: Scoon-Rogers and Lester (1995).

10.2.7 Number of persons in poverty

Important statistics related to income are the number and percent of persons in poverty. Currently, official annual estimates of these statistics are obtained from the CPS March Income Supplement. SIPP is an alternative data source that can provide these estimates.

Table 10.7 shows SIPP and CPS estimates of the percent of persons in poverty for 1984, 1985, 1990, and 1991. The 1984 and 1985 estimates are based on the full longitudinal file for the 1984 SIPP (McNeil et al., 1988). The 1990 and 1991 estimates are based on estimates from Shea (1995b). Both studies show that the SIPP estimates were below the CPS estimates for all of the demographic categories shown in the table. Williams (1987) discusses three possible reasons for the lower SIPP estimates:

- More complete reporting of transfer income in SIPP.
- The SIPP definition of self-employment income does not allow reporting of negative values, whereas the CPS definition does. Further, SIPP collects data on the "draw" from self-employment, rather than on the net profit as in CPS. See the discussion of self-employment earnings in Section 10.2.1.
- SIPP provides information on changes in household composition that occur during the reference year. Coder et al. (1987) showed that poverty estimates based on the 1984 SIPP Panel that took account of changes in household composition were lower than those that followed the CPS definition, which fixed the household composition as of March of the following year period.

Table 10.7 Percent of persons below poverty based on data from the CPS and SIPP, 1984, 1985, 1990, and 1991

Characteristics	1984		1985		1990		1991	
	CPS	SIPP	CPS	SIPP	CPS	SIPP	CPS	SIPP
All persons	14.4	11.5	14.0	11.0	13.5	10.1	14.2	10.6
Sex:								
Male	12.8	10.0	12.3	9.4	11.7	8.2	12.3	8.9
Female	15.9	12.9	15.6	12.4	15.2	11.9	16.0	12.2
Race:								
White	11.5	8.7	11.4	8.5	10.7	7.5	11.3	8.1
Black	33.8	30.4	31.3	28.3	31.9	27.0	32.7	27.1
Hispanic origin[1]	28.4	24.6	29.0	22.6	28.1	21.2	28.7	24.7
Age:								
Under 18	21.5	17.8	20.7	16.9	20.6	16.8	21.8	17.2
18 to 64	11.7	8.8	11.3	8.4	10.7	7.7	11.4	8.3
65 and over	12.4	10.8	12.6	10.9	12.2	8.1	12.4	8.5

[1]Persons of Hispanic origin may be of any race.
Source: McNeil et al. (1988) for 1984 and 1985 Panels; Shea (1995b) for the 1990 and 1991 Panels.

Subsequent analysis, however, has shown that the effect of fixing the household composition is relatively small (Shea, 1995b). For example, the CPS estimate of the poverty rate for 1990 was 13.5 percent while the SIPP estimate was 10.3 percent. When the SIPP household composition was fixed as of March 1991 to match the CPS universe, the SIPP estimate increased to 10.8 percent, still significantly less than the CPS estimate. Other analysts who have studied the effects of different poverty definitions on SIPP estimates include Hoppe (1988), Ruggles and Williams (1989), and Ruggles (1990).

10.3 Other estimates

This section summarizes evaluation studies on estimates other than income from SIPP including: labor force participation; assets, liability and wealth; vital events (birth, marriages and divorces); migration history; and other personal information (health insurance coverage and tax return filing status).

10.3.1 Labor force data

Employment status

Hill and Hill (1986) compared information on employment status from SIPP and PSID for the period July through December 1983, with emphasis on transitions between employment and unemployment and the length of spells of unemployment. They found that for comparable time periods and populations, SIPP produced estimates of unemployment incidence about 15 percent higher than those obtained in PSID, and they attributed this difference to the more frequent interviewing schedule of SIPP. The differences were larger for men than for women and Hill and Hill suggest that this may have resulted partly from a higher frequency of proxy reporting in SIPP (the idea being that proxy respondents might fail to report short periods of employment interspersed with extended periods of unemployment).

Working population

The SIPP is an alternative data source to CPS for producing estimates of the working population. As noted in Section 10.1, a primary goal of the CPS is the production of monthly labor force data whereas in SIPP, which focuses on income and program participation, the inclusion of labor force questions is primarily a means of ensuring accurate reporting of earnings. As explained in detail by Ryscavage and Bregger (1985), there are substantial conceptual and operational differences in the collection of labor force data in the two surveys.

Early comparisons of SIPP and CPS quarterly estimates of labor force status showed that SIPP estimates of persons looking for work (or on layoff) were consistently higher than CPS estimates of persons unemployed. Conversely, CPS estimates of persons not in the labor force were higher than SIPP estimates of persons with no labor force activity (Kasprzyk and Herriot, 1986).

Data on annual work experience are collected retrospectively in the CPS March Supplement. Coder et al. (1987) compared estimates from this source with estimates from Waves 1 to 4 of the 1984 SIPP Panel longitudinal research file. The results of this comparison, given in Table 10.8, show that SIPP yielded higher estimates of both the working population and the population looking for work or on layoff. Of particular interest is the larger SIPP estimate of persons not working at all but spending one or more weeks looking for work or on layoff. More detailed comparisons indicate that the two surveys identified similar numbers of year-round workers but that SIPP estimates of part-year workers were higher.

Ryscavage and Coder (1989) compared SIPP and CPS data on work experience in 1985. CPS data were from the Income and Work Experience Supplement conducted in March 1986. Overall estimates of the number of persons with any work experience during 1985 were similar for the two sources, but the comparisons showed several important differences in the distribution of annual work experience. Table 10.9 compares estimates of full- and part-time workers, by number of weeks worked during the year, from the two sources. For full-time workers, the SIPP estimates are smaller at the extremes of the distribution, but larger in the intermediate categories. Ryscavage and Coder hypothesize that the differences in reference periods for the two surveys are mainly responsible for the difference. The longer CPS reference period places a greater burden on respondents' memories.

Gross flows in labor force

Ryscavage and Feldman-Harkins (1988) compared SIPP and CPS estimates of month-to-month gross flows in labor force status during 1984. They concluded that:

> "As would be expected given the survey design of SIPP, its flows were generally smaller than those from the CPS, and more consistent with the net changes in its stock estimates than was the case with CPS flows and stocks. The quality of the SIPP flows, however, requires further investigation, specifically the effect of response error on the estimates. In addition, it would be useful to observe SIPP gross labor force flows in another phase of the business cycle."

Table 10.8 Estimates of work experience for persons ages 16 and older for 1983-1984, SIPP and CPS (in thousands)

Work experience	SIPP			CPS		
	Total	Male	Female	Total	Male	Female
Number						
Total persons ages 16 and older	175,862	83,467	92,396	176,711	83,746	92,965
Total with 1 or more weeks of work experience	130,190	70,261	59,929	122,810	66,792	56,018
Worked, total	122,969	67,403	55,566	119,362	65,236	54,126
With weeks looking or on layoff	23,738	13,113	10,625	19,200	11,491	7,709
Without weeks looking or on layoff	99,231	54,290	44,941	100,162	53,745	46,417
Did not work but had one or more weeks looking or on layoff	7,221	2,858	4,363	3,450	1,556	1,894
Total without any weeks of work experience	45,672	13,206	32,467	53,901	16,954	36,948
Percent						
Total persons ages 16 and older	100.0	100.0	100.0	100.0	100.0	100.0
Total with 1 or more weeks of work experience	74.0	84.2	64.9	69.5	79.8	60.3
Worked, total	69.9	80.8	60.1	67.5	77.9	58.2
With weeks looking or on layoff	13.5	15.7	11.5	10.9	13.7	8.3
Without weeks looking or on layoff	56.4	65.0	48.6	56.7	64.2	49.8
Did not work but had one or more weeks looking or on layoff	4.1	3.4	4.7	2.0	1.9	2.0
Total without any weeks of work experience	26.0	15.8	35.1	30.5	20.2	38.7

Source: Coder et al. (1987).

Table 10.9 SIPP and CPS work experience estimates (in thousands), by full-time and part-time status, 1985

Extent of employment	SIPP	CPS	Difference
Total	124,655	124,101	554
Full-time (35 or more hours/ week)	94,812	96,443	-1,631[1]
50-52 weeks	68,981	72,324	-3,343
40-49 weeks	9,986	8,082	1,904[1]
27-39 weeks	7,402	5,519	1,883[1]
14-28 weeks	5,047	5,770	-723[1]
13 weeks or less	3,397	4,747	-1,350[1]
Part-time (less than 35 hours/ week)	29,843	27,658	2,185[1]
50-52 weeks	10,444	10,205	239
40-49 weeks	4,319	3,329	990
27-39 weeks	5,292	3,335	1,957[1]
14-26 weeks	4,700	4,861	-161
13 weeks or less	5,088	5,927	-839[1]

[1]Significant at the .05 level.
Source: Ryscavage and Coder (1989).

10.3.2 Assets, liabilities, and wealth

Data on assets and liabilities were collected in topical modules for each SIPP panel (except the 1989 Panel). Table 10.10 shows a comparison of SIPP estimates with estimates derived from the Flow of Funds data of the Federal Reserve Board (FRB) for 1984 and 1988. Overall, the SIPP estimate of net worth for the household sector comes to 92 percent of the FRB estimate in 1984, but only 84 percent in 1988.

The overall level of agreement between SIPP and FRB is the net result of rather large offsetting differences in major categories of the balance sheet. SIPP estimates are substantially higher than those of the FRB for equity in owner-occupied housing and motor vehicles, and substantially below the FRB estimates for equity in noncorporate business, financial assets, and consumer debt. Analysts from the Census Bureau and other organizations have commented at some length on the 1984 comparisons. Census analysts concluded that the FRB estimate of home equity was not a good reference figure and that there was strong evidence that financial assets were underreported in SIPP (McNeil and Lamas, 1988). Other analysts agreed that the FRB home equity estimates were low and also pointed out that the FRB estimates of financial assets may have been too high because of the difficulty of distinguishing between household and business holdings of liquid assets (Curtin et al., 1989).

Table 10.10 Asset and liability estimates from SIPP and the Federal Reserve Board (FRB) balance sheet data for the household sector, for 1984 and 1988 (in billions)

Category	1984			1988		
	FRB balance sheet	SIPP	Ratio of SIPP to FRB balance sheet	FRB balance sheet	SIPP	Ratio of SIPP to FRB balance sheet
A. Equity in owner-occupied housing	$2,316.3	$2,823.6	1.22	$3,042.1	$3,628.6	1.19
Gross value	3,606.4	3,958.2	1.10	5,180.6	5,235.0	1.01
Debt	1,290.1	1,134.6	0.88	2,138.5	1,606.4	0.75
B. Equity in motor vehicles	287.0	410.5	1.43	424.3	490.3	1.16
Gross value	459.6	558.8	1.22	708.9	741.0	1.05
Debt	172.6	148.3	0.86	284.6	250.7	0.88
C. Equity in noncorporate business	2,235.1	1,680.2	0.75	2,410.7	1,764.9	0.73
Rental property	*	909.6	*	*	1,025.9	*
Other business equity	*	770.6	*	*	739.0	*
D. Financial assets	3,858.9	2,826.1	0.74	5,753.5	3,813.2	0.66
Interest-earning assets[1]	3,167.5	1,635.7	0.52	4,348.5	2,432.5	0.56
Corporate equities[2]	1,403.2	1,062.7	0.76	2,171.4	1,114.2	0.51
Other financial assets[3]	128.2	127.8	1.00	176.6	266.5	1.51
Less: Financial assets held by nonprofit sector or in personal trusts	(840.0)	N/A	N/A	(943.0)	N/A	N/A
E. Installment and other consumer debt[4]	379.9	241.5	0.64	409.1	245.8	0.60
F. Net Worth (A+B+C+D-E)	8,122.9	7,498.8	0.92	11,221.5	9,451.2	0.84

[1]Includes passbook savings accounts, money market deposit accounts, certificates of deposit, checking accounts, money market funds, and other interest-earning assets.
[2]Includes equities in stocks, mutual funds shares, and incorporated self-employed business or professions.
[3]Includes mortgages held by sellers and other financial assets not otherwise specified.
[4]Excludes debt for automobiles and mobile homes.
*Separate estimates not available.
N/A: Not applicable.
Source: Eargle (1990, Table D-2).

Curtin et al. (1989) compared data on wealth from SIPP, the PSID, and the 1983 Survey of Consumer Finances (SCF). The SIPP data used in the comparison were collected in Wave 4 of the 1984 Panel and represented averages of holdings at the end of August, September, October, and November 1984. The PSID data were collected in the 17th annual interview wave of the survey, conducted from March through September of 1984. The household sample for the 1983 SCF was interviewed from April through July of 1983, and a special set of interviews with pension providers was conducted from September to December 1983. The household sample for the SCF included a supplementary sample of high-income households in order to improve representation of the upper tail of the wealth distribution.

The comparisons made by Curtin et al. (1989) are extensive, including percent distributions of aggregate amounts by net worth size category for major asset types and total net worth, estimates of mean asset amounts by net worth size category for several asset types, and estimates of mean asset amounts by income size category for major asset types. Based on these comparisons, Curtin et al. conclude that:

> "Measured against the standards set by previous wealth surveys, all three of these data sets stand up quite well. They do not differ substantially among themselves when it comes to measuring total wealth and the distribution of wealth in the great bulk of the U.S. population."

They assert that the SCF had the highest overall potential for wealth analysis because of its supplemental sample of high-income households and its primary focus on obtaining detailed information about all types of assets and liabilities. The PSID collects less detail on individual asset types than SIPP. Curtin et al. argue that the PSID level of detail was sufficient to produce estimates of net worth useful as an independent variable in analysis.

10.3.3 Vital events

Estimates of vital events, including births, marriages and divorces, can be developed from SIPP in two ways. One way is to use cross-sectional retrospective data collected in fixed topical modules on marital, fertility, and migration history (for example, in Wave 8 of the 1984 Panel, Wave 4 of the 1985 Panel, and Wave 2 of 1992 Panel). The other is to use the longitudinal survey procedures, which record births as persons entering the sample and update marital status at each interview. Either type of estimate can be compared with the vital statistics compiled by the National Center for Health Statistics (NCHS) from registration certificates issued by the States. The estimates discussed in this section have been derived mainly from the topical modules.

Births

With respect to births, O'Connell (1988) reported that estimates of births from the Wave 8 topical module of the 1984 Panel were lower than vital statistics counts. McMillen (1989a) used a file of matched wave data from 1984 Panel cross-sectional files to estimate monthly births for 1984 and 1985. On this basis, he concluded that the SIPP estimates of births in that Panel were in fairly close agreement with the vital statistics data.

The SIPP estimate of first births from the 1992 Panel Wave 2 Fertility History Module is close to the CPS estimate from the June 1992 Fertility Supplement. Both survey estimates, however, are significantly below the number of births reported in the Vital Statistics records. Table 10.11 shows the SIPP and CPS comparison of fertility data for women who had a child within 1 year of the interview date. There are no significant differences between the survey estimates.

