2020 Survey of Income and Program Participation

Users' Guide

October 2021

Washington, D.C.

U.S. DEPARTMENT OF COMMERCE

ECONOMIC AND STATISTICS ADMINISTRATION

U.S. CENSUS BUREAU

This page intentionally left blank.

Acknowledgements

The *2019 Survey of Income and Program Participation (SIPP) Users' Guide* was prepared by Shelley Irving and Adam Smith of the SIPP Coordination and Outreach Staff in the Social, Economic, and Housing Statistics Division (SEHSD) of the U.S. Census Bureau. Many other staff from a number of divisions within the Census Bureau provided topic-specific content and useful comments, including analysts and subject-matter experts in SEHSD, the Demographic Statistical Methods Division (DSMD), and the SIPP survey team.

We also acknowledge the Census Bureau field representatives and staff, without whose dedication the SIPP would be impossible.

****

Because this edition of the *Users' Guide* builds on the previous editions, we also include the following acknowledgments, which appeared in the prior editions:

The first edition of the Survey of Income and Program Participation (SIPP) Users' Guide was prepared by Daniel Kasprzyk (then Office of the Director), Pat Doyle (Mathematica Policy Research, Inc.), Arnold Goldstein (Population Division), Patricia Kelly (Office of the Director), and David B. McMillen (then Office of the Director).

The second edition was prepared by the Data Access and Use Staff of the Data User Services Division. Geneva Burns coordinated the effort, assisted by Jackson Morton and J. Paul Wyatt. Andrea Meier of the Survey of Income and Program Participation Branch in the Statistical Methods Division prepared Chapter 8, "SIPP Cross-Sectional Weighting Procedures," under the direction of Rajendra P. Singh. We would like to thank our colleagues within the Census Bureau and our SIPP file users for their helpful comments.

The third edition of the *Survey of Income and Program Participation (SIPP) Users' Guide* was prepared for the U.S. Census Bureau by Westat. Charles T. Nelson was the Government Project Officer for the project within the Census Bureau, and Pat Doyle also provided invaluable support and guidance to the effort. Many other staff from a number of divisions within the Census Bureau shared their expertise and provided useful comments. In particular, we would like to thank Patrick Benton, John Boies, Judith Hubbard Eargle, Donald Keathly, Karen Ellen King, Gordon Lester, Stephen Mack, Mike McMahon, Thomas Palumbo, Donna Riccini, and Mahdi Sundukchi.

Chapters of the third edition were prepared by Louis Rizzo, Marianne Winglee, Alan Martinson, and Ilene France of Westat; Larry Radbill of Mathematica Policy Research, Inc.; Julie Sykes (then of Mathematica Policy Research, Inc.); and Elizabeth Sheley (Independent Consultant). Alan Martinson, Marty Franklin, Laurie Tomasino, and Carol Dominique of Westat provided editorial and production support; Julie Phillips (Independent Consultant) prepared the Index; and Ana Horton of Westat designed the cover. Garrett Moran served as the Westat Project Director.

CONTENTS

1 INTRODUCTION TO THE SURVEY OF INCOME AND PROGRAM PARTICIPATION 1

1.1 SIPP's Mission 1

1.2 A Brief History of SIPP 1

1.3 Uses of SIPP 3

1.4 Survey Content 4

1.5 Survey Organization 5

1.5.1 Panels 5
1.5.2 Waves 6
1.5.3 Rotation Groups (1984-2008 Panels) 7
1.5.4 Reference Periods 9
1.5.5 Core Content and Topical Modules (1984-2008 Panels) 9

2 SAMPLING DESIGN AND INTERVIEW PROCEDURES 14

2.1 Sample Design 14

2.1.1 Primary Sampling Units (PSUs) 14
2.1.2 Selection of Addresses in Sample PSUs 15

2.2 Locating Sample Address and Respondents 15

2.3 Following Movers 17

2.4 Difference Between Movers and Those Who Are Temporarily Away 18

2.5 Type 2 Persons 19

2.6 Interview Sequence 19

2.6.1 Income Screener Questions 21
2.6.2 Grouping Families (Clumps) 22

2.7 Feedback/Dependent Interviewing 23

2.8 Nonresponse 23

2.8.1 Type A Household Nonresponse 24
2.8.2 Type B Household Nonresponse 24
2.8.3 Type C Household Nonresponse 24
2.8.4 Type D Household Nonresponse 25
2.8.5 Type Z Person Nonresponse 25
2.8.6 Item Nonresponse 25

3 SIPP PUBLIC-USE DATA FILES 26

3.1 SIPP Data Files 26

3.1.1 2014 Panel Through 2020 SIPP 27
3.1.2 2014 SSA Supplement 28

3.1.3 1996 Through 2008 Panels *28*
3.1.4 1984 Through 1993 Panels *28*

3.2 CONFIDENTIALITY PROCEDURES FOR THE PUBLIC-USE FILES 29
3.2.1 Topcoding *29*
3.2.2 Suppression of Geographic Information *29*

3.3 STATUS FLAGS 30

3.4 UNDERSTANDING THE IDENTIFICATION (ID) VARIABLES IN SIPP 31
3.4.1 Sample Unit IDs *32*
3.4.2 Person Numbers *32*
3.4.3 Address IDs *32*
3.4.4 Month of Record IDs *33*

3.5 IDENTIFYING PERSONS AND THEIR RELATIONSHIPS: 2014 PANEL THROUGH 2020 SIPP 33
3.5.1 Determining Monthly Household Composition *34*
3.5.2 Monthly Household Relationships *38*
3.5.3 Identifying Parents, Children, Spouses, and Partners at a Monthly Level *39*
3.5.4 Determining Monthly Family Composition *39*
3.5.5 Type 2 Persons *40*

3.6 IDENTIFYING PERSONS AND THEIR RELATIONSHIPS: 1984 THROUGH 2008 PANELS 41
3.6.1 Determining Monthly Household Composition *41*
3.6.2 Determining Monthly Family Composition *41*
3.6.3 Determining Monthly Transfer Program Unit Composition *42*

3.7 IDENTIFYING PROGRAM PARTICIPATION IN GOVERNMENT TRANSFER PROGRAMS: 2014 PANEL THROUGH 2020 SIPP 42
3.7.1 Collecting Information for Program Units *43*
3.7.2 Identifying Program Participation in Government Transfer Programs *43*
3.7.3 Detailed Means-Tested Program Participation Variables *45*
3.7.4 Means-Tested Program Participation in a Sample Household *49*

3.8 GEOGRAPHY 52

4 TOPICS IN THE 2020 SIPP PUBLIC-USE DATA 54

4.1 DEMOGRAPHICS 54
4.1.1 Basic Demographics *54*
4.1.2 Educational Enrollment *55*
4.1.3 Educational Attainment *55*
4.1.4 Family and Household Relationships *56*
4.1.5 Fertility *58*
4.1.6 Language *59*
4.1.7 Marital Status and Marital History *60*
4.1.8 Nativity, Citizenship, and Parent Nativity *61*
4.1.9 Parent Mortality *62*
4.1.10 Residence *62*
4.1.11 Veteran Status *64*

4.2 ASSETS, EMPLOYMENT, AND EARNINGS 65
4.2.1 Assets *65*
4.2.2 Commuting and Work Schedules *70*
4.2.3 Employment and Earnings *72*

4.2.4 Parental Leave *76*

4.3 HEALTH AND WELL-BEING 77
4.3.1 Adult Well-Being *77*
4.3.2 Child Care *78*
4.3.3 Child Well-Being *79*
4.3.4 Dependent Care Expenses *79*
4.3.5 Disability *80*
4.3.6 Health Care Utilization and Medical Expenditures *81*
4.3.7 Health Insurance *82*

4.4 PROGRAM PARTICIPATION AND INCOME TRANSFERS 85
4.4.1 Annuity/Life Insurance Retirement Income *86*
4.4.2 Disability Income *87*
4.4.3 Energy Assistance *88*
4.4.4 General Assistance (GA) *89*
4.4.5 Lump Sum Severance Pay/Retirement Plan Income *90*
4.4.6 Miscellaneous Income *91*
4.4.7 Other Assistance *92*
4.4.8 Retirement Income *96*
4.4.9 School Meals *98*
4.4.10 Social Security Benefits for Oneself *99*
4.4.11 Social Security Benefits on Behalf of a Child *100*
4.4.12 Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) *101*
4.4.13 Supplemental Nutrition Assistance Program (SNAP) *103*
4.4.14 Supplemental Security Income (SSI) *105*
4.4.15 Support Paid *106*
4.4.16 Support Received *107*
4.4.17 Survivor Income *108*
4.4.18 Temporary Assistance for Needy Families (TANF) *109*
4.4.19 Tax Returns *110*
4.4.20 Unemployment Compensation Payment *112*
4.4.21 Veterans Benefits *113*
4.4.22 Workers' Compensation *114*

4.5 INCOME AND POVERTY RECODES 115
4.5.1 Income Recodes *116*
4.5.2 Poverty Recodes *123*

5 RESOURCES AND RESEARCH 126

5.1 SIPP WEBSITE 126
5.1.1 Data *126*
5.1.2 Technical Documentation *126*

5.2 RESEARCH PRODUCTS 127
5.2.1 P-70 Reports *127*
5.2.2 Table Packages *128*
5.2.3 Conference Papers and Presentations *128*

5.3 WEBINARS AND WORKSHOPS 128

6 DATA EDITING AND IMPUTATION 130

6.1 DATA EDITING 130

6.1.1 Types of Missing Data *130*
6.1.2 Goals of Imputation *131*
6.1.3 Assessing the Influence of Imputed Data on Analysis *131*

6.2 IMPUTATION 132
6.2.1 Model-Based Imputation *132*
6.2.2 Sequential Hot-Deck Imputation *135*

7 NONSAMPLING ERROR, SAMPLING ERROR, AND WEIGHTING 139

7.1 NON-SAMPLING ERROR 139
7.1.1 Effects of Non-sampling Error on Survey Estimates *140*

7.2 SAMPLING ERROR AND VARIANCE ESTIMATION 141
7.2.1 Direct Variance Estimation *142*
7.2.2 Variance Units and Variance Strata *142*
7.2.3 Replicate Weights *142*
7.2.4 Approximate Variance Estimates *144*
7.2.5 Variance Estimation with Imputed Data *145*

7.3 WEIGHTING 145
7.3.1 Choosing a Weight, 2018 Through 2020 SIPP *145*
7.3.2 How Weights Are Constructed *146*
7.3.3 Reference Month Weights *146*
7.3.4 Longitudinal and Calendar Year Weights *148*

APPENDICES 150

A. ACRONYMS 151

B. GLOSSARY 154

C. TOPICAL MODULE DESCRIPTIONS (1984 TO 2008 PANELS) 163

D. PUBLICATIONS IN THE P-70 SERIES 171

REFERENCES 177

Chapter 1

1 Introduction to the Survey of Income and Program Participation

This guide is intended as a reference for using 2020 Survey of Income and Program Participation (2020 SIPP[1]) data. Information on previous years of SIPP data is included; however, please see previous versions of the SIPP Users' Guide at <www.census.gov/programs-surveys/sipp/guidance/users-guide.html> for more details on historic data. Additional guidance about the survey is available on the SIPP website at <www.census.gov/sipp>, or by contacting the SIPP team at <census.sipp@census.gov>.

1.1 SIPP's Mission

Even as the face of SIPP has changed over the years, its purpose has remained the same: SIPP provides high-quality data related to income, labor force participation, social program participation and eligibility, and general demographic characteristics to measure the effectiveness of existing federal, state, and local programs. SIPP also serves to estimate future costs and coverage for government programs and to provide improved statistics on the distribution of income and measures of economic well-being in the country. The main objective of the SIPP has been, and continues to be, to provide accurate and comprehensive information about the income and program participation of individuals and households in the United States. The survey's mission is to provide a nationally representative sample for evaluating: 1) annual and sub-annual income dynamics; 2) movements into and out of government transfer programs; 3) family and social context of individuals and households; and 4) interactions among these areas.

1.2 A Brief History of SIPP

Until the advent of SIPP, the major source of data on income and program participation was the Current Population Survey's (CPS) March Income Supplement. The CPS continues to be the source of all official income and poverty statistics published by the Census Bureau; however, the CPS is designed primarily to obtain information on employment. Because income measurement was never the primary purpose of the CPS, it has certain gaps in that area. The CPS does not capture the impact of changes in household composition during the year, nor does the survey explicitly measure periods of program participation. Additionally, the CPS does not collect data on assets and liabilities, which are needed to completely measure a household's economic status and eligibility for program benefits. To add those items to the CPS questionnaire would dilute the focus of the survey and unduly increase survey participant burden. Finally, the CPS is designed to be a cross-sectional survey. During the 1970s, the increasing size of government programs and their interactions with the labor market led to a need for longitudinal data.

To address these data issues, the Department of Health, Education, and Welfare (HEW) initiated the Income Survey Development Program (ISDP) in the late 1970s. In developing ISDP content and procedures, HEW focused on questionnaire length, length of reference period, and linkage of survey data

[1] Beginning with 2018 SIPP, data are collected as yearly overlapping panels with new panels beginning each year. The 2020 SIPP includes Wave 3 of the 2018 Panel and Wave 1 of the 2020 Panel but is referred to as "2020 SIPP". The 2019 Panel was discontinued after Wave 1 and therefore not extended into subsequent years.

to program records. Much of the work centered around four experimental field tests that were conducted in collaboration with the Census Bureau. These tests examined different concepts, procedures, questionnaires, and recall periods.

Creation of SIPP (1984)

Based on experience obtained in the ISDP, planning began for implementation of a new survey known as the Survey of Income and Program Participation. The primary goals in designing SIPP were to improve reporting of income and other program-related data in a way that would allow the analysis of changes over time at a micro-level. The design also had to accommodate the collection of a large quantity of information in a flexible manner that allowed some information to be collected more frequently than other information. These goals were met principally by using a survey design in which the same people are interviewed more than once. Persons at households selected for a sample panel are interviewed about their income and other topics once every 4 months for approximately 2 1/2 years.

The first SIPP interviews were conducted in October of 1983 with a sample of over 20,000 households. At each interview, survey participants were asked to provide information covering the 4 months since the previous interview. The design of SIPP called for a nationally representative sample of individuals 15 years of age and older in households in the civilian non-institutionalized population. Those individuals, along with others who subsequently lived with them, were to be interviewed once every 4 months over a 32-month period.

The 1996 Redesign

In 1990, the Census Bureau asked the Committee on National Statistics (CNSTAT) at the National Research Council to undertake a comprehensive review of SIPP. The resulting report, *The Future of the Survey of Income and Program Participation* (Citro and Kalton, 1993), summarized the first 9 years of SIPP and provided recommendations for the future of the survey. Some of those recommendations were implemented in what is known as the 1996 redesign.

One of the goals of the 1996 redesign was to improve the quality of longitudinal estimates in order to provide better information for policy makers. Specific changes to the survey included:

- A larger sample,
- Interviewing participants over a full 4 years (48 months) instead of 32 months,
- The introduction of computer-assisted interviewing (CAI), which, among other improvements, permitted automatic consistency checks of reported data during the interview. Those checks reduced the level of post-collection edits and imputation and thus helped to maintain longitudinal consistency, and
- An oversampling of households from areas with high poverty concentrations.

Delayed by the 1995 and 1995-1996 federal government shutdowns and related furloughs, instead of in February 1996, the first interviews of the redesigned SIPP began in April 1996. Later in 1996, Congress passed the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA). That law significantly altered the nature of public transfer programs, shifting more responsibility to state governments, establishing new eligibility rules for a number of programs, and setting limits on who could receive benefits. The existing welfare program, Aid to Families with Dependent Children (AFDC), was

replaced with a new program, Temporary Assistance for Needy Families (TANF). Those changes came after interviewing for the 1996 Panel had already begun with a questionnaire designed for the array of transfer programs that existed before PRWORA was enacted. To accommodate program changes brought about by PRWORA, the Census Bureau began adapting transfer-program questions to reflect the current situation.

The 2014 Redesign (Re-engineered SIPP)

In 2006, the U.S. Census Bureau began its second major redesign of the SIPP that became known as 'Re-engineered SIPP' (or 'Re-SIPP'). The catalyst for this redesign was SIPP's imminent cancellation. Though SIPP was saved by support from various stakeholders, major changes were deemed necessary. The goals of the redesign were to reduce costs and respondent burden and to improve data quality and timeliness. The new survey instrument, called SIPP-EHC, was a complete redevelopment of the survey instrument built around changing the survey reference period from four months to one year. The new design also introduced an Event History Calendar (EHC) that assists the respondent's ability to recall events accurately over the longer reference period. The EHC also provides increased data quality and inter-topic consistency for information reported by survey participants. The first interviews with this new instrument were conducted in February 2014. The 2020 SIPP data, which are the focus of this guide, are based on the 2014 redesign.

1.3 Uses of SIPP

One of the most important reasons for conducting SIPP is to gather detailed information on participation in government transfer programs. Data from SIPP allow research on concurrent participation in multiple programs. SIPP data can also be used to address types of questions such as:

- How have changes in eligibility rules or benefit levels affected recipients?
- How have changes in the eligibility rules affected the program target population, that is, those eligible to receive benefits?
- How does income from other household members affect labor force participation and reasons for not working?
- How do wealth and income patterns differ for various age, sex, and racial groups?

SIPP's longitudinal features allow the analysis of selected dynamic characteristics of the population, such as changes in income, eligibility for and participation in transfer programs, household and family composition, labor force behavior, and other associated events. SIPP's design also allows cross-sectional data analyses.

Because SIPP is a longitudinal survey, capturing changes in household and family composition over a multiyear period, it can also be used to address questions like:

- What factors affect change in household and family structure and living arrangements?
- What are the interactions between changes in the structure of households and families and the distribution of income?
- What effects do changes in household composition have on economic status and program eligibility?

- What are the primary determinants of turnover in programs such as the Supplemental Nutrition Assistance Program (SNAP)?

1.4 Survey Content

While the main objective of SIPP is to provide accurate and comprehensive information about the income and program participation of individuals and households in the United States, SIPP also collects extensive data on other factors of economic well-being such as family dynamics, educational attainment, housing expenditures, asset ownership, health insurance, disability, child care, and food security. These data put the income and program recipiency of individuals and households into the family and social context.

Table 1-1 gives an overview of topics collected in SIPP, and Chapter 4 of this guide provides detailed information about how these topics appear in the 2020 SIPP public-use data. For more information about the collection of each topic prior to 2020, please visit the Content page of the SIPP website at <www.census.gov/programs-surveys/sipp/about/sipp-content-information.html>.

TABLE 1-1. TOPICS COLLECTED IN THE SURVEY OF INCOME AND PROGRAM PARTICIPATION

Demographic Characteristics	**Assets and Liabilities**	**Program Participation and Income Transfers**
Age, sex, race, and Hispanic origin	*Assets:*	Disability income payments
Educational enrollment	Interest-earning assets	EITC and tax filing status
Educational attainment	Other income-generating assets	Energy assistance
Family and household relationships	Other assets	General Assistance (GA)
Fertility	Retirement accounts	Life insurance retirement income
Language	*Liabilities:*	Lump sum severance pay/retirement plan income
Marital status and marital history	Debts secured by assets	Miscellaneous income
Nativity, citizenship, and parent nativity	Unsecured debt	'Other' assistance
Parent mortality	*Additional topics:*	Retirement income
Residence	Rent and mortgage payments	School meals
Veteran status	Utility payments	Social Security – child
		Social Security – self
Employment		Special Supplemental Nutrition Program for Women, Infants, and Children (WIC)
Commuting and work schedule	**Health and Well-Being**	Supplemental Nutrition Assistance Program (SNAP)
Earnings	Adult well-being	Supplemental Security Income (SSI)
Hours and pay changes	Child care	Support paid
Job characteristics	Child well-being	Support received
Labor force participation	Dependent care expenses	Survivor income benefit
Parental leave	Disability	Temporary Assistance for Needy Families (TANF)
Reasons for not working	Food security	Unemployment compensation
Time away without pay	Health care utilization and medical expenditures	Veterans benefits
	Health insurance	Workers' compensation

1.5 Survey Organization

This section describes the basic organizational structure of SIPP data collection, including *panels*, *waves, rotation groups,* and *reference periods*.

1.5.1 Panels

A SIPP panel is a group of households selected to be interviewed periodically over multiple years. A panel may run alone, such as the 2014 Panel, or concurrently with other panels, such as the 1990 through 1993 Panels. Since 1984, there have been 17 SIPP panels total. Table 1-2 shows historical SIPP panels and the time periods in which their interviews were conducted.

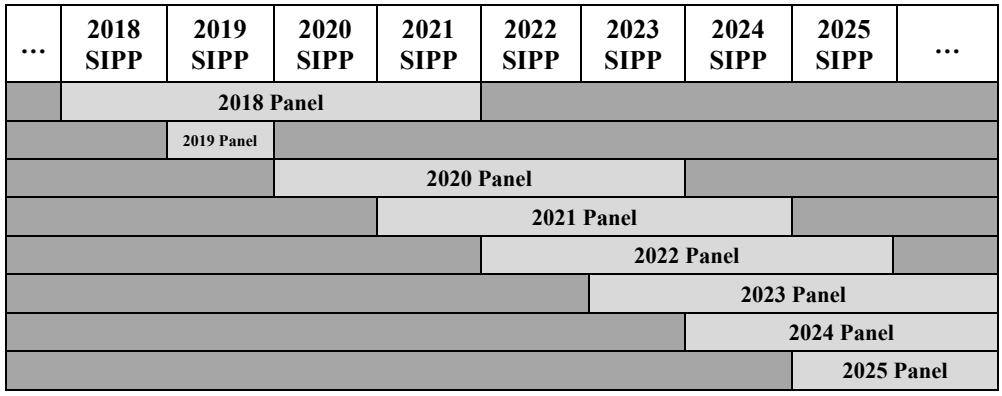
The first SIPP panel, the 1984 Panel, began interviews in October 1983. The 1985 Panel began in February 1985. Subsequent panels began in February of each calendar year, resulting in concurrent administration of the survey in multiple panels. Because of budget constraints, actual panel duration has varied. SIPP's original goal was to have panels of eight interviews covering 32 months. In several instances, due to budget and other issues, panels were terminated after seven interviews (28 months). Three panels were terminated even earlier: 1988 (six interviews), 1989 (three interviews), and 2000 (2 interviews). With certain exceptions, each panel overlapped part of the previous panel, with the result that there were two or three active panels at any given time. Overlapping panels allow data users to pool records from different panels, thus having larger samples (and lower standard errors) for cross-sectional analyses.[2] The overlapping feature of the SIPP design was dropped for the 1996 through 2014 Panels, but standard errors for non-overlapping panels remained small since at least 29,000 households were interviewed (considerably more than in the previous panels).

Similar to the 1990 through 1993 Panels, the 2018 Panels and beyond overlap with new panels that begin each subsequent year (2019, 2020, and so on). Figure 1-1 depicts this design. Although multiple panels are now collected during each annual interview cycle, they are referred to by their calendar year of collection (e.g., all panels collected in 2020 are collectively referred to as "2020 SIPP").

Note that the 2020 SIPP faced unique data collection complications because of the COVID-19 pandemic. Due to government restrictions, mandated stay-at-home orders, and health and safety concerns for both Census interviewers and sampled households, the 2020 SIPP switched to telephone-only interviewing on March 19, 2020. Telephone-only interviewing continued through June 30, the end of the 2020 data collection period. Because SIPP relies on in-person interviews to collect the best data from sampled households, telephone-only interviewing significantly disrupted Field Representatives' ability to reach and follow-up with respondents – especially those in their initial sample year. This led to lower-than-average unit response rates nationally. For technical documentation and more information about SIPP data quality, please visit the SIPP website's Technical Documentation page at <https://www.census.gov/programs-surveys/sipp/tech-documentation.html>.

[2] Combining data across panels also helps alleviate two types of bias common to longitudinal surveys: time-in-sample effects and attrition bias.

FIGURE 1-1. OVERLAPPING PANELS ILLUSTRATION



1.5.2 Waves

Within a panel, a full cycle of administering the questionnaire – i.e., conducting a round of interviews with the selected households – is a *wave*. The first time an interviewer contacts a household, for example, is Wave 1; the second time is Wave 2, and so forth.

Prior to the 2014 Panel, most SIPP panels consisted of between 8 and 16 waves. Because SIPP changed to interviewing annually instead of every 4 months, the 2014+ Panels consist only of 4 waves. There were 4 interviews for the 2014+ Panels instead of the 8 to 16 interviews in previous SIPP panels. As before, each round of interviews results in published public-use data files.

TABLE 1-2. SUMMARY OF THE 1984 TO 2020 SIPP PANELS

Panel	Date of First Interview	Date of Last Interview	Number of Wave 1 Eligible Households	Number of Waves
1984	Oct. 83	Jul. 86	20,897	9
1985	Feb. 85	Aug. 87	14,306	8
1986	Feb. 86	Apr. 88	12,425	7
1987	Feb. 87	May 89	12,527	7
1988	Feb. 88	Jan. 90	12,725	6
1989	Feb. 89	Jan. 90	12,867	3
1990	Feb. 90	Sep. 92	19,800	8
1991	Feb. 91	Sep. 93	15,626	8
1992	Feb. 92	May 95	21,577	10
1993	Feb. 93	Jan. 96	21,823	9
1996	April 96	Mar. 00	40,188	12
2001	Feb. 01	Jan. 04	50,500	9
2004	Feb. 04	Jan. 08	51,379	12

Panel	Date of First Interview	Date of Last Interview	Number of Wave 1 Eligible Households	Number of Waves
2008	Sept. 08	Dec. 13	52,031	16
2014	Feb. 14	Jun. 17	42,348	4
2018	Feb. 18	Ongoing	44,870	Ongoing
2019	Mar. 19	Jul. 19	24,500	1*
2020	Feb. 20	Ongoing	22,241	Ongoing

* Due to data collection issues during the 2019 SIPP, the 2019 Panel Wave 1 cases were de-prioritized during collection in favor of retaining the longitudinal capacity of the 2018 SIPP. The 2019 Panel was not extended into subsequent years, and data users of the 2019 SIPP data are advised to not analyze 2019 Panel Wave 1 cases alone.

1.5.3 Rotation Groups (1984-2008 Panels)

Prior to the 2014 Panel, sample members within each *wave* were divided into four subsamples of roughly equal size; each subsample was referred to as a *rotation group*. Each month, the members of one rotation group were interviewed. Over the course of four months, all rotation groups were interviewed, which provided data for the full set of four months and completed the wave. Figure 1-2 provides an illustration of the various rotation groups in each wave of the 2008 Panel.

FIGURE 1-2. ILLUSTRATION OF ROTATION GROUPS AND REFERENCE PERIODS, 1984-2008 PANELS

Reference Month	Rotation Group 1	2	3	4
May 08	W1 1			
June 08	W1 2	W1 1		
July 08	W1 3	W1 2	W1 1	
Aug. 08	**W1 4**	W1 3	W1 2	W1 1
Sept. 08	W2 1	**W1 4**	W1 3	W1 2
Oct. 08	W2 2	W2 1	**W1 4**	W1 3
Nov. 08	W2 3	W2 2	W2 1	**W1 4**
Dec. 08	**W2 4**	W2 3	W2 2	W2 1
Jan. 09	W3 1	**W2 4**	W2 3	W2 2
Feb. 09	W3 2	W3 1	**W2 4**	W2 3
Mar. 09	W3 3	W3 2	W3 1	**W2 4**
April 09	**W3 4**	W3 3	W3 2	W3 1
May 09	W4 1	**W3 4**	W3 3	W3 2
June 09	W4 2	W4 1	**W3 4**	W3 3
July 09	W4 3	W4 2	W4 1	**W3 4**
Aug. 09	**W4 4**	W4 3	W4 2	W4 1
Sept. 09	W5 1	**W4 4**	W4 3	W4 2
Oct. 09	W5 2	W5 1	**W4 4**	W4 3
Nov. 09	W5 3	W5 2	W5 1	**W4 4**
Dec. 09	**W5 4**	W5 3	W5 2	W5 1
Jan. 10	W6 1	**W5 4**	W5 3	W5 2
Feb. 10	W6 2	W6 1	**W5 4**	W5 3
Mar. 10	W6 3	W6 2	W6 1	**W5 4**
April 10	**W6 4**	W6 3	W6 2	W6 1
May 10	W7 1	**W6 4**	W6 3	W6 2
June 10	W7 2	W7 1	**W6 4**	W6 3
July 10	W7 3	W7 2	W7 1	**W6 4**
Aug. 10	**W7 4**	W7 3	W7 2	W7 1
Sept. 10		**W7 4**	W7 3	W7 2
Oct. 10			**W7 4**	W7 3
Nov. 10				**W7 4**

Reference Month	Rotation Group 1	2	3	4
Sept. 10	W8 1	*See Wave7 data in bottom*		
Oct. 10	W8 2	W8 1	*of first column*	
Nov. 10	W8 3	W8 2	W8 1	
Dec. 10	**W8 4**	W8 3	W8 2	W8 1
Jan. 11	W9 1	**W8 4**	W8 3	W8 2
Feb. 11	W9 2	W9 1	**W8 4**	W8 3
Mar. 11	W9 3	W9 2	W9 1	**W8 4**
April 11	**W9 4**	W9 3	W9 2	W9 1
May 11	W10 1	**W9 4**	W9 3	W9 2
June 11	W10 2	W10 1	**W9 4**	W9 3
July 11	W10 3	W10 2	W10 1	**W9 4**
Aug. 11	**W10 4**	W10 3	W10 2	W10 1
Sept. 11	W11 1	**W10 4**	W10 3	W10 2
Oct. 11	W11 2	W11 1	**W10 4**	W10 3
Nov. 11	W11 3	W11 2	W11 1	**W10 4**
Dec. 11	**W11 4**	W11 3	W11 2	W11 1
Jan. 12	W12 1	**W11 4**	W11 3	W11 2
Feb. 12	W12 2	W12 1	**W11 4**	W11 3
Mar. 12	W12 3	W12 2	W12 1	**W11 4**
April 12	**W12 4**	W12 3	W12 2	W12 1
May 12	W13 1	**W12 4**	W12 3	W12 2
June 12	W13 2	W13 1	**W12 4**	W12 3
July 12	W13 3	W13 2	W13 1	**W12 4**
Aug. 12	**W13 4**	W13 3	W13 2	W13 1
Sept. 12		**W13 4**	W13 3	W13 2
Oct. 12			**W13 4**	W13 3
Nov. 12				**W13 4**

Note: The cell entry *W1* 1 represents Wave 1, reference month 1. The last reference month of each wave is in boldface type. For rotation group 1, the reference months for Wave 1 were May 2008 through Oct. 2008.

1.5.4 Reference Periods

Each interview wave has a corresponding reference period that most of the data represent. For the 2014+ SIPP, the majority of interview questions ask about the preceding calendar year, and therefore these data in general refer to the previous calendar year. Notable exceptions to this are questions about disability status and parent-child relationships, which are asked as of the time of the interview. Prior to the 2014 Panel, the reference period for most interview questions was the preceding 4 months.

Annual interviews still provide key month-level data. For example, SIPP asks each respondent if he or she had a health insurance plan at any time between January 1 of the reference year and the time of interview. If the answer is yes, the event history calendar develops spell- and month-level data for the respondent's coverage over each of the months in the reference period. Over the course of a 4-wave panel, data are collected and released covering 48 consecutive months.

1.5.5 Core Content and Topical Modules (1984-2008 Panels)

Prior to the 2014 Panel, interview questions were split into *core content* and *topical module content*. Core content included questions that were asked at every interview, while topical module content included questions that were part of 'topical modules' and only asked in specific of waves. SIPP core content covered a variety of topics, including labor force status and employment, earnings, business ownership and earnings, asset ownership and income, program participation, child support received and paid, among others. SIPP core content also included small sections about health insurance ownership and coverage (Medicare coverage, Medicaid, private and employer-provided health insurance, and reasons for non-coverage), education (educational attainment and adult school enrollment), and other income and benefit sources.

During an interview, topical module questions generally followed the core questions (although not every wave included a topical module). These modules were designed to gather specific information on a wide variety of subjects. They provided a broader picture of the types of individuals who were responding to the survey and give SIPP some flexibility in collecting data on emerging issues. Table 1-3 shows a thematic grouping of topical modules that were administered before the 2014 redesign, and Appendix C provides a summary of each topical module.