Table 10.11 Number of births per 1,000 women, by selected characteristics, for women ages 15-44 years who had a child last year, 1992

Selected characteristics		Number of births per 1,000 women	
		SIPP 1992	CPS June 1992
Total		61.6	62.9
Age	15 to 29 years	83.8	85.9
	15 to 19 years	41.2	38.0
	20 to 24 years	99.0	103.4
	25 to 29 years	105.0	109.0
	30 to 44 years	41.7	42.9
	30 to 34 years	81.0	76.1
	35 to 39 years	31.9	38.0
	40 to 44 years	6.3	8.7
Race[1]	White	59.2	61.6
	Black	77.7	69.2
	Asian or Pacific Islander	54.2	63.5
	Hispanic	96.0	95.2
	Non-Hispanic	57.7	59.5
Marital status	Currently married	85.2	87.7
	Married, husband present	85.5	89.5
	Married, husband absent[2]	81.9	65.1
	Widowed or divorced	25.8	26.7
	Never married	36.6	35.1
Educational attainment	Less than high school	69.9	66.8
	High school, 4 years	62.2	65.1
	College: 1 or more years	57.2	59.7
	1 to 3 years	55.0	57.9
	4 or more years	60.0	62.4
Occupation	Managerial and professional	43.4	46.7
	Technical, sales, and admin. support	39.4	46.5
	Service occupations	42.8	45.7
	Farming, forestry, and fishing	51.5	43.7
	Precision production craft and repair	55.6	43.6
	Operators, fabricators, and laborers	41.8	31.2
Region of residence	Northwest	56.1	61.0
	Midwest	66.2	60.6
	South	60.9	62.5
	West	62.7	68.0
Metropolitan residence	Metropolitan	62.3	63.6
	In central cities	67.3	69.2
	Outside central cities	59.0	59.9
	Nonmetropolitan	59.2	60.4

[1]Persons of Hispanic origin may be of any race.
[2]Includes women separated from husbands.
Source: Unpublished Census tabulation: An Evaluation of SIPP Wave 2 Fertility History Module 1992 and June CPS, 1992.

Marriages and divorces

Table 10.12 shows estimates of the percent of persons ages 15 and older by current marital status in SIPP and CPS. The SIPP estimates were derived from the Wave 2 marital history topical modules of the 1987, 1988, and 1990 through 1992 Panels. The CPS estimates were derived from the March CPS of the corresponding years. Estimates of current marital status from both surveys are similar in all years.

Table 10.12 Percent of persons ages 15 and older by marital status from SIPP and CPS, 1987-1992

	Percent of persons ages 15 and older				
Marital status	SIPP 1987 Panel	SIPP 1988 Panel	SIPP 1990 Panel	SIPP 1991 Panel	SIPP 1992 Panel
Married, spouse present	75.5	75.1	74.8	74.4	74.6
Married, spouse absent					
Separated	3.5	3.3	3.2	3.3	3.2
Other	1.0	0.8	1.0	1.0	1.0
Widowed	9.5	9.2	9.7	9.7	9.1
Divorced	10.5	11.6	11.3	11.6	12.1
Number of Persons (thousands)	138,536	140,249	141,878	143,821	144,143
Marital Status	CPS March 1987	CPS March 1988	CPS March 1990	CPS March 1991	CPS March 1992
Married, spouse present	76.0	75.7	75.2	74.9	75.6
Married, spouse absent					
Separated	3.2	3.2	3.3	3.3	3.4
Other	1.2	1.3	1.0	1.1	1.1
Widowed	9.6	9.7	9.8	9.6	9.7
Divorced	9.9	10.1	10.7	11.1	11.5
Number of Persons (thousands)	137,503	138,956	141,570	142,217	144,096

Tables 10.13 shows a comparison of the number of marriages by year of marriage reported in each SIPP Panel between 1984 to 1988, 1990 to 1992, and the numbers in the NCHS, the June 1984 CPS, and June 1990 CPS. These numbers of marriages from SIPP were slightly lower than what was found in vital records from NCHS. Table 10.14 shows the number of divorces by year of divorce from the same sources. The numbers from the 1992 SIPP were again generally lower than the numbers provided by NCHS.

The reason for the apparent underestimation of marriages and divorces is not well understood. Initially it was felt that the use of mover weights (see Chapter 8) might be a major factor leading to underestimation of marriages (O'Connell, 1987a, 1987b; Hernandez, 1987). However, after further review of the estimation procedures, Singh (1988a) concluded that the mover weights "...obtain the correct probabilities of persons in the sample at a given time." He suggested that improved survey estimates of marriages might be achieved both by using different variables in the noninterview and second-stage ratio estimate adjustments and by improving the procedures used to capture data on marriages in the interviews.

Table 10.13 Number of marriages[1] by year of marriage for females, ages 15 and older: Comparisons of SIPP, vital statistics (NCHS), and CPS (in thousands)

	SIPP								NCHS	CPS	
Year	Wave 8 Panel 1984	Wave 4 Panel 1985	Wave 2 Panel 1986	Wave 2 Panel 1987	Wave 2 Panel 1988	Wave 2 Panel 1990	Wave 2 Panel 1991	Wave 2 Panel 1992		June 1985	June 1990[2]
1991	---	---	---	---	---	---	---	2,247	2,371	---	---
1990	---	---	---	---	---	---	2,588	2,329	2,448	---	---
1989	---	---	---	---	---	2,298	2,524	2,425	2,404	---	2,243
1988	---	---	---	---	---	2,621	2,284	2,284	2,389	---	2,266
1987	---	---	---	---	2,204	2,366	2,313	2,399	2,403	---	2,365
1986	---	---	---	2,369	2,348	2,195	2,049	2,430	2,407	---	2,244
1985	1,908	2,072	2,200	2,388	2,563	2,424	2,426	2,350	2,413	---	2,336
1984	1,862	2,058	2,579	2,510	2,492	2,353	2,128	2,416	2,477	2,363	2,155
1983	1,988	2,413	2,356	2,528	2,274	2,366	2,317	2,214	2,446	2,413	2,224
1982	2,396	2,282	2,477	2,296	2,230	2,415	2,244	2,145	2,456	2,379	2,352
1981	2,221	2,207	2,445	2,268	2,577	2,045	2,279	2,398	2,422	2,365	2,236
1980	2,345	2,467	2,377	2,387	2,570	2,320	2,449	2,516	2,390	2,416	2,474
1975	2,232	2,115	1,961	2,165	2,289	2,098	1,939	2,136	2,153	2,206	2,222
1970	2,304	2,098	1,952	2,057	1,956	2,057	2,195	2,007	2,159	2,234	2,282
1965	1,737	1,528	1,723	1,644	1,733	1,641	1,705	1,533	1,800	1,712	1,624
1960	1,500	1,463	1,566	1,492	1,347	1,349	1,442	1,466	1,523	1,577	1,429

[1]Number reflects only three marriages for women married more than three times.
[2]Data for women ages 15 through 65 only.
Source: Unpublished Census tabulation: A Review of the 1992 SIPP Wave 2 Marital History Topical Module.

Table 10.14 Number of divorces[1] by year of divorce for females, ages 15 and older: Comparisons of SIPP, Vital Statistics Records, and CPS (in thousands)

Year	SIPP								Vital Statistics Records	CPS	
	Wave 8 Panel 1984	Wave 4 Panel 1985	Wave 2 Panel 1986	Wave 2 Panel 1987	Wave 2 Panel 1988	Wave 2 Panel 1990	Wave 2 Panel 1991	Wave 2 Panel 1992		June 1985	June 1990[2]
1991	---	---	---	---	---	---	---	1,068	1,187	---	---
1990	---	---	---	---	---	---	1,039	917	1,175	---	---
1989	---	---	---	---	---	898	1,039	1,133	1,163	---	999
1988	---	---	---	---	---	904	980	929	1,183	---	1,012
1987	---	---	---	---	944	1,113	957	943	1,166	---	900
1986	---	---	---	869	944	844	898	1,037	1,178	---	917
1985	765	966	1,091	1,005	1,062	1,204	1,057	1,001	1,190	---	986
1984	665	869	867	1,126	1,082	956	938	920	1,169	1,062	963
1983	725	1,100	1,146	1,028	1,011	794	792	883	1,158	939	836
1982	912	857	1,089	1,048	1,034	1,159	816	894	1,170	918	992
1981	651	799	1,037	864	1,015	802	924	1,110	1,213	1,035	922
1980	770	1,109	1,029	930	1,116	1,016	1,076	1,052	1,189	1,075	1,057
1975	638	849	896	886	812	819	880	850	1,036	796	793
1970	440	579	594	582	481	549	559	517	708	561	538
1965	261	441	323	481	314	347	339	322	479	392	286
1960	245	220	252	257	368	225	334	229	393	284	180

[1]For women married more than three times, divorce data refer only to first, second, and last marriage.
[2]Data for women ages 15 through 65 only.
Source: Unpublished Census tabulation: A Review of the 1992 SIPP Wave 2 Marital History Topical Module.

10.3.4 Migration history

One source of migration data from SIPP is the topical module on migration history, which collects retrospective reports of migration. Using these data from the 1984 and 1985 Panels, DeAre (1988a,b) found that the SIPP estimates of the numbers of persons who had moved in the 12 months prior to interview about 10 percent below the corresponding CPS estimate.

Another measure of migration in SIPP can be developed through observing the movement of sample members over the course of a panel. Clark and Speare (1988) used data from the first four waves of the 1984 SIPP panel to calculate estimates of residential mobility and interstate migration, using data for persons who left the SIPP sample as well as those who remained in the sample for all 4 waves. The resulting SIPP estimates of residential mobility by age were in reasonably close agreement with the averages of CPS estimates for 1984 and 1985, taking into account the differences in reference periods and universe definitions for the two surveys. The results for interstate migration were less satisfactory; for persons under age 35, the SIPP estimates were below the CPS estimates.

Using the Wave 2 migration history module of the 1992 panel, an evaluation by Census staff found that the number of persons ages 15 years and older who changed residence between April 1991 and March 1992 was 31.6 million, according to the SIPP. The corresponding estimate from the CPS was 33.2 million. The estimated mobility rate (proportion of persons moving) was 16.2 percent from the SIPP, which is not significantly different from the CPS estimate of 17.0 percent. The characteristics of recent movers from the two surveys also are quite similar (Table 10.15). The SIPP estimates for both male and female recent movers ages 15 years and older were not statistically different from the CPS estimates. The percent of movers by age group also was similar between the surveys.

Table 10.16 shows the percent of recent movers ages 15 years of age and older from the SIPP and CPS, by type of mover (long or short distance, or from abroad). The CPS estimates included all persons from the March 1992 CPS who moved during the preceding 12 months. The SIPP estimates were based on those persons from the 1992 Panel whose last move was between April 1991 and March 1992. Geographic information in the SIPP module was only obtained for the most recent move, therefore, it was unavailable for about 10 percent of the SIPP movers in the period April 1991 to March 1992 who also made another move after March 1992. About 70 percent of the recent moves reported in the SIPP were within the same county, the CPS estimate was 61 percent. This difference was attributed to differences in data collection. The SIPP asks respondents whether the previous residence was in the same county, same state; whereas, the CPS asks respondents to provide the city and state or foreign county of previous residence.

Table 10.15 Sex and age distribution of movers in 1991-1992 from SIPP and CPS (in thousands)

Sex and age groups	Movers		Percent of Movers	
	SIPP	CPS	SIPP	CPS
Both sexes	31,570	33,202	100.0	100.0
15-19 years	2,528	2,892	8.0	8.7
20-24 years	6,024	6,529	19.1	19.7
25-29 years	6,805	6,654	21.6	20.0
30-34 years	5,264	5,065	16.7	15.3
35-44 years	5,693	5,992	18.0	18.0
45-54 years	2,634	2,846	8.3	8.6
55-64 years	1,178	1,570	3.7	4.7
65 years and over	1,443	1,654	4.6	5.0
Males	15,482	16,512	49.0	49.7
15-19 years	1,168	1,297	3.7	3.9
20-24 years	2,779	3,089	8.8	9.3
25-29 years	3,435	3,415	10.9	10.3
30-34 years	2,647	2,606	8.4	7.8
35-44 years	2,980	3,147	9.4	9.5
45-54 years	1,292	1,499	4.1	4.5
55-64 years	621	786	2.0	2.4
65 years and over	561	673	1.8	2.0
Females	16,088	16,686	51.0	50.3
15-19 years	1,360	1,595	4.3	4.8
20-24 years	3,245	3,439	10.3	10.4
25-29 years	3,371	3,239	10.7	9.8
30-34 years	2,617	2,458	8.3	7.4
35-44 years	2,713	2,845	8.6	8.6
45-54 years	1,342	1,347	4.3	4.1
55-64 years	557	783	1.8	2.4
65 years and over	881	980	2.8	3.0

Source: Unpublished Census tabulation: An Evaluation of 1992 Wave 2 Migration History Topical Module.

Table 10.16 Percent of recent movers ages 15 years and older from SIPP and CPS, by type of move: 1991-1992[1]

Type of move	Percent	
	1992 SIPP Wave 2	March 1992 CPS
From abroad	2.2	3.2
To different State	12.2	16.8
Within same State	85.6	80.0
Same county	69.6	61.2
Different county	16.0	18.8
Total	100.0	100.0

[1]Includes all persons from the March 1992 CPS who moved during the preceding 12 months; includes only those persons from the 1992 SIPP whose last move was between April 1991 and March 1992.
Source: Unpublished Census tabulation: An Evaluation of 1992 Wave 2 Migration History Topical module.

10.3.5 Other comparisons

Health insurance coverage

Among other comparisons of SIPP and CPS data, McNeil (1988) analyzed estimates of health insurance coverage status from the two surveys. CPS estimates were based on the 1985 March Supplement, and SIPP estimates were monthly averages for the first quarter of 1985. He found that SIPP estimates of persons without any health insurance coverage were 16 percent lower than the CPS estimates. He concluded that the SIPP estimates were more nearly correct and were of more analytical interest than the CPS estimates because of the reference period used for the latter.

Tax-return filing status

The topical module for Wave 6 of the 1984 Panel included questions on respondents' filing status for their 1984 Federal income tax returns. Table 10.17 compares distributions of persons filing returns based on SIPP data from the Property Income and Taxes module used in Wave 6 of the 1984 Panel, with distributions from the IRS Statistics of Income program. The SIPP estimates substantially understate the proportion of persons filing as unmarried heads of households and overstate the number of married joint filers (Coder, 1987b). Note, however, that the SIPP estimates use unweighted data.

Table 10.17 Comparison of the distributions of SIPP (unweighted) and IRS tax filers by type of return, 1984

Type of return	Percent		
	IRS	SIPP	Difference (IRS-SIPP)
Single	41.7	42.0	-0.3
Married, filing jointly	47.9	51.3	-3.4
Married, filing separately	0.8	1.5	-0.7
Unmarried head of household	9.4	5.0	4.4
Qualifying widow(er) with dependent children	0.1	0.3	-0.2
Total	100.0	100.0	-----

Source: Coder (1987b).

11. THE 1996 REDESIGN

In the early 1990s, the Census Bureau began an extensive review of the SIPP design in light of the experience gained with the SIPP panels to that time. That review led to a substantial redesign that was introduced in 1996, in conjunction with a sample redesign based on 1990 Census data. The main features of the SIPP redesign are as follows:

- Extension of the length of a panel to 4 years,
- Introduction of a new panel every 4 years, with minimal overlap between panels,
- Introduction of computer-assisted interviewing (CAI),
- Oversampling of low-income households, and
- Improvements in sample design.

This chapter describes the new design. Section 11.1 discusses the motives and justification behind the changes. Section 11.2 summarizes the sample design and selection. Section 11.3 describes the use of CAI in data collection and management. Section 11.4 discusses concerns about nonresponse with a panel of longer duration. Section 11.5 describes the weighting and imputation procedures for the 1996 Panel.

11.1 Overview of redesign

In its review of the SIPP design, the Census Bureau conducted a substantial outreach to collect views from both users and potential users of SIPP data. This activity was wide-ranging and included Federal agencies, universities, and research groups. The Bureau also asked the Committee on National Statistics (CNSTAT) of the National Research Council to convene a panel to carry out an independent review of SIPP. Weinberg and Petroni (1992) report some of the findings of the Bureau's review and Citro and Kalton (1993) present the findings and recommendations of the CNSTAT panel. The main issues identified by these efforts are as follows:

- *Panel length.* Prior to 1996, the typical panel length was 32 months. SIPP users found this period too short for many longitudinal analyses, such as measuring durations of spells of program participation, poverty, and health insurance coverage, and for comparing the living conditions of individuals before and after some transition such as divorce.
- *Sample size.* The sample size for a single panel with an overlapping panel design is relatively small, and often inadequate for detailed analysis of subgroups such as blacks, Hispanics, and the poor. Although overlapping panels can in principle be combined to increase sample size, the process of combining panels has presented a barrier to the use of this approach.
- *Timeliness.* Over time, the timeliness of data releases from SIPP has improved considerably. Since 1990, SIPP microdata have been distributed electronically and cross-sectional files are available about 6 months after data collection. Nevertheless, timeliness remains an important issue to users.
- *Attrition.* By Wave 8 the nonresponse rate rises to about 22 percent in a typical SIPP panel. Users are concerned about nonresponse bias.

In response to these concerns,

- The Census Bureau established the provision of data for longitudinal analysis as the primary goal for SIPP, and accordingly, extended the panel length to 4 years.

- To address sample size issues, the Bureau selected an essentially nonoverlapping design which concentrates resources in a single panel in the field at one time (the relative advantage of overlapping and nonoverlapping designs are discussed by Huggins and Fischer, 1994, and Citro and Kalton, 1993).

- The Bureau also decided to oversample persons with low incomes in order to improve the reliability of data for low-income related subgroups of policy interest.

- In common with its other surveys, the Bureau introduced computer-assisted personal interviewing (CAPI) and computer-assisted telephone interviewing (CATI) for data collection, a feature that should enhance the timeliness of data release.

A major concern about extending the life of the panel to 4 years is the increased nonresponse that may result. The Bureau has therefore adopted new procedures to follow movers and has experimented with the use of incentives. All these procedures are described in more detail in later sections of this chapter.

11.2 Sample design and selection

A sample redesign was introduced with the 1996 SIPP panel to reflect results of the 1990 Census. The panel started with a sample of about 37,000 households, including an oversampling of low-income households. The panel plans to follow members of those households for 4 years. Siegel and Mack (1995) provide an overview of the 1996 Panel sample design. New panels of a similar design are planned to start in 2000 and 2004. This section outlines the selection of sample PSUs (counties or groups of counties), the selection of addresses in sample PSUs, the method used for oversampling low-income households, and the population coverage of the 1996 sample.

Selection of Sample PSUs

The Census Bureau redefines PSUs every 10 years when population information becomes available after each decennial Census. The SIPP PSUs after the 1996 redesign are based on the 1990 Census; SIPP PSUs for panels between 1985 and 1993 are based on the 1980 Census. Siegel and Mack (1995) summarize the differences between the 1990 PSUs and the 1980 PSUs and Johnson (1990) gives details of changing PSU definitions for 1990.

The sample for the 1996 Panel is drawn from 322 PSUs, of which 112 are self-representing (SR) PSUs selected with certainty, and the remaining 210 PSUs are non-self-representing (NSR) units. The SR PSUs consist of 102 PSUs in 40 States. Each contained at least 196,513 housing units in 1990. For the remaining 10 States, the largest PSU is included with certainty, that is, treated as a SR PSU so that there is at least one certainty PSU per State. The NSR PSUs consist of a stratified sample selected from 105 strata, with two PSUs selected from each stratum. The initial stratification is by the four Census regions. Within each region, 15 further stratification variables are employed. The NSR PSUs are sampled with probability proportional to the numbers of housing units in the 1990 Census, using a linear programming algorithm that maximized the overlap between the 1980 and the 1990 sampled PSUs. Using this algorithm, the overlap attained between the 1990 sample of NSR PSUs and the 1980 sample PSUs is 76 percent, as compared with an overlap of about 30 percent had the algorithm not been used.

A major effect of the redesign is that the sample for the 1996 Panel contains more PSUs than the sample for earlier panels. Thus, the present sample is drawn from 322 PSUs, as compared with 230 PSUs for the 1985-1991 Panels and 284 PSUs for the 1992-1993 Panels. An advantage of the larger number of PSUs in the 1996 Panel is that the between PSU variance is smaller; in addition, there are more degrees of freedom for the estimation of variances for survey estimates.

Selection of addresses in sample PSUs

Within sample PSUs, living quarters are systematically selected from lists of addresses developed from the 1990 Census. Five separate frames are used: Unit, Area, New Construction, Group Quarters, and Coverage Improvement (see Weller et al., 1991; Hendrick, 1994). Similar frames from earlier Censuses were used in the earlier panels (see Section 3.2). The clusters formed in each frame are smaller than for earlier panels. This design change reduces the design effect and hence improves the precision of survey estimates. Within the unit frame, a single housing unit is selected. Within the area frame, the cluster consists of four "expected" housing units. In the group quarters frame, the cluster contains a single housing unit equivalent. In the new construction frame, half the clusters have four expected housing units and half have eight expected units.

Oversampling in the unit and area frames

Low-income households are oversampled in the unit and area frames, which together provided about 90 percent of the total SIPP sample. The purpose of oversampling is to increase the sample size of low-income households and reduce the variance for poverty estimates (see Weller et al., 1991). The method employed to oversample such households is based on a stratification approach described by Waksberg (1973). This method assigns units within sampled PSUs to either a high or low poverty stratum according to income or household characteristics associated with poverty.

In the unit frame, individual housing units are assigned to strata based either on household income or composition, as reported in the 1990 Census. Since income was collected on the Census long form, it is available for only a sample of housing units. For this sample of housing units, which comprises about one-sixth of the total, a housing unit is assigned to the high poverty stratum if the 1990 household income was below 150 percent of the poverty threshold; other housing units in the long form sample are assigned to the low poverty stratum. For other households, the assignment is based on household characteristics associated with poverty. If the 1990 household had any one of the following characteristics, the housing unit is assigned to the high poverty stratum:

- Female householder with children under 18 and no spouse present,
- Living in a central city of a Metropolitan Statistical Area (MSA) and renter with rent less than $300,
- Black householder and living in a central city of an MSA,
- Hispanic householder and living in a central city of an MSA,
- Black householder and householder less than age 18 or greater than age 64, and
- Hispanic householder and householder less than age 18 or greater than age 64.

In the area frame, stratification is based on block level information about the proportion of people in poverty. Blocks are ranked on the Census estimate of poverty or the household composition as listed above. The

assignment of blocks to the high or low poverty strata is done as an iterative process to optimize the gains of stratification.

The Bureau explored the possibility of developing separate sampling rates for the two strata at the PSU level but, based on its research, decided instead to establish a single overall rate for the high poverty stratum as well as for the low poverty stratum. Effectively, the high poverty stratum is oversampled at a rate of about 1.7 to 1.

To illustrate the effects of oversampling, Table 11.1 shows the numbers of households and persons attained from a self-weighting sample and a sample with oversampling, given a fixed overall sample size. The effective sample size under the oversample design is smaller, due to the design effect associated with the differentiated weighting resulting from the oversampling. The gain due to oversampling reflects the increase in the effective sample sizes of low-income related subgroups. At the household level, there is a 3 percent increase in the effective sample size for households in poverty below 150 percent of poverty level, a 17 percent increase for black households in poverty, and a 12 percent increase for Hispanic households in poverty. At the person level, the corresponding percentages are a 4 percent increase in persons in poverty, a 16 percent increase in black persons in poverty, and a 10 percent increase in Hispanic persons in poverty. The losses are in the high-income households. For households with income above $75,000, the effective sample size is reduced by 11 percent. The effective sample size for persons 55 and over is also reduced, by 7 percent.

Coverage

The concerns about the effects of survey undercoverage on the results derived from SIPP are discussed in Section 3.5. Table 11.2 shows the coverage ratios for the sample in the 1996 Panel for blacks and nonblacks by age. The coverage ratios are obtained by dividing the survey estimated population total (before second-stage weighting adjustment) by an independent population control total. The independently derived population totals for the 1996 Panel, as for the 1992 and 1993 Panels, are adjusted for the estimated undercount in the Census. For black males ages 15 and older, the coverage ratio is about 81 percent. For black females at the same ages, the coverage ratio is about 90 percent. The corresponding coverage ratios for nonblack males and nonblack females are 89 percent and 93 percent, respectively. These ratios are comparable to those from earlier panels (see Tables 3.4 to 3.6).

Table 11.1 Sample sizes under a self-weighting and an oversampling design

Characteristics of households and persons	Actual sample size		Effective sample size in oversampling design	Percentage increase (+) or decrease (-) in effective sample size
	Self-weighting design	Oversampling design		
Total Households (HHs)	36,730	36,730	34,980	
HHs <150 percent of poverty	9,182	9,929	9,456	3.0
Black HHs <150 percent	1,769	2,164	2,060	16.5
Female HHs <150 percent	5,522	6,055	5,766	4.4
Hispanic HHs <150 percent	1,233	1,447	1,378	11.8
HHs <100 percent of poverty	5,444	5,985	5,700	4.7
Black HHs <100 percent	1,190	1,471	1,400	17.7
Female HHs <100 percent	3,384	3,786	3,605	6.6
Hispanic HHs <100 percent	808	967	920	14.0
AFDC	1,289	1,511	1,439	11.6
Food Stamps	3,136	3,609	3,437	9.6
Social Security	10,526	10,488	9,988	-5.1
Supplemental Security	1,735	1,967	1,873	8.0
Black HHs	4,172	4,907	4,673	12.0
Female HHs	16,490	16,874	16,070	-2.5
Hispanic HHs	2,790	3,114	2,965	6.3
Female, no spouse, with relatives	276	293	279	1.1
Income $35,000 to 49,999	5,967	5,828	5,550	-7.0
Income $50,000 to 74,999	5,959	5,699	5,427	-8.9
Income $75,000 or more	4,886	4,590	4,371	-10.5
Total persons	95,402	95,402		
Persons <150 percent of poverty	23,466	25,584	24,365	3.8
Blacks <150 percent of poverty	4,970	6,066	5,777	16.2
Hispanics <150 percent of poverty	4,596	5,304	5,051	9.9
Persons <100 percent of poverty	14,118	15,660	14,914	5.6
Blacks <100 percent of poverty	3,367	4,162	3,963	17.7
Hispanics <100 percent of poverty	3,010	3,527	3,359	11.6
Persons 55 years and older	19,822	19,340	18,419	-7.1
Persons 65 years and older	12,187	11,881	11,315	-7.2
Blacks 16 years and older	7,951	9,283	8,840	11.2
Hispanics 16 years and older	6,606	7,312	6,963	5.4
Blacks, all ages	11,527	13,504	12,860	11.6
Hispanics, all ages	9,788	10,845	10,328	5.5
Nonblacks/Non-Hispanics, all ages	74,620	71,721	68,305	-8.5

Table 11.2 SIPP coverage ratios for January 1996

Age	Nonblack		Black	
	Male	Female	Male	Female
15	0.9175	1.1262	0.7427	0.7809
16-17	0.8640	0.9313	0.8322	0.9487
18-19	0.8582	0.8644	0.8693	0.8365
20-21	0.8849	0.8005	0.5960	0.9404
22-24	0.7859	0.8693	0.6024	0.6867
25-29	0.7984	0.8253	0.8512	0.8391
30-34	0.8792	0.9060	0.8839	0.8020
35-39	0.9175	0.9839	0.7076	0.9061
40-44	0.9063	0.9314	0.8126	0.9514
45-49	0.8991	0.9782	0.6948	0.7960
50-54	0.9667	0.9163	0.8945	1.2070
55-59	0.9468	0.9046	1.0196	0.9865
60-61	0.8405	0.9000	1.0210	0.9877
62-64	0.9869	1.0732	0.9914	0.9618
65-69	0.9276	0.9452	1.0787	0.7797
70-74	0.8870	0.9361	0.7870	1.3096
75-79	0.8852	1.0046	---------	0.9218
80-84	0.8974	0.9569	---------	---------
85+	0.9558	0.9780	---------	---------
15+	0.8906	0.9294	0.8080	0.8961

11.3 Data collection and processing

Computer-assisted interviewing (CAI) was introduced with the 1996 redesign. This change resulted from the computer-assisted survey information collection (CASIC) initiative at the Census Bureau. Computer automation offers convenient management, better data quality, and more rapid turnaround of data. The CAI experience of SIPP is summarized in U.S. Bureau of the Census (1998a) which reviews the experience of 12 demographic surveys conducted by the Bureau. This section describes the use of personal (CAPI) and telephone (CATI) interviews, the CAI instrument, changes in data collection and processing, the preparatory work undertaken for the change to CAI, and evaluations of the effects of CAI on SIPP estimates.

CAPI and CATI

With the 1996 redesign, data collection is conducted through computer-assisted personal interviews (CAPI) at Waves 1 and 2, and combinations of personal visits and telephone interviewing (CATI) for Waves 3 and beyond. The combinations of CAPI and CATI interviews to be used across waves are prespecified and vary for different subgroups of the sample according to the following scheme (Waite, 1996). Sample members are assigned (at the segment level) to one of three interviewing mode subgroups. Each subgroup is designated a pattern of interviewing modes to be repeated every three waves. Thus, for Waves 3, 4, and 5,

subgroup 1 is assigned the sequence CAPI-CATI-CATI, subgroup 2 the sequence CATI-CAPI-CATI and subgroup 3 the sequence CATI-CATI-CAPI. This scheme is applied with each rotation group. Under this scheme, one-third of the sample is interviewed by CAPI each wave and each month, and every household is interviewed in person once a year. The same sequence is repeated for Waves 6 and beyond, with a cycle of three waves.

CAI instrument

The development of the CAI instrument for SIPP involved only minor changes in content from the paper documents. Most of the changes occurred in the question wording, the structuring of the sections, and the skip patterns between questions. The CAI instrument includes items that previously were on the control card and now appear before the core interview. The core interview items are reorganized in three main parts: employment and earnings, other income, and additional questions (such as health insurance coverage). Chapter 3 of the *SIPP User's Guide* (U.S. Bureau of the Census, 1998) describes the new CAI instrument and how it differs from the paper-and-pencil version. Lamas et al. (1996) review the changes in the employment and earnings section.

The development of a CAI instrument is time-consuming, and once completed, changes to the instrument are neither straightforward nor simple. This may affect the survey's ability to handle last minute changes in survey content when the need arises.