TABLE 1-3. TOPICAL MODULES GROUPED THEMATICALLY

Category	Topical Module
Health, Disability, & Physical Well-Being	Adult Well-Being; Extended Measures of Well-Being; Children's Well-Being; Functional Limitations and Disability (Child and Adult); Health Status and Utilization of Health Care Services; Long-Term Care; Medical Expenses, Work Disability; Work Disability History
Financial	Annual Income and Retirement Accounts; Assets and Liabilities; Real Estate Property and Vehicles; Recipiency History; Retirement Expectations and Pension Plan Coverage; School Enrollment and Financing; Wealth and Eligibility, Selected Financial Assets; Shelter Costs and Energy Usage; Support for Non-household Members; Taxes

Category	Topical Module
Child Care & Financial Support	Child Care; Child Support Agreements; Child Support Paid; Support for Non-household Members
Education & Employment	Education and Training History; Employment History; School Enrollment and Financing; Work-Related Expenses; Work Schedule
Family Characteristics & Living Conditions	Household Extended Measures of Well-Being; Fertility History; Household Relationships; Marital History
Personal History	Education and Training History; Employment History; Fertility History; Marital History; Migration History; Recipiency History; Work Disability History
Welfare Reform	Eligibility for and Recipiency of Public Assistance; Job Search and Training Assistance; Job Subsidies; Transportation Assistance; Health Care; Food Assistance

Some topical modules were included in each panel, but unlike the core content are not in repeat waves. The frequency and timing of these modules also varied. For example, the Personal History topical modules were always administered once per panel, in Waves 1 or 2, while other topical modules were asked multiple times within the same panel; the Assets and Liabilities module, for example, was included three times within the 2001 Panel. In addition, some topical module questions looked like core questions. For example, the core questions asked about asset ownership, specifically whether or not the respondent had stocks or bonds and how much income they generated. The topical module questions on wealth asked about the amount held in stocks and bonds.

The reference periods for items in topical modules varied widely, ranging from the respondent's status at the time of the interview to the respondent's experience over his or her entire life. When working with data from the SIPP topical modules, it is necessary to check each question's wording carefully to ascertain the reference period. Question universes must also be checked for each question because topical modules are not uniformly asked of all respondents. For example, in the Retirement and Pension topical module, respondents 21 years of age or older were asked about their retirement and pension accounts with their current employer, while respondents 25 and older were asked questions relating to lump sums from a previous employer. Questions on shelter costs and energy usage were asked only of the reference person. In other modules, a screening question determined who was and was not asked the remainder of the module. In the Work Schedule module, for example, only those who worked during the previous month answered the entire set of questions.

Note that the relationship between topical module titles and content was not perfectly consistent. Over the history of SIPP, there have been situations in which either the topical module content changed with no change in title, or the topical module title changed with little change in content. In a few situations,

content floated from one topical module to another, and sometimes there was significant overlap in content between two topical modules with different titles.

Table 1-4 lists topical modules along with the panels and waves in which they were administered. For more details about SIPP's topical modules, please visit the SIPP website's *1984-2008 Topical Modules List* page at <www.census.gov/programs-surveys/sipp/tech-documentation/topical-modules.html>.

TABLE 1-4. SIPP TOPICAL MODULES, BY SUBJECT

Subject Areas	Panel and Wave*
Adult Well-Being	91-6, 92-3, 93-9, 96-8, 01-8, 04-5, 08-6, 08-9
Annual Income and Retirement Accounts	84-9, 85-5, 85-8, 86-5, 87-5, 88-5, 90-5, 90-8, 91-5, 91-8,92-5, 92-8, 93-5, 93-8, 96-4, 96-7, 96-10, 01-4, 01-7, 04-04, 04-07, 08-5, 08-8
Assets and Liabilities	84-4, 84-7, 85-3, 85-7, 86-4, 86-7, 87-4, 90-4, 91-7, 92-4,93-7, 96-3, 96-6, 96-9, 96-12, 01-3, 01-6, 01-9, 04-03, 04-06, 08-4, 08-7, 08-10
Basic Needs	93-9
Child Care Poverty - Included in Work Related Expenses/Child	01-3, 01-6, 01-9
Child Support Agreements	85-6, 86-3, 86-6, 87-3, 87-6, 88-3, 88-6, 89-3, 90-3, 90-6, 91-3, 92-6, 92-9, 93-3, 93-6, 93-9, 96-5, 96-11, 01-5, 01-8, 04-05, 08-06
Child Support Paid	96-3, 96-6, 96-9, 96-12, 01-3, 01-6, 01-9, 04-03, 04-06, 08-4, 08-7, 08-10
Child Care	84-5, 85-6, 86-3, 86-6, 87-3, 87-6, 88-3, 88-6, 89-3, 90-3,90-6, 91-3, 92-6, 92-9, 93-3, 93-6, 93-9, 96-4, 96-10, 01-4, 04-04
Children's Well-Being	92-9, 93-6, 93-9, 96-6, 96-11, 01-7, 04-03, 04-8, 08-4, 08-10
Disability Questions	96-4
Disability Status of Children	85-6, 86-3, 87-6, 88-3, 88-6, 89-3 88-6, 89-3
Earnings and Benefits	84-6
Education and Training History	86-2, 87-2, 88-2, 89-2, 90-2, 91-2, 92-2, 93-2, 96-2, 01-2, 04-02, 08-2
Education and Work History	84-3
Education and Training	84-6
Employer Provided Health Benefits	96-5, 01-5, 04-05, 08-6
Employment History	86-2, 87-2, 88-2, 89-2, 90-2, 91-2, 92-1, 93-1, 96-1, 01-1, 04-01, 08-1

Subject Areas	Panel and Wave*
Extended Measure of Well-Being (Consumer Durables, Living Conditions, Basic Needs)	91-6, 92-3, 93-9
Family Background	86-2, 87-2, 88-2
Fertility History	84-8, 85-4, 86-2, 87-2, 88-2, 89-2, 90-2, 91-2, 92-2, 93-2, 96-2, 01-2, 04-2, 08-2
Functional Limitations and Disability	90-3, 90-6, 91-3, 92-6, 93-3
Functional Limitations and Disability-Adults	92-9, 93-6, 96-5, 96-11, 01-5, 01-8, 04-05
Functional Limitations and Disability- Children	92-9, 93-6, 96-5, 96-11, 01-5, 01-8, 04-5, 08-6
Functional Activities	88-6,89-3
Health and Disability	84-3
Health Status and Utilization of Health Care Services	85-6, 86-3, 87-6, 88-3, 88-6, 89-3, 90-3, 90-6, 91-3, 92-6, 92-9, 93-3, 93-6, 96-3, 96-6, 96-9, 96-12, 01-3, 01-6, 01-9, 04-3, 04-6, 08-4, 08-7, 08-10
Home Based Self- Employment/Size of Firm	92-6, 93-3
Home Health Care	88-6, 89-3, 96-7
Household Relationships	84-8, 85-4, 86-2, 87-2, 88-2, 89-2, 90-2, 91-2, 92-2, 93-2, 96-2, 01-2, 04-02, 08-2
Housing Costs, Conditions, and Energy Use	84-4
Informing Care giving	01-07, 04-07, 08-9
Interest Earning Accounts	96-3, 96-6, 96-9, 96-12, 01-3, 01-6, 01-9, 04-03, 04-06
Job Offers	85-6, 86-3
Long-Term Care	85-6, 86-3, 87-6, 88-3, 88-6, 89-3
Marital History	84-8, 85-4, 86-2, 87-2, 88-2, 89-2, 90-2, 91-2, 92-2, 93-2, 96-2, 01-2, 04-2, 08-2
Medical Expenses and Work Disability	87-7, 88-4, 90-7, 91-4, 92-7, 93-4, 93-7, 96- 3, 96-6, 96-9, 96-12, 01-3, 01-6, 01-9, 04-3, 04-6, 08-4, 08-7, 08-10
Medical Expenses/ Utilization of Health Care - Adults and Children	01-3, 01-6, 01-9, 04-03, 04-06
Migration History	84-8, 85-4, 86-2, 87-2, 88-2, 89-2, 90-2, 91-2, 92-2, 93-2, 96-2, 01-2, 04-02, 08-02
Mortgages	96-3, 96-6, 96-9, 96-12, 01-3, 01-6, 01-09, 04-03

Subject Areas	Panel and Wave*
Other Assets	01-9, 04-03, 04-06
Other Financial Investments	96-3, 96-6, 96-9, 96-12, 01-3, 01-6, 01-9, 04-03, 04-09
Pension Plan Coverage	84-7, 86-8
Property Income and Taxes	84-6
Real Estate Property and Vehicles	84-7, 85-3, 85-7, 86-4, 86-7, 87-4, 90-4, 91-7, 92-4, 93-7
Real Estate, Shelter Costs, Dependent Care and Vehicles	87-7, 88-4, 90-7, 91-4, 92-7, 93-4, 93-7, 96-3, 96-6, 96-9, 96-12, 01-3, 01-6, 01-9, 04-03, 04-06, 08-4, 08-7, 08-10
Reasons for not Working/Reservation Wage	84-5
Recipiency History	86-2, 87-2, 88-2, 89-2, 90-2, 91-2, 92-1, 93-1, 96-1, 01-1, 04-01, 08-1
Rental Properties	96-3, 96-6, 96-9, 96-12, 01-3, 01-6, 01-9, 04-03, 04-06
Retirement Expectations and Pension Plan Coverage	84-4, 85-7, 86-4, 86-7, 87-4, 90-4, 91-7, 92-4, 93-9, 96-7, 01-7, 04-07, 08-3, 08-11
School Enrollment and Financing	84-9, 85-5, 85-8, 86-5, 87-5, 88-5, 90-5, 90-8, 91-5, 91-8, 92-5, 92-8, 93-5, 93-8, 96-5, 01-5, 04-05
Selected Financial Assets	87-7, 88-4, 90-7, 91-4, 92-7, 93-4, 96-3, 96-6, 96-9, 96-12, 01-3, 01-6, 01-9, 04-3, 04-6, 08-4, 08-7, 08-10
Shelter Costs/Energy Usage	84-4, 86-6, 87-3
Stocks and Mutual Fund Shares	96-3, 96-6, 96-9, 96-12, 01-3, 01-6, 01-9, 04-03, 04-06
Support for Non-household Members	84-3, 84-5, 84-8, 85-4, 85-6, 86-3, 86-6, 87-3, 87-6, 88-3, 88-6, 90-3, 90-6, 91-3, 92-6, 92-9, 93-3, 93-6, 93-9, 96-5, 01-5, 01-8, 04-05, 08-6
Taxes	84-9, 85-5, 85-8, 86-5, 87-5, 88-5, 90-5, 90-8, 91-5, 91-8, 92-5, 92-8, 93-5, 93-8, 96-4, 96-7, 96-10, 01-4, 01-7, 04-04, 04-07, 08-5, 08-8
Time Spent Outside Work Force	90-6
Value of Business	96-3, 96-6, 96-9, 96-12, 01-3, 01-6, 01-9, 04-03, 04-06
Utilization of Health Care Services-Children	92-9, 93-6, 96-5, 96-12
Utilization of Health Care Services	90-3, 90-6, 91-3, 92-6, 93-3
Utilization of Health Care Services-Adults	92-9, 93-6, 96-5, 96-12

* The number preceding the hyphen indicates the year of the panel, and the number following the hyphen indicates the wave number. Thus, 84-8 denotes that the information was collected in the 1984 Panel, during Wave 8.

Chapter 2

2 Sampling Design and Interview Procedures

SIPP uses a complex sample design rather than a simple random sample to determine which households are interviewed. The 2020 SIPP sampled 57,834 designated Living Quarters (LQs), which yielded 50,352 eligible LQs.[3] Of those eligible LQs, 21,989 households were interviewed between February and May 2020, resulting in 53,332 person interviews.

2.1 Sample Design

The Census Bureau employs a two-stage sample design to select the SIPP sample. The two stages are (1) selection of primary sampling units (PSUs) and (2) selection of address units within sample PSUs. Census Bureau interviewers follow an established procedure to identify sample members within the selected address units. This complex sample design has important implications for the estimation of standard errors. Because the SIPP design is not a simple random sample, the standard errors reported by most off-the-shelf statistical software will underestimate the true standard errors of estimates from SIPP.

The SIPP sample universe is the civilian, noninstitutionalized population of the United States.[4] The sampling universe is based on addresses from multiple sources – chiefly the 2010 Decennial Census – and contains approximately 304.4 million individuals. SIPP samples housing units from the current Master Address File (MAF), which is maintained by the U.S. Census Bureau and is the source of addresses for the American Community Survey, other demographic surveys, and the decennial census. The MAF is updated using the U.S. Postal Service's Delivery Sequence File and various automated, clerical, and field operations.

The 2020 SIPP sample is a multistage, stratified sample of 57,834 LQs from 686 sample areas designed to represent the civilian, noninstitutionalized population of the U.S.

2.1.1 Primary Sampling Units (PSUs)

PSUs comprise one or more contiguous counties. Single counties are used as long as the county has a population of 7,500 or more. When the population threshold is not met, adjacent counties are combined. Larger populated PSUs are identified as self-representing (SR) PSUs, while the remaining PSUs are identified as non-self-representing (NSR). Generally, PSUs with 100,000 or more housing units are classified as SR. SR PSUs are in the SIPP sample with certainty, while the NSR PSUs are stratified and selected with a probability proportionate to their size. During the stratification process, NSR PSUs are grouped according to their similarity on specified poverty measures. Given that SIPP uses a state-based sample design, all strata are formed within state boundaries. During the PSU selection process, two NSR

[3] An LQ may be ineligible for interview for many reasons, the most common being that the LQ is vacant or no longer a residential unit.

[4] Civilian status is determined at time of interview. Those who were non-civilians during the reference year but civilians at time of interview are in the sample universe.

PSUs are selected from each stratum with their probability proportionate to their size in relation to the entire stratum in which they belong. The 2020 SIPP contains 252 SR PSUs and 434 NSR PSUs in the sample.

2.1.2 Selection of Addresses in Sample PSUs

The universe of addresses within each sample PSU is divided into two strata, one with a higher concentration of low-income households and the other with a lower concentration of low-income households. Addresses are sorted by geographic and demographic variables, and a systematic selection of units is taken from each stratum. A higher sampling rate is used in the stratum with the higher concentration of low-income households, thereby resulting in an oversample of low-income households. For the 2020 SIPP, the sampling rate for the low-income stratum is 1.6 times the overall sampling rate in each PSU. This results in a 29 percent increase in the number of cases in and near poverty than without oversampling.

The 2014 Panel included a state expansion sample of 13,800 addresses selected from 16 states: Alabama, Arizona, Arkansas, Georgia, Illinois, Indiana, Kentucky, Louisiana, Michigan, Mississippi, New Mexico, North Carolina, Ohio, Pennsylvania, South Carolina, and Tennessee. The size of the expansion in each state was determined by the amount of sample needed to reach a target coefficient of variation (CV) of 6 percent on the estimate of low income. The four largest states – California, Florida, New York, and Texas - met this requirement without any additional sample.

2.2 Locating Sample Address and Respondents

During Wave 1, the interviewer visits the sampled address, compiles a household roster, and attempts to interview all members of the household who are aged 15 years and older. Information about household members under 15 years of age is collected via a proxy interview. Typically, a household member is someone who sleeps in the household the majority of the time. The SIPP instrument determines whether each person is a household member by asking a series of questions. Table 2-1 provides additional detail about who is, and is not, considered a household member in SIPP.

TABLE 2-1. SIPP HOUSEHOLD MEMBERSHIP DETERMINATION

Question	YES (Is Member of Household)	NO (Not Member of Household)
Person staying at SIPP address at time of interview		
Members of family, visitors, etc. ordinarily sleeps here	Y	
- here temporarily, no living quarters held elsewhere	Y	
- here temporarily, living quarters held elsewhere		N
In Armed Forces, stationed locally and sleeps here	Y	
In Armed Forces, stationed elsewhere and here on leave		N
Student temporarily attending school here, living quarters held elsewhere		N

Question	YES (Is Member of Household)	NO (Not Member of Household)
- married and accompanied by own family	Y	
- student nurse attending school nearby	Y	
Absent person who usually lives at SIPP address		
Inmate in an institutional special place regardless of whether living quarters are being held here		N
Temporarily on vacation, in hospital, and living quarters held	Y	
Absent for work, living quarters held here	Y	
Absent for work, living quarters held here and elsewhere but comes here infrequently		N
Unmarried college student working away from home during break, living quarters held here	Y	
In Armed Forces, stationed elsewhere	Y	
In school elsewhere, living quarters held-not married or with own family	Y	
- married and accompanied by own family		N
- attending school overseas		N
- student nurse living at school		N
Exceptions and doubtful cases		
Person with two residences, sleeps most often in other location		N
Person with two concurrent residences, sleeps here most often	Y	
Citizen of foreign country temporarily in U.S., living on premises of an embassy, ministry, legation, chancellery, or consulate		N
Citizen of foreign country temporarily in U.S.- studying here and no other usual residence in U.S.	Y	
- living and working here and no other usual residence in U.S.	Y	
- visiting or traveling in U.S.		N

While the Census Bureau prefers that all respondents who are present at the time of the interview answer for themselves, SIPP accepts proxy interviews from another household respondent when necessary. Within each household, the instrument identifies a reference person, typically the first person listed as the owner or renter of the housing unit.

For Waves 2+ cases, SIPP switches from a household survey to a person-based survey. All people who were interviewed in the first wave of the panel and any children subsequently born to or adopted by them are designated as original sample members.

When visiting the original household, the interviewer determines whether one or more original sample members are present. If so, the interviewer updates the household roster, listing all people living or

staying there, including anyone who may have joined the household, such as a new spouse or baby, and the dates they entered the household.

For those remaining at the same address, the interviewer verifies that certain previously collected information still applies, completes the questionnaire for each person aged 15 years and older, and collects certain information for children under age 15. Information is likewise collected for all new household members. They are interviewed each wave so long as they live with an original sample member. Also noted are people who left the household and their dates of departure, along with their new address (if known). Figure 2-1 illustrates a few examples of SIPP's following rules.

2.3 Following Movers

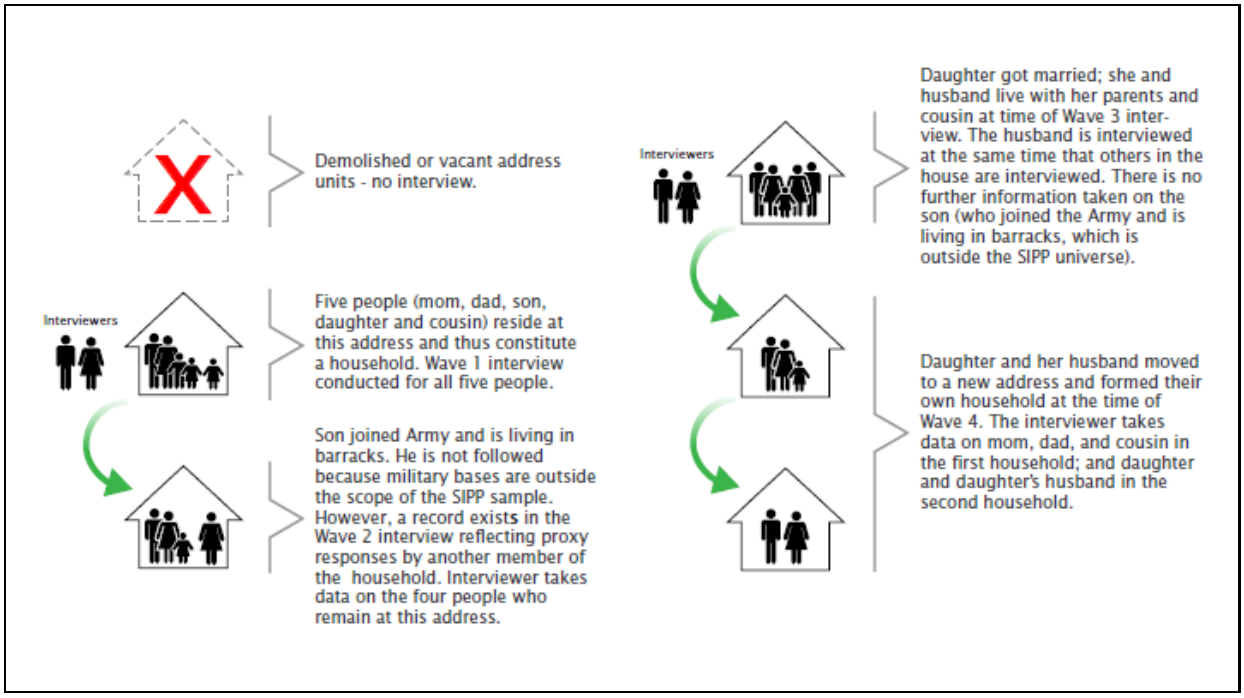
SIPP is a true longitudinal survey that tracks people over time. With few exceptions, original sample members are interviewed annually over the duration of the panel. When original sample members move to new addresses, interviewers attempt to locate and interview them every year.

Movers are interviewed at their new addresses, along with other household members who are living or staying there. When original sample members move into households with other individuals not previously in the survey, the new individuals become part of the SIPP sample for as long as they continue to live with an original sample member. Similarly, when new individuals move in with original sample people after the first interview, they too become part of the SIPP sample for as long as they continue to live with an original sample member. If no original sample members live at an address where a previous interview was conducted, SIPP does not collect information from the new occupants of that address.

If an entire household moves, the interviewer tries to find the original sample people and interview them at their new address(es). Should the entire household move more than 100 miles away from a SIPP PSU, Field Representatives (FRs) attempt to conduct the interview by telephone. If the household cannot be reached, the sample members are dropped from the survey.

SIPP does not interview original sample members if they move outside the United States, become members of the military living in barracks, or become institutionalized (e.g., nursing home residents, prison inmates). The Census Bureau attempts to track such individuals so they can be interviewed should they return to the noninstitutionalized resident U.S. population.

FIGURE 2-1. EXAMPLES OF SIPP FOLLOWING RULES



2.4 Difference Between Movers and Those Who Are Temporarily Away

There is an important difference between a mover and a person who is temporarily away. A mover no longer lives at the sample address. On the other hand, a person is temporarily away if the household is that person's usual place of residence, according to the membership rules given in Table 2-1, and specific living quarters are held for the person to which he or she is free to return at any time. The following two examples illustrate the distinction:

- A college student living on campus with a room held at home is still a household member at the sample address. In this case, the interviewer would try to interview that student or obtain a proxy interview with the household reference person. If the hypothetical college student originally lived in New York and, upon graduation, moved to Los Angeles to live on his or her own, the student would be considered to have moved as of the graduation date. The student's new address in Los Angeles would become his or her new household, and, if the student was an original sample member, he or she would be treated in the same way as any other original sample member who moved to the new address.
- If a household member is in the hospital following an operation but is expected to come home, that person is still a household member at the original address. If an individual interview is not feasible, the interviewer might do a proxy interview for that person. If, however, the person moved into a nursing home, he or she would not be eligible for a SIPP interview, whether individual or proxy. At each interview, the interviewer asks the status of any

primary sample member who entered an institution between Wave 1 and the current wave. If the interviewer learns that the person has returned to the noninstitutionalized population, an interview is attempted.

2.5 Type 2 Persons

One of the challenges of the SIPP is capturing data on individuals who were part of the household at some point in the reference year but left the household before the interview. Data on these individuals, referred to as "Type 2" people, are important to understand how household composition and income change over the year. The amount of information collected about Type 2 people varies by topic and is discussed more in Section 3.5.5, and within topics in Chapter 4. Note that a Type 2 can have been at any address a respondent lived at during the reference year, not just the interview address.

2.6 Interview Sequence

The interview sequence is listed in Figure 2-2. At the beginning of the instrument the household respondent is asked about the characteristics of the housing unit and the household roster is created. Next, the household respondent provides demographic information for all household members and provides information about household relationships.

At this point, personal interviews begin. Respondents provide a list of all Type 2 people, in addition to some basic demographic information about these people. Next, interviewers ask for additional demographic information about respondents, followed by a series of screener questions in an attempt to minimize respondent burden. Based on answers to these screener questions, the interviewer may skip over questions for certain topics. Section 3.4.1 provides more detail about income screeners.

The next part of the instrument is the Event History Calendar (EHC), which was developed to facilitate respondent recall over the longer reference period and allows for integrated reporting across the domains of residency, marital history, educational enrollment, employment/time not working, program participation (SSI, SNAP, TANF, WIC, and GA), and health insurance coverage. Data in the EHC are entered as spells and detailed information about each spell is then collected. Unlike the other sections of the instrument, the sequence within the EHC is not fixed. Interviewers may go through these topics in any order.

Once the EHC is complete, respondents go on to complete the rest of the interview (post-EHC). This includes questions about additional program participation, assets, health, and well-being.

The interview concludes with the interviewer getting contact information for each household member. This information can make it easier to find household members for future waves of interviews. Interviewers may pause an interview before it is complete and return to it at a later date. They can switch from one person's interview to another, or they can exit the interview completely. There is a section in the instrument that allows interviewers to set up an interview for another day or time.

FIGURE 2-2. THE 2020 SIPP INTERVIEW SEQUENCE

Household roster and demographics

- Address verification
- Sample address characteristics and coverage
- Roster creation
- Demographics (including date of birth, sex, race and ethnicity, relationships, educational attainment, and veteran status)

Start of personal interviews

- Type 2 Roster/Info
- Additional demographics (including certification and marital history)
- Income and program screener
- Event History Calendar
  - Residence
  - Marriage and cohabitation
  - School enrollment
  - Labor force (including jobs, businesses, contingent work, unemployment, time not in labor force, commuting, and work schedule)
  - Programs (GA, SNAP, SSI, TANF, WIC)
  - Health insurance (private, Medicare, Medicaid, military, other health insurance coverage, and coverage under ACA)
- Health insurance follow-up
- Dependent care expenses
- Social security
- Workers compensation
- Unemployment
- School meals
- Severance pay or lump sum payments
- Disability
- Retirement income
- Life insurance
- Child support paid
- Other support paid
- Tax returns and EITC
- Support received
- Assets owned
- Vehicles
- Primary residence
- Utility payments
- Debt
- Health (health status, utilization, medical expenditures, disability)
- Fertility (including multiple partner fertility)
- Parental leave
- Parent mortality and nativity

- Child care - Child well-being - Housing conditions - Neighborhood safety - Ability to meet essential expenses - Food security
Interview wrap-up - Follow-up visit/call info - Interview closeout
Contact History Instrument

FRs complete a Contact History Instrument (CHI) following each contact attempt. CHI is a standalone instrument and is not part of the SIPP. CHI is completed for both personal visit and telephone interviews. Using the CHI, FRs record contact attempts, outcomes of contact attempts, judgements of reluctance encountered from households, any strategies used with households, and more. The Neighborhood Observation Instrument (NOI) is part of the CHI. NOI is a set of observation questions that are completed during the first personal visit attempt and are not asked again. These questions ask about the condition of the neighborhood and ask for judgements of the FR on things like household age and presence of children.

2.6.1 Income Screener Questions

The SIPP instrument includes functionality to help reduce respondent burden. One example is with the income screener questions. This set of questions, which are asked in the pre-EHC portion of the instrument, limit the universe for some means-tested programs. When certain parameters are met, the instrument skips the SNAP, TANF, GA, and WIC questions in the EHC.

The income screener questions ask about:

- Annual income below a specified amount
- Monthly income below a specified amount
- A “just to be sure we didn’t miss anything” catch-all for programs

The income amounts are based on 200 percent of the poverty threshold for the number of people residing in the sample unit. Annual amounts are rounded to the nearest $1,000, while monthly amounts are rounded to the nearest $500. This conservative approach ensures that resource sharing between household members is captured. For example, in an unmarried partner household with one child, the father may earn more than 200 percent of the poverty threshold for an individual but not more than 200 percent of the poverty threshold for three people. In other words, if the household as a whole or one or more members could have been eligible for means-tested programs based on income, the intention is to ask about receipt of means-tested programs.

If respondents report having annual income above the specified threshold, they move on to the monthly question. If they then report having monthly income above the specified threshold, they move on to the

catchall program question ensuring they did not receive some type of means-tested assistance. If respondents report having income below a specified threshold or having some type of means-tested assistance, they remain in universe for the SNAP, TANF, GA, and WIC questions. Otherwise, if respondents satisfy all three program screener questions, they are not asked about receipt of these programs.

The income screener questions were designed to reduce respondent burden in the instrument. They are not included on the public use data file.

2.6.2 Grouping Families (Clumps)

In another effort to reduce respondent burden, SIPP groups certain related people into "clumps" whereby the household respondent's answer applies to all members of the clump. Thus, in the subsequent interviews of fellow "clump" members, certain questions will not be asked.

The clump is an internally derived instrument variable, and not part of SIPP's public-use data. When certain parameters are met, the SIPP instrument limits the number of people who are in universe for the income screener questions and the family- or household-based EHC screener questions (SNAP, TANF, and GA). The intention is to ask only one respondent the family- or household-based program questions when resources were shared during the entire reference period.

Who is included in the screener clump?

- Household respondent
- Spouse of the household respondent
- Children of the household respondent under age 22 who did not have dependents living with them

The clump always includes the household respondent, who is the first person interviewed. If the household respondent resided with a spouse during the entire reference period, then the spouse is included in the clump. Similarly, if the household respondent resided with their children under age 22 who did not have dependents of their own living with them during the entire reference period, then the children are included in the clump.

Who is not included in the screener clump?

- Non-nuclear families
- Parents of household respondents between the ages of 15 and 21

The clump intentionally excludes multigenerational households, unmarried partner households, and other types of non-nuclear family forms because of their complexity. This conservative approach allows for the income dynamics across different family forms to be captured. For example, in a multigenerational household, a daughter and granddaughter might receive TANF, while a father and mother did not. Additionally, the screener clump intentionally excludes the parents of a household reference person between the ages of 15 and 21. This conservative approach ensures the collection of the most accurate information from all adults in the household.

The clump is derived by the instrument to reduce respondent burden. This information is not included on the public use data file.

2.7 Feedback/Dependent Interviewing

Dependent interviewing is the process in which information from a previous interview carries forward into the current survey instrument to streamline the interviewing process and maximize data quality. In an effort to both improve data quality and reduce respondent burden, the SIPP instrument features the use of dependent data for Waves 2+ cases. Nearly 500 items are carried forward from wave to wave, allowing respondents to verify information for these items and reduce interview length.

Dependent data also mitigates the negative effects of seam bias. Seam bias is a common ailment of longitudinal surveys, in which event changes are reported disproportionately at the "seam" between waves. The key to alleviating seam bias is to create overlapping periods in which one wave's interview period includes a portion of the next year's reference period. Through dependent interviewing, the instrument already possesses data for the early part of the next wave's reference period, since each interview covers the previous calendar year plus the months leading up to the interview in the current calendar year. With this data from the previous wave, the instrument can tailor question wording to remind respondents of their situation during the previous wave. Therefore, the recall window shrinks, and respondents are less apt to report changes at the transition between two reference periods. When SIPP introduced increased use of dependent interviewing in the 2004 Panel, every area of the survey experienced a decrease in seam bias, with many topics cutting its effect in half (Moore, 2008).

Utilizing dependent data in a household survey, particularly one allowing for proxy respondents, creates the possibility of sharing information within a household that a respondent would rather keep private. To protect respondent confidentiality within a household, the Census Bureau instituted the Respondent Identification Policy (RIP) in 1998. RIP mandates that Census demographic surveys gain respondent consent to reference any previously collected data. Therefore, there is a small subsample of respondents for whom dependent data is not used.

2.8 Nonresponse

The Census Bureau distinguishes between household and person nonresponse. Household nonresponse occurs either when the interviewer cannot locate the household or when the interviewer locates the household but is unable to interview any adult household members. Person nonresponse occurs when the FR interviews at least one person in the household but cannot interview at least one other household member. This usually occurs because that person refuses to answer the questions or is unavailable and a proxy cannot provide the responses.

SIPP follows all original sample members from one wave to the next, with the exception of an original sample member under the age of 15 who moves out without an accompanying adult original sample member. Those original sample members who leave the original household and form new households are retained, and members of the new household are added to the SIPP sample. This expansion of original households can be estimated within the interviewed sample but is impossible to determine within the non-interviewed sample. Therefore, a growth factor based on the growth in the known sample is used to estimate the unknown expansion of the non-interviewed households. Growth factors account for the additional nonresponse stemming from the expansion of non-interviewed households. They are used to compute a more accurate estimate of the weighted number of non-interviewed household units (HUs) at

each wave. The Census Bureau categorizes household nonresponse as Types A, B, C and D, and person-level nonresponse as Type Z. Each type is discussed in more detail below.