Data collection and processing

The CAI environment also affects the process of data collection and post-data-collection processing. With CAI, some consistency checks are built into the instrument, allowing the interviewers to make corrections before concluding an interview. For Wave 2 and beyond, the responses from the preceding waves are available in CAI, and interviewers can remind the respondents about their previous responses. This feature is aimed at reducing the transition problems observed in earlier panels (see Section 6.1 for a discussion of the "seam" phenomenon).

The use of the CAI environment for case control and management is complex for SIPP because the survey follows persons rather than households over time. SIPP has not found any reduction in post-data-processing time.

Preparation and evaluation

In preparation for CAI, SIPP began the development of the data collection instrument in 1991, conducted a pretest in 1994, and completed a dress rehearsal in 1995. The pretest was conducted on about 300 households in three regional locations. Households were interviewed with the Wave 1 instrument, and cognitive interviewing techniques were employed to study responses to key questionnaire items on labor force participation, income from assets, and health insurance. Other aspects examined in the pretest included training, interviewing procedures, data collection and data transmission. The 1995 dress rehearsal tested all operational aspects included in the pretest as well as weighting and data tabulations. Interviews were conducted with a sample of over 7,000 households (Huggins and Fischer, 1994). The data collection for the dress rehearsal involved 130 supervisory field representatives and 360 field representatives (Hock and Winters, 1996).

Since the data collection for the dress rehearsal overlapped with Waves 7 and 8 of the 1993 Panel, it was possible to compare estimates based on data collected under the CAPI method with estimates based on data collected by the old paper-and-pencil method (PAPI). Hock and Winters (1996) compared estimates of the

number of program recipients, health insurance participants, and the number of workers from these two sources. Table 11.3 shows the estimates, the standard errors of the estimates, and *t*-score statistics for the significance of the difference between the estimates. Most estimates from the two data collection methods are not significantly different. For small subgroups, some significant differences are found. For example, Table 11.4 shows estimates of the numbers of full-time workers by sex and age. The CAPI estimate shows a much higher number of full-time workers over 75 years of age.

Table 11.3 Estimates of the number of persons (in thousands) from PAPI and CAPI[1]

	PAPI		CAPI		
	Estimate	S.E.	Estimate	S.E.	*t*-score[2]
Social Security	19,879	5,344	19,740	3,689	0.02
SSI	3,408	1,091	26,25	603	0.63
AFDC	2,803	1,115	2,407	706	0.30
WIC	2,918	1,260	2,125	531	0.58
Food Stamps	5,200	2,022	5,510	1,582	0.12
Medicare	18,553	4,943	18,082	3,428	0.08
Medicaid	17,287	6,278	16,019	4,571	0.16
Any kind of health insurance	123,622	36,385	120,807	27,557	0.06
Total workers	70,136	21,364	70,905	17,655	0.03
Full-time workers	55,344	17,315	46,603	11,646	0.42

[1]The PAPI estimates were based on Waves 7 and 8 of the 1993 Panel. The CAPI estimates were based on data from the dress rehearsal.
[2]*t*-score was calculated as (CAPI estimate – PAPI estimate)/$\sqrt{}$(CAPI S.E.^{2}+ PAPI S.E.^{2}).
Source: Hock and Winters (1996).

Table 11.4 Estimates of the number of full-time workers (in thousands) by sex and age from PAPI and CAPI[1]

		PAPI		CAPI		
	Age	Estimate	S.E.	Estimate	S.E.	*t*-score[2]
Males	15-24	3,552	1,324	2,859	1,097	0.40
	25-34	10,541	4,307	8,544	3,061	0.38
	35-44	9,491	2,591	8,011	1,641	0.48
	45-54	5,833	1,315	4,824	531	0.71
	55-64	3,067	7,51	2,831	579	0.25
	65-74	452	127	386	72	0.45
	75+	23	5	41	2	3.10
Females	15-24	2,694	1,099	2,110	747	0.41
	25-34	7,008	2,388	5,665	1,962	0.40
	35-44	6,297	1,802	5,329	1,099	0.42
	45-54	4,216	1,054	3,757	438	0.37
	55-64	1,965	512	1,883	364	0.12
	65-74	195	46	318	76	1.27
	75+	3	0.3	38	3	8.77

[1]The PAPI estimates were based on Waves 7 and 8 of the 1993 Panel. The CAPI estimates were based on data from the dress rehearsal.
[2]*t*-score was calculated as (CAPI estimate – PAPI estimate)/$\sqrt{}$(CAPI S.E.^{2}+ PAPI S.E.^{2}).
Source: Hock and Winters (1996).

Lamas et al. (1996) conducted similar comparisons of data from PAPI and CAPI for estimates of employment and earnings. Table 11.5 shows estimates of monthly earnings for persons ages 15 to 64 using data collected from PAPI and CAPI methods. The CAPI estimates of both aggregate and mean earnings are higher than the PAPI estimates in all categories. However, this study did not examine the statistical significance of the differences. Since comparative studies with other surveys and administrative data have typically found that SIPP estimates underestimate earnings (see Chapter 10), the higher estimates from CAPI was regarded as improvements in the new design.

Table 11.5 Estimates of mean monthly earnings for persons ages 15 to 64 from PAPI and CAPI[1]

Age and sex	Mean monthly earnings ($)	
	PAPI	CAPI
15-64 years old	2,378	2,641
15-64 years old, male	2,836	3,105
15-64 years old, female	1,828	2,091
15-17 years old	404	516
18-24 years old	1,176	1,279
25-34 years old	2,373	2,617
35-44 years old	2,673	3,027
45-54 years old	3,062	3,480
55-64 years old	2,627	2,724

[1]The PAPI estimates were based on Waves 7 and 8 of the 1993 Panel. The CAPI estimates were based on data from the dress rehearsal.
Source: Lamas et al. (1996).

11.4 Nonresponse

Nonresponse due to attrition is an increased concern with the 1996 Panel because of its longer duration. This section describes the rate of household nonresponse for Waves 1 through 6 of the 1996 Panel, and the Bureau's efforts to reduce nonresponse. Chapter 3 describes the nonresponse experienced in earlier panels. Mack and Petroni (1994) provide an overview of SIPP nonresponse research findings. Waite et al. (1997) discuss assessments of efforts to reduce nonresponse bias.

Rates and types of household nonresponse

Table 11.6 shows the rates of household noninterview and sample loss for Waves 1 to 6 of the 1996 Panel. The corresponding nonresponse rates for Panels 1984-1993 are shown in Table 5.1. The sample loss rate of 8.4 percent at Wave 1 of the 1996 Panel is comparable to the corresponding rates of around 9 percent for the 1992 and 1993 Panels. However, by Wave 6 the sample loss rate of the 1996 Panel rose to 27 percent. This level is appreciably higher than the corresponding rate of about 22 percent for Wave 6 of the 1992 and 1993 Panels. McMahon (1998) estimated that by Wave 5 of the 1996 Panel, about 17 percent of the sample loss is permanent.[1]

[1] By definition, a Type A (household refusal) is no longer eligible for interviewing if it is a Type A for two consecutive waves, and a Type D (unlocated movers) is no longer eligible if it is a Type D for three consecutive waves.

Table 11.6 Household noninterview and cumulative sample loss rates, Waves 1 to 6 of the 1996 Panel

Type of noninterview	Percent					
	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6
Type A	8.4	13.1	15.6	17.6	20.4	22.2
Type D	--	1.3	1.9	3.1	3.8	4.4
Sample loss[1]	8.4	14.5	17.8	20.9	24.6	27.4

[1]The sample loss rate consists of cumulative noninterview rates adjusted for unobserved growth in Type A noninterview units (created by splits).

By types of noninterview, the 1996 Panel shows lower rates of Type D noninterview but higher rates of Type A noninterview than earlier panels. Type A nonresponse consists of households occupied by persons eligible for interview, but who failed to complete an interview for reasons including: no one at home, refusal, and unable to locate a sample unit. Type D nonresponse consists of households of original sample persons who are living at unknown new addresses. The improvement in Type D noninterviews may be attributed to both improved and new procedures implemented to follow movers. To manage Type A noninterviews, the Census Bureau has experimented with the use of incentives to encourage cooperation. These procedures are discussed below.

Procedures to reduce Type D nonresponse

Several new procedures are employed in the 1996 redesign in an attempt to reduce the level of Type D nonresponse. They include the following:

- *Centralized locating activities.* One person in each regional office is now designated as the "locator." This person assists the local field representatives (FR) in tracking movers for SIPP as well as other Census Bureau surveys. When a FR determines that a case is a Type D, the locator is assigned to find the mover. The locator works with the FR to conduct both local and regional searches. Cases located in new sites are forwarded to another FR.

- *Extended tracking to 9 months.* The time allowed for tracking movers has been extended to 9 months for the 1996 Panel. The extension allows persons who have missed two consecutive waves of interviewing to continue with the survey. The old design allowed 5 months for tracking and tolerated only one missing wave per person.

- *Included Type D conversion in FR rating.* The FR's performance evaluation is revised to include a component on Type D conversion. This change is expected to provide more incentive for FRs to locate movers.

- *Automated transfer of Type D noninterviews between regional offices.* The CAI environment allows automatic transfer of Type D cases to another regional office, usually within the interviewing period for a given wave.

Incentives to reduce Type A nonresponse

The Census Bureau has experimented with using incentives to encourage cooperation with the 1996 Panel. Based on positive results from a Wave 1 experiment, the Bureau is conducting two more incentive experiments, at Wave 7, and at Waves 8 and 9. These incentive experiments are described in James (1997), Sundukchi (1998), and Winters (1998), and they are summarized below.

The incentive experiment at Wave 1 tested the effectiveness of a cash incentive of $20 or $10 against no incentive. About 20,000 sampled households received incentives; approximately one-half received a $20 incentive, and the other half received a $10 incentive. The remaining sample households were designated as the "no incentive" control group. For households that received an incentive, the incentive was introduced as a "token of appreciation." It was presented early in the interview, in the form of a voucher with a cash amount printed on it. The respondent was instructed to fill in a name, check the address, and return the voucher to the Census Bureau in a self-addressed and stamped envelope. Respondents were told that payments would be mailed in about 2 to 3 weeks, and that the payment would be made irrespective of whether they continued with the interview.

Table 11.7 shows the sample loss rates for the three incentive groups by poverty stratum for Waves 1 to 6. Households that received the $20 incentive had a lower sample loss rate than the other treatment groups in all waves in both strata. The improvement is especially noticeable among households in the high poverty stratum. It is encouraging to note that the benefit of an incentive given at Wave 1 appears to have persisted over waves. James (1997) and Mack et al. (1998) provide further details of the Wave 1 incentive study.

Table 11.7 Rate of sample loss by incentive group, Waves 1-6 of the 1996 Panel

Incentive at Wave 1	Poverty stratum	Cumulative sample loss (percent)					
		Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6
$20	High	5.9	11.4	14.4	16.9	21.1	23.0
$10		8.1	13.8	17.7	20.7	24.3	27.1
None		9.3	16.1	19.2	22.4	25.5	29.0
$20	Low	8.2	13.1	16.1	19.3	22.8	25.2
$10		9.5	14.4	18.2	21.3	24.2	26.7
None		9.1	14.9	18.1	21.2	24.5	27.3

Source: Mack et al. (1998).

At Wave 7 over 5,000 households received a second incentive of $20. This incentive was given to those households that had already received an incentive (either $20 or $10) at Wave 1 and the household income was less than or equal to 150 percent of the poverty threshold during Wave 1. A small number of nonpoverty households also received the incentive in order to provide the same treatment condition to neighboring households within a cluster. The noninterview rates of the Wave 7 incentive groups will be monitored until the end of the 1996 Panel.

The Waves 8 and 9 experiment concerns Type A refusal households. Some receive incentives of $40 or $20, and others receive nothing. The Type A refusal households from the previous wave, that is, Waves 7 and 8, receive an advance letter and an incentive voucher with the designated amount (if applicable) prior to the interview. Field representatives are also authorized to give incentives at the door if the letter and voucher failed to reach a respondent in one of the incentive groups. This experiment will also provide information on the effectiveness of a larger incentive amount.

Nonresponse study

A study of nonrespondents at Wave 1 of the 1996 Panel was conducted to provide information about their characteristics and to improve the Wave 1 nonresponse adjustments (Dajani and Winters, 1997). This study consisted of two components: a questionnaire mailed to the nonrespondents, and a field representative (FRs) questionnaire completed by the FRs based on their observations at the visit. The questions on the nonrespondent questionnaire included employment, income, program participation, savings, and health

insurance. The questions on the FRs questionnaire included race of the reference person, type of housing unit (i.e., public housing), number of adults in the household, and broad categories of income. The nonrespondent questionnaire was sent to all Type A nonrespondents; the response rate was about 22 percent (716 out of the 3,194 questionnaire sent). The results provided insights on the characteristics of the Wave 1 nonrespondents, and they have been used to refine the Wave 1 nonresponse weighting adjustment cells (see Section 11.5).

11.5 Imputation and weighting

The imputation and weighting procedures for earlier SIPP panels are discussed in Chapter 8. The changes adopted for the 1996 Panel are the result of the extensive research discussed in Section 8.4. Greater detail on these procedures is available in the *SIPP User's Guide* (U.S. Bureau of the Census, 1998) in chapters on Edits and Imputation, Use of Weights, and the Weighting Appendix. This section briefly outlines the new procedures.

Imputation

Longitudinal imputation was used in the 1991, 1992, and 1993 Panels to impute for one missing wave bounded by responding waves. Section 8.3.2 describes this imputation method which resulted in the retention of an additional 5 to 8 percent of the sample persons in the panel files. The 1996 Panel has expanded the use of longitudinal imputation to include persons with two consecutive missing waves bounded by responding waves in order to retain more sample persons in the longitudinal sample. By extending longitudinal imputation to two consecutive missing waves, the Census Bureau expects to retain an additional 1 percent of sample persons in the panel file.

Weighting

The main changes made in the weighting procedures for the 1996 Panel refine the adjustments of the weights to reduce variance and to compensate somewhat more effectively for nonresponse bias. The basic steps in constructing weights are the same as for the earlier panels (see Section 8.2 for the cross-sectional weights in the core wave files, and Section 8.3 for the longitudinal weights in the full panel file).

The new procedures for developing the cross-sectional weights are summarized in the Source and Accuracy Statement for the 1996 Panel Wave 1 Public Use File (Cahoon, 1998). They include the following:

- The Wave 1 cross-sectional weights do not include the first-stage ratio estimation factor used in earlier panels. This factor, as explained in Section 8.2.1, was applied to sample households in NSR PSUs to reduce the between PSU component of sampling error. It was dropped from the 1996 Panel because it had not been effective in reducing variance.

- The nonresponse adjustment classes were refined to include additional information that resulted in reducing nonresponse bias. At Wave 1, the nonresponse adjustment classes include the stratum classification for high and low poverty stratum within PSU. For Waves 2 and beyond, based on the research described in Section 8.4.4, more variables have been added, including household income, geographic division (Census subregion 1-9), and number of imputations for selected income and asset items (0, 1, 2 and more). The nonresponse adjustment classes were obtained using a tree search algorithm.

The Census Bureau plans to produce calendar year files from the 1996 Panel, at the end of each year of data collection. Cahoon (1998) provides the specifications for constructing the longitudinal weights for persons for the 1996 calendar year file based on data from Waves 1 through 3.

Hendrick (1995) describes the new nonresponse adjustment classes for the 1996 calendar year weight, and examines their effectiveness in reducing the nonresponse biases for poverty-related items. Table 11.8 shows estimates of nonresponse bias using the old and new adjustment classes. The results show significant reductions in the nonresponse biases for estimates of persons in poverty, persons receiving AFDC, and persons receiving Food Stamps.

Table 11.8 Difference in nonresponse bias between original and redesigned nonresponse adjustment classes

Poverty-related item	Estimate of nonresponse bias		*t*-test statistic
	Original adjustment classes	Redesigned adjustment classes	
Ratio of average monthly household income/average monthly poverty threshold	1.49	1.48	0.59
Average monthly household income	106.06	104.93	0.04
Average monthly person income	45.17	45.01	0.10
Percent persons in poverty	-1.85	-1.63	-5.66
Percent persons receiving AFDC	-0.11	-0.06	-2.27
Percent persons receiving Food Stamps	-0.16	-0.04	-3.52

Source: Hendrick (1995).

12. SUMMARY

This report has tried to give an objective and balanced presentation of the quality of SIPP data and the effects of sampling and nonsampling errors. We do not presume to assign any overall quality rating to SIPP data. Data quality is too complex to summarize with a simple rating or numerical score. We prefer that SIPP data users make their own judgements based on the best information available.

There have been important changes in design, data collection, and processing procedures since the start of the survey. Some of these changes attempted to deal with sources of error that became evident in the course of the survey. Other changes, such as the reductions in sample size in the mid-1980s, were necessary because of reductions in funding. These changes affected the quality of data from the different SIPP panels. Sampling errors for panels with reduced sample sizes were somewhat larger than those for other panels. More recent panels are likely to have smaller nonsampling errors than the early panels as a result of continual procedural improvements.

This chapter presents some suggestions to users on what they can do to use SIPP data effectively and to minimize the effects of error on their analyses and interpretations of the survey results. We identify what we consider to be the main sources of error that affect the quality of SIPP estimates based on the 1984 to 1993 Panels. We also mention research projects now underway, or about to begin, which will examine the quality of SIPP estimates based on the 1996 design.

12.1 Suggestions for users

This section contains some general guidelines for making effective use of SIPP data products. We begin with some suggestions that apply to all users, whether they are working with SIPP publications or microdata files.

- **Become familiar with the broad features of the survey design.** Key features (see Chapters 2 and 11) include: the definition of the survey population; the sample structure, which prior to the redesign consisted of overlapping panels; the rotation groups within panels; the scheduling of interview waves for each panel and the reference periods used for each wave; and the rules for following sample persons during the life of a panel. The *SIPP User's Guide* (U.S. Bureau of the Census, 1998) is a good source for further information. Be alert to changes, especially between panels, in content, procedures, sample sizes, and scheduling of interviews.

- **Be alert to the effects of sampling and nonsampling errors on the estimates.** The effects of sampling errors are discussed in Chapter 9. Analytical statements in the texts of SIPP publications have been tested for statistical significance. Users of published data can perform their own tests, using the generalized variance functions and instructions provided in each publication. The effects of nonsampling error are less easy to quantify and their effects depend on how the data are used. Key sources of nonsampling error—nonresponse, differential undercoverage, and measurement error (especially the seam phenomenon)—are identified in Section 12.2 and were discussed at length in earlier chapters.

Use of SIPP microdata files opens up many possibilities not available to those who work only with published data. There are three categories of microdata files: internal-use files, public use files, and special research files. The greatest amount of detail is found on the Census Bureau's internal files, which are available for use only by Census Bureau employees. For the other two kinds of files, some of the detail is omitted in order to minimize the possibility that any sample person could be identified, which would violate the Census Bureau's legal obligations, and its assurances of confidentiality to persons who participate in the survey. Specifically, some income and asset amount items are top-coded and some low-level geographic codes are

omitted or scrambled. The distinction between special research files and public use files is that the former include data that are considered to be preliminary. The data have been developed using imputation and estimation methods that have not yet been fully evaluated. Special research files have been made available for research purposes to users who agree in writing to observe certain requirements relating to redistribution of the files and publication of findings based on them (McMillen, 1989b). No restrictions are placed on the use of public use files.

The *SIPP User's Guide* (U.S. Bureau of the Census, 1998) provides detailed information on file content and structure of microdata files. We present here only a few broad suggestions concerning the use of data files:

- **Obtain and review all pertinent documentation for the files with which you are working.** In addition to the *SIPP User's Guide***,** purchasers of files receive documentation specific to the files purchased, including copies of the questionnaires used to collect the data on the files. A copy of the *SIPP lnterviewers' Manual* and lists of research reports and SIPP working papers may also be obtained from the Census Bureau's Data User Services Division. Errors are sometimes discovered in the data files themselves or in the documentation. The Bureau sends file-specific user notes to all purchasers to inform them of problems. Another source of information is the SIPP Supplement that is periodically included with the Newsletter of the Association of Public Data Users (APDU), 87 Prospect Avenue, Princeton, NJ 08544, (609) 452-6025. Since the end of 1988, APDU has had a SIPP Committee, comprised of experienced users of SIPP microdata files, that provides advice to the Census Bureau on the design, content, and documentation of SIPP data products. Additional relevant information is available at the SIPP web site <http://www.sipp.census.gov/sipp>.

- **Avoid the use of unweighted data.** Use of unweighted data for descriptive or analytical purposes can lead to incorrect conclusions. This is especially true for the 1985 and later SIPP panels, for which the base weights vary by primary sampling unit. In the 1990 Panel, further variations in weights were introduced by the carryover of extra households from the 1989 Panel. In the 1996 Panel, low-income households are overrepresented in the sample. The weights provided in the microdata files reflect the sample selection procedures actually used and are also designed to minimize biases resulting from nonresponse and other causes. A review of the variability of the sampling weights in some panels (Table 8.1) indicates that distributions and other statistics based on correctly weighted data may differ substantially from those based on unweighted data. For example, if persons ages 65 or older have weights near the maximum, the use of weights would provide an estimate of the proportion of persons ages 65 or older which is larger than would be obtained without weights.

- **If possible, calculate sampling errors directly from the data for the key estimates in your analysis.** The generalized variance functions provided with the file documentation can be used for rough estimates of sampling errors for specific items. However, more accurate estimates of standard errors can be obtained for the estimates of interest by applying variance estimation methods that are appropriate for the complex SIPP sample design. Procedures for applying these methods are described in the Sampling Error chapter of the *SIPP User's Guide* (U.S. Bureau of the Census, 1998). It almost goes without saying that either of these approaches to the estimation of sampling errors is preferable to the calculation of sampling errors as though the estimates were based on a simple random sample. In general, that assumption would lead to substantial underestimates of sampling error for some kinds of items.

Finally, all readers, particularly those who are users of SIPP microdata, are urged to let the Census Bureau know what they have learned about the quality of SIPP data and to pass on their ideas for improving the quality of SIPP data products. The Bureau depends on feedback from users to set its research priorities and guide its efforts to improve the quality of the data. A Census Bureau contact list is provided in Section 1.4.

12.2 Major sources of error

The main sources of error in SIPP are similar to those found in many other household surveys and are relatively easy to identify. Quantifying their contributions to the total error of SIPP estimates is much more difficult. Except for the results from the SIPP Record Check Study (see Section 6.3.4), there is very little direct information on the size of error components arising from different sources. Furthermore, the effects of errors vary, depending on the ways in which the data are used. Errors that have little effect on cross-sectional analyses may create significant problems for longitudinal analyses.

Nevertheless, even indirect knowledge about important sources of error is useful as a guide to the allocation of available resources to the different phases of survey operations and to evaluation and quality assurance activities, methodological research and experimentation.

Nonresponse

As in most panel surveys, nonresponse is a major concern in SIPP. As explained in Chapter 5, nonresponse occurs at several different levels: household nonresponse at the first wave and thereafter; person nonresponse in interviewed households; and item nonresponse, including complete nonresponse to topical modules. At the household level, sample loss for the 1991 Panel rose from about 8.4 percent at Wave 1 to 21.4 percent by Wave 8 (Table 5.1). At the personal level for the same panel, 23.0 percent of the original sample persons who participated in Wave 1 missed one or more interviews for which they were eligible in later waves (Table 5.3), and in addition there were nonrespondents at Wave 1. Nonresponse is higher for persons who move than for those who do not (Table 5.4). Nonresponse rates increased in the 1992 and 1993 Panels over earlier panels, and they have increased further in the 1996 Panel.

Item nonresponse has been minimal for core items on labor force activity, income recipiency, and asset ownership. It has been somewhat higher for income amounts, especially self-employment earnings and interest (Table 5.6). For interviewed persons, nonresponse to topical modules ranged from 4.0 to 6.5 percent in the 1991 Panel (Table 5.9). In the topical modules, especially high item nonresponse rates occurs for questions on asset amounts (Table 5.8).

From the sampling error point of view, the cumulative effect of these different types of nonresponse is to reduce effective sample sizes and, therefore, to increase sampling error, especially for longitudinal estimates. As explained in Chapter 8, a combination of imputation and weighting procedures is used to minimize the biases resulting from differences between the characteristics of respondents and nonrespondents. The development of these procedures has been informed by several research studies, with emphasis on the treatment of wave nonresponse (see Section 8.4).

Differential undercoverage

A second major area of concern is differential undercoverage of demographic subgroups, as indicated by the variation in coverage ratios shown in Tables 3.4 through 3.6 and Table 11.2. The group most affected is young adult black males. Moreover, the coverage ratios shown prior to 1992 understate the size of the undercoverage for young adult black males. This is because the ratios until that time were based on projections from Census counts and did not include any adjustment for Census undercoverage, which is known to be above average for this population subgroup.

The use of the second-stage ratio adjustments in both cross-sectional and longitudinal estimates to compensate for undercoverage is believed to reduce both sampling error and bias of the estimates. The effect of the weighting on sampling error is demonstrable, but little is known about its effect on biases associated with undercoverage.

Measurement error

The third major area of concern is measurement error, with particular emphasis on the seam phenomenon, that is, a pronounced tendency of survey respondents to report month-to-month changes for months in adjacent waves at substantially higher rates than for adjacent months in a single wave. The seam effect has been observed in aggregate data (see Tables 6.1 to 6.3) and the SIPP Record Check Study has provided more direct evidence of its existence (see Section 6.3.4). Because of the rotation group design used in SIPP, cross-sectional estimates are not likely to be seriously affected by this reporting pattern, but it can affect estimates of the covariance structure and may have significant adverse effects on multivariate analyses dealing with transitions or length of spells (Singh et al., 1989; Kalton et al., 1992).

As reported in Chapter 6, many evaluation studies and experiments have been carried out in connection with SIPP and its predecessor, the Income Survey Development Program, to gain a better understanding of how key design factors affect the levels of nonsampling error. Until now, no other sources of error stand out so clearly as the three just described.

12.3 Current research

The cumulative results of research and investigations on the quality of SIPP data have been incorporated in developing the 1996 redesign. As the Census Bureau introduced the new design in 1996, plans were also made to continue the process of evaluation and improvement. This section summarizes suggestions and plans for evaluation of the SIPP estimates based on the 1996 redesign. Most of the material in this section is based on Huggins and Fischer (1994).

Reducing nonresponse

Starting with the major sources of error identified in the previous section, two types of research activities are focused on nonresponse. One is concerned with improving cooperation. Since the Wave 1 incentive study for the 1996 Panel found that giving a cash incentive to households at their first interview reduced household nonresponse, two further experiments with cash incentives have been conducted (see Section 11.4). One, conducted at Wave 7, involves providing an incentive to low-income households. The other, conducted in Waves 8 and 9, involves offering an incentive to nonresponding households. In both cases the effects of these incentives will be monitored through the end of the panel.

The second type of nonresponse research aims to compensate for the biases resulting from nonresponse. The weighting and imputation research (Section 8.4) provides a great deal of information about the extent and effect of attrition bias. The 1996 Panel research focuses on estimating attrition bias as well as improving weighting, imputation, and other analytical methods to estimate or adjust for attrition bias. Particular emphasis is placed on weighting for longitudinal analysis. Continuing research in this area is considered critical since the goal of SIPP is now much more focused on longitudinal statistics.

Survey coverage

The problem of differential undercoverage of population subgroups is similar for all of the Census Bureau's household surveys. Several studies have provided information about the causes and magnitude of survey undercoverage for major Census-based household surveys (Fay, 1989a; Shapiro, 1992; Shapiro et al., 1993). The roster research and the Living Situation Survey discussed in Section 3.5 have provided useful insights into the nature of within-household enumeration of individuals. The results of that research and of ongoing research of Decennial Census coverage will be reviewed for possible indications of how survey coverage can be improved in SIPP.

Seam phenomenon

Many projects studying the seam phenomenon have been completed (Section 6.3). The SIPP Record Check Study (Section 6.3.4) has provided some direct evidence on individual errors in reporting when transitions in recipiency or amounts occur, as well as several other kinds of information that are helpful in understanding the nature of response variance and bias for SIPP data on participation and benefit amounts for major public income transfer programs. However, the reasons for the occurrence and extent of the seam phenomenon are not yet well understood. With computer-assisted interviewing (CAI) in the 1996 Panel, respondents are reminded about their responses in the preceding wave. The impact of respondent feedback on the seam phenomenon has yet to be explored.

Variance reduction

Several research and development activities aim at variance reduction for SIPP estimates. The feasibility of reducing variances by using IRS data as controls in the second-stage ratio estimation procedure shows promise (Section 8.4.6). Similar use of administrative data from other sources will continue to be investigated. The broad goal for this area of research is to develop estimation procedures for SIPP that make effective use of auxiliary data available from administrative sources and also from the Current Population Survey.

Continued evaluations of the effectiveness of the design for oversampling low-income households in the 1996 SIPP Panel are planned. These evaluations will be important in reaching decisions as to whether the current oversampling methodology should be met in subsequent panels or whether alternative methodologies should be investigated. There will also be further investigation of methods for oversampling other population subgroups of special interest, such as aged persons in poverty.

Evaluation of estimates

Comparisons of SIPP estimates with data from external sources, including both program and administrative data and data from other surveys, will continue. Evaluation of SIPP estimates at the beginning of the 1996 Panel is critical for benchmarking and data quality assessment, as well as for identifying problems that need to be corrected. In addition to comparisons with administrative estimates, estimates from the 1996 redesign will be compared with estimates from the 1993 Panel to examine the effects of the redesign.

As in the past, the preliminary and final results of these research activities will be made public in a variety of ways. The SIPP working paper series and papers presented at professional meetings will be major modes of distribution. The bibliography of SIPP publications on the Internet serves as a major source of information about the availability of methodological research findings from SIPP. Persons with questions about the current status of specific projects may write to the SIPP Branch, Statistical Methods Division, Bureau of the Census, Washington, DC 20233 or call (301) 457-4192.

REFERENCES

Allen, T. M. (1991). Probing Suggestions. Internal Census Bureau memorandum to R. P. Singh, December 17th.

Allen, T. M. (1993). Analysis of Cost and Performance Changes Since the Introduction of Telephone Interviewing. Internal Census Bureau memorandum, February 26th.