2.8.1 Type A Household Nonresponse

Type A household nonresponse occurs when the interviewer finds an eligible household address but cannot obtain interviews. Examples of Type A nonresponse include:

- The interviewer finds no one at home despite repeated visits.
- All eligible household members are away during the entire interview period (e.g., an extended vacation).
- Household members refuse to participate in the survey.
- Interviews cannot be conducted because of language barrier.

When this type of household nonresponse occurs in Wave 1, SIPP does not attempt to interview the household members at subsequent waves. For Type A nonresponse that occurs in Waves 2+, however, interviewers try to obtain interviews in the following waves.

2.8.2 Type B Household Nonresponse

Type B non-interviews occur when a sample unit is currently vacant, unfit for residence, or occupied by people who are ineligible to be interviewed. Examples of Type B non-interviews include:

- The sample unit is under construction and is not ready to be occupied.
- The sample unit is to be demolished (due to fire, deterioration, etc.).
- All household members age 15+ are on active duty in the Armed Forces.

When this type of non-interview occurs in Wave 1, SIPP does not attempt to interview the household in subsequent waves. For Type B non-interviews that occur in Waves 2+, SIPP will attempt to interview household members in following waves if the entire household was in the Armed Forces, institutionalized, or designated as a "Type B – Other" household. Type B non-interviews do not affect the response rate.

2.8.3 Type C Household Nonresponse

Type C non-interviews occur when a sample unit does not exist or is permanently unable to be occupied, or when there are no eligible household members remaining in a Wave 2+ interview. Examples of Type C non-interviews include:

- The sample unit address includes a unit designation that is not in use.
- The sample unit was converted to a permanent business.
- All household members are deceased (Waves 2+).
- All household members moved (Waves 2+).

SIPP does not attempt to interview Type C non-interviews in subsequent waves. Similar to Type B non-interviews, Type C non-interviews do not affect the response rate.

2.8.4 Type D Household Nonresponse

Type D household nonresponse occurs when original sample members move to an unknown address or an address where the respondent is unable to be interviewed. Type D nonresponse applies only to Wave 2 and beyond.

2.8.5 Type Z Person Nonresponse

Type Z person-level nonresponse occurs when a sample person is part of the household on the date of the interview but refuses to answer or is not available for the interview and a proxy interview is not obtained. While household nonresponse is usually handled by weighting adjustments, Type Z cases are handled by imputation. (See discussion of imputation and weighting in Chapters 6 and 7, respectively.)

2.8.6 Item Nonresponse

Item nonresponse is an additional source of missing data; it occurs when respondents do not answer one or more survey questions, even though they complete most of the questionnaire. Respondents might refuse to answer a particular question or set of questions, or not have the information requested. Although interviewers are trained to attempt to persuade respondents to answer all applicable questions and will call back if a respondent can provide data later, those efforts are not always successful. Item nonresponse can also result from the post-interview data editing process when respondents provide inconsistent information or when an interviewer incorrectly records a response. In many cases, the Census Bureau handles item nonresponse by imputation, that is, by assigning values for the missing items. Chapter 6 provides more information about how SIPP handles missing data.

Chapter 3

3 SIPP Public-Use Data Files

This chapter provides an introduction to the SIPP public-use data files. For more information about these data files, please visit the *Survey of Income and Program Participation Datasets* webpage at <www.census.gov/programs-surveys/sipp/data/datasets.html>. Because variables in the data do not replicate the survey instrument, data users should keep in mind the following:

1. The variables on the data files do not correspond one-to-one with the questionnaire items - the variables are listed in a different order, some variables are not included in the core wave files at all, and some variables are created from a combination of other variables
2. The range of possible values of the variables on the data files does not always correspond one-to-one with the response categories shown on the survey instrument or in the data dictionary

3.1 SIPP Data Files

The Census Bureau has published the following file types over time, as discussed in Sections 3.1.1 through 3.1.4:

Primary Data Files (2014+): Person-month-level data file that includes all of the survey's topical content. Includes up to 12 records for every person. Household-level variables are copied across all members of a household; person-level variables are copied across all months for a person; spell-level variables are copied across all months of a spell; month-level variables vary month-to-month.

Replicate Weight Files: Person-month-level replicate weights for records on the corresponding primary or core data file. The 2014+ data have 240 replicate weights. The 2004 and 2008 Panels have 120 replicate weights. The 1996 and 2001 Panels have 108 replicate weights.

Longitudinal Weight Files: Weights for the cohort of respondents for whom data were collected every month of the prescribed longitudinal period (two-year, three-year, and four-year periods) regardless of the panel they originate from. These cohorts also include those who left the target population through death or because they moved to an ineligible address (institution, foreign living quarters, military barracks), as well as those for whom data were imputed for missing months. The Census Bureau makes nonresponse adjustments to the longitudinal weights to compensate for attrition and post stratification adjustments to make the weighted sample totals conform to population totals for key variables.

Longitudinal Replicate Weight Files: Replicate weights for those present on the Longitudinal Weight files.

Core Data Files (1984-2008 Panels): Data file for the survey's 'core' content – content that was asked in every wave of a panel. Prior to the 1990 Panel, core wave files had one record per person. Each record contained four occurrences of each monthly variable. Beginning with the 1990 Panel, the core wave files were published in person-month format, with one record per person for each month of the 4-month

reference period a person is in the sample. Includes up to 4 records for every person. Household-level variables are copied across all members of a household; person-level variables are copied across all months for a person; month-level variables vary month-to-month.

Topical Module Data Files (1984-2008 Panels): Person-level data file for the survey's 'topical module' content – content that was asked during specific waves of the survey. Topical module files for the 1996 - 2008 Panels contain one record for each person who was in the sample with a completed (or imputed) interview in the fourth month of the wave's reference period (the month immediately prior to the interview). This arrangement is similar to the person-month format of the core wave files, but only records for month four are included in the topical module files. Prior to the 1996 Panel, the topical module files contained one record for each person who was interviewed or for whom an interview was attempted in that wave. In general, each topical module file contains data for all of the topical module subject areas administered during a particular wave. Each topical module file also contains selected information from the SIPP core; thus, for some analyses, those files can be used independently from the core wave and full panel data files. When more detailed information from the SIPP core file is needed, data from the topical modules must be merged with data from the core wave or full panel files. For more information on using the topical module files, please see the 2008 version of this guide at <www.census.gov/programs-surveys/sipp/guidance/users-guide.html>.

Full Panel Longitudinal Research File (1984-1993 Panels): A rectangular (flat) file with one record for each person who, at any time during the life of a panel, was a member of an interviewed household.

Table 3-1 describes the general variable naming conventions for variables within the 1996+ SIPP data files. Generally speaking, pre-1996 data files did not have specific variable prefixes.

TABLE 3-1. VARIABLE NAMING CONVENTIONS, 1996+ SIPP FILES

Variable Prefix	Description
E	Basic edited variable
A	Universe and imputation status flag for edited and recoded variables
R	Recoded variable
T	Topcoded or bottomcoded variable
G	Sample or geography identifier
S	Record ID variable
W	Weight variable

3.1.1 2014 Panel Through 2020 SIPP

The Census Bureau publishes a **Primary Data File**, a **Replicate Weights File, Longitudinal Weights Files** (where applicable)**, and Longitudinal Replicate Weights Files** (where applicable) for every

calendar year of data collection. These files are issued in person-month format, therefore there are up to 12 records for each respondent on the file.[5]

3.1.2 2014 SSA Supplement

The Census Bureau also conducted a supplementary interview of special questions on behalf of the Social Security Administration in October of 2014. This supplementary interview is known as the "SSA Supplement." A primary data file and replicate weight file are also available for this interview. For more information about SSA Supplement data, please visit <www.census.gov/programs-surveys/sipp/data/datasets/2014-panel/ssa.html>.

3.1.3 1996 Through 2008 Panels

The Census Bureau published **Core Data Files**, **Topical Module Data Files**, and **Replicate Weight Files** for each wave of the 1996 through 2008 Panels. The core data files are in person-month format and contain up to four records for each primary sample member and each person who lived with a primary sample member at any time during the 4-month reference period covered by the wave. Each record corresponds to one of the four reference months covered by the wave.

Topical module files for each wave of these panels contain one record for each person who was a sample responding (or Type Z nonresponding) member of a SIPP household during the fourth month of the reference period for the wave. The specific topical content included in each wave is discussed in Section 1.5.5. There are no topical module files after the 2008 Panel because most topical module content was rolled into the annual interview with the 2014 redesign, and therefore all response data reside on a single file.

Each wave of the 1996 through 2008 Panels includes **Longitudinal Weight Files** and **Longitudinal Replicate Weight Files**. Both files include longitudinal *panel* weights as well as longitudinal *calendar year* weights.

3.1.4 1984 Through 1993 Panels

Similar to the 1996 through 2008 Panels, the 1984 through 1993 Panels have **Core Data Files** and **Topical Module Data Files** for each wave. Prior to the 1990 Panel, however, core data files were issued with a single record for each person. Each record contained data for all four of the reference months covered by the wave. Earlier editions of this Users' Guide provide details about these files.

The Census Bureau also published **Full Panel Longitudinal Research Files** for these panels. Full panel files were released after interviewing for a panel was completed, and contain one record for each original sample member, all children, and all adults who entered the sample after Wave 1. People who were not interviewed for one or more months over the course of the panel either have their data imputed or are

---

[5] In some circumstances, respondents will not have 12 person-month records on the file. Infants who are born during the reference year will have person-month records starting at the month of birth through the end of the reference year. In Waves 2+, respondents may leave the sample if they no longer meet the survey frame universe (e.g., if they die or become institutionalized) or if they are unable to be located for a follow-up interview. In these cases, respondents will have person-month records from the beginning of the reference year until the months they left the survey sample.

identified as not in the sample, although their records remain in the file. Variables within each record correspond to the information that was collected in the core content sections of the interviews. Different variables occur with different frequency, depending upon how often certain questions were asked. For example, because a sample member's sex, date of birth, and race are unlikely to change, the variables corresponding to those attributes occur only once in each record. On the other hand, some questions from the core content, such as those about income and program participation, are asked for each month of the panel; the number of corresponding variables will reflect that fact. Similarly, SIPP-generated information can occur once (e.g., person number) or many times (e.g., monthly interview status) on each record.

3.2 Confidentiality Procedures for the Public-Use Files

Before the files are released for public use, they undergo additional editing to protect the confidentiality of respondents. Among the procedures used are the topcoding of selected variables (e.g., income, assets, and age) and the suppression of geographic information. Because of these procedures, estimates based on data from the public-use files will differ slightly from the Census Bureau's published estimates.

3.2.1 Topcoding

One piece of information that might reveal a respondent's identity is very high income or assets. For that reason, the Census Bureau topcodes these variables and other continuous variables before making their information publicly available. This is done in different ways for different variables. For incomes and amounts with a skewed distribution, such as annual salary or home values, the amount is replaced by an average of topcoded values.[6] For other income variables, such as program income for which a threshold is determined by institutional rules, amounts over a certain maximum value are recoded to that maximum. In other words, income on the public-use data files has a ceiling value for these variables. As in the past, summary income variables for persons, families, and households are the sum of the component variables after they have been topcoded. Although income is the primary variable that is topcoded, other variables that may disclose a respondent's identity, such as age, are also topcoded. Others, such as occupation, may be aggregated into broader categories. A few variables, such as starting dates for employment and variables with negative values may be bottomcoded if they pose a disclosure risk.[7]

3.2.2 Suppression of Geographic Information

Geographic information that can be used to directly identify survey respondents, such as an address, is removed from the public-use files. In addition, individual metropolitan areas and specific non-

---

[6] For hourly wages, the median above the threshold is output on the data file instead of the mean. For most variables using this methodology, the mean output is the within-cell average based on person-characteristics. More information about these characteristics and the distributions above the topcoding threshold, such as mean, median and variance, is available. For specific variables and waves, see the SIPP webpage user notes at <www.census.gov/programs-surveys/sipp/tech-documentation/user-notes/2018-usernotes.html>.

[7] In the case of bottomcoded earnings and asset variables, the set of outliers is replaced with the threshold value to minimize the difference between the SIPP's internal and public-use poverty rates. The mean and median values below the bottomcode are also provided to users wishing to implement a different procedure. For a discussion of specific variables, see the SIPP webpage user notes at www.census.gov/programs-surveys/sipp/tech-documentation/user-notes.html.

metropolitan areas (such as counties outside of metropolitan areas) are never identified. SIPP does identify metropolitan and nonmetropolitan status for respondents in states where (1) both the metropolitan and non-metropolitan populations are over 250,000 or (2) the state's population is entirely metropolitan. Respondents living in states that do not fit these criteria are coded as non-identified on metropolitan status.

3.3 Status Flags

Nearly all variables on the SIPP file have a status flag (previously referred to as "allocation flag" or "imputation flag" in 2008 and earlier Panels). As in previous panels, these variables continue to start with the letter 'A.' For example, the status flag for the ERACE variable is ARACE. For the 2020 SIPP data, the status flags contain 10 possible values. A status flag with a value of 0 indicates that the item is "Not in universe," which means that item did not meet the criteria to get data. Items where the data are as reported by the respondent have values of 1 for the status flag. Status flags values between 2 and 6 indicate the item underwent a type of imputation. Status flag values of 7 and 8 are most common among recoded variables and indicate that an item underwent a combination of imputation methods either with reported data (values of 7) or without (values of 8). Finally, some status flags have a value of 9 meaning that data users need to look at the status flags of the individual components of this recode.

TABLE 3-2. STATUS FLAG VALUES

Status Flag Value	Meaning	Description
0	Not in universe	The item is not imputed, and no data was reported by the respondents.
1	In universe, as reported	The item is not imputed, and data was reported by the respondents.
2	Statistical imputation (hot-deck)	The item is imputed a value from hot-deck imputation.
3	Logical imputation	The item is imputed logically.
4	Model-based imputation	The item has gone through a model-based imputation procedure.
5	Cold-deck value imputation	The item is imputed to a cold-deck value. This cold-deck value is what each cell in the hot-deck matrix gets initialized to. It is the most common reported value a respondent would be likely to report for that particular cell. So, if an item could not get a hot-deck value, this initialized value is used to set the value for that item.
6	Imputed from a range	The item is imputed based on a range.
7	Combination of 1 and 2/3/5/6	The item is imputed using reported value and any combination of imputation method values 2, 3, 5, and 6 (described above).
8	Combination of 2/3/5/6	The item is imputed using any combination of imputation method value 2, 3, 5, and 6 (described above).

Status Flag Value	Meaning	Description
9	Can be determined from the status flags for the components of this recode	Used for recodes that are entirely made up of variables on the public-use file.

3.4 Understanding the Identification (ID) Variables in SIPP

Because different files contain different information, the capacity to identify people across those files is important. SIPP is a longitudinal survey designed to allow researchers to track people over time; other critical functions include identifying individuals over time and identifying when a person is present in the sample. Finally, because the relationships among people change over time, identification of those relationships at any specific time is important. The key to these tasks lies in understanding how SIPP ID variables are used to identify persons, families, and households.

Table 3-3 describes SIPP's highest-level ID variables. Sections 3.4.1 through 3.4.4 further detail these variables. Additional variables are used to identify people who are members of related subfamilies, unrelated subfamilies (also known as *secondary families*), and transfer program units such as food stamp units. Sections 3.5 and 3.7 discuss these other variables. For more information about 1984-2008 Panel ID variables, please see previous versions of this guide at <www.census.gov/programs-surveys/sipp/guidance/users-guide.html>.

TABLE 3-3. SIPP ID VARIABLES

File Type	Sample Unit ID	Current Address ID	Person Number	Month of record ID
Panels Prior to the 1990 Panel				
Core Wave Person-Month Files	SUID	ADDID	PNUM	N/A
Topical Module Files	ID	ADDID	PNUM	N/A
Full Panel (and Partial-Panel) Longitudinal Research Files	PP-ID	HH-ADDID	PP-PNUM	N/A
1990 - 1996 Panels				
Core Wave Person-Month Files	SUID	ADDID	PNUM	MONTH
Topical Module Files	ID	ADDID	PNUM	N/A
Full Panel (and Partial-Panel) Longitudinal Research Files	PP-ID	HH-ADDID	PP-PNUM	N/A
1996 - 2008 Panels				
Core Wave Person-Month Files	SSUID	SHHADID	EPPPNUM	RHCALMN

File Type	Sample Unit ID	Current Address ID	Person Number	Month of record ID
Topical Module Files	SSUID	SHHADID	EPPPNUM	N/A
2014 Panel Through 2020 SIPP				
Primary Person-Month Data File	SSUID	ERESIDENCEID	PNUM	MONTHCODE

3.4.1 Sample Unit IDs

During the initial interview (Wave 1), each sampled dwelling unit is assigned an SSUID identifier, a unique (random) sample-unit ID (SUID). The same SSUID is assigned to each person within the Wave 1 household and never changes throughout the panel. If an original sample member moves to a different address, the person keeps the same SSUID. If new people join an original sample member at either the original or a new address, they are assigned the same SSUID as the original member and become secondary sample members. At the conclusion of the panel, all people who have ever resided with an original sample member share the same SSUID. This SSUID is their common link to the original sample unit.

For the 1996 Panel through 2020 SIPP, the variable SSUID is the sample-unit identifier. Prior to the 1996 Panel, the sample-unit identifier is one of PP-ID, ID, or SUID, depending on the file type.

Between the 2014 Panel and the 2018 SIPP, SSUID increased from 12 to 14 digits. The additional two digits reflect the SIPP Panel that the sample unit joined the SIPP sample. For example, all sample units that are part of the 2018 SIPP Panel have '18' as the last two digits in their SSUID, whereas those that are part of the 2020 Panel have '20' as the last two digits in their SSUID.

3.4.2 Person Numbers

All sample members are assigned a person number when they first enter the SIPP panel. The person number differentiates persons within the sample unit and does not change throughout the panel. The first digit indicates the wave in which a person joined the SIPP sample, and the next two digits correspond to the person's place on the household roster.

For the 2014 Panel through 2020 SIPP, the variable PNUM is the person-number identifier. For the 1996 through 2008 Panels, it is EPPNUM. Prior to the 1996 Panel, the person-number identifier is either PNUM or PP-PNUM, depending on the file type.

3.4.3 Address IDs

2014 Panel Through 2020 SIPP: Residence ID (ERESIDENCEID)

Residence ID (ERESIDENCEID) is a unique six-digit identifier for addresses within original sample units (SSUIDs). While the sample unit identifier (SSUID) and person number (PNUM) values remain the same for the entire panel, movers both within and across waves are identified by a change in the ERESIDENCEID field. The first digit of ERESIDENCEID identifies the first wave in which the address appears in the data.

Addresses entered in Wave 1 start with 1, those entered in Wave 2 start with 2, and so on. The second and third digits, which previously identified spinoff households in the 2014 Panel, will always be '00' for ERESIDENCEID values in the 2018 through 2020 SIPP data. Spinoff households are new households formed from at least one member of the originally sampled household. The last three digits of ERESIDENCEID represent the unique numbering of addresses that are associated with an original sample unit (SSUID) and that enter the sample in the same wave.

While the Wave 1 interview addresses are usually assigned values of 100001, this does not have to be the case. There are a small number of cases with values other than 100001 for Wave 1 interview addresses.

1984-2008 Panels: Current Address ID (SHHADID, ADDID) and Entry Address ID (EENTAID, ENTRY)

The Current Address ID (SHHADID, ADDID) identifies each housing unit occupied by one or more original sample members in any given month. Current address IDs are assigned within sample units (they are unique only when combined with the sample unit ID variable), and they have two parts. The first part (one digit for all but the 1992 and 1996 Panels; two digits for the 1992 and 1996 Panels) identifies the wave in which one or more original sample members were first scheduled to be interviewed at the address. The second part of the ID is one digit, and it is used to sequentially number addresses for households that split into two or more households as a result of a move to a different location by original sample persons. All Wave 1 households have a current address ID of 11. Any new addresses that are occupied in Wave 2 are numbered 21, 22, and so on; new addresses occupied during the Wave 3 reference period are numbered 31, 32, 33, and so on. The current address ID is a monthly variable, the value of which changes in the month in which an individual moves to a new address.

The Entry Address ID (EENTAID, ENTRY) is the current address ID that a sample member occupied when he or she first entered the SIPP sample. It is used in conjunction with the person number to uniquely identify persons within the sample unit and does not change even if the person moves.

3.4.4 Month of Record IDs

As previously discussed, many SIPP files are published in person-month format – each record (row) refers to a specific reference month. As a result, a single person is described by multiple monthly records. For the 2014 Panel through 2020 SIPP, the variable MONTHCODE identifies a record's calendar reference month. For example, records with MONTHCODE=1 refer to January, MONTHCODE=2 are records for February, and so on. For the 1996 through 2008 Panels, the variable RHCALMN identifies the calendar reference month of a record on the core data files. For the 1990 through 1993 Panels, it is the variable MONTH.

3.5 Identifying Persons and Their Relationships: 2014 Panel Through 2020 SIPP

Each person in the 2014 Panel through 2020 SIPP data can be uniquely identified by the combination of a sample unit ID (SSUID) and a person number (PNUM). These ID variables are crucial when linking the records for a single person across multiple SIPP data files. Table 3-4 demonstrates the relationship between SSUID and PNUM.

TABLE 3-4. HOW TO UNIQUELY IDENTIFY PERSONS, 2014 PANEL THROUGH 2020 SIPP

Sample Unit ID (SSUID)	Person Number (PNUM)	Notes
2014 SIPP Panel		
123456789123	101	Original sample member
123456789123	102	Original sample member
123456789123	301	Enters SIPP sample in Wave 3
123456789123	401	Enters SIPP sample in Wave 4
123456789123	402	Enters SIPP sample in Wave 4
321456789123	101	Original sample member
321456789123	102	Original sample member
321456789123	201	Enters SIPP sample in Wave 2
2018 SIPP		
12345678912318	101	Original sample member in 2018 Panel
12345678912318	102	Original sample member in 2018 Panel
32145678912318	101	Original sample member in 2018 Panel
32145678912318	102	Original sample member in 2018 Panel
2020 SIPP		
12345678912318	101	Original sample member in 2018 Panel
12345678912318	201	Enters SIPP sample in Wave 2 of 2018 Panel
12345678912318	301	Enters SIPP sample in Wave 3 of 2018 Panel
12345678912320	101	Original sample member in 2020 Panel
32145678912320	101	Original sample member in 2020 Panel
32145678912320	102	Original sample member in 2020 Panel

3.5.1 Determining Monthly Household Composition

A *household*, as the term is used in Census Bureau publications, consists of all people who occupy a housing unit, regardless of their relationships to each other. For many purposes, a household can be thought of as people living at a common address. A person's current address ID for a given month, together with his or her sample unit ID, identifies the household in which that person is a member for that month.

A key factor in the strength of the SIPP longitudinal survey framework is collecting residence history information for all respondents during the reference year and following sample members when they move between waves. Table 3-5 shows the variables necessary for identifying sample members who have moved.

TABLE 3-5. VARIABLES FOR IDENTIFYING MOVERS, 2014 PANEL THROUGH 2020 SIPP

Name	Description	Values
SSUID	Sample Unit Identifier: Uniquely identifies each initially sampled dwelling unit. Every person sampled is either an original member of one of those units (an original sample member, or OSP) or lives with someone who was a member of an initially sampled dwelling unit. A person's connection to that unit is an attribute of that person and does not change over time. This means that as people move from address to address, their SSUID stays the same. As new people join the homes of original sample members, they receive the SSUID of the original sample members.	2014 SIPP 000000000000: 999999999999 2020 SIPP 00000000000018: 99999999999920
PNUM	Person Number: Differentiates persons within the sample unit. Does not change throughout panel. The first digit indicates the wave in which a person joined the SIPP sample; the second two correspond to the person's place on the household roster.	101:499
ERESIDENCEID	Current Address ID: Identifies the specific residence where the person lived during each month. When members report living at or moving to previously un-sampled households, it changes to signify new (spinoff) sample households and the wave added to the survey.	100001:400999

Table 3-6 provides example residence history data for three original sample members and a fourth sample member who joined the household, interviewed together at an original sample address in Wave 2, as the data would appear on the 2020 SIPP public-use file. Each person, denoted by the values 101, 102, 103, and 301 on the variable PNUM, has twelve monthly records of residence data (MONTHCODE). The corresponding sample unit (SSUID), monthly Residence ID (ERESIDENCEID), interview month Residence ID (EHRESIDENCID), and mover flag (TMOVER) variables may be used to identify residence history, identify mover status, and construct households over the twelve-month period. Note that the interview month Residence ID (EHRESIDENCID) does not vary across months for a respondent for a given wave and that respondents carry the same sample unit identifier (SSUID) across the entire panel.

In the example household:

- All four people lived together during the interview month (all four people have the same SSUID and EHRESIDENCID values).
- Person 1 lived at the interview address for all twelve months of the reference period

(ERESIDENCEID=EHRESIDENCID for all twelve months of data).

- Person 2 and Person 3 lived at the interview address with Person 1 for months 9 to 12 (September-December). All three people have the same ERESIDENCEID value for months 9 to 12 (September-December), and this value is equal to the interview month Residence ID value, EHRESIDENCID.
- Person 2 and Person 3 moved to the interview address in month 9 (September). The ERESIDENCEID value for both people changed from '100002' in month 8 to '100001' in month 9 (September), indicating that the two respondents changed residences. Alternatively, the mover flag value (TMOVER = 2) for month 9 (September) on both respondents' records also indicates that the respondents moved during month 9 (September), and that this move is to a new residence in the same county and same state as the previous residence.
- Person 2 and Person 3 lived together for months 1 through 8 (January-August) at a different residence than the interview month residence (ERESIDENCEID values for both people are equal for months 1 through 8 (January-August), but this value differs from the EHRESIDENCID value).
- Person 4 did not live with Person 1, Person 2, or Person 3 for any of the twelve months of the reference period (monthly ERESIDENCEID values are never equal to the ERESIDENCEID values for Person 1, Person 2, or Person 3 in the same month during the reference period).
- Person 4 moved to the interview address, with Person 1, Person 2, and Person 3, at some point after the reference period ended and before or during the interview month (Person 4 has the same SSUID and EHRESIDENCID values, despite having a different ERESIDENCEID value than each of the other three people for every month of the reference period).

TABLE 3-6. IDENTIFYING HOUSEHOLDS AND MOVERS

MONTHCODE	SSUID	PNUM	ERESIDENCEID	EHRESIDENCID	TMOVER
(Person 1)					
1	10111110312318	101	100001	100001	*missing* (Not in universe)
2	10111110312318	101	100001	100001	1 (non-mover)
3	10111110312318	101	100001	100001	1 (non-mover)
4	10111110312318	101	100001	100001	1 (non-mover)
5	10111110312318	101	100001	100001	1 (non-mover)
6	10111110312318	101	100001	100001	1 (non-mover)
7	10111110312318	101	100001	100001	1 (non-mover)
8	10111110312318	101	100001	100001	1 (non-mover)
9	10111110312318	101	100001	100001	1 (non-mover)
10	10111110312318	101	100001	100001	1 (non-mover)
11	10111110312318	101	100001	100001	1 (non-mover)
12	10111110312318	101	100001	100001	1 (non-mover)
(Person 2)					
1	10111110312318	102	100002	100001	*missing* (Not in universe)
2	10111110312318	102	100002	100001	1 (non-mover)

MONTHCODE	SSUID	PNUM	ERESIDENCEID	EHRESIDENCID	TMOVER
3	10111110312318	102	100002	100001	1 (non-mover)
4	10111110312318	102	100002	100001	1 (non-mover)
5	10111110312318	102	100002	100001	1 (non-mover)
6	10111110312318	102	100002	100001	1 (non-mover)
7	10111110312318	102	100002	100001	1 (non-mover)
8	10111110312318	102	100002	100001	1 (non-mover)
9	10111110312318	102	100001	100001	2 (mover)
10	10111110312318	102	100001	100001	1 (non-mover)
11	10111110312318	102	100001	100001	1 (non-mover)
12	10111110312318	102	100001	100001	1 (non-mover)
(Person 3)					
1	10111110312318	103	100002	100001	*missing* (Not in universe)
2	10111110312318	103	100002	100001	1 (non-mover)
3	10111110312318	103	100002	100001	1 (non-mover)
4	10111110312318	103	100002	100001	1 (non-mover)
5	10111110312318	103	100002	100001	1 (non-mover)
6	10111110312318	103	100002	100001	1 (non-mover)
7	10111110312318	103	100002	100001	1 (non-mover)
8	10111110312318	103	100002	100001	1 (non-mover)
9	10111110312318	103	100001	100001	2 (mover)
10	10111110312318	103	100001	100001	1 (non-mover)
11	10111110312318	103	100001	100001	1 (non-mover)
12	10111110312318	103	100001	100001	1 (non-mover)
(Person 4)					
1	10111110312318	301	300001	100001	*missing* (Not in universe)
2	10111110312318	301	300001	100001	1 (non-mover)
3	10111110312318	301	300001	100001	1 (non-mover)
4	10111110312318	301	300001	100001	1 (non-mover)
5	10111110312318	301	300001	100001	1 (non-mover)
6	10111110312318	301	300001	100001	1 (non-mover)
7	10111110312318	301	300001	100001	1 (non-mover)
8	10111110312318	301	300001	100001	1 (non-mover)
9	10111110312318	301	300001	100001	1 (non-mover)
10	10111110312318	301	300001	100001	1 (non-mover)
11	10111110312318	301	300001	100001	1 (non-mover)
12	10111110312318	301	300001	100001	1 (non-mover)

3.5.2 Monthly Household Relationships

SIPP data include a number of recoded variables that denote the monthly relationship of each household member to every other household member. Previous SIPP panels only showed the relationship of each person to the household reference person, except for the Wave 2 Household Relationship Topical Module. For help analyzing relationships in prior SIPP panels, see the 2014 SIPP Users' Guide at <www2.census.gov/programs-surveys/sipp/tech-documentation/methodology/2014-SIPP-Panel-Users-Guide.pdf>.

SIPP allows for a possible 30 persons on the roster (up to 20 interviewed household members plus up to 10 Type 2 people). Consequently, for each month there are 30 RREL variables (RREL1-RREL30) and 30 corresponding PNUM identifiers (RREL_PNUM1-RREL_PNUM30) on each person's record. The vector of monthly relationships (RREL[x]) must be used in conjunction with the vector of person numbers (RREL_PNUM[x]) in that month to determine monthly household relationships. The household relationships are listed in Table 3-7.

TABLE 3-7. MONTHLY HOUSEHOLD RELATIONSHIP VALUES

99 = self	10 = half siblings
1 = opposite-sex spouse	11 = stepsiblings
2 = opposite-sex unmarried partner	12 = adoptive siblings
3 = same-sex spouse	13 = other siblings
4 = same-sex unmarried partner	14 = parent-/child-in-law
5 = biological parent/child	15 = brother-/sister-in-law
6 = stepparent/stepchild	16 = aunt/uncle/niece/nephew
7 = adoptive parent/child	17 = other relative
8 = grandparent/grandchild	18 = foster parent/ child
9 = biological siblings	19 = other nonrelative

Table 3-8 shows an example of the matrix values as they would appear on the public-use file for a household containing a couple and their child for months 1 and 2. All three individuals live together at month 1, while PNUM=102 moves to a different residence at month 2. This residence change is reflected in the RREL and RREL_PNUM vectors.

TABLE 3-8. MONTHLY HOUSEHOLD RELATIONSHIP MATRIX

SSUID	PNUM	ERESID ENCEID	MONTH CODE	RREL1	RREL2	RREL3	RREL_ PNUM1	RREL_ PNUM2	RREL_ PNUM3
12345…	101 (Husband)	100001	1	99	1	5	101	102	103
12345…	102 (Wife)	100001	1	1	99	5	101	102	103
12345…	103 (Child)	100001	1	5	5	99	101	102	103
12345…	101 (Husband)	100001	2	99	5	*missing*	101	103	*missing*
12345…	102 (Wife)	100002	2	99	*missing*	*missing*	102	*missing*	*missing*
12345…	103 (Child)	100001	2	5	99	*missing*	101	103	*missing*

3.5.3 Identifying Parents, Children, Spouses, and Partners at a Monthly Level

RREL and the relationship matrix allow for the identification of various relationships between household members, but a simpler way exists for identifying a person's parent(s), child(ren), spouse, or partner. Four variables (EPNSPOUS_EHC, EPNCOHAB_EHC, RPNPAR1_EHC, and RPNPAR2_EHC) identify the person number of the spouse, person number of the partner, person number of the first parent, and person number of the second parent, respectively. In each case, the person is identified only if she or he is living at the same address as the record holder in that month.