Allen, T. M., and Petroni, R. J. (1994). Mover Nonresponse Adjustment Research for the Survey of Income and Program Participation. *Proceedings of the Section on Survey Research Methods, American Statistical Association,* 662-667.

Altman, M. (1989). SIPP Record Usage Rates for the 1984 and 1985 Panels. Annual Round-Up Topical Modules. Internal Census Bureau memorandum to C. G. Bowie, June 30th.

An, A., Breidt, F., and Fuller, W. (1994). Regression Weighting Methods for SIPP Data. *Proceedings of the Section on Survey Research Methods, American Statistical Association,* 434-439.

Bailey, L., Moore, T., and Bailar, B. (1978). An Interviewer Variance Study for Eight Impact Cities of the National Crime Survey Cities Sample. *Journal of the American Statistical Association,* 73, 16-23.

Bollinger, C. R., and David, M. H. (1993). *Modeling Food Stamp Participation in the Presence of Reporting Errors.* SIPP Working Paper No. 9307, U.S. Bureau of the Census.

Bollinger, C. R., and David, M. H. (1995). Sample Attrition and Response Error: Do Two Wrongs Make A Right? *Proceedings of the Bureau of the Census Annual Research Conference, U.S. Bureau of the Census,* 289-310.

Bollinger, C. R., and David, M. H. (1996). Modeling Discrete Choice Using Attrition as a Proxy for Response Error. Unpublished technical report, University of Wisconsin.

Bowie, C. G. (1984). Resolving Wave 1 Problems in the Assignment of GRIN Codes for SIPP. Internal Census Bureau memorandum to B. Fink, July 20th.

Bowie, C. G. (1986). SIPP Topical Module response rates. Internal Census Bureau memorandum, April 8th.

Bowie, C. G. (1988a). SIPP Topical Module response rates. Internal Census Bureau memorandum, March 17th.

Bowie, C. G. (1988b). Household Survey Profile for the Survey of Income and Program Participation. Prepared for the Task Force on Household Survey Redesign, U.S. Office of Management and Budget.

Bowie, C. G. (1990). SIPP Household Sample Loss. Internal Census Bureau memorandum, February 21st.

Bowie, C. G., and Kasprzyk, D. (1987). *A Review of the Use of Administrative Records in the Survey of Income and Program Participation.* SIPP Working Paper No. 8721, U.S. Bureau of the Census.

Brooks, C., and Bailar, B. (1978). *An Error Profile: Employment as Measured by the Current Population Survey.* Statistical Policy Working Paper 3, Office of Federal Statistical Policy and Standards, U.S. Department of Commerce.

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*, 351-356.

Butler, D. (1991). SIPP '87: Gift Experiment Results. Internal Census Bureau memorandum to R. P. Singh, April 2, 1991.

Butler, D. (1993). SIPP '88, '89, '90 and '91 Coverage Ratios. Internal Census Bureau memorandum, October 14, 1993.

Bye, B., and Gallicchio, S. (1988). A Note on Sampling Variance Estimates for Social Security Program Participants from the Survey of Income and Program Participation. *Social Security Bulletin*, 51 (10), 4-21.

Bye, B., and Gallicchio, S. (1989). *Two Notes on Sampling Variance Estimates from the 1984 SIPP Public-Use Files.* SIPP Working Paper No. 8902, U.S. Bureau of the Census.

Cahoon, L. S. (1998). *SIPP 96: Source and Accuracy Statement for 1996 Panel Wave 1 Public Use Files.* Internal Census Bureau memorandum to Walter Odom, January 8th.

Cantor, D. (1991). Results of SIPP Interviews Used for Preparation for Milwaukee Pretest. Unpublished Westat Memorandum to K. H. Marquis, K. Bogen, and J. Moore, September 5, 1991.

Cantor, D., Brandt, S., Green, J. (1991). Results of First Wave of SIPP Interviews. Unpublished Westat memorandum to C. G. Bowie, February 21, 1991.

Cantor, D., Brandt, S., Green, J., and Moesinger, K. (1992). Results from Wave 2 of SIPP. Unpublished Westat memorandum to D. Lampe, January 20, 1992.

Cantor, D., and Edwards, C. (1992). *Testing Alternative Household Roster Questions for the Survey of Income and Program Participation.* SIPP Working Paper No. 9203, U.S. Bureau of the Census.

Carmody, T., Meier, A., and Fischer, D. (1988). Survey of Income and Program Participation Experiments and Their Results. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 331-336.

Chakrabarty, R. (1987). Composite Estimation for SIPP Annual Estimates. Internal Census Bureau memorandum, September 23rd.

Chakrabarty, R. (1988). The Effects of Time-in-Sample Bias on SIPP Estimation. Internal Census Bureau memorandum, October 12th.

Chakrabarty, R. (1989a). *Composite Estimation for SIPP Annual Estimates.* SIPP Working Paper No. 8911, U.S. Bureau of the Census.

Chakrabarty, R. (1989b). Time-in-Sample Biases in the Survey of Income and Program Participation. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 309-314.

Chapman, D. (1988). Minutes of SIPP Eighth Interview Study Debriefing Sessions. Internal Census Bureau memorandum, October 13th.

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. F., Hernandez, D., and Herriot, R. (1986). Longitudinal Household Concepts in SIPP: Preliminary Results. *Proceedings of the Bureau of the Census Second Annual Research Conference,* 598-619. (Also available as SIPP Working Paper No. 8611, U.S. Bureau of the Census.)

Citro, C. F., Hernandez, D., and Moorman, J. (1986). Longitudinal Household Concepts in SIPP. *Proceedings of the Social Statistics Section, American Statistical Association*, 361-365.

Citro, C. F., and Kalton, G. (1993). *The Future of the Survey of Income and Program Participation*. Washington, DC: National Academy Press.

Clark, R., and Speare, A. (1988). Measuring Geographical Mobility Using Panel Data from SIPP. *Proceedings of the Business and Economics Statistics Section, American Statistical Association*, 376-381.

Coder, J. (1980). Some Results from the 1979 ISDP Research Panel. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 540-545.

Coder, J. (1987a). Some Notes on the Quality of SIPP Unemployment Compensation Data. Internal Census Bureau memorandum to G. Green, April 7th.

Coder, J. (1987b). Additional Tabulations of Wave 6 Tax Data. Internal Census Bureau memorandum to G. Green, October 29th.

Coder, J., Burkhead, D., Feldman-Harkins, A., and McNeil, J. (1987). *Preliminary Data from the SIPP 1983-84 Longitudinal Research File.* SIPP Working Paper No. 8702, U.S. Bureau of the Census.

Coder, J., and Scoon-Rogers, L. S. (1996). *Evaluating the Quality Income Data Collection in the Annual Supplement to the March Current Population Survey and the Survey of Income and Program Participation.* SIPP Working Paper No. 9604, U.S. Bureau of the Census.

Curtin, R., Juster, F. T., and Morgan, J. (1989). Survey Estimates of Wealth: An Assessment of Quality. *In The Measurement of Saving, Investment and Wealth* (R. E., Lipsey and H. S. Tice., eds.). National Bureau of Economic Research, Studies of Income and Wealth, 52, University of Chicago Press, Chicago.

Dajani, A. (1996). Generalized Variances for the 1990 SIPP Panel: Revision. Internal Census Bureau memorandum, September 30, 1996.

Dajani, A., and Winters, F. (1997). Adjusting for Wave 1 Unit Nonresponse in the 1996 SIPP Panels. *Proceedings of the Section on Survey Research Methods, American Statistical Association,* 383-388.

Davey, E. H., (1996). Noninterview Rates for Selected Major Demographic Household Surveys. Internal Census Bureau memorandum, July 17th.

David, M. H. (ed.). (1983). *Technical, Conceptual and Administrative Lessons of the Income Survey Development Program*. New York: Social Science Research Council.

DeAre, D. (1988a). SIPP Wave 8 Data on Migration. Internal Census Bureau memorandum to A. Norton, January 7th.

DeAre, D. (1988b). 1985 SIPP Wave 4 Data on Migration. Internal Census Bureau memorandum to A. Norton, June 22nd.

Deville, J. C., Sarndal, C. E., and Sautory, O. (1993). Generalized Raking Procedures in Survey Sampling. *Journal of the American Statistical Association*, 88, 1013-1020.

Dorinski, S. M. (1995). *Continuing Research on Use of Administrative Data in SIPP Longitudinal Estimation*. SIPP Working Paper No. 9506, U.S. Bureau of the Census.

Dorinski, S. M., and Huang, H. (1994). Use of Administrative Data in SIPP Longitudinal Estimation. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 656-661.

Doyle, P., and Dalrymple, R. (1987). The Impact of Imputation Procedures on Distribution Characteristics of the Low Income Population. *Proceedings of the Bureau of the Census Third Annual Research Conference*, 483-508. (Also available as SIPP Working Paper No. 8710, U.S. Bureau of the Census.)

Duncan, G., and Hill, M. (1985). Conceptions of Longitudinal Households: Fertile or Futile? *Journal of Economic and Social Measurement*, 13, 361-376.

Durant, S., and Gbur, P. (1988). *Testing Telephone Interviewing in the Survey of Income and Program Participation and Some Early Results*. SIPP Working Paper No. 8824, U.S. Bureau of the Census.

Eargle, J. (1990). Household Wealth and Asset Ownership: 1988. Current Population Reports P70-22, U.S. Bureau of the Census.

Ernst, L., and Gillman, D. (1988). Excluding Sample that Misses Some Interviews from SIPP Longitudinal Estimates. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 348-353. (Also available as SIPP Working Paper No. 8825, U.S. Bureau of the Census.)

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*, 682-687. (Also available as SIPP Working Paper No. 8406, U.S. Bureau of the Census.)

Fasteau, H. (1973). CPS Redesign-Evaluation of Encoding Subsystem. Internal Census Bureau memorandum to M. Boisen, April 5th.

Fay, R. (1989a). An Analysis of Within-household Undercoverage in the Current Population Survey. *Proceedings of the Bureau of the Census Fifth Annual Research Conference*, 156-175.

Fay, R. (1989b). Theory and Application of Replicate Weighting for Variance Calculations. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 212-217.

Fay, R., Passel, J., Robinson, J., and Cowan, C. (1988). *The Coverage of Population in the 1980 Census*. Evaluation and Research Reports, PHC80-E4, U.S. Bureau of the Census.

Folsom, R., and Witt, M. (1994). Testing a New Attrition Nonresponse Adjustment Method for SIPP. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 428-433.

Gbur, P. (1987). SIPP: Limitations Due to Sample Size. Internal Census Bureau memorandum to R. P. Singh, July 14th.

Gbur, P., and Petroni, R. J. (1989). Preliminary Evaluation of Maximum Telephone Interviewing on the SIPP. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 320-325.

Gift Experiment Workgroup. (1988). SIPP '87: Gift Experiment Results Through Wave 3. Internal Census Bureau memorandum to R. P. Singh, July 6th.

Goudreau, K., Oberheu, H., and Vaughan, D. (1984). An Assessment of the Quality of Survey Reports of Income from the AFDC Program. *Journal of Business and Economic Statistics*, 2, 179-186.

Guarino, J. (1996). ISS Record Study: Profile of Usage Progression. Internal Census Bureau memorandum to SIPP Research and Evaluation Steering Committee, March 12th.

Haber, S. (1989). Findings from the SIPP Fringe Benefits Feasibility Study: Response Rates and Data Quality. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 326-330.

Hainer, P., Hines, C., Martin, E., and Shapiro, G. M. (1988). Research on Improving Coverage in Household Surveys. *Proceedings of the Bureau of the Census Fourth Annual Research Conference*, 513-539.

Heeringa, S., and Lepkowski, J. (1986). Longitudinal Imputation for the SIPP. *Proceedings of the Section on the Survey Research Methods, American Statistical Association*, 206-210. (Also included in SIPP Working Paper No. 8704, "The Treatment of Person-Wave Nonresponse in Longitudinal Surveys.")

Hendrick, M. (1994). Documentation of the Sampling and Code Assignment for the HUCS (Housing Unit Coverage Study) Sample. Internal Census Bureau memorandum for Hendrick for documentation, December 22.

Hendrick, M. R. (1995). The Creation and Evaluation of Longitudinal Nonresponse Cells for the 1996 Survey of Income and Program Participation. *Proceedings of the Section on Survey Research Methods*, Volume I, American Statistical Association.

Hernandez, D. (1987). Shortfalls in Recent Marriages Estimated from the SIPP Wave 8 Marital History Module and Evidence from the PSID. Internal Census Bureau memorandum to A. Norton, January 2nd.

Herriot, R. (1983). The Use of Administrative Records in Social and Demographic Statistics. *Bulletin of the International Statistical Institute*, Vol. L, Book 1, 263-276.

Hill, D. (1986). An Additive Model of Recall Error: Analysis of SIPP Data. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 226-230.

Hill, D. (1987). Response Errors in Labor Surveys: Comparisons of Self and Proxy Reports in the Survey of Income and Program Participation (SIPP). *Proceedings of the Bureau of the Census Third Annual Research Conference*, 299-319. (Also available as SIPP Working Paper No. 8706, U.S. Bureau of the Census.)

Hill, D. (1989). A Poisson Model of Response and Procedural Error Analysis of SIPP Reinterview Data. *Proceedings of the Section on Survey Research Methods, American Statistical Association.* (Also available as SIPP Working Paper No. 8913, U.S. Bureau of the Census.)

Hill, M., and Hill, D. (1986). Labor Force Transitions: A Comparison of Unemployment Estimates from Two Longitudinal Surveys. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 220-225.

Hirschberg, D., Yuskavage, R., and Scheuren, F. (1977). The Impact on Personal and Family Income of Adjusting the Current Population Survey for Undercoverage. *Proceedings of the Social Statistics Section, American Statistical Associatio*n, 70-80.

Hock, E. C., and Winters, F. (1996). Comparing Certain Effects of Redesign on Data from the Survey of Income and Program Participation. *Proceedings of the Section on Survey Research Methods,* Volume I, American Statistical Association, 605-610.

Holden, K. (1988). The Income Shock of Widowhood: Findings from the Survey of Income and Program Participation. Report to the American Association of Retired Persons.

Hoppe, R. (1988). *Defining and Measuring Nonmetro Poverty: Results from the Survey of Income and Program Participation.* SIPP Working Paper No. 8822, U.S. Bureau of the Census.

Huang, H. (1993). Report on SIPP Recall Length Study. Internal Census Bureau memorandum to SIPP Executive Committee, May 12th.

Hubble, D. (1986). SIPP 1984: Declining Hispanic Estimates. Internal Census Bureau memorandum, January 6th.

Huggins, V. J. (1987). Evaluation of Missing Wave Data from the Survey of Income and Program Participation (SIPP). *Proceedings of the Section on Survey Research Methods, American Statistical Association,* 205-209.

Huggins, V. J. (1994). Effect of Telephone Interviewing on Type A and Record Usage Rates. Internal Census Bureau memorandum to S. L. Durant, September 20th.

Huggins, V. J., and Fay, R. (1988). Use of Administrative Data in SIPP Longitudinal Estimation. *Proceedings of the Section on Survey Research Methods, American Statistical Association,* 354-359.

Huggins, V. J., and Fischer, D. (1994). The Redesign of the SIPP. *Proceedings of the Section on Survey Research Methods, American Statistical Association,* 668-673.