With these variables, users can identify a variety of family configurations. For example, these variables can be used to identify households containing three generations. Table 3-9 shows an example household containing a husband, wife, their two adult children, and their grandchild.

TABLE 3-9. IDENTIFYING, PARENTS, CHILDREN, AND SPOUSES

Person	Relationship to John Doe	PNUM	EPNSPOUS_EHC	RPNPAR1_EHC	RPNPAR2_EHC
John Doe	Self	101	102	*missing*	*missing*
Jane Doe	Spouse	102	101	*missing*	*missing*
John Doe Jr.	Child	103	*missing*	101	102
Joan Doe	Child	104	*missing*	101	102
Jimmy Doe	Grandchild	105	*missing*	104	*missing*

3.5.4 Determining Monthly Family Composition

The term *family*, as used at the Census Bureau, refers to a group of two or more people related by birth, marriage, or adoption who reside together; all such people are considered members of one family. For example, a grandparent living with a grandchild would be considered a family, as would a pair of cousins, a parent with an adopted child, an aunt with a nephew, or a blended family of married partners with a mix of biological and step children.

SIPP groups related persons into families via the variable RFAMNUM by using the monthly household relationships (RRELs). This variable will have the same numeric value for everyone in the household that month who is related to the family reference person (RFAMREF) by birth, marriage, or adoption. If there is a second group of people in the household who are related to each other, but not to the first family, then they will have a value of '2' on RFAMNUM (and are known as a *subfamily*). People who are unrelated to anyone else in the household will have a unique value of RFAMNUM and will not have a value for type of family (RFAMKIND).

Rules for Assigning RFAMNUM and RFAMKIND

When grouping sample members into RFAMNUM groups and assigning RFAMKIND, the SIPP processing edits institute the following process through the household roster:

1. The edit checks the roster for the first married couple and picks one of the spouses as the family reference person (RFAMREF). The edit then looks for all persons related to the selected spouse

and assigns them all RFAMNUM=1. The edit then moves through the household roster, looking for a second married couple not yet assigned to a family and initiates the same grouping process, grouping members in RFAMNUM=2. The edit continues through the roster, scanning for married couples who are not part of a previously assigned RFAMNUM group. Each person in the group with a married couple will have the family type variable, RFAMKIND, set to a value of 1 for "married couple."

2. Once all married couples are in families, the edit looks for the first mother-child pair not yet assigned to a family, selects the mother as the family reference person (RFAMREF), and groups anyone related to the mother (and not already assigned to a family) into a RFAMNUM group. Each person in that group will have the RFAMKIND value of 2, for "female, no spouse present."
3. Once pairs are in families, the edit looks for the first father-child pair not yet assigned to a family, selects the father as the family reference person (RFAMREF), and groups anyone related to the father (and not already assigned to a family) into a RFAMNUM group. Each person in that group will have the RFAMKIND value of 3, for "male, no spouse present."
4. Once all father-child pairs are in families, the edit looks for remaining household members not yet assigned to a family. If any people have any relatives not already assigned to a family, they are grouped in the same RFAMNUM group. Those with no relatives are grouped in their own RFAMNUM group. Each person in a family of more than one will have RFAMKIND set to a value of 2 for "female, no spouse present" or 3 for "male, no spouse present" according to the sex of the family reference person (RFAMREF).

Table 3-10 provides an example of how household members are assigned to families.

TABLE 3-10. IDENTIFYING A FAMILY

Person	Relationship	SSUID	PNUM	ERESID ENCEID	RFAM NUM	RFAM KIND	RFAM REF
John Doe	Self	12345678998720	101	100001	1	1	101
Jane Doe	Spouse	12345678998720	102	100001	1	1	101
Jimmy Doe	Child	12345678998720	103	100001	1	1	101
James Moe	Unrelated friend	12345678998720	104	100001	2	3	104
Joseph Moe	James' cousin	12345678998720	105	100001	2	3	104
Joan Roe	James' girlfriend	12345678998720	106	100001	3	*missing*	106

3.5.5 Type 2 Persons

As mentioned in Section 2.5, a Type 2 person is someone who lived with a respondent at any address (not just the interview address) during the reference year but not at the time of interview. While Type 2 people are never interviewed, and as such do not have a person-record in the data, SIPP respondents are asked to provide some information about Type 2 people. The SIPP data files provide the following information for each Type 2 person: sex, age, whether he/she worked for pay during the time lived together, and annual income. This information is available on the records of interviewed respondents. For example, the variables ET2_SEX1, TT2_AGE1, ET2_WORK1, and TT2INC1 show the sex, age, employment status,

and income for the first Type 2 person reported for a SIPP respondent. This information will only be populated for the months in which the SIPP respondent resided with the Type 2 person.

Type 2 people are also eligible to be selected as recipients of respondents' programs, such as health insurance or TANF. Additionally, a version of monthly household poverty and income estimates include Type 2 people. Each SIPP respondent may report up to 10 Type 2 persons. Though Type 2 people do not have a person-record in the data, they are given PNUM values of 60-69. If you see values of 60-69 in variables that reference PNUMs (e.g., RREL_PNUM), know that Type 2 people are being pointed to.

3.6 Identifying Persons and Their Relationships: 1984 Through 2008 Panels

Each person in the 1984 through 2008 Panels can be uniquely identified by the combination of a sample unit ID, an entry address ID,[8] and a person number. These ID variables, shown in Table 3-3, are necessary when linking the records for a single person across multiple SIPP data files within a panel. They also contain substantive information that may be useful in some situations. For additional information on identifying persons and their relationships beyond what is described in Sections 3.6.1 through 3.6.3, please see previous versions of this Users' Guide at <www.census.gov/programs-surveys/sipp/guidance/users-guide.html>.

3.6.1 Determining Monthly Household Composition

A person's current address ID in any given month, together with his or her sample unit ID, identifies the household in which that person is a member for that month. Members of the same household in a given month always have an interview status of 1 and share the same sample unit ID and current address ID.

3.6.2 Determining Monthly Family Composition

Methods for identifying the interrelationships among the household members that define these groups vary, depending on the data file being used. The topical module files do not contain any of the information needed to directly identify the different types of families.[9] When it is necessary to identify family membership in an analysis that uses information from a topical module, it is also necessary to merge data from the topical module file with either a core wave file or a longitudinal file.

Identifying family membership is easiest when working with the person-month-format core wave files. The Census Bureau has two principal methods for distinguishing families. The first method defines a family as all persons who are related and living together. The family ID variable RFID is used with this definition. RFID groups the household reference person with all related household members by assigning them the same ID number. This family group corresponds to the Census Bureau's definition of a primary

[8] For the 1996-2008 Panels, the entry address is not necessary to uniquely identify individuals in SIPP. Its use will not create any problems; it just provides additional information.

[9] The one exception is the Wave 2 topical module that collects detailed information about all of the relationships among all of the people who are household members at the time of the Wave 2 interview.

family. RFID groups members of each unrelated subfamily (and primary and secondary individuals) separately.

The second method is similar to the first in defining a family, but the family excludes members of related subfamilies. The family ID variable RFID2 is used with this definition. RFID2 equals 0 for members of related subfamilies. RFID2 groups members of each unrelated subfamily (and primary and secondary individuals) in the same way, as RFID each group has a unique number.

Data users who want to analyze multi-generational families would use RFID2 (FID2) and the variable RSID (SID). RSID (SID) treats related subfamilies as distinct family units by assigning members of related subfamilies nonzero values. Data users can easily distinguish unrelated subfamilies from other family units when they use these variables and numbering schemes.

3.6.3 Determining Monthly Transfer Program Unit Composition

Some analyses involve summarizing data for units other than households or families. The SIPP core data contain sufficient information to identify program units for participants in a range of transfer programs, including Medicare; Medicaid; Temporary Assistance for Needy Families (TANF);[10] General Assistance (GA); Railroad Retirement; Social Security; Veterans Compensation and Pensions; Food Stamps; and the Special Supplemental Nutrition Program for Women, Infants, and Children's Nutrition Program (WIC).

The SIPP data contain fields for each adult and child, indicating whether the individual received benefits (either directly or by virtue of his or her relationship to another person designated as the benefit owner) from each of these programs in each month. The SIPP data also contain information that permits identification of program units within households. One person in each program unit is identified as the benefit owner, and variables identifying that benefit owner are included on the records of the people who are part of the program unit. People who are members of a common program unit in a given month can then be identified as those who are in the sample in that month with common values of the sample unit ID, the current address ID, and the benefit owner ID.

3.7 Identifying Program Participation in Government Transfer Programs: 2014 Panel Through 2020 SIPP

SIPP collects data on a number of government transfer programs, including the following:

- Disability benefits
- Energy assistance
- Food assistance
- Foster/child/spousal support
- General Assistance (GA)

[10] In August 1996, the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA) was signed into law. This legislation replaced AFDC with TANF. In the 1996 Panel, the questions for income type 20 referred to the AFDC program prior to Wave 4 and to the TANF program beginning in Wave 4. In Wave 9, the questions were expanded somewhat to capture the larger array of program types that could exist under TANF.

- Retirement income
- School breakfast/lunch
- Supplemental Nutrition Assistance Program (SNAP), formerly called food stamps
- Social Security (adult)
- Social Security (on behalf of a child)
- Special Supplemental Nutrition Program for Women, Infants, and Children (WIC)
- Supplemental Security Income (SSI)
- Survivor benefits
- Temporary Assistance for Needy Families (TANF)
- Transportation assistance
- Unemployment/workers' compensation
- Veterans benefits

Most of these data are reported at the annual and monthly level for each individual, but some are reported only annually and for the household (i.e., Energy Assistance). In the next section, several programs are used to highlight the ways in which programs can be received across persons in households over time. Each program, regardless of how it is collected, is also discussed in depth later in this document.

3.7.1 Collecting Information for Program Units

Transfer programs vary in a number of ways that make the collection of benefit data a challenge. Some programs are means-tested, requiring applicants' resources to fall below a defined threshold to qualify for benefits. Non-means-tested programs waive income and asset restrictions but may have other eligibility restrictions such as employment or job loss. Some programs are designed to provide benefits to an individual, while others determine eligibility at the family or household level. Numerous instrument efficiencies allow some information to be collected at a household or family level to reduce respondent burden. Editing of the collected data assigns information to covered persons in most instances, although for EHC programs there is also information on the record of the benefit "owner" who may or may not be covered.

3.7.2 Identifying Program Participation in Government Transfer Programs

Two key concepts for understanding how to measure program participation in the SIPP are **coverage** and **ownership**. The coverage variable on the public-use file that has the same value for the entire reference period is a recode indicating coverage in at least one month of the reference period.

The coverage and ownership variables on the public-use file that have the same value for each month of a spell include a variable identifying *benefit coverage this month* and *benefit ownership this month*.

Table 3-11 shows how to identify coverage and ownership for several means-tested programs. For example, benefits from SSI cover an individual, while benefits from some other means-tested programs (SNAP, TANF, GA, and WIC) cover a "program unit," which includes all household members who meet the eligibility criteria. Program units may include just one person, more than one person but not all household members, or all household members; they vary based on program rules and household

composition. As a result, there are slight variations in the way variables are assigned values between SSI and the other programs.

For SSI, for example, information about spells is available on each monthly record included in a spell for the respondent covered by the benefit. There are two types of coverage indicators. A reference period coverage indicator (which has the same value in each month of the reference period) shows whether a respondent was covered by SSI benefits in one or more months of the reference year (RSSI_YRYN). The monthly coverage indicator identifies who was covered by an SSI benefit in a given month (RSSI_MNYN). The value of this variable may change across the reference period. If the respondent is under age 18, ESSI_OWN points to the owner of the benefit. Unlike other EHC programs, all of the information about SSI spells is available on the record of the child covered by the benefit (see Table 3-12 and Table 3-13).

TABLE 3-11. MEASURING COVERAGE AND OWNERSHIP IN MEANS-TESTED PROGRAMS IN THE 2018 THROUGH 2020 SIPP

	Reference Period Variable	**Monthly Variables**	
	Coverage	**Coverage**	**Ownership**
Supplemental Security Income (SSI)			
Benefit owner			
Covered	RSSI_YRYN=1	RSSI_MNYN=1-12	ESSI_OWN=missing*
Not covered	RSSI_YRYN=2	RSSI_MNYN=2	ESSI_OWN=*missing*
Not Benefit Owner			
Covered	RSSI_YRYN=1	RSSI_MNYN=1	ESSI_OWN=PNUM
Not covered	RSSI_YRYN=2	RSSI_MNYN=2	ESSI_OWN=*missing*
Supplemental Nutrition Assistance Program (SNAP)			
Benefit owner			
Covered	RSNAP_YRYN=1	RSNAP_MNYN=1	ESNAP_OWN=PNUM
Not covered	RSNAP_YRYN=2	RSNAP_MNYN=2	ESNAP_OWN=PNUM
Not Benefit Owner			
Covered	RSNAP_YRYN=1	RSNAP_MNYN=1	ESNAP_OWN=PNUM†
Not covered	RSNAP_YRYN=2	RSNAP_MNYN=2	ESNAP_OWN=*missing*
Temporary Assistance for Needy Families (TANF)			
Benefit owner			
Covered	RTANF_YRYN=1	RTANF_MNYN=1	ETANF_OWN=PNUM
Not covered	RTANF_YRYN=1	RTANF_MNYN=2	ETANF_OWN=PNUM
Not Benefit Owner			
Covered	RTANF_YRYN=1	RTANF_MNYN=1	ETANF_OWN=PNUM†
Not covered	RTANF_YRYN=2	RTANF_MNYN=2	ETANF_OWN=*missing*
General Assistance (GA)			

	Reference Period Variable	Monthly Variables	
	Coverage	**Coverage**	**Ownership**
Benefit owner			
Covered	RGA_YRYN=1	RGA_MNYN=1	EGA_OWN=PNUM
Not covered	RGA_YRYN=2	RGA_MNYN=2	EGA_OWN=PNUM
Not Benefit Owner			
Covered	RGA_YRYN=1	RGA_MNYN=1	EGA_OWN=PNUM†
Not covered	RGA_YRYN=2	RGA_MNYN=2	EGA_OWN=*missing*
Special Supplemental Nutrition Program for Women, Infants, and Children (WIC)			
Benefit owner			
Covered	RWIC_YRYN=1	RWIC_MNYN=1	EWIC_OWN=PNUM
Not covered	RWIC_YRYN=2	RWIC_MNYN=2	EWIC_OWN=PNUM
Not Benefit Owner			
Covered	RWIC_YRYN=1	RWIC_MNYN=1	EWIC_OWN=PNUM†
Not covered	RWIC_YRYN=2	RWIC_MNYN=2	EWIC_OWN=*missing*

* This information is collected for SSI recipients under the age 18. Once SSI recipients turn age 18, they may receive benefits in their own name.

† For people covered by benefits who are not the owner, the ownership variables (e.g., ESNAP_OWN) hold the PNUM value of the benefit owner.

For GA, TANF, SNAP, and WIC a respondent aged 15 or older in each program unit is identified as the benefit owner for each spell. The ownership variable is available on the records of the benefit owner in every month of the spell and on other household members' records each month they had benefit coverage (e.g., ESNAP_OWN in Table 3-11). Detailed information about the spell is located on the record of the benefit owner in the months covered by a spell, regardless of whether this person was covered by the benefit (see Table 3-12 and Table 3-13).

Table 3-11 above shows some of the variable names for the two types of coverage indicators. A reference period coverage indicator identifies whether a respondent was covered by the benefit in one or more months of the reference period (e.g., RSNAP_YRYN in Table 3-11). The monthly coverage indicators identify which household members were covered by a benefit in a given month (e.g., RSNAP_MNYN). The value of this variable may change across the reference period.

3.7.3 Detailed Means-Tested Program Participation Variables

Several variables on the public-use file provide information about when a program spell began and ended. These variables, which have the same value for each month of a spell, include:

- Begin month of benefit receipt
- End month of benefit receipt
- Year benefit receipt began if receiving in January of the reference period in a Wave 1 interview
- Continuation flag for spells that include the last month of the reference period

As Table 3-12 shows for SSI, the months receipt began and ended, the year benefit receipt began, and the continuation flag are available on the monthly records included in a spell for the person receiving the SSI benefit. For SNAP, TANF, GA, and WIC, the months receipt began and ended, the year benefit receipt began, and the continuation flag are available on the monthly records included in a spell for the owner of a benefit. Other programs are described in depth later in this document under the header for those programs.

TABLE 3-12. IDENTIFYING SPELL DURATION IN MEANS-TESTED PROGRAMS IN THE 2018 THROUGH 2020 SIPP

	Spell begin month	Spell end month	Continuation flag*	Left-censored year†
Supplemental Security Income (SSI)				
Benefit owner				
Covered	ESSI_BMONTH=1-12	ESSI_EMONTH=1-12	RSSI_CFLG=1-3	ESSI_LCYR ≥ 0
Not covered	ESSI_BMONTH=*missing*	ESSI_EMONTH=*missing*	RSSI_CFLG=*missing*	ESSI_LCYR=*missing*
Not Benefit Owner				
Covered	ESSI_BMONTH= *missing*	ESSI_EMONTH=*missing*	RSSI_CFLG= *missing*	ESSI_LCYR=*missing*
Not covered	ESSI_BMONTH=*missing*	ESSI_EMONTH=*missing*	RSSI_CFLG=*missing*	ESSI_LCYR=*missing*
Supplemental Nutrition Assistance Program (SNAP)				
Benefit owner				
Covered	ESNAP_BMONTH=1-12	ESNAP_EMONTH=1-12	RSNAP_CFLG=1-3	RSNAP_LCYR ≥ 0
Not covered	ESNAP_BMONTH=1-12	ESNAP_EMONTH=1-12	RSNAP_CFLG=1-3	RSNAP_LCYR ≥ 0
Not Benefit Owner				
Covered	ESNAP_BMONTH=*missing*	ESNAP_EMONTH=*missing*	RSNAP_CFLG=*missing*	RSNAP_LCYR=*missing*
Not covered	ESNAP_BMONTH=*missing*	ESNAP_EMONTH=*missing*	RSNAP_CFLG=*missing*	RSNAP_LCYR=*missing*
Temporary Assistance for Needy Families (TANF)				
Benefit owner				
Covered	ETANF_BMONTH=1-12	ETANF_EMONTH=1-12	RTANF_CFLG=1-3	RTANF_LCYR ≥ 0
Not covered	ETANF_BMONTH=1-12	ETANF_EMONTH=1-12	RTANF_CFLG=1-3	RTANF_LCYR ≥ 0
Not Benefit Owner				
Covered	ETANF_BMONTH=missing	ETANF_EMONTH=*missing*	RTANF_CFLG=*missing*	RTANF_LCYR=*missing*
Not covered	ETANF_BMONTH=*missing*	ETANF_EMONTH=*missing*	RTANF_CFLG=*missing*	RTANF_LCYR=*missing*
General Assistance (GA)				
Benefit owner				
Covered	EGA_BMONTH=1-12	EGA_EMONTH=1-12	RGA_CFLG=1-3	RGA_LCYR ≥ 0
Not covered	EGA_BMONTH=1-12	EGA_EMONTH=1-12	RGA_CFLG=1-3	RGA_LCYR ≥ 0
Not Benefit Owner				
Covered	EGA_BMONTH=missing	EGA_EMONTH=missing	RGA_CFLG=missing	RGA_LCYR=*missing*
Not covered	EGA_BMONTH=*missing*	EGA_EMONTH=*missing*	RGA_CFLG=*missing*	RGA_LCYR=*missing*

	Spell begin month	Spell end month	Continuation flag*	Left-censored year†
Special Supplemental Nutrition Program for Women, Infants, and Children (WIC)				
Benefit owner				
Covered	EWIC_BMONTH=1-12	EWIC_EMONTH=1-12	RWIC_CFLG=1-3	RWIC_LCYR≥0
Not covered	EWIC_BMONTH=1-12	EWIC_EMONTH=1-12	RWIC_CFLG=1-3	RWIC_LCYR≥0
Not Benefit Owner				
Covered	EWIC_BMONTH=*missing*	EWIC_EMONTH=missing	RWIC_CFLG=*missing*	RWIC_LCYR=*missing*
Not covered	EWIC_BMONTH=*missing*	EWIC_EMONTH=*missing*	RWIC_CFLG=*missing*	RWIC_LCYR=*missing*

* This variable is only available for spells that end in the last month of the reference period (e.g., ESSI_EMONTH=12). Like other detailed spell variables, it is present for each month of the spell.

† This variable is only available for spells that include January of the reference period (e.g., ETANF_BMONTH=1) the first time a respondent is interviewed. Like other detailed spell variables, it is present for each month of the spell.

Additional detailed program participation variables include:

- Reasons benefit receipt started
- Reasons benefit receipt stopped if receipt ended during the reference year
- Whether the benefit covered only children, only adults, or both (available for TANF and WIC)
- Monthly benefit amount

The reasons benefit receipt started, reasons benefit receipt stopped, and type of benefit coverage (for TANF and WIC) have the same value for each month of a spell but may change values between spells. For SSI, the reasons benefit receipt started and stopped are located on the monthly records included in a spell for the person covered by the SSI benefit. For other EHC programs, the reasons benefit receipt started and stopped and the type of benefit coverage (for TANF and WIC) are available on the benefit owner's monthly records included in a spell, regardless of whether this respondent was covered by the benefit.

Benefit amounts may vary from month to month within a spell. Amount variables contain the monthly payment received by the individual or program unit. For SSI, the amount variable is on the monthly records included in a spell for the person covered by the SSI benefit. When an amount is collected for a program unit (SNAP, TANF, GA, or WIC), the value is placed on the benefit owner's monthly records included in a spell to avoid double counting income when creating family and household-level income recodes (see Table 3-13).

TABLE 3-13. IDENTIFYING DETAILED SPELL INFORMATION FOR MEANS-TESTED PROGRAMS IN THE 2018 THROUGH 2020 SIPP

	Why began*	Why ended†	Coverage type	Amount‡
Supplemental Security Income (SSI)				
Benefit owner				
Covered	ESSI_BRSN=1-4	ESSI_ERSN=1-7 or *missing*	n/a	TSSI_AMT≥0
Not covered	ESSI_BRSN=*missing*	ESSI_ERSN=*missing*	n/a	TSSI_AMT=*missing*

	Why began*	Why ended†	Coverage type	Amount‡
Not Benefit Owner				
Covered	ESSI_BRSN=1-4	ESSI_ERSN=1-7 or *missing*	n/a	TSSI_AMT≥0
Not covered	ESSI_BRSN=*missing*	ESSI_ERSN=*missing*	n/a	TSSI_AMT=*missing*
Supplemental Nutrition Assistance Program (SNAP)				
Benefit owner				
Covered	ESNAP_BRSN=1-9	ESNAP_ERSN=1-7 or *missing*	n/a	TSNAP_AMT≥0
Not covered	ESNAP_BRSN=1-9	ESNAP_ERSN=1-7 or *missing*	n/a	TSNAP_AMT≥0
Not Benefit Owner				
Covered	ESNAP_BRSN=*missing*	ESNAP_ERSN=*missing*	n/a	TSNAP_AMT=*missing*
Not covered	ESNAP_BRSN=*missing*	ESNAP_ERSN=*missing*	n/a	TSNAP_AMT=*missing*
Temporary Assistance for Needy Families (TANF)				
Benefit owner				
Covered	ETANF_BRSN=1-9	ETANF_ERSN=1-9 or *missing*	ETANF_COVTYP=1-3	TTANF_AMT≥0
Not covered	ETANF_BRSN=1-9	ETANF_ERSN=1-9 or *missing*	ETANF_COVTYP=1-3	TTANF_AMT≥0
Not Benefit Owner				
Covered	ETANF_BRSN=*missing*	ETANF_ERSN=*missing*	ETANF_COVTYP =*missing*	TTANF_AMT=*missing*
Not covered	ETANF_BRSN=*missing*	ETANF_ERSN=*missing*	ETANF_COVTYP =*missing*	TTANF_AMT=*missing*
General Assistance (GA)				
Benefit owner				
Covered	EGA_BRSN=1-9	EGA_ERSN1=1-9 or *missing*	n/a	TGA_AMT≥0
Not covered	EGA_BRSN=1-9	EGA_ERSN1=1-9 or *missing*	n/a	TGA_AMT≥0
Not Benefit Owner				
Covered	EGA_BRSN=*missing*	EGA_ERSN=*missing*	n/a	TGA_AMT=*missing*
Not covered	EGA_BRSN=*missing*	EGA_ERSN=*missing*	n/a	TGA_AMT=*missing*
Special Supplemental Nutrition Program for Women, Infants, and Children (WIC)				
Benefit owner				
Covered	EWIC_BRSN=1-9	EWIC_ERSN=1-7 or *missing*	EWIC_COVTYP=1-3	TWIC_AMT≥0
Not covered	EWIC_BRSN=1-9	EWIC_ERSN=1-7 or *missing*	EWIC_COVTYP=1-3	TWIC_AMT≥0
Not Benefit Owner				
Covered	EWIC_BRSN=*missing*	EWIC_ERSN=*missing*	EWIC_COVTYP=*missing*	TWIC_AMT=*missing*
Not covered	EWIC_BRSN=*missing*	EWIC_ERSN=*missing*	EWIC_COVTYP=*missing*	TWIC_AMT=*missing*

* For each reported spell, respondents are asked why benefit receipt began. Respondents select reasons from a mark all that apply list. Since most respondents select a single reason, the public-use file only has one variable that corresponds to the first reason reported.

† For each reported spell that ends during the reference period, respondents are asked why benefit receipt stopped. Respondents select reasons from a mark all that apply list. Since most respondents select a single reason, the public-use file only has one variable that corresponds to the first reason reported.

‡ When program benefits cover multiple household members, the benefit amount is only recorded on the owner's record to avoid double counting amounts when calculating family and household-level income.

3.7.4 Means-Tested Program Participation in a Sample Household

Table 3-14 shows a sample household consisting of four related household members. A husband (person 101) and wife (person 102) own the house. Additionally, their adult daughter (person 103) and grandson (person 104) lived with them for the duration of the reference period. Three examples are presented to demonstrate:

- SNAP receipt that covers everyone in the household with a change in payment amount (Table 3-15)
- SSI receipt for an individual (Table 3-16)
- Multiple TANF spells across the reference period with a change in program unit (Table 3-17)

TABLE 3-14. SAMPLE HOUSEHOLD

	Husband	Wife	Adult daughter	Grandson
PNUM	101	102	103	104
TAGE	60	53	20	2

SNAP Example

The entire household was covered by SNAP benefits from January through December of the reference period (Table 3-15). From January through June of the reference period (months 1 to 6), the benefit amount was $340. Due to an increase in household income the benefit amount decreased to $295 for the rest of the reference period (months 7 to 12). The benefit owner was the wife (person 102). As Table 3-15 shows, the begin month of benefit receipt, the end month of benefit receipt, and the benefit amount are only present on the wife's (person 102) monthly records. The other household respondents point to person 102 as the owner (ESNAP_OWN=102), indicating that the full detailed spell information can be found on that person's record. This example includes a subset of SNAP variables (see Table 3-11, Table 3-12, and Table 3-13 for a full listing). Note that the amount variable is structured similarly to other detailed spell variables in terms of whose records are in universe and have values (see Table 3-17 for an example using the full set of spell variables).

TABLE 3-15. IDENTIFYING MONTHLY PROGRAM PARTICIPATION: SNAP

Month	Variables					
	RSNAP_YRYN*	ESNAP_ BMONTH	ESNAP_ EMONTH	RSNAP_MNYN	ESNAP_OWN	TSNAP_AMT
Husband (101)						
1	1	*missing*	*missing*	1	102	*missing*
2	1	*missing*	*missing*	1	102	*missing*
3	1	*missing*	*missing*	1	102	*missing*
4	1	*missing*	*missing*	1	102	*missing*
5	1	*missing*	*missing*	1	102	*missing*

Month	Variables					
	RSNAP_YRYN*	ESNAP_ BMONTH	ESNAP_ EMONTH	RSNAP_MNYN	ESNAP_OWN	TSNAP_AMT
6	1	*missing*	*missing*	1	102	*missing*
7	1	*missing*	*missing*	1	102	*missing*
8	1	*missing*	*missing*	1	102	*missing*
9	1	*missing*	*missing*	1	102	*missing*
10	1	*missing*	*missing*	1	102	*missing*
11	1	*missing*	*missing*	1	102	*missing*
12	1	*missing*	*missing*	1	102	*missing*
Wife (102)						
1	1	1	12	1	102	340
2	1	1	12	1	102	340
3	1	1	12	1	102	340
4	1	1	12	1	102	340
5	1	1	12	1	102	340
6	1	1	12	1	102	340
7	1	1	12	1	102	295
8	1	1	12	1	102	295
9	1	1	12	1	102	295
10	1	1	12	1	102	295
11	1	1	12	1	102	295
12	1	1	12	1	102	295
Adult Daughter (103)						
1	1	*missing*	*missing*	1	102	*missing*
2	1	*missing*	*missing*	1	102	*missing*
3	1	*missing*	*missing*	1	102	*missing*
4	1	*missing*	*missing*	1	102	*missing*
5	1	*missing*	*missing*	1	102	*missing*
6	1	*missing*	*missing*	1	102	*missing*
7	1	*missing*	*missing*	1	102	*missing*
8	1	*missing*	*missing*	1	102	*missing*
9	1	*missing*	*missing*	1	102	*missing*
10	1	*missing*	*missing*	1	102	*missing*
11	1	*missing*	*missing*	1	102	*missing*
12	1	*missing*	*missing*	1	102	*missing*
Grandson (104)						
1	1	*missing*	*missing*	1	102	*missing*
2	1	*missing*	*missing*	1	102	*missing*
3	1	*missing*	*missing*	1	102	*missing*
4	1	*missing*	*missing*	1	102	*missing*
5	1	*missing*	*missing*	1	102	*missing*
6	1	*missing*	*missing*	1	102	*missing*
7	1	*missing*	*missing*	1	102	*missing*
8	1	*missing*	*missing*	1	102	*missing*
9	1	*missing*	*missing*	1	102	*missing*
10	1	*missing*	*missing*	1	102	*missing*

Month	Variables					
	RSNAP_YRYN*	ESNAP_ BMONTH	ESNAP_ EMONTH	RSNAP_MNYN	ESNAP_OWN	TSNAP_AMT
11	1	*missing*	*missing*	1	102	*missing*
12	1	*missing*	*missing*	1	102	*missing*

* RSNAP_YRYN has the same value in each month because it is a reference period indicator.

SSI Example

The wife (person 102) reported receiving $700 in SSI benefits during each month of the reference period (Table 3-16). Since SSI provides benefits to individuals, the example only includes the records for person 102. This example includes a subset of SSI variables (see Table 3-11, Table 3-12, and Table 3-13 for a full listing).

TABLE 3-16. IDENTIFYING MONTHLY PROGRAM PARTICIPATION: SSI

Month	Variables					
	RSNAP_YRYN*	RSSI_MNYN	ESSI_ BMONTH	ESSI_ EMONTH	ESSI_OWN†	TSSI_AMT
	Wife (102)					
1	1	1	1	12	*missing*	700
2	1	1	1	12	*missing*	700
3	1	1	1	12	*missing*	700
4	1	1	1	12	*missing*	700
5	1	1	1	12	*missing*	700
6	1	1	1	12	*missing*	700
7	1	1	1	12	*missing*	700
8	1	1	1	12	*missing*	700
9	1	1	1	12	*missing*	700
10	1	1	1	12	*missing*	700
11	1	1	1	12	*missing*	700
12	1	1	1	12	*missing*	700

* RSNAP_YRYN has the same value in each month because it is a reference period indicator.

† ESSI_OWN only has values for respondents under age 18.