Huggins, V. J., Samuel, M., and Weidman, L. (1985). *Methodologies for Imputing Longitudinal Survey Items.* SIPP Working Paper No. 8506, U.S. Bureau of the Census.

Huggins, V. J., and Winters, F. (1995). Income and Poverty Time Series Data from the SIPP. SIPP Working Paper No. 9504, U.S. Bureau of the Census.

Ikeda (1993). Comments on August 24, 1993 Memorandum Estimates for Telephone Interviews. Internal Census Bureau memorandum to R. J. Petroni, November 24th.

Jabine, T. B., King, K. E., and Petroni, R. J. (1990). Survey of Income and Program Participation Quality Profile. 2nd edition. U.S. Bureau of the Census.

James, T. (1992). Record Usage Rates for a Telephone Interview Wave. Internal Census Bureau memorandum, November 4th.

James, T. (1993). SIPP: Monitoring the Effects of Telephone Interviewing - An analysis of Type A Rates. Internal Census Bureau memorandum, September 30, 1993.

James, T. (1994). Monitoring the Effects of Telephone Interviewing. Internal Census Bureau memorandum to Research Evaluation Committee, March 8th.

James, T. L. (1997). Results of the Wave 1 Incentive Experiment in the 1996 Survey of Income and Program Participation. *Proceedings of the Section on Survey Research Methods,* American Statistical Association, 834-839.

Jean, A., and McArthur, E. (1987). *Tracking Persons Over Time.* SIPP Working Paper No. 8701, U.S. Bureau of the Census.

Johnson, N. (1990). Recommended Changes in PSU Definitions for the 1990 SIPP and the CPS Redesign (Doc # 2.2-R-05) Internal Census Bureau Memorandum from PSU Definition Subgroup for 1990 Redesign Budget and Planning Subgroup and Research Agenda Subgroup.

Jones, C. E. (1983a). Survey of Income and Program Participation - Quality Control Specifications for the Verification of Data Keying. Internal Census Bureau memorandum to L. Love, June 10th.

Jones, C. E. (1983b). Cross-Sectional Weighting Specifications for the First Wave of the 1984 Panel of the Survey of Income and Program Participation (SIPP). Internal Census Bureau memorandum to T. Walsh, November 25th.

Jones, C. E. (1985a). SIPP 85: Cross-Sectional Weighting Specifications for Wave I - Revision. Internal Census Bureau memorandum to T. Walsh, November 21st.

Jones, C. E. (1985b). SIPP - Publishing Central City Estimates. Internal Census Bureau memorandum to R. Herriot, December 4th.

Jones, C. E. (1993). Comparison of Record Usage Rates Between Personal Visit Waves and Telephone Interviews Waves. Internal Census Bureau memorandum, September 23rd.

Kalton, G. (1986). Handling Wave Nonresponse in Panel Surveys. *Journal of Official Statistics,* 2, 303-314.

Kalton, G., and Brick, M. (1994). Weighting Schemes for Household Panel Surveys. *Proceedings of the Section on Survey Research Methods, American Statistical Association,* 785-790.

Kalton, G., Hill, D., and Miller, M. E. (1990). *The Seam Effect in Panel Surveys.* SIPP Working Paper No. 9011 (116), U.S. Bureau of the Census.

Kalton, G., Kasprzyk, D., and McMillen, D. B. (1989). Nonsampling Errors in Panel Surveys, in Kasprzyk, D., Duncan, G., Kalton, G., and Singh, R. P. (eds.). *Panel Surveys*. Wiley.

Kalton, G., and Lepkowski, J. (1985). Following Rules in SIPP. *Journal of Economic and Social Measurement,* 13, 319-330.

Kalton, G., Lepkowski, J., Heeringa, S., Lin, T., and Miller, M. E. (1987). *The Treatment of Person-Wave Nonresponse in Longitudinal Surveys.* SIPP Working Paper No. 8704, U.S. Bureau of the Census.

Kalton, G., Lepkowski, J., and Lin, T. (1985). Compensation for Wave Nonresponse in the 1979 ISDP Research Panel. *Proceedings of the Section on Survey Research Methods, American Statistical Associatio*n, 372-377. (Also included in SIPP Working Paper No. 8704, "The Treatment of Person-Wave Nonresponse in Longitudinal Surveys.")

Kalton, G., McMillen, D., and Kasprzyk, D. (1986). Nonsampling Error Issues in SIPP. *Proceedings of the Bureau of the Census Second Annual Research Conference,* 147-164. (Also available as SIPP Working Paper No. 8602, U.S. Bureau of the Census.)

Kalton, G., and Miller, M. E. (1991). The Seam Effect with Social Security Income in the Survey of Income and Program Participation. *Journal of Official Statistics,* 7, 235-245.

Kalton, G., Miller, D. P., and Lepkowski, J. (1992*). Analyzing Spells of Program Participation in the SIPP.* SIPP Working Paper No. 9210 (171), U.S. Bureau of the Census.

Kasprzyk, D. (1983). Social Security Number Reporting, the Use of Administrative Records, and the Multiple Frame Design in the ISDP. *In Technical, Conceptual and Administrative Lessons of the Income Survey Development Program*. M. David, ed. New York: Social Science Research Council.

Kasprzyk, D. (1988). *The Survey of Income and Program Participation: An Overview and Discussion of Research Issues.* SIPP Working Paper No. 8830, U.S. Bureau of the Census.

Kasprzyk, D. (1989). Sample Counts. Internal Census Bureau memorandum to M. Martin, May 10th.

Kasprzyk, D., Duncan, G., Kalton, G., and Singh, R. P. (eds.). (1989). *Panel Surveys*. New York: John Wiley & Sons.

Kasprzyk, D., and Herriot, R. (1986*). Some Aspects of the Survey of Income and Program Participation.* SIPP Working Paper No. 8601, U.S. Bureau of the Census.

Kasprzyk, D., and McMillen, D. (1987). SIPP: Characteristics of the 1994 Panel. *Proceedings of the Social Statistics Section, American Statistical Association*, 181-186.

Kearney, A., Tourangeau, R., Shapiro, G. M., and Ernst, L. (1993). Coverage Improvement From Experimental Residence Questions, Contractor report prepared by NORC for U.S. Bureau of the Census.

King, K. E. (1990). SIPP: Restructuring the 1989 and 1990 Panels. Internal Census Bureau memorandum, January 31st.

King, K. E. (1995). SIPP 1986-1993: Monitoring the Effects of Telephone Interviewing - for Analysis of Type A and Type D Noninterview Rates. Internal Census Bureau memorandum, February 16th.

King, K. E., Petroni, R. J., and Singh, R. P. (1987). *Quality Profile for the Survey of Income and Program Participation.* SIPP Working Paper No. 8708, U.S. Bureau of the Census.

Kish, L. (1992). Weighting for Unequal P_i. *Journal of Official Statistics,* 8, 183-200.

Klein, B., and Vaughan, D. (1980). Validating Recipiency Reporting of AFDC Mothers in a Pilot Household Survey. *In J. Olson* (ed.) (1980). Reports from the Site Research Test.

Kobilarcik, E., and Singh, R. P. (1986). SIPP: Longitudinal Estimation of Persons' Characteristics. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 214-219.

Kominski, R. (1990). The SIPP Event History Calendar: Aiding Respondents in the Dating of Longitudinal Events. *Proceedings of Survey Research Methods Section American Statistical Association,* 553-558.

Lamas, E., and McNeil, J. (1987). An Analysis of the SIPP Asset and Liability Feedback Experiment. *Proceedings of the Social Statistics Section, American Statistical Association,* 195-199. (Also available as SIPP Working Paper No. 8725, U.S. Bureau of the Census.)

Lamas, E., Palumbo, T., and Eargle, J. (1996). *The Effect of the SIPP Redesign on Employment and Earnings Data.* SIPP Working Paper No. 9606, U.S. Bureau of the Census.

Lepkowski, J., Kalton, G., and Kasprzyk, D. (1989). Weighting Adjustments for Partial Nonresponse in the 1984 SIPP Panel. *Proceedings of the Section on Survey Research Methods, American Statistical Association,* 296-301.

Lepkowski, J., Kalton, G., and Kasprzyk, D. (1990). Measurement Errors in Panel Surveys. Paper prepared for the International Conference on Measurement Errors in Surveys. Tucson, Arizona Institute for Social Research, University of Michigan, Ann Arbor, Michigan.

Lepkowski, J., Miller, D., and Luis, E. (1993). Imputation for Wave Non-response in the SIPP. Institute for Social Research, University of Michigan, Ann Arbor, Michigan.

Little, R., and Su, H. (1989). Item Nonresponse in Panel Surveys. In *Panel Surveys*, Kasprzyk, D., Duncan, G. J., Kalton, G., and Singh, R. P., eds. New York: John Wiley and Sons.

Loudermilk, C. (1989). Data on Discontinued Offices. Internal Census Bureau memorandum to P. J. Petroni, March 29, with attachment, Permit office boundary problems, February 14th.

Mack, S. P., Huggins, V. J., Keathley, D., and Sundukchi, M. (1998). Do Monetary Incentives Improve Response Rates in the Survey of Income and Program Participation. *Proceedings of the Section on Survey Research Methods, American Statistical Association* (in press).

Mack, S. P., and Petroni, R. J. (1994). *Overview of SIPP Nonresponse Research Data.* SIPP Working Paper No. 9414, U.S. Bureau of the Census.

Mack, S. P., and Waite, P. J. (1995*). Nonresponse Research Plans for the Survey of Income and Program participation.* SIPP Working Paper No. 9503, U.S. Bureau of the Census.

Mack, S. P., Williams, T., and Tin, J. (1995). SIPP Nonresponse Highlights. Internal Census Bureau memorandum, March 29, 1995.

Marquis, K. H. (1990). Report of the SIPP Cognitive Interviewing Project. Unpublished report. U.S. Bureau of the Census. August 1990.

Marquis, K. H. (1995). The SIPP Measurement Quality Experiment Beyond: Basic Results and Implementation. *Proceedings of the U.S. Bureau of the Census 1995 Annual Research Conference.*

Marquis K. H., and Moore, J. C. (1989a). Response Errors in SIPP: Preliminary Results. *Proceedings of the Bureau of the Census Fifth Annual Research Conference,* 515-536.

Marquis K. H., and Moore, J. C. (1989b). Some Response Errors in SIPP-With Thoughts About Their Effects and Remedies. *Proceedings of the Section on Survey Research Methods, American Statistical Association,* 381-386.

Marquis, K. H., and Moore, J. C. (1990). Measurement Errors in SIPP Program Reports. *Proceedings of the U.S. Bureau of the Census' 1990 Annual Research Conference,* 721-745.

Marquis, K. H., Moore, J. C., and Bogen, K. (1991). A Cognitive Approach to Redesigning Measurement in the Survey of Income and Program Participation. *Proceedings of the Section on Survey Research Methods, American Statistical Association,* 413-418.

Marquis, K. H., Moore, J. C., and Huggins, V. J. (1990). Implications of SIPP Record Check Results for Measurement Principles and Practice. *Proceedings of the Section on Survey Research Methods. American Statistical Association,* 564-569.

Martin, E. (1996). Household Attachment and Survey Coverage. *Proceedings of the Section on Survey Research Methods. American Statistical Association,* 526-531.

McArthur, E. (1987). Measurement of Attrition Through the Completed SIPP 1984 Panel: Preliminary Results. Internal Census Bureau memorandum to the SIPP Nonresponse Working Group, November 20th.

McCormick, M. K., Butler, D. M., and Singh, R. P. (1992). Investigating Time in Sample Effect for the Survey of Income and Program Participation. *Proceedings of the Section on Survey Research Methods. American Statistical Association,* 554-559.

McMahon, M. F. (1998). Sample Loss Rates Through Wave 5 for the 96 Panel. Internal Census Bureau memorandum to The SIPP Record 98-01, February 24th.

McMillen, D. (1989a). The Measurement of Demographic Transitions in the Survey of Income and Program Participation. *Proceedings of the Social Statistics Section, American Statistical Association,* 363-368.

McMillen, D. (1989b). Census Programs for Special Release of SIPP Data. SIPP Supplement to APDU Newsletter, No. 2, January.

McMillen, D., and Herriot, R. (1985). Toward a Longitudinal Definition of Households. *Journal of Economic and Social Measurement.* 13, 504-509. (Also available as SIPP Working Paper No. 8402, U.S. Bureau of the Census.)

McNeil, J. (1988). CPS and SIPP Estimates of Health Insurance Coverage Status. Internal Census Bureau memorandum, May 3rd.

McNeil, J. (1989). Quarterly Estimates of Core Characteristics: 1984, 1985, and 1986 SIPP Panels. Internal Census Bureau memorandum for the record, July 21st.

McNeil, J. (1990). Panel Bias in SIPP. Internal Census Bureau memorandum, June 22nd.

McNeil, J. (1991). Quarterly Estimates from SIPP. Internal Census Bureau memorandum.

McNeil, J., and Lamas, E. (1988). *Year-apart Estimates of Household Net Worth from the Survey of Income and Program Participation.* SIPP Working Paper No. 8804, U.S. Bureau of the Census.

McNeil, J., Lamas, E., and Harpine, C. (1988). Moving Into and Out of Poverty: Data from the First SIPP Panel File. *Proceedings of the Social Statistics Section, American Statistical Association,* 279-283.

Meier, A. (1988). SIPP '85: Results of Debriefing Respondents. Internal Census Bureau memorandum to R. P. Singh, September 26th.

Montfort, E. (1988). Mobile Home New Construction for 1985 AHS-National. Internal Census Bureau memorandum to D. Schwanz, February 2nd.

Moore, J. C. (1988). Self/proxy Response Status and Survey Response Quality - A Review of the Literature. *Journal of Official Statistics,* 4, 155-172.

Moore, J. C., Bogen, K., and Marquis, K. H. (1993). A Cognitive Interviewing Approach for the Survey of Income and Program Participation: Development of Procedures and Initial Test Results. *Proceedings of Statistics Canada Symposium*, 92, 31-40.

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*. 726-731. (Also available as SIPP Working Paper No. 8407, U.S. Bureau of the Census.)

Moore, J. C., and Marquis K. H. (1989). The Quality of Self and Proxy Data in a Census Bureau Panel Survey. Presented at the annual meeting of the American Association for Public Opinion Research, St. Petersburg, Florida, May 18-21st.

Moore, J. C., Marquis, K. H., and Bogen, K. (1996). The SIPP Cognitive Research Evaluation Experiment: Basic Results and Documentation. Unpublished U.S. Bureau of the Census report, January 11, 1996.

Nelson, C. T. (1993). *The Quality of Census Bureau Survey Data Among Respondents with High Income.* SIPP Working Paper No. 9306, U.S. Bureau of the Census.

Nelson, D., Bowie, C. G., and Walker, A. (1987). SIPP Sample Loss and the Efforts to Reduce It. *Proceedings of the U.S. Bureau of the Census Third Annual Research Conference*, 629-643. (Also available as SIPP Working Paper No. 8709, U.S. Bureau of the Census.)

Nelson, D., McMillen, D., and Kasprzyk, D. (1985). *An Overview of the SIPP, Update 1.* SIPP Working Paper No. 84-01, U.S. Bureau of the Census.

Neter, J., and Waksberg J. (1965). *Response Errors in Collection of Expenditures Data by Household Interview: An Experimental Study*. Technical Paper No. 11, U.S. Bureau of the Census.