TANF Example

The adult daughter (person 103) reported receiving TANF benefits in January through April of the reference period that covered her and her son (months 1 to 4) but stopped receiving benefits for a few months (months 5 to 8) (Table 3-17). In September, she reported receiving TANF benefits that only covered her son (person 104) through the remainder of the reference period (months 9 to 12). Detailed information about each spell, including the benefit amount, is available on the benefit owner's record

(person 103). Since the adult daughter (person 103) and grandson (person 104) are the only people covered by the TANF benefit, this example only includes the records for these two household members.

TABLE 3-17. IDENTIFYING MONTHLY PROGRAM PARTICIPATION: TANF

Month	Variables											
	RTANF_ YRYN*	RTANF_ MNYN	PNUM	ETANF_ BMONTH	ETANF_ EMONTH	ETANF_ OWN	TTANF_ AMT	ETANF_ LCYR	RTANF_ CFLG	ETANF_ COVTYP	ETANF_ BRSN	ETANF_ ERSN
Adult Daughter (103)												
1	1	1	103	1	4	103	190	2012	*missing*	3	4	6
2	1	1	103	1	4	103	190	2012	*missing*	3	4	6
3	1	1	103	1	4	103	190	2012	*missing*	3	4	6
4	1	1	103	1	4	103	190	2012	*missing*	3	4	6
5	1	2	103	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*
6	1	2	103	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*
7	1	2	103	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*
8	1	2	103	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*
9	1	2	103	9	12	103	90	*missing*	3	1	9	*missing*
10	1	2	103	9	12	103	90	*missing*	3	1	9	*missing*
11	1	2	103	9	12	103	90	*missing*	3	1	9	*missing*
12	1	2	103	9	12	103	90	*missing*	3	1	9	*missing*
Grandson (104)												
1	1	1	104	*missing*	*missing*	103	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*
2	1	1	104	*missing*	*missing*	103	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*
3	1	1	104	*missing*	*missing*	103	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*
4	1	1	104	*missing*	*missing*	103	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*
5	1	2	104	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*
6	1	2	104	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*
7	1	2	104	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*
8	1	2	104	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*
9	1	1	104	*missing*	*missing*	103	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*
10	1	1	104	*missing*	*missing*	103	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*
11	1	1	104	*missing*	*missing*	103	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*
12	1	1	104	*missing*	*missing*	103	*missing*	*missing*	*missing*	*missing*	*missing*	*missing*

* RTANF_YRYN has the same value in each month because it is a reference period indicator.

3.8 Geography

The geographic data released on the SIPP Public-Use File include state and metropolitan/nonmetropolitan status variables. Starting with the 2004 SIPP Panel, the sample is designed to be representative at the state level. However, for each state estimate the coefficient of variation should be calculated to determine if the estimate is reliable. For the 2014 SIPP Panel, the sample was expanded in 16 states to obtain a coefficient

of variation of 6 percent or lower for the estimate of low-income individuals. No such expansion has been done since.

The SIPP sample has never been designed to produce metropolitan-area (MSA) level estimates. However, MSA identifiers and additional geographic data are collected and may be available for access at Census Research Data Centers for researchers with Special Sworn Status and approved projects.

Identifying States in the 2014 Panel Through 2020 SIPP

The variables TEHC_ST and TST_INTV, which are based on the Federal Information Processing Standards (FIPS) State Code, identify the 50 states and District of Columbia. TST_INTV identifies the state of residence for the interview month, while TEHC_ST identifies the state for each of a respondent's residences during the reference period. TEHC_ST also identifies the world region for respondents who lived outside of the United States for some months during the reference year.

Identifying Metropolitan Areas in the 2014 Panel Through 2020 SIPP

The variables TEHC_METRO and TMETRO_INTV identify households that are located in either a metropolitan or non-metropolitan area. TMETRO_INTV identifies the metropolitan status for the interview month residence, while TEHC_METRO identifies the metropolitan status for each of a respondent's residences during the reference period. TEHC_METRO and TMETRO_INTV have three categories: "metro," non-metro," and "not identified." Metro or non-metro areas with a population below 250,000 are recoded to "not identified." No other reassignment is done for this variable.

See Section 3.2.1 for more details on how the topcoding for TEHC_METRO and TMETRO_INTV is implemented in the 2014 Panel through the 2020 SIPP.

SIPP's *Metropolitan and Micropolitan* webpage at <www.census.gov/programs-surveys/metro-micro/about.html> provides more detail about metropolitan area delineations over time.

4 Topics in the 2020 SIPP Public-Use Data

This chapter details topical content in the 2020 SIPP public-use data files. Much of the 2020 SIPP topical content and guidance overlaps with the 2018 and 2019 SIPP. Significant differences, if present, are noted within each topic. For guidance on content in previous panels, please visit the Technical Documentation page of the SIPP website at <www.census.gov/programs-surveys/sipp/tech-documentation.html>.

The terms household, person, and person-month are used throughout this chapter to indicate the record level for variables. Household-level variable's values do not vary among household members. Each household member has the same value for the variable on all person-month records. Person-level variable's values have the same value on all person-month records. Person-month-level variable's values may vary across months for persons. That is, each person may have a different value for the variable on each person-month record. Additional definitions for terms used throughout this chapter – such as coverage unit, reference person, continuation flag, etc. – are available in the Glossary of this guide.

While the main objective of SIPP is to provide accurate and comprehensive information about the income and program participation of individuals and households in the United States, SIPP also collects extensive data on many additional facets of economic well-being. Employment status and income are not the sole determinants of households' and individuals' financial situations. To create a more complete representation of national well-being, SIPP also collects extensive information concerning family dynamics, educational attainment, housing expenditures and conditions, asset ownership, health insurance, disability, childcare, and food security. These data put income, program recipiency, and program eligibility into the household, family, and social context. Thus, SIPP is the only federal data source that can facilitate the examination of the ways in which these factors interact to influence financial well-being and movement into or out of government assistance programs.

4.1 Demographics

4.1.1 Basic Demographics

SIPP collects basic demographic information about the whole household during each interview, including each person's date of birth, sex (ESEX), race, and Hispanic origin (EORIGIN). Age variables include each individual's birth year (TDOB_BYEAR), as well as his or her age at the time of the interview (TAGE) and at each of the 12 months of the reference period (TAGE_EHC). Race variables include a variable with categories for "White alone," "Black alone," "Asian alone," and "Residual" (ERACE), and a more detailed race variable (TRACE).

Level of information the Basic Demographics data provide

The Basic Demographics data are person-level variables, except for TAGE_EHC, which is a person-month variable.

4.1.2 Educational Enrollment

The Educational Enrollment section collects information from respondents aged 3 or older on the months of enrollment, grade, type, full-time or part-time status, credential worked toward, and whether or not a grade was repeated. This information is important because school enrollment is related to labor force participation for adults and can be an outcome measure for children.

Level of information the Educational Enrollment data provide

Most Educational Enrollment variables provide person-month information. All information should be consistent within the same spell of enrollment, and variables EEDGREP and EEDENROLL do not vary across months.

Other concepts related to the Educational Enrollment data

Educational Enrollment questions refer to school enrollment during the individual months of the reference year. "Enrollment spell" refers to a length of consecutive month(s) during which the respondent is continuously enrolled.

4.1.3 Educational Attainment

The demographic Educational Attainment content collects information for adults aged 15 and older about an individual's highest level of educational attainment. There is also information on whether high school graduates got their diplomas by graduating from high school or by passing a GED exam.

The Educational Attainment content also includes data on alternative credentials. This section includes information on whether respondents have a professional certification, state or industry license, and/or educational certificate. Respondents who report a professional certification or license also report what organization or institution awarded the credential.

This information is important because it provides details on the education, training, qualifications, and skill development of the adult population. Educational attainment is an important outcome in itself and can be used to show how education affects a variety of socioeconomic outcomes.

Level of information the Educational Attainment data provide

All Educational Attainment items are person-level items. This information can be used to examine differences in educational attainment, credentials, and skill development across subgroups of the population. The information on educational attainment and alternative credentials can be used with other SIPP topics such as labor force involvement to examine how attainment is related to a variety of socioeconomic outcomes.

Additional guidance for using the Educational Attainment data

The educational attainment and GED questions for all household members are asked of the household respondent in the Demographics section. The questions related to professional certifications, licenses, and educational certificates are part of each person's interview.

Other concepts related to the Educational Attainment data

The measures of credentials were developed by the Federal Interagency Working Group on Expanded Measures of Enrollment and Attainment (GEMEnA), a group tasked with improving federal data collection on education, credentials, and job training. After an extensive development process, the GEMEnA group defined the concepts accordingly:

Professional certification: A credential awarded by a certification body based on an individual demonstrating through an examination process that he or she has acquired the designated knowledge, skills, and abilities to perform a specific job. The examination can be written, oral, or performance based. Certification is a time-limited credential that is renewed through a recertification process.

License: A credential awarded by a licensing agency based on pre-determined criteria. The criteria may include some combination of degree attainment, certifications, certificates, assessment, apprenticeship programs, or work experience. Licenses are time-limited and must be renewed periodically.

Educational certificate: A credential awarded by an educational institution based on completion of all requirements for a program of study, including coursework and test or other performance evaluations. Certificates are typically awarded for life (like a degree). Certificates of attendance or participation in short-term training (e.g., one-day) are not counted as educational certificates.

4.1.4 Family and Household Relationships

Relationship to reference person, parent pointers (variables that identify the co-resident parents of the record holder), and parent type (whether biological, adoptive, or step) were collected in the 2020 SIPP. The relationship of everyone in the household to everyone else was not directly collected; however, by using information collected in other questions, data users can identify household relationships for every month of the reference period, allowing for a more nuanced study of changing household relationships over time. These monthly relationship variables are used as the basis for family and household recodes. As was the case in the 2018 and 2019 SIPP, the 2020 SIPP data also indicate cohabiting relationships that do not include the reference person throughout all months. Family and household information allows us to understand differences in poverty, program participation, work status, etc., by household composition. Below is detailed information regarding these topics.

Relationship to reference person

Since the 2014 Panel, the relationship to reference person question has included 18 categories. Relationship to reference person is a person-level variable (ERELRPE) collected in the front of the instrument with the basic demographics. This question is only asked relative to the reference person.

Monthly relationship of everyone in the household

For every month of the reference period, SIPP creates a monthly (person-month) recode of the relationship of each respondent in the household that month to every other person in the household that month (RREL1-RREL30). Each person with a record on the file has a set of variables that show how the other household members are related to the record holder during the months they lived together. This recode uses a combination of residence ID, relationship to reference person, marital status, spouse

pointers, cohabiting partner pointers, parent pointers, parent type, Type 2 relationships, and biological child lists from the fertility section to code monthly relationships. The 'pointer' variables contain the person number of the record holder's parent, spouse, or cohabiting partner. The monthly relationships are reciprocal; for example, a grandparent and grandchild will both have an 8 as their relationship to one another (grandparent/grandchild). Parent pointers and age variables must be used to determine which person is the grandparent and which is the grandchild. Parent pointers and age are both available at the person-month level.

There is an array of variables indicating which household member is referenced in each monthly relationship variable (RREL_PNUM1-RREL_PNUM30). These variables contain person numbers. It is important to use these variables in combination with the monthly relationship variables to map household relationships.

For more information about residence ID, see the Residence section (Section 3.4.3). For more information about marital status, spouse pointers, and cohabiting partner pointers, see the Marital Status and Marital History section (Section 4.1.7). See the Fertility content (Section 4.1.5) for further information about the biological child lists.

Parent and child pointers

Respondents are asked if each person in the household has a parent present at the time of interview. If they answer yes, they are asked if this person has a second parent present. If a parent is reported, the type of relationship to that parent (biological, step, or adoptive) is determined. Moving from asking about mothers and fathers to parents allows us to capture two-parent families in same-sex relationships rather than showing these families as single-parent families. This improves our ability to collect information about diverse family types without losing previously collected information. Parent pointers are available for people of all ages. The 'pointer' variable contains the person number of the household member identified as the record holder's parent.

Interview month parent pointers (EPNPAR1, EPNPAR2) and parent type (EPAR1TYP, EPAR2TYP) are person-level variables on the child's record. In the 2020 SIPP, the sex of adults identified as parents in the parent pointers is provided as a person-level recode on the child's record (RPAR1SEX, RPAR2SEX). Mothers and fathers can be identified by using parent pointers in conjunction with recodes of the sex of parent(s).

In addition, monthly versions of parent pointers (RPNPAR1_EHC, RPNPAR2_EHC), parent type variables (RPAR1TYP_EHC, RPAR2TYP_EHC), and parent sex variables (RPAR1SEX_EHC, RPAR2SEX_EHC) are recodes based on monthly relationships, parent pointers, residence, and spouse pointers.

On the parent's record, users will find child pointers (RPNCHILD1-RPNCHILD19) and child type (RCHTYP1-RCHTYP19), which are recodes of parent pointers and parent type. The vector of child pointers contains a list of the person number(s) of the record holder's children who are present in the household during the interview month.

Household and family recodes

The 2020 SIPP has a recoded household-level variable for the number of people in the household (RHNUMPER). There are additional household recodes, including the number of people in the household who are 65 years and over (RHNUM65OVER) and the number of people in the household who are under 18 years old (RHNUMU18). Each of these variables has two versions: one counting only interviewed household members and one counting interviewed household members and Type 2 people (persons living in the household that month, but not at the time of interview) which have the suffix "WT2" added at the end of the variable name (e.g., RHNUMPERWT2).

The family recodes are monthly recodes that show family number (RFAMNUM), family type (RFAMKIND), family reference person (RFAMREF), number of people in a family (RFPERSONS), and number of related people under 18 years in a family (RFRELU18). Because the 2020 SIPP does not collect a 'reference person' for each month of the reference period, family recodes do not have a reference person available to define a primary family. Instead, all people in the household that month are assigned to a family, and each family is assigned a family reference person. Like household recodes, there are two sets of family recodes: one that includes only interviewed family members and one that includes interviewed family members and Type 2 people who are related to those family members and are present that month, which have the suffix "WT2" added at the end of the variable name (e.g., RFAMNUMWT2).

Level of information the Family and Household data provide

Relationship to the household respondent, child pointers, and child type are person-level variables only and are based on interview month information. Parent pointers, parent type variables, and parent sex variables are available at both the person-level and person-month-level. Relationships of everyone in the household and family and household composition variables, are available at the person-month-level.

4.1.5 Fertility

The SIPP instrument asks for complete fertility histories from all respondents aged 15 years and older, including:

- The total number of children ever born/fathered (TCEB)
- The month and year of birth for each child for up to 20 births (TCBYR_1-TCBYR_20; NOTE: The public-use file only releases year of birth and caps the maximum number of births based on current disclosure protection guidelines. Month of birth is not released on the public-use file)
- A direct question about multiple partner fertility (EMPF), asked of all adults with more than one child
- Whether they are a grandparent (EGRNDPR)

Based on this information, additional recode variables are created for other important concepts related to fertility, including:

- Age at first birth (TAGE_FB)
- Whether their current union (married or cohabiting) is a childbearing one (RCB_UNION)
- Number of childbearing unions (TNUM_CBU)

Level of information the Fertility data provide

Fertility data are available at the person level. For instance, all respondents aged 15 years or older have a TCEB value indicating the number of children they have given birth to/fathered, if any. RCB_UNION indicates whether the respondent's current union is a childbearing one and is included on both partners' records with a value of 1 if the couple share at least one child.

Additional guidance for using the Fertility data

The public-use file also includes a parent screener variable (EPAR_SCRNR), which is asked in the beginning of the survey and before the Fertility questions. Individuals who are labeled as parents in this variable may be parents of biological, step, or adopted children. The remaining Fertility questions are only asked in reference to biological children. Therefore, a person may have an affirmative value in the parent screener variable but have zero children ever born (TCEB).

The Fertility questions are collected in the back end of the survey, after Assets and Disability, but before any of the Child Care or Child Well-Being questions. Household relationships, demographics, marital history, and child well-being are other topics in the SIPP with direct ties to the fertility content. As fertility and the timing of family formation are key predictors of economic well-being, fertility also has important implications for employment and social program use.

4.1.6 Language

Language content includes information on whether people aged 5 and older speak a language other than English at home. For people who speak a language other than English at home, there is additional information on what language they speak and how well they speak English. The data file also contains a recode variable that indicates households in which no one aged 14 or older speaks English exclusively at home, or in which no one speaks English "very well."

The content is important because language use and English ability are indicators of possible membership in a community with certain resources, while lack of English-speaking ability can sometimes be a barrier to accessing public services or educational and job opportunities. Additionally, language data allow researchers and policy makers to examine the association between English language proficiency and important socioeconomic outcomes, as well as look at language use and ability of the U.S. population.

Level of information the Language data provide

The Langue data are provided as person-level variables.

Additional guidance for using the Language data

If the interview is conducted in a language other than English, the household respondent is asked if others in the household speak a language other than English at home. If the interview is conducted in English, the household respondent is asked whether anyone speaks a language other than English at home.

RLNGISOL is a variable indicating whether the respondent lives in a limited English-speaking household. It is set to 1 for households in which no one aged 14 or older speaks English, as well as for households in which another language is spoken at home and no one speaks English very well. The language content can be used to describe the languages spoken in the U.S. and to examine how language

affects various outcomes. It can also be used to explore whether people who speak a language other than English at home differ from English-only speakers across a variety of socioeconomic outcomes.

Other concepts related to the Language data

Limited English-Speaking Households: No one in the household 14 and over speaks English only or speaks English "very well" (RLNGISOL).

4.1.7 Marital Status and Marital History

Marital Status and Marital History information collected in SIPP is more detailed than the information collected in other Census Bureau surveys. It can be used in conjunction with other content in the survey such as employment, education, or poverty by demographic characteristics, as well as to examine changes in household composition over time.

Current marital status is reported in the variable EMS and is collected about every person aged 15 years and older living in the household during the interview month. Values of this variable include never married; married, spouse present; married, spouse absent; widowed; separated; or divorced. For household members who are or have been married, SIPP also collects:

- Identification of the spouse in the household (EPNSPOUSE)
- Number of times married (EXMAR)
- Whether they have ever been widowed or divorced (EVERWID, EVERDIV)
- The year of current marriage (TYRCURRMARR, asked if the person is currently married)
- The year of first marriage (TYRFIRSTMARR, directly asked only if the person is not currently in his or her first marriage)

Marital status *changes* during the reference period are collected in the EHC for each respondent. Up to three spells of marital status can be recorded per person. All respondents aged 15 and older have a monthly marital status (EMS_EHC). Additionally, SIPP collects data to identify respondents' spouses and partners, namely:

- Respondents married during any month, and living with their spouse, have a spouse pointer for that month indicating the person number of their spouse (EPNSPOUS_EHC).
- Respondents who do not have a spouse present are asked whether they had a boyfriend, girlfriend, or partner in the household and, if yes, they have a cohabiting partner pointer indicating the person number of their cohabiting partner (EPNCOHAB_EHC; EPNCOHAB if the boyfriend, girlfriend, or partner was present at interview month).

Major changes in the Marital Status and Marital History data between the 2019 SIPP and the 2020 SIPP

Starting with the 2020 SIPP, respondents were not asked whether they were in a registered domestic partnership or civil union with their cohabiting partner.

Level of information the Marital Status and Marital History data provide

Marital status, spouse pointers, and cohabitating partner pointers are available at both the person and person-month levels for each person aged 15 years and older. Times married, ever widowed, ever divorced, year of first marriage, and year of current marriage are all person-level variables.

Current marital status and spouse pointer information is collected in the Demographics section and asked of the reference person about every person aged 15 years and older. Other person-level information is collected in a series of marital history questions prior to the EHC.

Marital status changes are asked in the EHC section of each person aged 15 years and older. This spell - level information is then recoded into the monthly marital status, spouse pointers, and cohabitation pointers. Only one cohabiting partnership can be recorded for each marital history spell. If a person is identified as a cohabiting partner, he/she is edited as a partner in all months during that spell in which both persons do not have a spouse present and are themselves present in the household.

4.1.8 Nativity, Citizenship, and Parent Nativity

The Nativity and Citizenship questions are located in the Demographics section and are asked of the reference person about everyone in the household. The Parent Nativity questions are located in the Parent section of the survey and are asked of all respondents whose biological parent(s) is/are not present in the household. Questions include:

- Whether a respondent was born in the United States
- If so, state of birth
- Nativity status and country of origin for the respondent's biological parents (asked of each respondent)

For foreign-born respondents:

- Country of birth
- Citizenship status
- How the respondent became a citizen
- Immigration status on arrival
- Year of entry

Level of information the Nativity, Citizenship, and Parent Nativity data provide

All nativity and citizenship questions are collected at the person level. Questions on parents' nativity status can be used to calculate generational status of respondents.

Additional guidance for using the Nativity, Citizenship, and Parent Nativity data

Nativity and Citizenship are asked of the reference person for every person in the household in the demographics section. Biological mother's and biological father's birth in the U.S. and country of birth are asked during the back of the survey for each respondent whose biological parent(s) are not present in the household.

4.1.9 Parent Mortality

Respondents who do not have their biological mother/father in the household are asked the parent mortality questions, including:

- Whether a respondent's biological mother/father is still alive
- If a parent is deceased, the month and year of death
- The month and year that the respondent's biological mother/father was born

Since Wave 2 of the 2014 Panel, only data regarding whether a respondent's biological mother/father is still alive (EBMOM, EBDAD) have been released on the public-use files.

Level of information the Parent Mortality data provide

All parent mortality questions reflect status at the time of interview and are person-level variables.

Additional guidance for using the Parent Mortality data

These items are asked during the Parent section at the back end of the survey of each respondent whose biological parent(s) is/are not present in the household. For respondents whose biological parent(s) is/are present in the household, the parent(s) is/are marked as still alive. Thus, the parent mortality variables on the released file contain information on the biological parents of all respondents.

4.1.10 Residence

The Residence section collects information about an individual's residence history from the start of the reference year through the interview month, including:

- Where the respondent lived
- How long the respondent lived at the residence
- Whom the respondent lived with
- Why the respondent moved to the residence
- Tenure status
- Type of living quarters
- Receipt of housing assistance

As with other data collected in the EHC, residence information is collected in reverse chronological order. That is, each respondent's residence history is collected beginning with the most current residence and working backward through the reference period. The instrument collects information on up to five residences at which the respondent lived during the reference year.

For the residence spell that includes January of the reference period, information on the year and month the respondent moved to the residence is collected along with the tenure status of the residence the respondent lived in prior to the residence that includes January of the reference period.

The instrument stores address information for each new residence entered within a wave. This information is then available for other respondents to select when reporting residence histories, or for reports of multiple residence spells at the same address. This improves the reliability and consistency of

the reported address information. For each residence spell, respondents report whom they lived with. The tenure status, type of living quarters, housing assistance receipt, and reason for move information may vary across spells.

In some instances where respondents in a household report living together for an entire reference period, residence history information for the reference period is copied from the first respondent to subsequent respondents who lived with the first respondent. The exception is when respondents report living together during January of the reference month. In these situations, each respondent is asked the month and year of move and the reason for move to the residence, as these may vary across respondents. When respondents do not report living together for the entire reference period, each respondent's residence history information is reported separately.

The residences where respondents lived are assigned unique residence IDs (ERESIDENCEID, unique within SSUID), which can be used to identify households across the reference year and across panels. Residence IDs are assigned in processing and may not correspond to the chronological ordering of residences for any given respondent. The address information is fed back in subsequent waves and available for respondents to select.

To help data users, recode variables are constructed, including a monthly mover flag and geographic recodes (such as region of residence and metropolitan status). The monthly mover flag (TMOVER) allows data users to easily identify when a respondent changed residences and the type of move (for example, within the same county, to a new county in the same state, or to a new state).

Level of information the Residence data provide

Residence data are collected at the spell level in the EHC and are provided to data users in person-month format. Characteristics of the interview address (region, state, metropolitan status, tenure, and receipt of housing subsidies) are available as household-level variables. This information is also provided as person-month variables for each residence during the reference year.

Additional guidance for using Residence data

All residence information is available at a monthly level. The ERH_BMONTH and ERH_EMONTH variables indicate the beginning and ending months of the residence spell (months beyond reference period are capped at 12). Geographic information for each residence, including region, state of residence, and metropolitan status, are available at a monthly level. Type of living quarters, tenure, rent subsidy, and housing voucher receipt are available for each residence at a monthly level.

The mover flag (TMOVER) identifies when a respondent changes residences and the type of move, while the reason for move (EEHC_WHY) is collected for each residence during the reference period. For the residence during January of the reference period, year and month moved in are collected. Tenure status of the unit lived in prior to January of the reference period is also collected.

The 2020 SIPP includes several variables that can be used to identify if the respondent moved during the reference period, the month of the move, and the type of move. These variables include:

- ERESIDENCEID - identifies unique households within an original sample unit. A change in ERESIDENCEID from one month to the next indicates a change in residences.

- ERH_BMONTH - the beginning month of the residence spell.
- ERH_EMONTH - the end month of the residence spell.
- TMOVER – a mover flag that indicates if the respondent moved during a given month and the type of move. The categories of TMOVER are:
    1. Same house in the U.S., non-mover
    2. Different house in the U.S., same state and county
    3. Different house in the U.S., same state, different county
    4. Different house in the U.S., different state in the Northeast
    5. Different house in the U.S., different state in the Midwest
    6. Different house in the U.S., different state in the South
    7. Different house in the U.S., different state in the West

For moves that occurred during January of the reference year of the first wave the respondent was interviewed or moves that are not captured in the survey reference period, the year (TEHC_MVYR) and month (EEHC_MVMO) of move are collected, but not the type of move. In waves 2+, returning sample members will have valid data on TMOVER for January of the reference year, as the edited December data from the previous wave are used to create the TMOVER variable. New respondents in wave 2+ (and all respondents in wave 1) will be blank on TMOVER for January of the reference year.

The variables TEHC_MVYR, EEHC_MVMO, ERH_BMONTH, and ERH_EMONTH are used to create a monthly duration of residence variable (TRESDUR).

The residence data are edited for consistency so that ERESIDENCEID is used in the creation of TRESDUR and TMOVER, and TRESDUR is used in the creation of TMOVER. In most instances, these variables will provide consistent information when identifying a move. However, in a small number of cases a change in ERESIDENCEID from one month to another may not be reflected on TMOVER and TRESDUR. In these instances, a respondent has changed residences from one month to another but reports always living at the residence he or she has moved to on the reason for move variable (EEHC_WHY). Because respondents' own definitions of what constitutes a move may differ from how the SIPP survey identifies a move, we have chosen to leave these apparent inconsistencies in the data.

Other concepts related to the Residence data

Residence ID: Identifies unique households within a given original sample unit (ERESIDENCEID).

Mover flag: Identifies when a respondent changed residences and whether the move was within the same county, to a new county in the same state, or to a new state (TMOVER). In other words, the variable flags the first month in a new residence.

4.1.11 Veteran Status

The Veteran Status content collects information on respondents aged 17 or older about whether they have ever served in the United States Armed Forces and their periods of service.

Respondents are asked whether they have ever served on active duty in the United States Armed Forces. Veterans are men and women who have served (even for a short time), but are not currently serving, on active duty in the U.S. Army, Navy, Air Force, Marine Corps, or Coast Guard, or who served in the U.S. Merchant Marine during World War II.

Respondents who indicated that they ever served on active duty in the past or were currently on active duty are asked to indicate the period or periods in which they served, even if just for part of the period. The periods were determined by the Department of Veterans Affairs (VA) and generally distinguish veterans who served during wartime periods from those whose only service was during peacetime. These service periods include:

- September 2001 to present
- August 1990 to August 2001 (including Persian Gulf War)
- May 1975 to July 1990
- Vietnam Era (August 1964 to April 1975)
- February 1955 to July 1964
- Korean War (July 1950 to January 1955)
- January 1947 to June 1950
- World War II (December 1941 to December 1946)
- November 1941 or earlier

Level of information the Veterans data provide

The Veterans data are available at the person level, except for EAF_HHLD, which is a household-level variable.

4.2 Assets, Employment, and Earnings

4.2.1 Assets

The Assets content provides detailed information on assets and liabilities for individuals and households. Three types of data are collected:

1. Asset ownership during the reference period, as well as type of ownership (joint ownership and/or individual ownership).
2. Value of assets and any debts held against these assets as of the last day of the reference period.
3. Income received from each asset during the reference period.

Most questions are asked of all household members who were aged 15 years and older at the time of the interview (referred to as person-level variables), while a few questions are asked only of the household reference person (referred to as household-level variables).

The Assets content produces a wide range of national estimates for wealth and debt, including the distribution of net worth. Additionally, person-level and household-level wealth and income can be used to model eligibility for various government programs. The level of detail within SIPP also provides data users the flexibility to construct their own units of analysis (individuals, families, households, etc.) and study how changes in household structure and other conditions affect wealth over time.

Table 4-1 lists topics covered within the Assets content, as well as the microdata abbreviation used for naming the associated variables. Indicators for income questions and type of ownership questions (where applicable) are included as well. Additionally, starting with the 2018 SIPP, many asset variables are imputed with statistical model-based imputation, as described in Section 6.2.1 of this Users' Guide. The table also lists which variables were imputed using model-based-imputation.

TABLE 4-1. TOPICS COVERED IN THE 2020 SIPP ASSETS CONTENT

Topic	Microdata abbreviation	Level of data	Income questions	Type of ownership questions	Model-based imputation of ownership	Model-based imputation of value
ASSETS*						
Retirement Accounts						
IRA and KEOGH accounts	IRAKEO	Person			✓	✓
401(k), 403(b), 503(b), and Thrift Savings Plan accounts	THR401	Person			✓	✓
Defined-benefit or cash balance plan	PENSION	Person			✓	
Interest-Earning Assets						
Government securities	GOVS	Person	✓	✓	✓	
Checking accounts	CHK	Person	✓	✓	✓	
Savings accounts	SAV	Person	✓	✓	✓	
Money market accounts or funds	MM	Person	✓	✓	✓	
Certificates of deposit	CD	Person	✓	✓	✓	
Municipal and corporate bonds	MCBD	Person	✓	✓	✓	
Educational savings accounts	ESAV	Household				
Other Income-Generating Assets						
Stocks	ST	Person	✓	✓	✓	
Mutual funds	MF	Person	✓	✓	✓	
Rental property	RP	Person	✓	✓	✓	
Annuities	ANN	Person	✓		‡	
Trusts	TR	Person	✓		‡	
Other Assets						
Other real estate	RE	Person		✓	✓	
Businesses owned as a job	BSJ(1-7)	Person		†	✓	
Businesses owned as an investment only	BSI(1-3)	Person			✓	
Life insurance policies	LIFE	Person			✓	
Primary residence (for non-mobile homes)	PR	Household		†		
Primary residence (for mobile homes)	MH	Household		†		
Cars, trucks, and vans	VEH	Household		†		
Recreational vehicles	RECV	Household		†		
Other financial investments (such as coins, collectibles, jewelry, artwork, mortgages paid to him/her, other loans owed to him/her, and royalties)	OINV	Person	✓		✓	

Topic	Microdata abbreviation	Level of data	Income questions	Type of ownership questions	Model-based imputation of ownership	Model-based imputation of value
LIABILITIES						
Debts Secured by Assets						
Primary residence debt (for non-mobile homes)	PR	Household		†		
Primary residence debt (for mobile homes)	MH	Household		†		
Rental Property Debt	RP	Person		✓		
Other real estate	RE	Person		✓		
Vehicle debt	VEH	Household		†		
Recreational vehicle debt	RECV	Household		†		
Debt on businesses owned as a job	BSJ(1-7)	Person		†		
Debt on businesses owned as an investment only	BSI(1-3)	Person				
Debts Not Secured by an Asset (Unsecured Debt)						
Credit card debt and store bills	CC	Person		✓	✓	
Student loans and educational-related expenses	ED	Person		✓	✓	
Medical Debt	MED	Person			✓	✓
Other debt (includes loans obtained through a bank or credit union, money owed to private individuals, debt held against mutual funds or stocks)	OT	Person		✓	✓	
ADDITIONAL TOPICS						
Rent and mortgage payments	RENTMORT	Household				
Utility payments	UTILS	Household				

*All reported assets are those held outside of other pooled assets. For example, reported mutual funds are those held outside of retirement accounts.

†Joint ownership can be inferred from the reported list of owners.

‡2020 SIPP uses model-based imputation to impute whether respondents owned any annuities, trusts, or both (ANNTR).

As with the 2014 Panel through 2019 SIPP, variable names are designed to make it easier to identify related variables using consistent stems for topics and characteristics. For example, EOWN_CHK and TJSCHKVAL both refer to checking accounts. TJSCHKVAL and TJSSAVVAL both refer to the value of an account held jointly with a spouse as of the last day of the reference period (for a checking account and

a savings account, respectively). Figure 4-1 details the convention used to name variables within the section:

FIGURE 4-1. NAMING CONVENTIONS FOR ASSETS AND LIABILITIES VARIABLES

Person-level assets and liabilities:

Ownership of asset or liability:	**E**	+	**OWN_** *(for assets)* *or* **DEBT_** *(for liabilities)*	+	***Variable Abbreviation***		
	E *or* **T**	+	***Joint/Own Prefix***	+	***Variable Abbreviation***	+	***Characteristic***
Details for asset or liability:	*(***T** *when variable is topcoded)*		*(**JS**, **JO**, or **O**, where applicable)*		*(such as **THR401**, **SAV**, **CHK**)*		*(**INC** for income, **VAL** for value, **DEBTVAL** for debt value, etc.)*

Household-level assets and liabilities:

Ownership of asset:	**EOWN_**	+	***Variable Abbreviation***				
Details for asset: (includes details for liabilities)	**E** *or* **T** *(***T** *when variable is topcoded)*	+	***Variable Abbreviation*** *(such as **VEH**, **PR**, **ESAV**)*	+	***Asset Number*** *(when multiple assets of one type are reported)*	+	***Characteristic*** *(**VAL** for value, **DEBTVAL** for debt value, etc.)*

Vehicle Value and Status Flag Assignment

In SIPP, respondents are asked to report the year, make, and model of their vehicles. Using this information, values are assigned from a dataset created by J.D. Power, which is one of the major providers of vehicle trade-in value data. For respondents who report a year, make, and model, the vehicle values are usually assigned a status flag value of 1. For respondents who provide slightly less usable vehicle information, such as reporting a year and make but not a model, or reporting a very new vehicle that doesn't have sufficient resale history, these car values are usually assigned a status flag of 5. This indicates that their vehicle values involved more complex calculations, which may result in less accurate data than vehicles with a status flag of 1. For respondents who provided so little vehicle information that their values had to be imputed via hot-deck, their status flag is assigned a value of 2.

Respondent Burden Reduction

For person-level assets that can be owned jointly, the survey is designed to refrain from asking identical questions to spouses. This is to reduce respondent burden. The survey tailors question text to ensure that proper values are collected. For example, if a married couple owns an asset jointly with a non-household member, the first spouse is asked to report only "your and your spouse's share" of an asset's value, debt, and income. Then, the second spouse skips these questions. In editing, reported values are divided by two and copied to both spouses' records. In all cases, both spouses are asked about individual ownership (ownership in own name).

Major changes in the Assets data between the 2019 SIPP and the 2020 SIPP

In previous years, questions about assets held jointly with a spouse or partner were asked of married couples and respondents in a registered domestic partnership or civil union. Since respondents are no longer asked about registered domestic partnerships (EREGDOMPART) in the 2020 SIPP, only married couples are in universe for assets held jointly with a spouse (EJS or TJS variables). Respondents who are not married and own assets jointly with another household member or with a non-household member are in universe for EJO or TJO variables.

Level of information the Assets data provide

Most asset questions are asked of all household members who were aged 15 years and older at the time of the interview and are available as person-level variables. A few topics are collected at the household level, including educational savings accounts; primary residence (both mobile and non-mobile homes); cars, trucks, and vans; recreational vehicles; and rent, mortgage, and utilities payments.

Additional guidance for using the Assets data

Asset Value, Debt, and Income

All asset and debt values are as of the last day of the reference period (December 31st of the preceding year). An asset value of zero means that either the value was zero or the asset was no longer owned as of the last day of the reference period. A debt value of zero means that the respondent did not hold the debt as of the last day of the reference period. Reported income is the total amount of income received between the first and last days of the reference period.

Educational Savings Accounts

The household reference person is asked whether anyone in the household was the owner or a beneficiary of an educational savings account. If so, the survey then collects detailed information on up to three accounts with the largest balances, including the owner and the beneficiary of each account and the value of each account as of the last day of the reference period. Only those educational savings accounts whose owner resides in the household are included in the calculations of net worth.

Lump Sum or Regular Distribution Payments

The variable TDRAW_AMT was added to the 2019 SIPP as a follow-up to ERET_LUMPSUM. TDRAW_AMT lists the amount of lump sum or regular distribution payments received from an IRA,

Keogh, 401k, 403b, 503b, thrift plan, defined-benefit pension, and/or cash balance plan that was not reported in ERETTYP1YN-ERETTYP8YN.

The purpose of the TDRAW_AMT is to capture unreported retirement income among respondents 59 years of age or older as of the last day of the reference period who had never been retired or had been retired but did not report receiving retirement income in ERETANY, but did report receiving retirement income in ERET_LUMPSUM. Respondents that report receiving a lump sum payment from a retirement plan in both ELMPTYP1YN and ERET_LUMPSUM are excluded.[11]

This variable is imputed via Sequential Regression Multivariate Imputation (SRMI). For more information about SRMI procedures, see Section 6.2.1 of this guide.

Recreational Vehicles

SIPP asks the household reference person whether anyone in the household owned recreational vehicles and, if so, the type. It then collects information on the owners of each vehicle, vehicle value, and vehicle debt as of the last day of the reference period. If two of the same type of recreational vehicle are owned (such as owning two motorcycles), characteristics for the second recreational vehicle of that type are put into the 'other recreational vehicle' variables.

Other concepts related to the Assets data

Net Worth and Wealth: Net worth is the sum of asset values minus the sum of liabilities for a person or household. Wealth is typically defined as only the sum of asset values for a person or household.

Primary Residence: Primary residence is the residence at which household members live most of the time. In the SIPP Assets content, primary residence variables are split into *primary residence that is not a mobile home* and *primary residence that is a mobile home*. This is to capture differences between the two home types. In general, SIPP collects information on the value of the home or mobile home, the number of mortgages or loans (if any) on the property, and additional details pertaining to each mortgage or loan (such as the interest rate, whether it is fixed or variable, and if the debt is for the site if it is a mobile home).

4.2.2 Commuting and Work Schedules

The Commuting data include means of transportation to work, distance to work, minutes to work, parking and toll expenses, additional commuting expenses, and other job-related expenses. The Work Schedule data consist of days of the week worked, days of the week worked entirely at home, the start and end times of work, the type of schedule worked, and the reason for working said schedule.

The Commuting and Work Schedule data are collected at the spell level for each job the respondent reports in the EHC, up to seven jobs for the entire calendar year. This means that the data associated with a given job will be copied to every monthly record for the specific spell of work at the job. As with other

[11] For additional details about how ELMPTYP1YN, TLMPAMT, and TDRAW_AMT can be used together to create a more complete picture of retirement income, please refer to the user note, 2020 SIPP: Retirement Income at <https://www.census.gov/programs-surveys/sipp/tech-documentation/user-notes.html>.

data in the EHC, the commuting and work schedule information is collected in reverse chronological order. That is, each respondent's information is collected and recorded beginning with the most current job.

The Commuting section begins by asking the mode of transportation used to get to work (such as drove alone, took the bus, worked from home, etc.). The specific mode of transportation determines the subsequent commuting questions that each respondent is asked. If the respondent drove his or her own vehicle or a company car, the respondent is asked about miles driven to work, miles reimbursed, and parking/tolls. If the respondent commutes using another mode (and excluding worked from home), he or she is asked about other commuting expenses. With the exception of those who worked from home, workers are asked the one-way travel time to work in minutes. Finally, all workers are asked if they have any other job-related expenses not already reported, and, if so, the annual amount of these expenses.

The Work Schedule section immediately follows the commuting questions and begins by asking which days of the week the respondent worked for the given job, the start and end time of the job, and if there were any days worked entirely at home. If the respondent did work from home, he or she is asked to report the specific days worked from home. The section concludes with two questions on the type of work schedule and the reason for the work schedule that are asked of all respondents who report a job.

The same information is collected, for each respondent, for up to seven jobs, and for up to two spells per job for the entire calendar year. Prior to the 2014 SIPP Panel, the commuting and work schedule data were collected in separate topical modules that were not asked during the same wave. The commuting data were collected for up to three jobs held during the reference month, while the work schedule data were collected for up to two jobs held during the reference month.

Level of information the Commuting and Work Schedule data provide

Commuting and work schedule data are collected at the spell level in the EHC and provided to data users in person-month format.

Additional guidance for using the Commuting and Work Schedule data

The Commuting and Work Schedule variables are asked for each job spell held by a respondent and are edited in person-month format. Place of work geography is collected as an address for each job spell but is not released to the public. The Commuting and Work Schedule data also provide information about working at home.

Respondents can mark multiple modes of transportation. Among respondents who indicate multiple modes, a follow-up question identifies the primary mode of transportation. Subsequent questions relate to the primary mode of transportation. Respondents have the option of selecting 'Worked at home' as a mode of transportation

Questions about commuting distance and time both refer to the daily one-way trip and exclude respondents who selected 'Worked at home' as their primary mode of transportation. Commuting distance is collected in miles, and commuting time is collected in minutes.

Work schedule information includes days usually worked, usual times beginning and ending work, and reason for working the schedule. Respondents also receive a question about working at home at least one full day per week.

Multiple questions measure expenses. Commuters who drive are asked how many of the daily one-way miles driven to work are reimbursed by their employer. Respondents are also asked about the daily amount of parking expenses, the daily amount of other commuting expenses, and the annual amount of job-related expenses.

4.2.3 Employment and Earnings

Changes in labor force status, earnings, or usual hours of employment can all impact eligibility for and receipt of benefits. The longitudinal nature of SIPP allows it to measure movements in and out of the labor force, movement between jobs, and changes in earnings, making the labor force data critical. The detailed weekly nature of the employment data also allows for observation of short duration spells of employment and non-employment.

SIPP collects information about an individual's work history from the beginning of the reference year through the interview month, including:

- When a job was held
- Different types of employment earnings
- Employment and business characteristics
- Reasons the respondent did not hold a job
- Whether the respondent looked for work

The data are collected at a spell level, instead of the person-month format in which they ultimately appear. Most of the information about an employment or non-employment spell does not vary over the time it was held. This means that much of the information about a job spell will be copied to every monthly record in which the job was held. The exceptions are earnings sources that vary by month—bonuses, commissions, overtime, and tips—and recoded variables created in processing. In the 1990 to 2008 Panels, data were also collected at a spell level, but much of the information about a job spell was copied to every monthly record in the reference period, regardless of when that job was held.

For all of the variables regarding timing of an event about a job, the dates are collected in reverse chronological order to improve recall. However, on the output file the variables are available chronologically. So, each respondent's jobs are collected beginning with the most current job, but on the data file they start with the first job held in the reference year and end with the current job. The respondent may hold multiple jobs concurrently.

The Employment section begins by asking whether the respondent currently works for pay. If not, it asks whether the respondent worked at all since January of the reference year. Basic information about the job, such as beginning and ending dates, the type of arrangement (job for employer, self-employed business, or other), and the reason for the job ending (when applicable) are collected first. The next questions are characteristics of the job/business such as industry, occupation, union status, the number of employees, and incorporation status.

For industry and occupation, respondents are asked to provide kind of business or industry, kind of work, and usual activities at work. Industry describes the kind of business conducted by a person's employing organization, or own business if self-employed. Occupation describes the kind of work the person does on the job. Questions administered by Census personnel are open-ended, and respondents' answers about industry and occupation are recorded verbatim. These responses are converted to standard codes by Census Bureau coding staff during processing.[12]

To account for changes in the labor force, the Census Bureau periodically revises the code lists used to classify industry and occupation in the SIPP and other surveys. As a result, data users interested in comparing industry and occupation over time in the SIPP should be aware that different code lists are used depending on the SIPP panel. For instance, the 2014 Panel uses the 2012 Census Industry Code List and the 2010 Census Occupation Code List, while the 2020 SIPP uses the 2017 Census Industry Code List and the 2018 Census Occupation Code List.[13] Because different code lists are used between SIPP panels, the data are not directly comparable unless conversion rates for industry and occupation are first applied.[14]

Due to an unexpected issue in the 2018 SIPP, select employment variables could change across the reference year in a way that was inconsistent with internal survey specifications. Specifically, industry (TJBx_IND and TNJ_IND) and occupation (TJBx_OCC and TNJ_OCC) all could vary between January and December. However, industry should remain the same across all months of a job while occupation should remain the same across all months of a job spell according to survey specifications.[15] This issue was corrected in the 2019 data, which may result in fewer cases with industry and/or occupation codes.

In a small number of cases values for class of worker (EJBx_CLWRK) also changed across the year. Like occupation, the survey specifications indicate that class of worker can only change between job spells. Further, changes in class of worker are only possible between private jobs by profit/not-for-profit status, and between self-employed jobs by incorporation status. Additional improvements were made in the 2019 SIPP to enhance consistency of class of worker between January and December.

Next, the questionnaire asks about the types of earnings the respondent received (wages, salaries, commissions, tips, overtime, or bonuses) and the amounts earned. For wage or salary earnings, SIPP asks about current or most recent pay rate. Respondents can report up to two changes in pay rate over the job spell. For example, respondents may report changes in pay rate due to an annual raise or due to a

---

[12] For more information on how responses to the industry and occupation questions are coded, see the "General Overview of the Alphabetical Indexes of Industries and Occupations" page at <www2.census.gov/programs-surveys/demo/guidance/industry-occupation/overview2019.pdf>

[13] Additional information concerning the Census industry and occupation code lists, including documentation, crosswalks, and PUMS code lists are available online at <www.census.gov/topics/employment/industry-occupation/guidance/code-lists.html>.

[14] See *Recent Changes in the Census Industry and Occupation Classification Systems* for additional information about the Code Lists, conversion rates, and comparability between panels. This document is available at <www.census.gov/library/publications/2020/demo/acs-tp78.html>

[15] For more information about the issue in the 2018 SIPP, see the User Note on Inconsistent Industry and Occupation Codes <https://www.census.gov/programs-surveys/sipp/tech-documentation/user-notes/2018-usernotes/inconsistent-codes.html>

movement in and out of part-time status. SIPP also ask for the number of hours worked per week, and similarly allow the respondent to report up to two changes in the number of hours worked. If earnings from commissions, tips, overtime, and bonuses are received every month, then SIPP collects the data at the spell level. For example, a waiter who received tips in every month of his job would report how much he usually earned in tips. If earnings from commissions, tips, overtime, and bonuses are *not* received every month, then SIPP collects the monthly amounts. For example, a real-estate agent who received commissions in only some months on the job would report the commissions she earned in each month. Finally, SIPP asks respondents to report any time they were away from the job without pay within the reported job spell.

The same questions are asked for each job, for up to seven jobs. This practice is consistent with SIPP's data collection practice from the 2014 Panel but differs from the 1984 to 2008 Panels. In the 1984 to 2008 Panels, SIPP only collected data for the first two jobs and sorted them by which one was held for more weeks or hours. Beginning with the 2014 Panel, no "main" job is determined, and jobs are sorted on the public-use file by the earliest month they were held within the reference period. Ties are broken by which job was reported first by the respondent. Similarly, moonlighting is no longer distinguished separately from any other job. Respondents with a main job and an additional part-time job will simply have two jobs listed concurrently. In addition, the classification of "contingent workers" is no longer used. The questions about whether a job has an "other" work arrangement and whether that arrangement was definite are still present, but they are applied to a particular job and not to an individual. In this way, SIPP can identify occasional or indefinite work that happens in addition to any main job held.

If an individual held more than seven jobs during the reference period, some summary information is collected about the additional jobs. That information is included with the employment status recode and the earnings recodes (e.g., RMESR, TPEARN) but is not included on the edited file directly.

Jobs may be linked across waves using EJB(n)_JOBID. If a job was held at the time of the interview, then during the subsequent interview the respondent will answer questions about the job and earnings during the next wave. Some responses are carried forward (such as industry and work arrangement), while others are asked again (such as occupation and earnings).

Finally, during periods when respondents are not employed, SIPP collects information about their labor force status. This includes information about why they were not working: unpaid work in a related household member's family business or farm for 15 hours or more per week,[16] time spent on layoff, and time spent looking for work. To facilitate consistency in reporting, if the respondent held a job during part of the year, the survey instrument will calculate gaps in employment automatically and collect information about each spell of non-work separately. Note that it is possible for separate spells of non-employment to be present in a single month if there is a short job spell within the month. Because the data are mapped to a person-month format, this can lead to multiple spells of non-employment being mapped to the same month. In these rare cases, identified by the variable ROVERLAPMN, the output file will have some characteristics from both spells of non-employment, and the timing variables will relate to the earlier spell.

[16] Questions about industry and occupation also are administered to unpaid family workers.

For data users' convenience, the public-use data file provides both recoded variables and the components used to create those recodes. Examples of recoded variables include weekly (RWKESR1-5) and monthly employment status (RMESR), weekly wage and salary earnings (TJB(n)_WKSUM1-5), monthly total earnings from each job (TJB(n)_MSUM and TJB(n)_MSUMALT), monthly total earnings from all jobs (TPEARN and TPEARN_ALT), hours worked per week across all jobs (TWKHRS1-5), and monthly average hours worked per week across all jobs (TMWKHRS).

Level of information the Employment and Earnings data provide

All data related to employment status are available at the weekly level. All information about spells of employment and non-employment are included on the monthly records, and so monthly statistics can be calculated for all items.

Most Employment and Earnings data is provided at the person-month level. However, the job screener flags indicating the presence of job (n) during the reference year (EJB(n)_SCRNR) are person-level variables.

Additional guidance for using the Employment and Earnings data

All of the labor force data are collected in the EHC. Each job gets its own line in the calendar, and periods of time not working are recorded on a separate line.

The Away Without Pay question specifies that the time away should be two weeks or more. However, if the respondent reports only a single week away without pay, we do include that information on the file.

Other concepts related to the Employment and Earnings data

Class of Worker: The ownership type of a respondent's employer, or the respondent's own business if self-employed. This variable identifies whether the respondent is self-employed, works in the private sector, or in government. The class of worker category is, in most cases, independent of industry and occupation.

Commission: A form of payment based on the amount of services the employee performed. For example, in sales positions, many times the sales employee is paid a percentage of the amount the customer paid or a percentage of the profit received by the firm. SIPP also asks respondents to include broader pay for unit of service when reporting commission amounts, which could include payments more commonly known as piece-rates. It could also include other types of pay schemes, such as painters who are paid by the room or truck drivers who are paid by the mile. Commission can be in lieu of wage and salary payments ("straight commission") or in addition to these payments.

Earnings: Remuneration (pay, wages) of a worker for services performed during a specific period of time. SIPP focuses on reporting monthly earnings, but for wage and salary earnings, data are available at the weekly level.

Employed Persons: Persons who report holding a job, business, or doing any other work for pay during the time period specified. This includes those who are temporarily absent from a job, whether the absence was paid or unpaid. Also included as employed are those who worked 15 hours or more unpaid in a related household member's family business or farm.

Full-Time Workers: Those who usually work 35 hours or more per week.

Industry: Describes the kind of business conducted by a person's employing organization, or own organization, if self-employed. Industry data are derived from questions about employer/business name, kind of industry.

Labor Force: The sum of those classified as either unemployed or employed.

Occupation: Describes the kind of work that the individual does on the job. Occupation data are derived from questions about kind of work and usual activities at work.

On Layoff: Persons are considered to be "on layoff" if they are currently unemployed due to slack work, shortages, or other business-related concerns (rather than personal misconduct). Persons on layoff who are expecting to be recalled to their positions are considered unemployed whether they looked for work or not during the layoff.

Part-Time Workers: Those who usually work fewer than 35 hours per week.

Unemployed Persons: Persons who had no employment during the week and were either on layoff awaiting recall or looking for work.

Wages and Salaries: Payments per a unit of time, generally not tied to performance. Wages are payment per hour of work; salaries are usually given as an annual amount. Both wages and salaries can be paid in a number of ways; SIPP allows them to be reported as weekly, biweekly, monthly, or twice per month, as well as reporting the hourly wage, annual salary, or gross annual amount. Any other pay schedules are asked to report monthly averages.

Weekly Hours: The number of hours the respondent usually (more than half the time) worked per week at their job during the period specified. Respondents whose hours vary are instructed to attempt to give an average amount.

4.2.4 Parental Leave

The SIPP instrument collects parental leave information for all respondents who have given birth to or fathered at least one child. Available data for these respondents include:

- Whether they worked for pay at any time during the pregnancy (EPREGWORK)
- Whether they continued working right up to the birth (EPREGSTOP)
- What type(s) of leave they used after the child was born and up to the time the baby was 12 weeks old (EBIRTHRSN1-10)
- How long (in months) prior to the birth they stopped working (EPREBIRTHINT)
- What type(s) of leave (if any) they used prior to the birth (EPREGRSN1-10)
- Whether they worked at any time after the birth (EBIRTHWORK)
- How long (in months) after the birth they started working (TPSTBIRTHINT)

Universes for individual variables vary based on responses to prior items in the question sequence – for example, EBIRTHRSN1-10 are only in universe for those that continued working right up to the birth.

Level of information the Parental Leave data provide

Parental leave data are available at the person level for all respondents who have given birth to or fathered at least one child (TCEB>0). All data are collected in reference to the birth of the respondent's first child.

Additional guidance for using Parental Leave data

The 2008 SIPP panel and other prior SIPP panels included questions regarding maternal leave. Parental leave questions were not included in the 2014 SIPP panel or in the 2018 SIPP. The revised question sequence outlined above was added to the survey beginning in 2019 SIPP and was asked of both men and women. There are few nationally representative sources that track use of parental leave in the United States, which does not have a federally guaranteed parental leave policy.

Parental leave questions are only asked in reference to biological children. Parental leave information is collected in the back end of the survey, after Assets, Disability, and Fertility, but before any of the Child Care or Child Well-Being questions.

4.3 Health and Well-Being

4.3.1 Adult Well-Being

SIPP collects information on certain characteristics of the household (e.g., are there cracks in the ceiling or walls or holes in the floor), surrounding neighborhood (e.g., is the neighborhood safe from crime), and food security (e.g., did the household ever not have enough food). These questions are intended to be answered by the household reference person; however, any household member aged 15 years and older may act as a proxy. These measures offer additional insight into living conditions that more common measures (such as income or wealth) are unable to provide.

Level of information the Adult Well-Being data provide

All Adult Well-Being questions are asked of the reference person and are available at the household level, meaning the answers provided by the household respondent apply to the entire household. These questions refer to conditions experienced during the reference year. For respondents who lived in multiple locations during this period, the questions refer to the residence in which the reference person lived the majority of the time.

Additional guidance for using Parental Leave data

The information on adult well-being and food security can be used to examine differences in living conditions, financial hardships, and food security among the population. The food security component of these data includes six yes or no questions about the household's access to food, which are used to produce two recode variables. The recode variable RFOODS counts the number of these questions to which the reference person responded 'yes', indicating a challenge to food security. The recode variable RFOODR categorizes these counts into three categories: low (0-2 yes responses), medium (3-4 yes responses), or high (5-6 yes responses) food insecurity.

4.3.2 Child Care

SIPP collects information on child care arrangements for children under the age of 14. The 2014 SIPP redesign minimizes respondent burden by streamlining the question format. Instead of running through a series of questions for each child, the questionnaire is organized by arrangement type. For example, respondents are first asked whether they used grandparent care for any child; if so, they are then asked which child used grandparent care. Data are edited to store child care data on the parent and the child's record.

The Child Care and Child Well-Being sections of the survey utilize a concept labeled 'reference parent.' This is the person who will answer the questions about the children. In households in which both parents are present, the mother is the reference parent. If the mother is not available for an interview, the father of the child can give proxy responses for her. Priority is given to biological mothers, then other mothers, then biological fathers, and then other fathers. In single-parent families, the resident parent is the reference parent. If neither parent is in the household, the householder is the reference parent. The child care questions ask the respondent to identify the child care arrangements used while the reference parent was working or in school. If the reference parent was not working or enrolled in school in the fall (September-December) of the reference year, then the respondent is asked to identify the child care arrangements used while the reference parent was otherwise busy.

Note that Child Care data from the 2014 Panel through 2020 SIPP are not comparable to data from the 2008 and prior SIPP panels due to changes in data collection and data processing. In 2008 and prior SIPP panels, the parent was asked about each form of child care used for each individual child in the household with the reference period being a typical week in the previous month. In addition, the parent was asked about the number of hours the child spent in each arrangement. Beginning with the 2014 SIPP Panel, for each type of child care arrangement, the reference parent was asked whether any child living in the household regularly used the reported type of care in a typical week in December of the reference year, which could be up to 6 months before the interview. If a parent had more than one child in universe for the reported type of child care, a follow up question asked them to identify all children using that arrangement. Beginning with the 2019 SIPP, the reference parent was asked about child care during the fall of the reference year, which could be up to 11 months before the interview.

Level of information the Child Care data provide

The data are edited so that data are available on the reference parent's record as well as the child's record. They are available at the person level. From the reference parent's record, measures can be extrapolated to the couple level, family level, household level, and national level. The child care questions are asked about care arrangements used in the Fall (September-December) of the reference year.

Additional guidance for using Child Care data

The child care variables starting with an E are the edited variable stored on the parent's record. Variables starting with an R are recoded variables that are stored on the child's record and are based on the original variable stored on the parent's record. The variable ERP identifies an adult as a reference parent (1=yes, 2=no), and EREFPAR is the child-level variable that identifies the PNUM of the reference parent.

Other concepts related to the Child Care data

Child Care: Child can be in parental, relative, and nonrelative child care arrangements.

Child care assistance: Either federal or employer assistance.

Head Start: A federally subsidized child care program.

4.3.3 Child Well-Being

The Child Well-Being section covers a variety of topics such as shared meals, reading, school engagement, outings, and activities with nonresident parents. If the mother is not available for an interview, the father of the child can give proxy responses for her. Priority is given to biological mothers, then other mothers, then biological fathers, and then other fathers. In single-parent families, the resident parent is the reference parent. If neither parent is in the household, the householder is the reference parent.

Level of information the Child Well-Being data provide

Child well-being data is stored on the reference parent's record and then a recode is created and placed on the child's record. Variables regarding meals, outings, reading, and religious service attendance are global questions and cover engagement with all children, not engagement child by child. Therefore, whatever value is on the parent's record is carried directly over to the child's record. For example, for the variable EDINRPAR, the reference parent with two children ages 5 and 10 tells us that she eats dinner with her children 5 nights a week. The value on EDINRPAR is then copied on to each child's record as a recoded variable (RDINRPR).

The child well-being variables are available at the person level, and they can be used at a child level, reference parent level (household level), or national level.

Additional guidance for using the Child Well-Being data

The Child Well-Being questions fall after the Event History Calendar, following the Child Care section. The universe of respondents consists of adults who are parents of children under 18 years old (ERP). In households in which both parents are present, the mother is the reference parent.

Variables starting with an E are the edited variable stored on the reference parent's record. Variables starting with an R are recoded variables that are stored on the child's record and are based on the original variable stored on the parent's record. The variable ERP identifies an adult as a reference parent (1=yes, 2=no), and EREFPAR is a child-level variable that identifies the PNUM of the reference parent.

4.3.4 Dependent Care Expenses

In multi-person households, the Dependent Care Expenses section asks the household reference person if during December of the reference year anyone in the household paid for the care of a child or a disabled person so that a household member could attend work, attend training, or look for a job.

Level of information the Dependent Care Expenses data provide

All dependent care expense questions are asked of the household reference person and are available at the household level.

4.3.5 Disability

The Disability content covers whether the respondent has difficulty performing certain activities due to a physical, mental, or emotional condition. The section contains six core questions about difficulty with daily tasks, along with six additional questions specific to children or the working-age population. The six core questions ask about hearing difficulty (EHEARING), seeing difficulty (ESEEING), cognitive difficulty (ECOGNIT), ambulatory difficulty (EAMBULAT), self-care difficulty (ESELFCARE), and independent living difficulty (EERRANDS). This set of six questions is consistent with the standard disability question set implemented across multiple government surveys, including the American Community Survey (ACS) and the Current Population Survey (CPS). Consistent with other surveys, the age universes for the core disability items differ. While EHEARING and ESEEING apply to all individuals eligible for an interview, ECOGNIT, EAMBULAT, and ESELFCARE only apply to individuals aged 5 and older and EERRANDS is only asked of individuals aged 15 and older.

The additional child disability questions ask about developmental delays (for children 1 to 5 years old), difficulty playing with other children (5 to 14 years old), and difficulty with schoolwork (5 to 14 years old).

The work-limiting disability questions apply to individuals aged 15 years and older and cover difficulty finding or keeping a job (for individuals under age 70), limitations in the kind or amount of work possible, and being prevented from working at all.

To help data users, two disability recode variables are constructed as summary measures of disability:

- RDIS is a standard summary measure of disability that has been used across government surveys since 2008, including the ACS and the CPS. The measure is based on responses to the six core questions about difficulty with daily tasks: ESEEING, EHEARING, ECOGNIT, EAMBULAT, ESELFCARE, and EERRANDS. RDIS=1 indicates that the respondent has difficulty with at least one of the six tasks.
- RDIS_ALT is an alternative summary measure of disability that incorporates the child disability questions, EDDELAY, EPLAYDIF, and ESCHOOLWK, and the work-limiting disability questions, EFINDJOB, EJOBCANT, and EDISABL. RDIS_ALT=1 indicates that the respondent either has difficulty with one or more of the six daily tasks used for the standard summary measure of disability OR has a value of 1 for at least one of the six child or work-limiting disability variables.

Level of information the Disability data provide

All data related to disability status are available at the person level and represent status at the time of the interview.

Additional guidance for using the Disability data

Most of the disability questions are located within the same section of the interview. However, one work-limiting disability question, limitations in the “kind or amount of work” a person can do (EDISABL), is asked prior to the main Disability section as a lead-in to questions about the receipt of disability income.

The format of the main Disability section is as follows: first, respondents are asked if they have any difficulty seeing or hearing. All other questions are then presented in a series based on the age category to which the respondent belongs. While disability data are collected for both children and adults, disability data are not collected for individuals within the household under age 1.

Additionally, the concept of work disability may be captured through the variables EFINDJOB, EJOBCANT, and EDISABL.

4.3.6 Health Care Utilization and Medical Expenditures

The *Health Care Utilization* content collects information about:

- Respondent's health status (excellent, very good, good, fair, or poor)
- Number of days sick in bed and hospitalized
- Prescription medication use
- Number of visits to dentists and medical providers
- For uninsured respondents, visits to medical providers or dentists, and source (e.g., emergency room)

The *Medical Expenditures* content collects information about respondents' medical expenses, including:

- Health insurance premiums
- Medical care, prescriptions, or medical supplies
- Over-the-counter (OTC) medical items
- Whether the respondent had a Flexible Spending Account

Level of information the Health Care Utilization and Medical Expenditures data provide

All data related to health care utilization and medical expenditures are available annually at the person level. Expense and utilization questions refer to the reference calendar year.

Additional guidance for using the Health Care Utilization and Medical Expenditures data

It is important to understand the significant improvements to the Medical Out-of-Pocket questions, beginning in 2018 SIPP, in order to use these data correctly.

- The question on amount paid for health insurance premiums (THIPAY) was split into two separate questions, in order to separate out amounts paid for comprehensive health insurance premiums (THIPAYC) versus supplemental health insurance premiums (THIPAYS).
- The framing of the health care and OTC expenditure questions changed to align with the premium questions. In 2014, the health care and OTC questions asked the amounts paid *for you, by anyone* in the household. In 2018, these now ask for the amounts paid *by you, for anyone* in the household.
- Four questions were added to capture who in the household these payments were for (TWHIPAYC, TWHIPAYS, TWMDPAY, TWOTCMDPAY). After each amount question, the respondent is asked: "Who in the household did you pay for? (Mark all that apply)"
- Question text on what medical costs included were moved from the question text into optional text for the interviewer to read if clarification was needed.

- Consistent with Panel 2014, the universe for all amounts questions is essentially all persons (see SIPP ecodebook for minor participation constraints). This broad universe is intended to align with other surveys such as the CPS. Persons who would not be expected to have paid expenses, either because of age (under age 15) or because the family was uninsured all year (in the case of premium questions only), have values set to $0 and allocation flags set to '3: logical assumption'. As a result of this broader universe, the total imputation rate will appear higher.