O'Connell, M. (1987a). Analytical Problems Encountered When Using Wave 8 Data for Cross Sectional Analysis. Internal Census Bureau memorandum, December 3rd.

O'Connell, M. (1987b). Further Analysis of SIPP Wave 8 Data on Marital Events. Internal Census Bureau memorandum, December 18th.

O'Connell, M. (1988). Review of SIPP 1985 Panel Wave 4 Fertility History Module. Internal Census Bureau memorandum to D. Kasprzyk, August 22nd.

Olson, J. (ed.) (1980). Reports from the Site Research Test: Income Survey Development Program. Washington, DC: Social Security Administration.

Parmer, R. (1985). Documentation of the Decision to Use Census Misses from the Housing Unit Coverage Studies (HUCS) for Coverage Improvement of the Current Surveys. Internal Census Bureau memorandum to D. Schwanz, January 10th.

Pennell, S. (1993). *Cross-Sectional Imputation and Longitudinal Editing Procedures in the Survey of Income and Program Participation.* SIPP Working Paper No. 9314, U.S. Bureau of the Census.

Pennell, S. G., and Lepkowski, J. M. (1992). Panel Conditioning Effects in the Survey of Income and Program Participation. *Proceedings of Survey Research Methods Section American Statistical Association*, 566-571.

Pennie, K. (1988). The Survey of Income and Program Participation: Results of the Reinterview Program for 1987. Internal Census Bureau memorandum, November 7th.

Petroni, R. J. (1989). SIPP Objectives. Internal Census Bureau memorandum, July 11th.

Petroni, R. J. (1994). SIPP: Coverage of Persons Entering and Leaving a Household During a Wave. Internal Census Bureau memorandum, March 2nd.

Petroni, R. J., Huggins, V. J., and Carmody, T. (1989). Research and Evaluation Conducted on SIPP. *Proceedings of the Bureau of the Census Fifth Annual Research Conference*, 308-338.

Petroni, R. J., and King, K. (1988). Evaluation of the Survey of Income and Program Participation's Cross-Sectional Noninterview Adjustment Methods. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 342-347.

Rawlings, S. W., and Saluter, A. F. (1995). Household and Family Characteristics: March 1994, U.S. Bureau of the Census, Current Population Reports, P20-483.

Riccini, D. (1984). Imputation of Current Household Size Variables for Mover Households-Revised. Internal Census Bureau memorandum, September 27th.

Rizzo, L., Kalton, G., and Brick, M. (1996). A Comparison of Some Weighting Adjustment Methods for Panel Nonresponse. *Survey Methodology*, 22(1), 43-53.

Rizzo, L., Kalton, G., Brick, M., and Petroni, R. (1994). Adjusting for Panel Nonresponse in the SIPP. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 422-427.

Roebuck, M. (1985). SIPP Variance Estimation for 3rd Quarter 1983. Internal Census Bureau memorandum, May 28th.

Roebuck, M. (1986). 1985 SIPP Variance Estimation: Choosing An Estimator. Internal Census Bureau memorandum, March 17th.

Roman, A., and O'Brien, D. (1984). The Student Follow-up Investigation of the 1979 ISDP. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 732-737. (Also available as SIPP Working Paper No. 8407, U.S. Bureau of the Census.)

Ruggles, P. (1990). Drawing the Line: Alternative Poverty Measures and Their Implications for Public Policy. Washington, DC: Urban Institute Press.

Ruggles, P., and Williams, R. (1989). *Measuring the Duration of Poverty Spells.* SIPP Working Paper No. 8909, U.S. Bureau of the Census.

Ryscavage, P., and Bregger, J. (1985). *New Household Survey and the CPS: A Look at Labor Force Differences. Monthly Labor Review*, 108(9), 3-12. (Also available as SIPP Working Paper No. 8507, U.S. Bureau of the Census.)

Ryscavage, P., and Coder, J. (1989). *Longitudinal and Retrospective Measures of Work Experience.* SIPP Working Paper No. 8903, U.S. Bureau of the Census.

Ryscavage, P., and Feldman-Harkins, A. (1988). *A Comparison of Gross Changes in Labor Force Status from SIPP and CPS.* SIPP Working Paper No. 8816, U.S. Bureau of the Census.

Sae-Ung, S., and Winters, F. (1998). Analysis of Nonresponse Effects on Income and Poverty Time Series Data from SIPP. Draft paper.

Scheuren, F. (1983). Design and Estimation for Large Federal Surveys Using Administrative Records. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 377-381.

Schwede, L. (1993). An Empirical Exploration of Residence Rules: The Living Situation Survey. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 144-149.

Scoon-Rogers, L., and Lester, G. H. (1995). Child Support for Custodial Mothers and Fathers: 1991 Current Population Reports Consumer Income Series, 60-187. August 1995.

Shapiro, G. M. (1972). Current Survey Segmenting Error Involving Two Addresses with Identical House Numbers. Internal Census Bureau memorandum, August 15th.

Shapiro, G. M. (1983a). NCHS/SIPP/GPS/QHS Redesign - Procedure for Computer Segmentation of Address-type Enumeration Districts (EDs) and Reservation of Measures. Internal Census Memorandum to R. O'Reagan, October 12th.

Shapiro, G. M. (1983b). NCS/SIPP/CPS/HIS/QHS Redesign - General Sampling Specifications for Old Construction in Nonrotating PSUs. Internal Census Bureau memorandum to R. O'Reagan, June 17th.

Shapiro, G. M. (1983c). NCS/SIPP/GPS Redesign: Area Sampling Specifications Using 1980 Census Materials. Internal Census Bureau memorandum to R. O'Reagan, October 6th.

Shapiro, G. M. (1984). NCS/SIPP/GPS/HIS/QHS/AHS - National General Sampling Specification for Old Construction in Nonrotating PSUs - Final Revision. Internal Census Bureau memorandum to R. O'Reagan, June 1st.

Shapiro, G. M. (1992). Whole Household Undercoverage vs. Within Household Undercoverage. Internal Census Bureau memorandum to Coverage Research Committee, February 27th.

Shapiro, G. M., Diffendal, G., and Cantor, D. (1993). Survey Undercoverage: Major Causes and New Estimates of Magnitude. Internal Census Bureau Memorandum.

Shapiro, G. M., and Kostanich, D. (1988). High Response Error and Poor Coverage are Severely Hurting the Value of Household Survey Data. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 443-448.

Shea, M. (1995a). SIPP Data Quality. Internal Census Bureau memorandum to V. J. Huggins, February 10th.

Shea, M. (1995b). Dynamics of Economic Well-Being: Poverty 1990-1992. Current Population Reports: Household Economic Studies, 70-112. January 1995.

Siegel, P., and Mack. S. P. (1995). Overview of Redesign Methodology for the SIPP. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 227-232.

Singh, R. P. (1985). SIPP: Research on SIPP Oversampling/Subsampling. Internal Census Bureau memorandum to G. M. Shapiro and D. Kasprzyk, August 12th.

Singh, R. P. (1986). SIPP 84: Deviations from Cross-Sectional Weighting Specifications. Internal Census Bureau memorandum to G. M. Shapiro, December 22nd.

Singh, R. P. (1987). SIPP 84: Extension of Gross Flows Studies. Internal Census Bureau memorandum to G. M. Shapiro, December 22nd.

Singh, R. P. (1988a). SIPP Estimates of Number of Persons Marrying Per Year. Internal Census Bureau memorandum to G. M. Shapiro, February 1st.

Singh, R. P. (1988b). SIPP 85: Evaluation of the Employer-provided Benefits Study. Internal Census Bureau memorandum to G. M. Shapiro, October 2nd.

Singh, R. P. (1989a). SIPP: Decisions on Ways to Improve Data Quality. R. P. Singh, to the Research and Evaluation Steering Committee, February 7, 1989. Internal Census Bureau memorandum to C. G. Bowie, August 2nd.

Singh, R. P. (1989b). SIPP: Decisions on Ways to Improve Data Quality. Internal Census Bureau memorandum to Research and Evaluation Steering Committee, February 7th.

Singh, R. P. (1991). Results from the SIPP Field Representatives Questionnaire. Internal Census Bureau memorandum, to G. M. Shapiro, November 4th.

Singh, R. P., and Petroni, R. J. (1988). *Nonresponse Adjustment Methods for Demographic Surveys at the U.S. Bureau of the Census.* SIPP Working Paper No. 8823, U.S. Bureau of the Census.

Singh, R. P., Weidman, L., and Shapiro, G. M. (1989). Quality of SIPP Estimates. In *Individuals and Families in Transition: Understanding Change Through Longitudinal Data*. Papers presented at the Social Science Research Council Conference in Annapolis, Maryland, March 16-18, 1988, U.S. Bureau of the Census. (Also available as SIPP Working Paper No. 8901, U.S. Bureau of the Census.)

SIPP Nonresponse Workgroup. (1988). SIPP '84: Characteristics of Initially Interviewed Households by Response Status. Internal Census Bureau memorandum, February 10th.

Statt, R., Vacca, E. A., Wolters, C., and Hernandez, R. (1981). Problems Associated with Using Building Permits As a Frame of Post-census Construction: Permit Lag and Identification. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 226-231.

Sundukchi, M. (1998). SIPP 96: Wave 7 Incentives. Internal Census Bureau memorandum from Baer to Kirkendall, April 1st.

Sutterlin, R. (1993). Estimates for Telephone Interviews. Internal Census Bureau memorandum to R. J. Petroni, August 24th.

Sweet, E. M. (1994). Roster Research Results from the Living Situation Survey. Internal Census Bureau memorandum.

Tepping, B., and Boland, K. (1972). *Response Variance in the Current Population Survey.* Working Paper No. 36, U.S. Bureau of the Census.

Tourangeau, R., Shapiro, G. M., Kearney, A., and Ernst, L. (1993). Who Lives Here? Survey Undercoverage and Household Roster Questions. Contractor report prepared by NORC for Census Bureau.

Tin, J. (1996). *Program Participation and Attrition: The Empirical Evidence*. SIPP Working Paper No. 9611, U.S. Bureau of the Census.

Tremblay, A. (1994). Longitudinal Imputation of SIPP Food Stamp Benefits. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 809-814.

U.S. Bureau of the Census. (1968). *Evaluation and Research Program of the U.S. Census of Population and Housing, 1960: Effects of Interviewers and Crew Leaders*. Series ER 60, No. 7. Washington, DC: U.S. Government Printing Office.

U.S. Bureau of the Census. (1978). *The Current Population Survey: Design and Methodology*. Technical Paper 40. Washington DC: U.S. Government Printing Office.

U.S. Bureau of the Census. (1985). *Economic Characteristics of Households in the United States*. Current Population Reports, Series P-70, Nos. 2 to 6. Washington, DC: U.S. Government Printing Office.

U.S. Bureau of the Census. (1986). *Economic Characteristics of Households in the United States*. Current Population Reports, Series P-70, No. 6. Washington, DC: U.S. Government Printing Office.

U.S. Bureau of the Census. (1988). SIPP 5010 Interviewer's Manual. Washington, DC: U.S. Bureau of the Census.

U.S. Bureau of the Census. (1998). *Survey of Income and Program Participation User's Guide*. Washington, DC: U.S. Bureau of the Census.

U.S. Bureau of the Census. (1998a). *Automated Data Collection: How Have the Demographic Surveys Fared Thus Far?* Washington, DC: U.S. Bureau of the Census.

Vaughan, D. (1989a). Reflections on the Income Estimates from the Initial Panel of the Survey of Income and Program Participation (SIPP). In *Individuals and Families in Transition: Understanding Change Through Longitudinal Data*. Papers presented at the Social Science Research Council Conference in Annapolis, Maryland, March 16-18, 1988, U.S. Bureau of the Census. (Also available as SIPP Working Paper No. 8906, U.S. Bureau of the Census.)

Vaughan, D. (1989b). An Initial Look at Subannual Estimates of Net Income from Self-employment in the Survey of Income and Program Participation (SIPP). Informal note to addressees, March 20th.

Vaughan, D. (1989c). Development and Evaluation of a Survey-based Type of Benefit Classification for the Social Security Program. *Social Security Bulletin*, 52(l)**,** 12-26. (Also available as SIPP Working Paper No. 8920, U.S. Bureau of the Census.)

Waite, P. J. (1988). SIPP: Data Keying Verification Results (July-September 1988). Internal Census Bureau memorandum to T. Walsh, October 26th.

Waite, P. J. (1990). SIPP: Maximum Telephone Interviewing Experiment - Preliminary Analysis. Internal Census Bureau memorandum to S. C. Courtland, June 26th.

Waite, P. J. (1995a). SIPP Nonresponse. Internal Census Bureau memorandum to M. J. Longini, June 8th.

Waite, P. J. (1995b). SIPP 1996: Creating Useful Direct Variances and Generalized Variance Parameters. Internal Census Bureau memorandum to D. Weinberg, and A. Norton, February 24th.

Waite, P. J. (1996). SIPP 96: Specifications for Interview Mode Flag. Internal Census Bureau memorandum to Chester Bowie, May 17th.

Waite, P.J., Huggins, U.J., and Mack, S.P. (1997). Assessment of Efforts to Reduce Nonresponse Bias: 1996 Survey of Income and Program Participation. Paper presented at the 8th International Workshop on Household Survey Nonresponse, Mannheim, Germany, Sept. 24-26.

Waksberg, J. (1973). The Effects of Stratification with Differential Sampling Rates on Attributes of Subsets of the Population. *Proceedings of the Section on Social Statistics, American Statistical Association*, 429-434.

Weidman, L., King, K., and Williams, T. (1988). Evaluation of the SIPP Asset Feedback Experiment. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 337-341.

Weinberg, D. and Petroni, R. J. (1992). The Survey of Income and Program Participation in the 1990's. SIPP Working Paper No. 9206. U.S. Bureau of the Census.

Weller, G. (1991). Obtaining Within-PSU Stratification and Sampling Rates for SIPP Oversampling. Internal Census Bureau memorandum to B. Walter, June 7th.

Weller, G., Huggins, V. J., and Singh, R. P. (1991). Oversampling the Low-Income Population in the Survey of Income and Program Participation (SIPP). *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 544-549.

Williams, R. (1987). *Measuring Poverty with the SIPP and the CPS.* SIPP Working Paper No. 8723, U.S. Bureau of the Census.

Williams, T. (1993). SIPP Variances for the First Quarter of the 1990 Calendar Year. Internal Census Bureau memorandum.

Williams, T. (1997). SIPP Calendar Year Estimation Research: Procedures and Results. Internal Census Bureau memorandum to SIPP R&E Committee.

Williams, T., and Bailey, L. (1996). Compensating for Missing Wave Data in the SIPP. SIPP Working Paper No. 9605, U.S. Bureau of the Census.

Winters, F. (1998). SIPP:96 Incentives for Reducing Attrition. Census Bureau memorandum from Tupek to Kirkendall, May 29th.

Ycas, M., and Lininger, C. (1981). The Income Survey Development Program: Design Features and Initial Findings. *Social Security Bulletin*, 44(11)**,** 13-19.

Young, N. (1989). Wave Seam Effects in the SIPP. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 393-398. (Also available as SIPP Working Paper No. 8921, U.S. Bureau of the Census.)

INDEX