Note on Infants

In the 2020 SIPP, infants (defined as household members less than one year old as of December of the calendar year) do not receive proxy interviews. As a general strategy, infants without data are imputed. However, for the Health Care Utilization and Medical Expenditures section, infants are expected to have unique health care needs, since they have unusually high medical utilization and expenditures compared to other groups. Due to these differences between infants and the rest of the population, we do not impute health care utilization and medical expenditure information for those who are less than one year old. Values for these variables will be set to 'not in universe' except for the two exceptions described below.

1. The questions for the amount paid out-of-pocket for medical expenses indirectly collects information on infants, by asking for amounts paid *for others in the household* for health insurance premiums, health care expenditures, or over-the-counter products.

2. For the variable EHLTSTAT, we impute values for infants. This variable collects self-reported (or proxy-reported) health status. We impute values for this variable because we do not have any reason to believe that reported health status of infants is fundamentally different from that of other young children.

4.3.7 Health Insurance

The 2020 SIPP data provide information on health insurance coverage. These data can be used to measure and track the percentage of the population with private health coverage, medical assistance (Medicaid), Medicare, military coverage, and other types of health insurance. Among other things, these data provide information on who the plan holder is, whether the employer pays for premiums, and whether the plan was purchased in a marketplace/exchange. It also allows users to look at how health insurance coverage changes over time, either on its own or in conjunction with other variables.

For each type of coverage, SIPP asks whether a respondent is currently covered by health insurance or was covered at any point during the reference year. If so, the person is asked to provide a beginning and ending month for that time (spell) of coverage. SIPP captures different spells of coverage within the year, each with complete plan details. For example, private insurance questions ask about up to two simultaneous sources of coverage in any given month, with a follow-up question about the more specific type and/or source. Starting in Wave 2, the survey may bring forward health insurance coverage data from the prior wave. Respondents who had ongoing insurance coverage as of Wave 1 may be reminded of their

previous responses and can indicate whether the insurance spell they reported at the last interview is still ongoing.

Respondents who had direct purchase, Medicaid, or 'other type' of insurance may report that their coverage was obtained through <www.Healthcare.gov> or a state-based exchange, and whether that coverage had a premium and subsidy. Marketplace coverage is classified as private direct-purchase coverage, but the data collection methodology acknowledges that some respondents perceive Marketplace coverage as a public program, while others perceive it as private health insurance. We do not constrain respondent reports of exchange-based coverage by eligibility criteria. For example, a respondent who reported income above their state's Medicaid eligibility requirements may still report that they have Medicaid obtained through <www.Healthcare.gov> or a state-based exchange. Data users may use their own judgement in how to interpret these data.

What level of information does the Health Insurance content provide?

Health insurance coverage information is available at the person-month level. This information can be used to measure health insurance rates or health insurance changes over time either on its own or in conjunction with other variables.

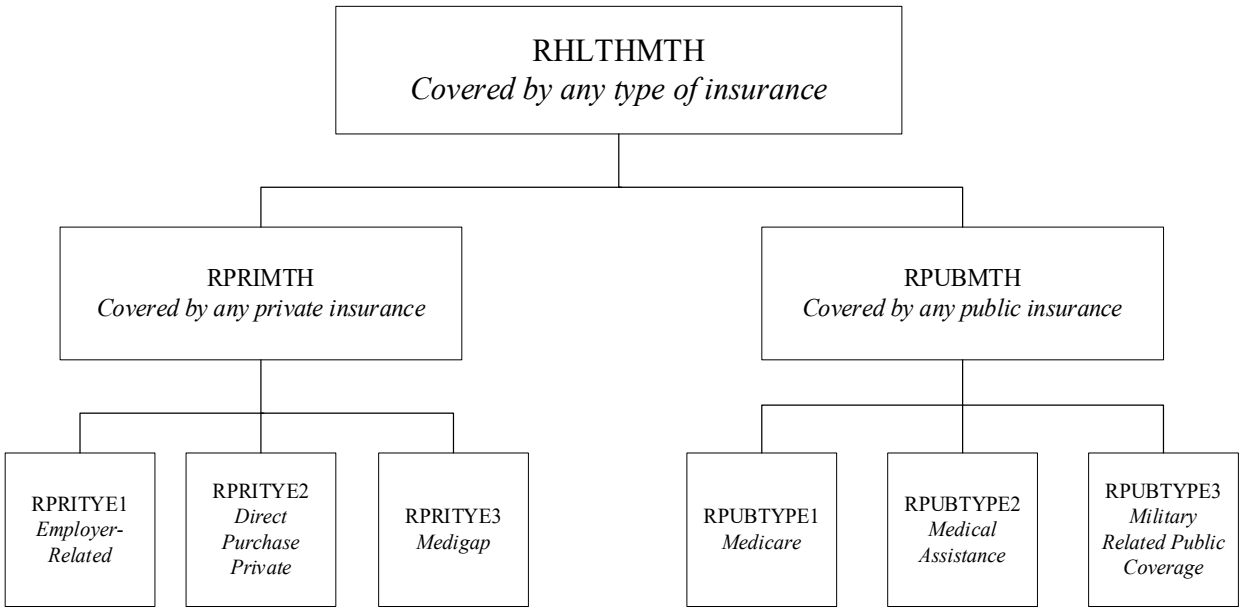
Additional guidance for using the Health Insurance data

Health insurance coverage variables in the 2020 SIPP

There are numerous health insurance variables in the 2020 SIPP data, corresponding to the many health insurance questions asked within the survey. For convenience, the dataset also includes nine high-level recode variables that indicate who had any health insurance, whether that insurance was public or private, and what type of insurance it was.

As Figure 4-2 below shows, variable RHLTHMTH measures all types of comprehensive health insurance, while RPRIMTH and RPUBMTH measure private and public coverage, respectively. Six additional variables, RPRITYPE1-RPRITYPE3 and RPUBTYPE1-RPUBTYPE3, help data users identify any type of coverage a person may have had, be it employer-related, direct purchase, Medigap, Medicare, Medical Assistance and military related public coverage (e.g., CHAMPVA and VA Care). Indian Health Services (IHS) and single service plans such as prescription drug, vision, or dental plans are not considered comprehensive health insurance plans and are not included in any of these nine variables.

FIGURE 4-2. HEALTH INSURANCE COVERAGE IN THE 2020 SIPP: NINE HIGH-LEVEL VARIABLES



Health Insurance Recodes for the 2018 Through 2020 SIPP

Table 4-2 shows how the coverage types collected in the EHC correspond to health insurance recode variables.

TABLE 4-2. HEALTH INSURANCE RECODES BY TOPIC AREA IN THE 2020 SIPP

EHC Categories			2018-2020 Recode Variables	
Topic	**Type**		**Private/Public**	**Private/Public Type**
Private Coverage	1	(Current) Employer	RPRIMTH	RPRITYPE1
	2	Former Employer		
	3	Union/Association		
	4	Bought it directly		RPRITYPE2
	5	School		
Medicare	1	Part A (Hospital)	RPUBMTH	RPUBTYPE1
	2	Part B (Medical)	RPUBMTH	
	3	Medigap (Supplement)	RPRIMTH	RPRITYPE3
	4	Part C (Advantage Plan)	RPUBMTH	RPUBTYPE1
	5	Part D (Rx Coverage)	N/A	N/A
Medicaid	N/A		RPUBMTH	RPUBTYPE2

EHC Categories			2018-2020 Recode Variables	
Topic	**Type**		**Private/Public**	**Private/Public Type**
Military	1	TRICARE or CHAMPUS	RPRIMTH	RPRITYPE1
	2	CHAMPVA	RPUBMTH	RPUBTYPE3
	3	VA		
Other	1	Gov't/public assistance	RPUBMTH	RPUBTYPE2
	2	Through someone else (parent/relative)	N/A	ROTHCOVMTH
	3	Indian Health Services	N/A	RIHSMTH

Note on Infants

Previously, infant coverage was only captured if they were reported under another household member's plan. Beginning in 2018, questions were added to capture health insurance information about infants (individuals who are less than one year old by December of the reference year) via a proxy adult. The new infant questions ask whether each infant is currently covered by any health insurance, and if so, what type(s). As a result of asking about infant coverage directly, the imputation rates have improved due to more reported data available.

4.4 Program Participation and Income Transfers

Program participation information collected in the EHC (GA, SNAP, SSI, TANF, and WIC) is asked in reverse chronological order. That is, respondents are first asked if they or someone in their family are currently receiving the program. If a respondent reports current receipt, the month of the reference period or interview year receipt began is collected. If no one in the respondent's family is currently receiving the program, respondents are then asked if they received it at any time since the first month of the reference period. If receipt is reported during that time, the months that receipt began and ended are collected. If a program spell includes January of the reference period the first time a respondent is interviewed, information on the year the respondent first began receiving benefits is collected.

A continuation flag for respondents who reported receiving the program in the last month of the reference period provides information on whether the spell: (1) ended in the last month of the reference period; (2) ended after the last month of the reference period but before the interview month; or (3) was ongoing as of the interview month.

For each spell, respondents are asked to report the person in whose name the benefit is received (benefit owner), who was covered by the benefit, why receipt began, why receipt ended (if applicable), and payment amount, as this information may vary across spells.

Respondents may report up to four different payment amounts in a spell to allow for changes in payments resulting from factors such as cost of living adjustments and earnings variation. The amounts are recoded into a single monthly variable.

4.4.1 Annuity/Life Insurance Retirement Income

SIPP collects person-level data for respondents aged 30 or older who reported currently being retired, ever being retired, or receiving retirement income from a paid-up life insurance policy or annuity during the reference year.

Specifically, the Annuity/Life Insurance Retirement Income content collects:

- Whether the payment was received as a lump-sum
- Each month of receipt
- The total amount received during the reference year

Level of information the Annuity/Life Insurance Retirement Income data provide

Annuity/Life Insurance Retirement Income is collected at the person level for all household members who were age 30 or older and who reported currently being retired or ever being retired during the reference year. Individuals under age 30 or who reported not ever being retired are not in universe for the Annuity/Life Insurance Income Benefit section. The Annuity/Life Insurance Retirement Income data are available as person-level variables.

Additional guidance for using the Annuity/Life Insurance Retirement Income data

The Annuity/Life Insurance Income Benefit section collects up to two types of life insurance income benefits, total amount received, and the month(s) when the payments were received.

Other concepts related to the Annuity/Life Insurance data

Annuity: An insurance product purchased by a policyholder at his or her retirement or during his or her working years to guarantee a regular income after retirement. Some annuity contracts provide guaranteed distribution for a set number of payments, and some provide payment until death. SIPP includes all such annuities purchased by the person, except those already reported as part of the employee's pension.

Paid-up Life Insurance Policy: A whole life insurance policy, universal life insurance policy, or permanent policy that provides a benefit upon death of the policyholder and accumulates cash value over time, enabling benefits to be paid in the event that the policyholder voluntarily terminates the insurance policy before death or the insured event occurred. The paid-up value is the sum of money that the insurance company will pay to the policyholder in the event that the policyholder voluntarily terminates the life insurance policy.

Retirement Income: Payments received based on length of service at retirement, age of retirement, and other employer retirement requirements. Examples of retirement income include U.S. Government Railroad Retirement; pension from a company, union, federal civilian employee pension; U.S. military retirement pay; retainer pay; National Guard retirement, Reserve Forces retirement; state government pension; local government pension; and any pension type other than Social Security.

Retirement Income Life Insurance: Individuals with a sufficient number of years of service or who have a disability may retire before reaching age 65 with a pension plan that includes life insurance. Individuals employed in hazardous occupations such as law enforcement may be eligible to retire with a pension that includes a paid-up life insurance policy or annuity as early as age 40 if they have completed 20 or more years of service.

4.4.2 Disability Income

SIPP collects person-level data about whether an individual between the ages of 15 and 69 received disability income at any time during the reference year. The data collected in the 2020 SIPP include the type of disability payment and the amount received during each month of the reference year. The disability income content includes the monthly amount received by benefit type.

The Disability Income section begins by asking if the respondent received any income due to a health condition during the reference year. If so, respondents are then asked which type(s) of disability benefit was/were received:

1. Payments from a sickness, accident, or disability insurance policy
2. Employer disability payments
3. Pension from a company or union including income from a profit-sharing plan
4. Federal Civil Service or other Federal civilian employee pension
5. State government pension
6. Local government pension
7. U.S. Military retirement pay
8. U.S. Government Railroad Retirement
9. Black Lung benefits
10. Other disability income

Level of information the Disability Income data provide

Disability Income data is collected at the person-month level for all household members who were between the ages of 15 and 69 during the reference year. Household members under the age of 15 at the end of the reference year and those over the age of 69 at the beginning of the reference year are not in universe for the disability income section. The Disability Income data are available as person-level variables.

Other concepts related to the Disability Income Data

Black Lung Benefits: The United States Government program enacted in 1973 to provide monthly payments and medical benefits to coal miners who have disability from pneumoconiosis as a result of employment in and around the United States' coal mines. The act also provides monthly benefits to a miner's dependent survivors if black lung disease caused the miner's death.

Disability: An impairment that may be physical, cognitive, mental, hearing, vision, sensory, emotional, developmental, chronic disease, or some combination thereof.

Disability Insurance: An insurance policy that provides the policyholder benefits in the event that he or she becomes sick or injured.

Federal Civilian Employee Pension, or Federal Employees Retirement System (FERS): A civilian retirement pension for civilian federal employees who began their employment after 1987. It may also provide disability benefits.

Federal Civil Service Pension: A retirement pension for civilian federal employees who began their civilian federal employment between 1920 and 1987. It may also provide disability benefits.

Local Government Pension: An employee retirement plan for vested local government employees. Such plans often provide disability benefits.

Military Retirement Pay: The retirement pension plan provided for active-duty military personnel with benefits that start when the employee retires. It potentially provides disability benefits.

National Guard Retirement: An Army National Guard retirement plan provided for employees who have 20 years of service in the National Guard. By age 60, this retirement pension may be combined with any additional income or retirement. National Guard Retirement potentially provides disability benefits.

Other Disability Income: Supplementary income in the event of illness or accident that prevents the insured from working at their employment.

Pension: An employee benefits plan in which employees and employers pay into a group trust fund. The payment benefits are based on a formula factoring an employee's final pay and employment tenure with an employer.

Reserve Forces Retirement: A retirement pension plan for service members who complete 20 years of service and who become eligible for retirement from the Reserve Forces at age 60. Reserve service members may begin retirement benefits before age 60 if they deploy for war or national emergency. Effective January 2008, for every consecutive 90 days that service members are mobilized, their annuity start date is reduced three months.

Sickness and Accident Insurance: A type of disability insurance that provides coverage when the policyholder becomes sick, injured, or dies from an accident.

State Government Pension: An employee retirement plan for vested state government employees. Such plans often provide disability benefits.

U.S. Government Railroad Retirement: A federally administered retirement program established for railroad employees and employers. The program provides retirement, unemployment, sickness, disability, spousal, and survivor benefits. Tier 1 benefits take the place of Social Security and are first payable at age 62 or after 30 years of service for early retirement. The first full retirement benefits are payable for retired employee between the ages of 65 and 67. The Tier 2 benefits are similar to a private pension retirement plan.

4.4.3 Energy Assistance

The Energy Assistance section collects information from the household respondent about the payment of utilities and receipt of energy assistance at any time during the reference period, including:

- Whether the household paid separately for utilities (for households receiving a rental subsidy or housing voucher; EUTILITIES)
- Whether the household received energy assistance (EENERGY_ASST)
- Whether assistance was in the form of checks sent to the household (EENERGY_PMT1); coupons or vouchers sent to the household (EENERGY_PMT2); and/or payments sent directly to the utility company, fuel dealer, or landlord (for households that reported receipt of energy assistance; EENERGY_PMT3)

Level of information the Energy Assistance data provide

Energy assistance data are available at the household level. The reporting period is the entire reference period.

Other concepts related to the Energy Assistance data

Energy Assistance: Government energy assistance programs that help pay heating and cooling costs. This assistance is often paid directly to the electric company, fuel dealer, or landlord. Eligibility varies slightly from state to state, but all recipients have limited incomes.

4.4.4 General Assistance (GA)

The General Assistance (GA) section collects information about benefit receipt from the start of the reference period through the interview month. For each GA benefit unit, one adult in the household is considered the owner of the GA benefits (the person in whose name the benefits are received). The GA owner may or may not be covered by the GA benefit.

As shown below, all people are assigned values for the reference period and monthly coverage indicators. People who are covered but are not the owner are assigned a value for an additional variable that points to the benefit owner and provides a link to details about the spell. Then, the benefit owner has values assigned for all the variables applicable to the spell.

<u>Universe: All people</u>

- Reference period indicator (RGA_YRYN)
- Monthly coverage indicator (RGA_MNYN)

<u>Universe: Benefit owner and people who are covered by the benefit</u>

- Benefit owner (EGA_OWN)

<u>Universe: Benefit owner</u>

- Begin and end month of spell (EGA_BMONTH and EGA_EMONTH)
- Year receipt began if receiving in January of the reference period of a Wave 1 interview (RGA_LCYR)
- Continuation flag for spells that include the last month of the reference period (RGA_CFLG)
- Reason receipt began (EGA_BRSN)
- Reason receipt ended (EGA_ERSN)

- Payment amount (TGA_AMT)

Data users should be aware that variable names were standardized across the EHC programs (SNAP, WIC, SSI, TANF, and GA) starting in the 2018 SIPP. To see a side-by-side comparison of GA variable names in the 2014 SIPP Panel and the 2018 SIPP, please see Section 4.4.4 in the 2018 Survey of Income and Program Participation Users' Guide.

Level of information the GA data provide

GA data are collected at the spell level in the EHC and are provided to users in a person-month format. While all data are in a monthly format, the value of the variables may stay the same or vary over time.

Same value for the reference period

RGA_YRYN

Same value for duration of spell

RGA_MNYN, EGA_BMONTH, EGA_EMONTH, EGA_OWN, RGA_LCYR, RGA_CFLG, EGA_BRSN, EGA_ERSN

Value may vary within a spell

TGA_AMT

Other concepts related to the General Assistance data

General Assistance or General Relief Programs: General Assistance programs are funded and administered at the state and local government levels. Some states do not have General Assistance programs, and some states that have General Assistance programs allow localities to choose whether or not to implement the programs. States might also have different names for their programs. The programs serve people who do not qualify for federal assistance. The populations these programs serve vary by locality. Examples include the disabled, the elderly, the unemployed, children, families with children, and employable individuals without children. The needs of each recipient may also vary. Assistance may be a one-time emergency payment or a regular monthly benefit and/or payment. Payments received may be in cash or in the form of a voucher. Also, the benefit and payment received may be intended for a specific need, such as medical expenses, burial expenses, or to meet general needs.

4.4.5 Lump Sum Severance Pay/Retirement Plan Income

SIPP collects person-level data from respondents aged 15 years and older about receipt of lump sum severance pay and retirement plan income at any time during the reference year.

For respondents who report receiving lump sum severance pay/retirement plan income, SIPP collects the type of payment:

- Pension/retirement pay
- Severance pay
- Deferred payment/final paycheck

- Something else (respondents are then asked to specify the type)

Respondents are then asked to report the total amount received during the reference year. Then respondents are asked if any of that money was rolled over to an IRA or other kind of retirement plan. For respondents who rolled over money (or who plan to), the amount rolled over is then collected.

Level of information the Lump Sum Severance Pay/Retirement Plan Income data provide

Lump Sum Income Benefit information is collected at the person level for all household members who were aged 15 years and older during the reference year. Children under the age of 15 by the end of the reference year are not in universe for the Lump Sum Income Benefits section. The Lump Sum Severance Pay/Retirement Plan data are available in the form of person-level variables.

Additional guidance for using the Lump Sum Severance Pay/Retirement Plan Income data

The Lump Sum Income Benefits section collects up to four types of lump sums, total dollar amount received from the lump sum income benefits, two roll over lump sum income benefit types, and the total dollar amount rolled over.

Other concepts related to the Lump Sum Severance Pay/Retirement Plan Income data

Deferred payment: An employer or employee arrangement in which the compensation is paid after the 1- or 2-week span that the income is earned. Pensions and retirement plans are examples of deferred payment plans.

Individual Retirement Account (IRA): A personal retirement plan that allows employees, self-employed individuals, and other individuals to set aside money each year for retirement in a tax-deferred account. Earnings on all amounts contributed to any IRA accumulate on a tax-deferred basis.

Pension: An employee benefits plan in which employees and employers pay into a group trust fund. The payment benefits are based on a formula factoring an employee's final pay and employment tenure with an employer.

Retirement plan: A financial arrangement offered by employers, insurance companies, and unions that replaces employment income during retirement.

Severance pay: Employer-offered compensation to newly unemployed employees. The payment may be offered to retired, resigned, laid-off, or dismissed with cause employees and may waive the employee's right to pursue a legal claim against the former employer or work for a competitor of the employer.

4.4.6 Miscellaneous Income

SIPP collects person-level data from respondents aged 15 or older regarding miscellaneous income receipt at any time during the reference year. The section begins by asking if the respondent received money or income from any of the following sources during the reference year:

- Community or religious charity
- Family or friends

- Roomers or boarders
- Estates
- Incidental or casual earnings
- Miscellaneous cash income, such as lottery winnings
- National Guard or Reserve Pay

If a respondent reports receiving any of the seven miscellaneous income sources, the respondent is then asked how much income was received during the reference year from all the miscellaneous sources.

Level of information the Miscellaneous Income data provide

Miscellaneous Income is collected for all household members who were age 15 or older during the reference year. Children under the age of 15 by the end of the reference year are not in universe for the Miscellaneous Income section. The Miscellaneous Income data are available in the form of person-level variables.

Other concepts related to the Miscellaneous Income data

Community Charity: An organization formed for the purpose of collecting and distributing benefits to support individuals, families, and children. Community charities include churches, social organizations, and government assistance programs.

Estates: An individual's property, entitlements, and obligations.

Incidental or Casual Earnings: Income or profit generated from informal, occasional, or irregular periods of receipt or employment.

Miscellaneous Cash Income: Any cash income that does not fall into a specific category.

National Guard or Reserve Pay: Any pay received while on duty (active or temporary) as a member of the National Guard or the Guard Reserve of any branch of the Armed Forces.

Religious Charity: A for-profit or not-for-profit organization whose purpose, membership, and leadership regard the observance of faith or vow of faith. A religious charity may manage a house of worship, provide payment to religious leadership, or provide payment to a religious organization.

Roomers or Boarders: Housing unit residents who rent a housing unit or section for one night or more. The housing rental agreement may include meals and dining facilities.

Short-Term Cash Assistance: Any assistance received from a government agency, charitable organization such as a church, community organization, or family and friends. This does not include TANF, General Assistance, or long-term assistance from any source.

Trusts: A contract administered by a trustee for the distribution of money or property to a beneficiary.

4.4.7 Other Assistance

Although the programs in this section are not necessarily means-tested, SIPP only collects these data from respondents aged 15 years and older who answered yes to at least one of the income screener questions. The Other Assistance questions focus on informal types of assistance (food, transportation, other clothing

and housing), and training) that fall outside the government's main social welfare programs. For each kind of assistance or topic, variables identify the type of assistance received during the reference period, the source(s) of assistance received during the reference period, and the months in which assistance was received during the reference period.

Food Assistance

SIPP asks respondents whether they received various types of food assistance (excluding SNAP, WIC, and school meals, if collected previously) during the reference period:

- Money, vouchers, or certificates to buy groceries
- Bags of groceries or packaged foods
- Meals from a shelter, soup kitchen, Meals-on-Wheels, or other charity
- Something else

Respondents who report receiving food assistance are asked about the source(s) of assistance during the reference period:

- Government social service agency
- Community or religious charitable organization
- Family or friends
- Someplace else

Respondents are then asked to identify the months they received food assistance during the reference period.

Transportation Assistance

SIPP asks respondents whether they received various types of transportation assistance during the reference period:

- Gas vouchers
- Bus or subway tokens or passes
- Help repairing, registering, or insuring a car
- Rides to a doctor's office or medical appointment
- Some other kind

Respondents who report receiving gas vouchers or bus or subway tokens or passes are asked about the source(s) of assistance during the reference period:

- Government social service agency
- Someplace else

Respondents are then asked to identify the months they received transportation assistance during the reference period.

Clothing and Housing Assistance

SIPP asks respondents whether they received various types of clothing assistance during the reference period:

- Free or reduced-price clothes
- Money or vouchers
- Both free or reduced-price clothes and money or vouchers

Respondents who report receiving clothing assistance are asked about the source(s) of this assistance during the reference period:

- Government agency
- Community or religious charity
- Family or friends
- Employer
- Someplace else

SIPP asks respondents whether they received housing assistance to help pay for housing during the reference year (not including energy assistance, if collected previously).

Respondents who report clothing and/or housing assistance are then asked to identify the months they received clothing and/or housing assistance during the reference period.

Training Assistance

This section asks respondents ages 18-64 with a household income under 200 percent of the poverty line whether they participated in any of the following types of training programs to improve or enhance their job readiness skills during the reference period:

- Classes or training to improve basic reading or math skills
- Job readiness training to learn about resume writing, job interviewing, or building self-esteem
- Job search programs or job clubs, or job resource centers to find out about jobs, to schedule interviews, or to fill out applications
- Training to learn a specific job skill, such as computers, car repair, nursing, day care work, or some other job skill

Respondents who report receiving TANF during the reference period are asked:

- Whether the training or use of job search resources was a requirement of the state or county welfare office, their choice, or both.

Respondents are then asked to identify the months they participated in training assistance programs during the reference period.

Level of information the Other Assistance data provide

The Other Assistance data are provided to users in the form of person-level and person-month-level variables.

Additional guidance for using the Other Assistance data

For each topic, SIPP has an initial question that asks whether the respondent received different types of during the reference period. Each type is edited as a separate yes/no variable. To identify whether a respondent received assistance in at least one month of the reference period, each yes/no variable for that topic must be checked.

To determine the months of receipt during the reference period, monthly yes/no variables are available for each topic. If respondents report more than one type or source of assistance for a single assistance program (e.g., reporting transportation assistance for repairing a car and rides to medical appointments), users are not able to identify which months of receipt were associated with each specific type or source of assistance.

While all data are in a person-month format, the value of some variables is the same for entire reference period (person-level variables). The value of other variables may vary from month to month (monthly-level variables).

<u>Person-level variables</u>

- Type(s) of assistance
    - EFOOD_TYPE1 – Money, vouchers, or certificates to buy groceries
    - EFOOD_TYPE2 – Bags of groceries or packaged food
    - EFOOD_TYPE3 – Meals from a shelter, soup kitchen, Meals-on-Wheels, or other charity
    - EFOOD_OTH – Some other kind of food assistance
    - ETRANS_TYPE1 – Gas vouchers
    - ETRANS_TYPE2 – Bus or subway tokens or passes
    - ETRANS_TYPE3 – Help repairing, registering, or insuring a car
    - ETRANS_TYPE4 – Rides to a doctor's office or medical appointment
    - ETRANS_OTH – Some other kind of transportation assistance
    - ECLTH_TYPE – Clothes, money or vouchers, or both clothes and money or vouchers
    - EHOUSE_ANY – Assistance to help pay for housing
    - EWELACTV1 – Classes or training to improve basic reading or math skills
    - EWELACTV2_1 – Job readiness training to learn about resume writing, job interviewing, or building self-esteem
    - EWELACTV2_2 – Job search programs or job clubs, or job resource centers to find out about jobs, to schedule interviews, or to fill out applications
    - EWELACTV2-3 – Training to learn a specific job skill, such as computers, car repair, nursing, day care work, or some other job skill
    - EWELACTV3 – Training or use of job search resources was a requirement of the state or county welfare office, their choice, or both because the state or local welfare office required it, or they chose to do it, or for both reasons
    - EWELACTV4 – Participated in any work experience program, such as community service job, in exchange for TANF

- Source(s) of assistance
    - EFOOD_SR1YN – Food source was through a government agency
    - EFOOD_SR2YN – Food source was through community or religious charitable organization

- EFOOD_SR3YN – Food source was through family or friends
- EFOOD_SR4YN – Food source was through someplace else
- EGAS_SOURCE – Source of gas vouchers was a government social service agency
- ETOKEN_SOURCE – Source of bus or subway tokens or passes was a government social service
- ECLTH_SR1YN – Source of clothes or money or vouchers for clothes through a government agency
- ECLTH_SR2YN – Source of clothes or money or vouchers for clothes through a community or religious charity
- ECLTH_SR3YN – Source of clothes or money or vouchers for clothes through family or friends
- ECLTH_SR4YN – Source of clothes or money or vouchers for clothes through an employer
- ECLTH_SR5YN – Source of clothes or money or vouchers for clothes through someplace else

Monthly-level variables

- Month(s) of assistance
  - EFOOD_MNYN – Received any food assistance during months of reference period
  - ETRANS_MNYN – Received transportation assistance during months of reference period
  - EOTHAS_MNYN – Received any other type of assistance during months of reference period
  - EWELAC_MNYN – Attended training during months of reference period

4.4.8 Retirement Income

SIPP collects person-level data regarding whether an individual aged 30 or older received retirement income during the reference year, including the type of income and amount received during each month of the reference year.

For respondents who report receiving retirement income, SIPP collects which type(s), specifically:

- Federal Civil Service or other Federal Civilian employee pension
- Local government pension
- Military retirement pay
- National Guard or Reserve Forces retirement
- Other retirement income
- Pension from a company or union including income from a profit-sharing plan
- State government pension
- U.S. Government Railroad Retirement

SIPP then collects separate monthly amounts for each type of retirement income. Respondents can report up to four amounts received and start months for each of the retirement benefits. If the respondent received more than four different monthly amounts throughout the reference year for any one retirement

income benefit type, the instrument asks the amount of money that the respondent received during the first start month.

Level of information the Retirement Income data provide

Retirement Income information is collected at the person-month level for all respondents who report ever being retired. The Retirement Income data are available in the form person-month-level variables on months of receipt and amount received. Additionally, these data include person-level reference period indicators and continuation flags.

Other concepts related to the Retirement Income data

Federal Civilian Employee Pension or Federal Employees Retirement System (FERS): A civilian retirement pension for civilian federal employees who began their employment after 1987.

Federal Civil Service Pension: A retirement pension for civilian federal employees who began their civilian federal employment between 1920 and 1987.

Local Government Pension: A pension plan for vested local government employees.

Military Retirement Pay: Retirement pension plan provided for active-duty military personnel with benefits that start when the employee retires.

National Guard Retirement: Army National Guard employee retirement plan provided for employees who have 20 years of service while serving the Guard by age 60 and may be combined with any additional income or retirement pension plan.

Other Retirement Income: Income provided for employee retirement, such as an Individual Retirement Account (IRA).

Pension: An employee benefits plan in which employees and employers pay into a group trust fund. The payment benefits are based on a formula factoring an employee's final pay and employment tenure with an employer.

Reserve Forces Retirement: A retirement pension plan for service members who complete 20 years of service and who become eligible for retirement from the Reserve Forces at age 60. Reserve service members may begin retirement benefits before age 60 if they deploy for war or national emergency. Effective January 2008, for every consecutive 90 days that service members are mobilized, their annuity start date is reduced three months.

State Government Pension: A pension plan for vested state government employees.

U.S. Government Railroad Retirement: A federally administered retirement program established for railroad employees and employers. The program provides retirement, unemployment, sickness, disability, spousal, and survivor benefits. Tier 1 benefits take the place of Social Security and are first payable at age 62 or after 30 years of service for early retirement. The first full retirement benefits are payable for retired

employee between the ages of 65 and 67. The Tier 2 benefits are similar to a private pension retirement plan.

4.4.9 School Meals

The School Meals section collects information about the receipt of school lunches and breakfasts at any point during the reference year, including:

- Whether any children in the family usually got the school lunch/breakfast that their school provided
- Whether the lunches/breakfasts received were free or reduced-price
- Indicator of receipt of school lunches/breakfasts is available on the record of eligible children

Level of information the School Meals data provide

Information about the receipt of school meals is asked of designated parents of children between the ages of 5 and 18 who have not yet graduated from high school. Information about receipt of school meals is stored on the parent's record. Additionally, all children between the ages of 5 and 18 who have not yet graduated from high school have indicators of receipt of school breakfasts and lunches on their records. The reporting period is the entire reference period. The School Meals data are available as person-level variables.

Additional guidance for using the School Meals data

While all data are available in a monthly format, these variables hold the same value over the entire reference period. The following variables are available on the record of designated parents of eligible children:

- Indicator that one or more children in the family usually got the lunch their school provided (ESCHOOLLUNCH)
- Indicator that one or more children in the family usually got the breakfast their school provided (ESCHOOLBREAK)
- Whether the lunches received were free or reduced-price because the family qualified for the School Lunch Program, or full price because the family did not qualify for the School Lunch Program (EFREE_LUNCH)
- Whether the breakfasts received were free or reduced-price because the family qualified for the School Breakfast Program, or full price because the family did not qualify for the School Breakfast Program (EFREE_BREAK)

The following variables are available on the record of children between the ages of 5 and 18 who have not yet graduated from high school:

- Whether respondent got free lunches, reduced-price lunches, full-price lunches, or did not receive school lunches (RLUNCH_CHLD)
- Whether respondent got free breakfasts, reduced-price breakfasts, full-price breakfasts, or did not receive school breakfasts (RBREAK_CHLD)

Other concepts related to the School Meals data

Federal School Breakfast/Lunch Program: The federal school breakfast and lunch program is a program sponsored by the Food and Nutrition Service of the U.S. Department of Agriculture. This program subsidizes the cost of all school meals for all children in many school districts across the country through the National School Lunch Program. Recipients of free or reduced-price meals are children in school who live in households with limited incomes or who live in districts in low-income areas participating in the Community Eligibility Provision. The benefit generally comes in the form of a discounted or free lunch and/or breakfast every school day during the school year, but in some localities the program is extended through the summer months.

4.4.10 Social Security Benefits for Oneself

The Social Security Benefits for Oneself section collects person-level data for each household member aged 18 or older who received Social Security benefits at any time during the reference year.

The section begins by asking about any Social Security benefit receipt for oneself during the reference year. If the respondent reports any Social Security benefit receipt, SIPP then asks why the respondent began receiving Social Security benefits. Social Security benefits for oneself can be received for five reasons:

1. Retirement – the respondent has stopped working due to age or length of job tenure and receives Social Security benefits based on the respondent's own earnings record.
2. Disability – the respondent has an impairment that limits the respondent's ability to work.
3. Widowed – a surviving spouse receives the Social Security income of their decedent spouse.
4. Spouse – a beneficiary receives Social Security income based on the record of a living spouse, whether the beneficiary and spouse are currently married or divorced.
5. Other – any reason that the respondent receives Social Security income for themselves, exclusive of reasons 1-4.

If the respondent reports a disability as the reason for receipt, SIPP then asks at what age the respondent began receiving Social Security due to his or her disability.

For married respondents, SIPP asks whether the Social Security benefits were received jointly with the spouse.

Amounts are collected by asking about the amount received in the most recent month of receipt, which can be the interview month. Previous amounts are collected by moving backwards chronologically from the interview month to the start of the reference period. If the start month for the amount reported is after the month in which benefit receipt started, the instrument then asks how much was received prior to the most current amount and when that amount was received.

Spells and amounts are output with two types of variables. First, person-month-level indicator variables show whether the respondent received Social Security benefits for oneself during the month. Second, where respondents report receipt for that month, person-month-level amount variables show the amount.

Level of information the Social Security Benefits for Oneself data provide

Monthly Social Security benefit amounts are collected at the person level for all household members aged 18 or older during the reference year. Respondents under the age of 18 at the end of the reference year are

not in universe for the Social Security Self Benefits for Oneself section. The Social Security Benefits for Oneself data are available in the form person-month-level variables on months of receipt and amount received. Additionally, these data include person-level reference period indicators and continuation flags.

Additional guidance for using the Social Security Benefits for Oneself data

The Social Security Self section collects amount(s) received and when the amount was received. It collects up to four amount changes during the reference year.

Other concepts related to the Social Security Benefits for Oneself data

Disability: An impairment that may be physical, cognitive, mental, hearing, vision, sensory, emotional, developmental, chronic disease, or some combination thereof.

Retired: Employment status or point in time when employees terminate their employment or reduce their employment hours due to their age or tenure. Retirement is usually between the ages of 50 and 70. However, early The Social Security Administration specifies a cohort-based retirement can be at any age, before the age or tenure needed for eligibility for support and funds from the employer and government age at which beneficiaries may claim full benefits. Benefits may be claimed early (no earlier than 62) for a lower benefit level, or they may be claimed after the full retirement age for a potentially higher benefit level.

Social Security on behalf of oneself: Social Security is managed by the Social Security Administration and is designed to provide retirement, healthcare, disability, and survivors' insurance for U.S. workers. As insured employees work, they pay Social Security taxes and earn credits that count toward eligibility for Social Security benefits. Most workers need 10 years of work to qualify for benefits. Social Security benefits may include monthly income benefits, health insurance (Medicare, Medicaid, or both), lump sum payments, and cost of living adjustments. Workers at least age 62 can retire with reduced benefits and can receive higher benefits at age 65, 66, or 67, depending on their birthdate. Current, former, and surviving spouses can receive Social Security benefits for themselves based on their spouse's earnings record.

4.4.11 Social Security Benefits on Behalf of a Child

The Social Security Benefits on Behalf of a Child section collects person-level data from respondents who received Social Security benefits on behalf of a child at any time during the reference year. Monthly Social Security child benefit amounts are collected at the person level for all household members aged 18 or older with a child who resided in the house at any time during the reference year. Respondents under the age of 18 at the end of the reference year are not in universe for items asking whether the respondent received Social Security income on behalf of a child. Persons receiving Social Security income on behalf of a child must be 18 or older. Parents under 18 years old are not asked these questions.

The section begins by asking about any Social Security benefit receipt on behalf of a child during the reference year. If the respondent reports any receipt and there is more than one child in the household, SIPP asks the respondent to identify the child(ren) in whose name(s) benefits were received. SIPP next asks why the respondent began receiving Social Security benefits.

Monthly amounts are collected by asking about the amount received in the most recent month of receipt, which can be the interview month.

Previous amounts are collected by moving backwards chronologically from the interview month to the start of the reference period. If the start month for the amount reported is after the month in which benefit receipt started, the instrument then asks how much was received prior to the most current amount.

Spells and amounts are output with two types of variables. First, person-month-level indicator variables show whether the respondent received Social Security benefits on behalf of a child during the month. Second, where respondents report receipt for that month, person-month-level amount variables show the amount.

Three variables can be used to link adults receiving Social Security income on behalf of a child with the child beneficiary. First, ESSKIDCOVFLG is a yes/no indicator for coverage on the child's record. Second, ESSCADLTPTR is a variable on the child's record that contains the PNUM of the adult who receives Social Security income on the child's behalf. Third, ESSKIDCTR is a variable on adults' records that shows the number of children on whose behalf the adult receives Social Security income. Note that some adults receiving Social Security income on behalf of a child may not be linked to a child. In most instances this is because the child is not living with the adult for all or part of the reference period.

Level of information the Social Security Benefits on Behalf of a Child data provide

Monthly Social Security child benefit amounts are collected at the person level for all household members aged 18 or older with a child who resided in the house at any time during the reference year. Respondents under the age of 18 at the end of the reference year are not in universe for items asking whether the respondent received Social Security income on behalf of a child. Persons receiving Social Security income on behalf of a child must be 18 or older. Parents under 18 years old are not asked these questions.

The Social Security Benefits on Behalf of a Child data provide person-month level information on months of receipt, amount received, and coverage. Additionally, these data include person-level reference period indicators and continuation flags.

Additional guidance for using the Social Security Benefits on Behalf of a Child data

The Social Security Benefits on Behalf of a Child section collects amount(s) received (ESSCAMT) and when the amount was received (ESSCMNYN). It collects up to four amount changes during the reference year.

Other concepts related to the Social Security Benefits on Behalf of a Child data

Child: Refers to family relationship and/or the age of the person. In family relationships, a child references a parent's biological, step, adopted, or guardianship daughter or son. Child also references anyone under age 18.

Disability: An impairment that may be physical, cognitive, mental, hearing, vision, sensory, emotional, developmental, chronic disease, or some combination thereof.

4.4.12 Special Supplemental Nutrition Program for Women, Infants, and Children (WIC)

The WIC section collects information about benefit receipt from the start of the reference period through the interview month. The WIC section does not ask who the owner of the WIC benefit is, so the WIC

owner variable, EWIC_OWN, is fully imputed using edited information about the family relationships in the household. For each WIC benefit unit, one adult in the household is considered the owner of the WIC benefits (the person in whose name the benefits are received). WIC benefits may be provided to pregnant and postpartum women, infants, and children up to the age of five who are low-income and are at nutritional risk. The WIC owner may or may not be covered by the WIC benefit. In cases in which only children are covered, an adult in the house may be the owner, but not covered by the benefit.

As shown below, all people are assigned values for the reference period and monthly coverage variables. People who are covered but are not the owner are assigned a value for an additional variable that points to the benefit owner and provides a link to details about the spell. Then, the benefit owner has values assigned for all of the variables applicable to the spell.

The WIC section does not ask about the amount of money received, since WIC provides access to purchase a basket of foods and formula and is not a cash transfer program. For this reason, all amounts are fully imputed using edited information about the number and type of recipients in the benefit unit together with information from the USDA on the average expenditure per person of benefit packages.

Variables on the public-use file provide the following information for the reference period:

Universe: All people

- Reference period indicator (RWIC_YRYN)
- Monthly coverage indicator (RWIC_MNYN)

Universe: Benefit owner and people who are covered by the benefit

- Benefit owner (EWIC_OWN)

Universe: Benefit owner

- Begin and end month of spell (EWIC_BMONTH and EWIC_EMONTH)
- Year receipt began if receiving in January of the reference period in a Wave 1 interview (RWIC_LCYR)
- Continuation flag for spells that include the last month of the reference period (RWIC_CFLG)
- Reason receipt began (EWIC_BRSN)
- Reason receipt ended (EWIC_ERSN)
- Payment amount (TWIC_AMT)
- Type of WIC coverage (EWIC_COVTYP)

WIC can be received to support children under age 5 (no adults), only a pregnant or postpartum woman (no children), or both a woman and a child or children. Postpartum women can be covered whether they are breast-feeding or not, provided they are at nutritional risk.

Level of information the WIC data provide

WIC data are collected at the spell level in the EHC and are provided to users in a person-month format. While all data are in a monthly format, the value of the variables may stay the same or vary over time.

Same value for the reference period

RWIC_YRYN

Same value for the spell

RWIC_MNYN, EWIC_BMONTH, EWIC_EMONTH, EWIC_OWN, RWIC_LCYR, RWIC_CFLG, EWIC_BRSN, EWIC_ERSN

Value may vary within a spell

TWIC _AMT, EWIC_COVTYP

Other concepts related to the WIC data

Farmers' Market Nutrition Program (FMNP): A federal grant program that began in 1992. The FMNP program provides locally grown food to WIC participants with the goal of expanding awareness and use of farmers' markets for low-income women (up to 6 months after birth or after pregnancy ends), infants, and children up to age 5 who are at nutrition risk.

WIC: The Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) A federal grant program administered by the state and at the federal level by the Food and Nutrition, U.S. Department of Agriculture in 1974. The WIC program provides supplemental food, healthcare referrals, and nutrition education for low-income women, infants, and children up to age 5 who are at nutritional risk. Pregnant, breast-feeding, and postpartum women are eligible for WIC if they are at nutritional risk. Recipients of WIC receive monthly benefits that may include checks, vouchers, or electronic cards to purchase food. WIC benefits also include supplemental foods, nutrition education, and nutrition counseling, as well as screening and referrals for health, welfare, and social services.

4.4.13 Supplemental Nutrition Assistance Program (SNAP)

The Supplemental Nutrition Assistance Program (SNAP) section collects information about benefit receipt from the start of the reference period through the interview month. For each SNAP benefit unit, one adult in the household is considered the owner of the SNAP benefits (the person in whose name the benefits are received). The SNAP owner may or may not be covered by the SNAP benefit.

As shown below, all people are assigned values for the reference period and monthly coverage indicators. People who are covered but are not the owner are assigned a value for an additional variable that points to the benefit owner and provides a link to details about the spell. Then, the benefit owner has values assigned for all of the variables applicable to the spell.

Universe: All people

- Indicates benefit receipt in at least one month of the reference period (RSNAP_YRYN)
- Indicates benefit receipt in this month (RSNAP_MNYN)

Universe: Benefit owner and people who are covered by the benefit

- Benefit owner (ESNAP_OWN)

Universe: Benefit owner

- Begin and end month of spell (ESNAP_BMONTH and ESNAP_EMONTH)
- Year receipt began if receiving in January of the reference period in a Wave 1 interview (RSNAP_LCYR)
- Continuation flag for spells that include the last month of the reference period (RSNAP_CFLG)
- Reason receipt began (ESNAP_BRSN)
- Reason receipt ended (ESNAP_ERSN)
- Payment amount (TSNAP_AMT)

Data users should be aware that variable names were standardized across the EHC programs (SNAP, WIC, SSI, TANF, and GA) starting in the 2018 SIPP. To see a side-by-side comparison of SNAP variable names in the 2014 SIPP Panel and the 2018 SIPP, please see Section 4.4.13 of the 2018 Survey of Income and Program Participation Users' Guide.

Level of information the SNAP data provide

SNAP data are collected at the spell level in the EHC and are provided to users in a person-month format. While all data are in a monthly format, the value of the variables may stay the same or vary over time.

Same value for the reference period

RSNAP_YRYN

Same value throughout the spell

RSNAP_MNYN, ESNAP_BMONTH, ESNAP_EMONTH, ESNAP_OWN, RSNAP_LCYR, RSNAP_CFLG, ESNAP_BRSN, ESNAP_ERSN

Value may vary within a spell

TSNAP_AMT

Additional guidance for using the SNAP data

Benefit Owner: A monthly variable that identifies the person number of the individual in whose name the benefit is.

Continuation Flag: A person-level variable that indicates whether the year-end spell (1) ended at the end of the reference year, (2) continued at the end of the reference year but ended before the interview month, or (3) ongoing interview month.

Spell: The duration of a benefit or event reported by a respondent in the Event History Calendar section of the survey that is defined by a start month, end month, and a continuous span between the start and end month.

Other concepts related to the SNAP data

Supplemental Nutrition Assistance Program (SNAP): This federal aid program, administered by each state and funded by the U.S. Department of Agriculture, provides food purchasing assistance for low-income residents. SNAP is commonly known by its former name, the Food Stamp Program**.** This program was enacted by Congress in the Food Stamp Act of 1964 and later renamed to SNAP (Supplemental Nutrition Assistance Program) by the Farm Bill of 2008, when the distribution system of coupon and stamp was replaced and updated with the Electronic Benefit Transfer (EBT) System. The amount of benefits received depends on the size, income, and expenses of the recipient's household. The benefits are distributed monthly via direct deposit to the EBT. Recipients may use the EBT to pay for food at participating food retail markets, which may include supermarkets, grocery stores, farmers' markets, and farmers' market roadside stands.

4.4.14 Supplemental Security Income (SSI)

The SSI section collects information about benefit receipt from the start of the reference period through the interview month. If a child under age 18 receives SSI, a person aged 18 or older is designated as the owner of the SSI benefits (the person in whose name the benefits are received).

As shown below, all people are assigned values for the reference period and monthly coverage indicators. People who are covered by SSI have values assigned for all of the variables applicable to the spell.

Universe: All people

- Reference period indicator (RSSI_YRYN)
- Monthly coverage indicator (RSSI_MNYN)

Universe: People who receive SSI

- Begin and end month of spell (ESSI_BMONTH and ESSI_EMONTH)
- Benefit owner if the person receiving benefits is under age 18 (ESSI_OWN)
- Year receipt began if receiving in January of the reference period in a Wave 1 interview (RSSI_LCYR)
- Continuation flag for spells that include the last month of the reference period (RSSI_CFLG)
- Reason receipt began (ESSI_BRSN)
- Reason receipt ended (ESSI_ERSN)
- Payment amount (TSSI_AMT)
- Payment source (ESSI_SRC1 and ESSI_SRC2)

Respondents may report whether they receive their monthly SSI benefit in one or two payments. For each payment, the monthly payment amounts are collected first followed by the payment source. Respondents may report up to five different payment amounts for each source they report in a spell to allow for changes in payments resulting from factors such as cost of living adjustments and earnings variation. The source of each payment is available for analysis, but the amounts are recoded into a single monthly variable. The primary reason for this is respondents who received a combined federal and state payment are not asked to report which portion of the benefit was a federal payment and which portion of the benefit was a supplemental state payment.

Data users should be aware that variable names were standardized across the EHC programs (SNAP, WIC, SSI, TANF, and GA) starting in the 2018 SIPP. To see a side-by-side comparison of SSI variable names in the 2014 SIPP Panel and the 2018 SIPP, please see Section 4.4.14 of the 2018 Survey of Income and Program Participation Users' Guide.

Level of information the SSI data provide

SSI data are collected at the spell level in the EHC and are provided to users in a person-month format. While all data are in a monthly format, the value of the variables may stay the same or vary over time.

Same value for the reference period

RSSI_YRYN

Same value for the spell

RSSI_MNYN, ESSI_BMONTH, ESSI_EMONTH, ESSI_OWN, RSSI_LCYR, RSSI_CFLG, ESSI_BRSN, ESSI_ERSN

Value may vary within a spell

TSSI_AMT, ESSI_SRC1, ESSI_SRC2

Other concepts related to the SSI data

SSI: Supplemental Security Income (SSI) is a means-tested cash assistance program administered by the Social Security Administration with states having the option to provide supplemental payments. To qualify, a person must meet at least one categorical eligibility requirement for being blind, disabled, or aged 65 or older, as well as have limited income and assets.

4.4.15 Support Paid

SIPP collects the following information from parents of children under 21 years of age living outside their household:

- The number of children under 21 years of age living outside their household
- If the parents made payments to support those children
- How much support was paid during the previous calendar year
- How often the respondent spent time with the children during the previous calendar year

All respondents aged 15 years and older are asked if they provided financial support to people living outside their household, including:

- Their parent or parents
- Children 21 years and older
- Other related person(s)
- Current or ex-spouse(s)
- Other unrelated people.

For each category, the number of persons supported is collected, as well as the amount paid during the previous calendar year.

Level of information the Support Paid data provide

The support paid data are collected at the person level. Amounts of support paid are available at the person level.

4.4.16 Support Received

SIPP collects person-level data about monetary payments that the respondent received for foster child care, child support, and alimony (spousal support) at any time during the reference year.

- Foster Child Care – Asked of guardians aged 15 and older, about whether they received foster child support payments for their foster child(ren) under 18 years of age, from the state in which they lived.
- Child Support – Asked of parents or legal guardians aged 15 and older, who did not previously report receiving pass-through child support payments about whether they received child support payments for their child(ren) under 21 years of age from the child(ren)'s parent(s) who lived outside their household.[17] Respondents are asked if child support payments were ever court ordered or informally agreed upon.
- Alimony – Asked of respondents aged 15 and older, who are currently divorced or separated or who have ever been divorced, about whether they received alimony or spousal support payments from a former spouse.

The collection of Support Received information is important as we attempt to measure the amounts of inter-household transfer payments that occur, as well as whom those payments are supporting. Collecting amounts of support received allows us to compare and contrast with amounts reported by respondents who pay support for their child(ren) and other persons living outside their household. The collection of monthly Support Received income amounts allows users to identify regular and irregular receipt of these income sources.

Level of information the Support Received data provide

Support Received content is collected at the person level. Amounts of support received are available at the person-month level.

Other concepts related to the Support Received data

Child Support: Financial help received from the other parent(s) may include payments made directly or indirectly to provide support for child(ren) in the form of rental or mortgage payments, medical costs or health insurance paid, costs associated with child care, school, or camp, or other tangible items to provide for the needs of the child(ren). The SIPP child support battery does not explicitly collect detailed

[17] In some cases when a person receives TANF or state program payments, the state agency tries to establish a child support order from the absent parent and becomes the collector of child support on behalf of the person receiving TANF benefits. In some states a portion or all of these collected payments are sent to the custodial parent. These payments are called 'pass through', 'disregard', or 'bonus' child support.

information about in-kind support, although it can be seen indirectly in a variety of measures throughout the instrument.

4.4.17 Survivor Income

SIPP collects data about whether an individual aged 15 years and older received survivor income benefits at any time during the reference year. Respondents report the type(s) received:

- Black Lung benefits
- Federal Civil Service or other federal civilian employee pension
- Income from a paid-up life insurance policy or annuity
- Local government pension
- Military retirement pay
- National Guard or Reserve Forces retirement
- Other (asked to specify)
- Payments from an estate or trust
- Pension from a company or union, including income from a profit-sharing plan
- State government pension
- U.S. Government Railroad Retirement
- Veterans' compensation/pension
- Workers' Compensation

Level of information the Survivor Income data provide

Survivor Income Benefit information is available at the person-month level for each adult household member 15 years and older who is a widow/widower.

Other concepts related to the Survivor Income data

Annuity: An annuity is a pension plan purchased by a policyholder at his or her retirement or during his or her working years to guarantee a regular income after retirement. Some annuity contracts provide guaranteed distribution for a set number of payments and some provide payment until death. SIPP includes all such annuities purchased by the person, except those already reported as part of the employee pension.

Black Lung Benefits: Black Lung Disability refers to the federal program beginning in 1973 that provides pension to coal miners who have disability from pneumoconiosis or to their dependent survivors.

Estates: An individual's property, entitlements, and obligations.

Federal Civil Service Pension: The Civil Service Retirement Act of 1920 provided a retirement system for federal employees who began their civil service employment between 1920 and 1987. CSRS is a contributory retirement system and employees share in the expense of annuities.

Federal Civilian Employee Pension: The Federal Employees Retirement System (FERS) is the retirement system for federal employees who began their civil service employment after 1987. FERS provides retirement income benefits from three sources: a basic benefit plan, Social Security, and Thrift

Savings Plan. Two of the three parts of FERS can go with employees to their next job if they leave the Federal Government. The federal agency withholds the cost of the basic benefit and Social Security from their pay as payroll deductions.

Local Government Pension: Also known as a public pension plan, this is the local government employee retirement plan for vested local government employees.

4.4.18 Temporary Assistance for Needy Families (TANF)

The TANF section collects information about benefit receipt from the start of the reference period through the interview month. For each TANF benefit unit, one adult in the household is considered the owner of the TANF benefits (the person in whose name the benefits are received). The TANF owner may or may not be covered by the TANF benefit.

As shown below, all people are assigned values for the reference period and monthly coverage indicators. People who are covered but are not the owner are assigned a value for an additional variable that points to the benefit owner and provides a link to details about the spell. Then, the benefit owner has values assigned for all of the variables applicable to the spell.

<u>Universe: All people</u>

- Reference period indicator (RTANF_YRYN)
- Monthly coverage indicator (RTANF_MNYN)

<u>Universe: Benefit owner and people who are covered by the benefit</u>

- Benefit owner (ETANF_OWN)

<u>Universe: Benefit owner</u>

- Begin and end month of spell (ETANF_BMONTH and ETANF_EMONTH)
- Year receipt began if receiving in January of the reference period in a Wave 1 interview (RTANF_LCYR)
- Continuation flag for spells that include the last month of the reference period (RTANF_CFLG)
- Reason receipt began (ETANF_BRSN)
- Reason receipt ended (ETANF_ERSN)
- Payment amount (TTANF_AMT)
- Coverage type (ETANF_COVTYP)
- Receipt of pass-through child support (ETANF_PTCS)

For each TANF spell, respondents are asked to report the person in whose name the benefit is received (benefit owner), who was covered by the benefit, why receipt began, why receipt ended (if applicable), and payment amount as this information may vary across spells. They are also asked about whether the benefit covers covered only children, only a pregnant woman, or both children and adults.

If the TANF benefits cover children, then receipt of pass-through child support is asked about.

Data users should be aware that variable names were standardized across the EHC programs (SNAP, WIC, SSI, TANF, and GA) starting in the 2018 SIPP. To see a side-by-side comparison of TANF variable names in the 2014 SIPP Panel and the 2018 SIPP, please see Section 4.4.18 of the 2018 Survey of Income and Program Participation Users' Guide.

Level of information the TANF data provide

TANF data are collected at the spell level in the EHC and are provided to users in a person-month format. While all data are in a monthly format, the value of the variables may stay the same or vary over time.

Same value for the reference period

RTANF_YRYN

Same value for the spell

RTANF_MNYN, ETANF_BMONTH, ETANF_EMONTH, ETANF_OWN, RTANF_LCYR, RTANF_CFLG, ETANF_BRSN, ETANF_ERSN

Value may vary within a spell

TTANF_AMT, ETANF_COVTYP, ETANF_PTCS

Other concepts related to the TANF data

Temporary Assistance for Needy Families (TANF): TANF – which depending on where someone lives may have a different state or local name – is an income supplement program serving low-income families with dependent children. Depending on state rules, benefits may cover pregnant women, single parent families, married-couple families, and only children (child-only cases occur when a child's guardian does not qualify for benefits). TANF is funded at the federal level through block grants but administered at the state level.

While women receive TANF, any court-ordered child support goes directly to the state, which in turn decides whether and how much of the payment goes to the mother.[18] This is called pass-through child support, bonus child support, or disregard payments. Some states do not pass any along, some states pass along a portion but subtract that from the benefit, so the payment remains the same, and some states pass along a portion of the payment.

4.4.19 Tax Returns

The Tax Returns sections asks about taxes filed during the reference year, meaning that these questions reference taxes paid on income in the calendar year prior to the reference year.[19] SIPP collects person-

[18] For state level information regarding pass through policy see: <www.ncsl.org/research/human-services/state-policy-pass-through-disregard-child-support.aspx>.

[19] If someone is interviewed in 2019, the reference year is calendar year 2018, and tax data will be collected regarding their filing for tax year 2017.

level data from respondents aged 15 or older regarding federal tax returns filed during the reference year, specifically:

- Whether the individual filed a tax return during the reference year (for the tax year of the reference year minus one)
- Whether the individual plans to file a return (if no return was filed) for the tax year of the reference year minus one
- Filing status for the tax year of the reference year minus one
    - Single
    - Married filing jointly
    - Married filing separately
    - Head of household
- Whether anyone claimed the respondent as a dependent (respondents between age 15 and 25) for tax year of the reference year minus one
- Whether the individual received an Earned Income Tax Credit (EITC) during the reference year

Level of information the Tax Returns data provide

Tax Return information is collected at the person level for all household members age 15 or older during the reference year. Children under the age of 15 are not in universe. The Tax Return data are available as person-level variables.

Other concepts related to the Tax Returns data

Earned Income Tax Credit: A tax credit earned by an individual with a tax return filing status of single, head of household, or married filing jointly who has earned income and adjusted gross income within certain limits.

Head of Household: An individual who is unmarried or considered unmarried (unmarried as of the last day of the tax year) who lived with one or more family members and who has paid more than half the cost of keeping the home for the tax year.

Married Filing Jointly: An individual who is married or considered married (married as of the last day of the tax year). To file jointly both the individual and their spouse must agree to file a joint return. On a joint return, you combine income and deduct your combined allowable expenses. A couple can file a joint return even if one had no income or deductions. A joint return tax may be lower than your combined tax for other filing statuses.

Married Filing Separately: A married individual who is responsible only for his or her own taxes or because it results in less tax than filing a joint return.

Single: An individual who is unmarried, considered unmarried (unmarried as of the last day of the tax year), or a widow/widower who does not qualify as head of household, married filing separately, or married filing jointly.

4.4.20 Unemployment Compensation Payment

SIPP collects person-level data from respondents aged 15 or older regarding Unemployment Compensation (UC) receipt, also known as unemployment insurance (UI), at any time during the reference year. Additional questions categorize UC receipt into three different types: regular, supplemental, and other.

The UC section begins by asking whether the respondent received any income from unemployment compensation during the reference year. If so, then the instrument asks about which types of UC were received. The instrument collects data for three types of compensation. The first type is regular, which is the most common type, received from state unemployment agencies and the federal government. The second is supplemental, which includes unemployment insurance paid for privately. The third type is denoted as other in the questionnaire and corresponds to union benefits for which respondents may be eligible if they are members of a union.

For each benefit type reported, the instrument asks for the months of receipt and amount received. For respondents who received multiple types of UC, once the data about the first type are collected, the instrument moves on to ask about the next type.

Level of information the Unemployment Compensation Payment data provide

Unemployment Compensation data are collected at the person level for all household members who were 15 and older during the reference year. Teenagers who were aged 15 and older during the reference year are in universe for the Unemployment Compensation section. UC amounts are available at the person-month level.

Other concepts related to the Unemployment Compensation Payment data

Covered Employment: "Covered" here refers to being covered by state unemployment insurance programs. Most work for wages falls under the category of covered employment. Exceptions vary by state, but can include independent contractors, sole proprietors or members of a partnership, railroad workers subject to the Federal Railroad Unemployment Insurance Act, some real estate brokers, some insurance brokers, some agricultural workers, elected officeholders, some church employees, and other exceptions on a state-by-state basis.

Other Unemployment Insurance: This category of unemployment insurance is a catch-all for unemployment compensation not from regular or supplemental unemployment insurance programs. The most common "other" programs correspond to union benefits for which respondents may be eligible if they are members of a union.

Supplemental Unemployment Insurance: This refers to all unemployment insurance plans paid for privately. Some of these are paid for by employees themselves, or by the self-employed. They can be used as an additional wage replacement to regular unemployment compensation, because state UI payments are generally capped at a weekly or monthly amount that may be lower than workers' full wage rate. These programs are also provided by employers in sectors in which employment is not covered by regular UI plans. Employer-sponsored plans in non-covered employment make up the bulk of supplemental unemployment insurance plans.

Unemployment Insurance (UI)/Unemployment Compensation (UC): This is the name of the programs administered by every state and the District of Columbia that provide assistance to jobless people who are looking for work. In these programs, employers are required to pay a certain amount to the state unemployment agency monthly. The amount paid is tied to each worker's wage rate and hours worked. These payments fund the unemployment insurance agency. Not all types of employment are covered by UI.

For workers who lose their job, file for UI, and are determined to be eligible for it, these programs provide a payment equal to a percentage of previous earnings. The amount of the payment varies by state. The number of weeks that a worker can receive the benefit also varies by state. During recessions, the federal government often temporarily expands and or extends UC payments. In SIPP, payments from this type of program are coded as regular Unemployment Compensation payments.

4.4.21 Veterans Benefits

The Veterans Benefits section collects person-level data from household members aged 17 or older during the reference year who reported previously serving in the U.S. Armed Forces but are not enlisted at the time of the interview.

SIPP collects the type of veterans benefits and amount received during each month in the reference year. The section begins by asking whether the respondent received veterans benefits at any point during the reference year.

For respondents who report receiving veterans benefits, SIPP asks what type of benefits were received during the reference year. Respondents may report receipt of five separate benefit types:

1. Service-connected disability compensation plan
2. Veterans pension
3. Other VA payments
4. G.I. Bill benefits
5. Insurance proceeds

SIPP asks each VA benefit recipient if he or she was required to complete an Eligibility Verification Report, or EVR, during the reference year in order to receive VA benefits.

For respondents receiving service-connected disability compensation benefits, SIPP collects their disability rating:

1. 0 percent
2. 10-20 percent
3. 30-40 percent
4. 50-60 percent
5. 70+ percent

The instrument collects monthly receipt amounts separately for each benefit type.

Level of information the Veterans Benefits data provide

Monthly benefit amounts for each type of benefit reported are collected at the person level for all household members aged 17 or older during the reference year who reported previously serving in the U.S. Armed Forces but are not currently enlisted. VA benefit amounts are available at the person-month level.

Additional guidance for using the Veterans Benefits data

The Veterans Benefits section collects the type of benefit(s), amount(s) received, and when the amount was received for up to four amount changes during the reference year.

Other concepts related to the Veterans Benefits data

Military Retirement Pay: A retirement pension plan provided for active-duty military personnel with benefits that start when the employee retires.

National Guard Retirement: An Army National Guard employee retirement plan provided for employees who have 20 years of service in the National Guard. By age 60, this retirement plan may be combined with additional income or pensions.

Pension: A benefit plan in which employees and employers pay into a group trust fund. The pension benefits are based on the employee's final pay and the employment tenure with the employer.

Veterans Benefits: A monthly pension administered by the United States Department of Veteran Affairs to active-duty military branch members, their dependents, and their survivors.

Veterans Compensation: A monthly payment administered by the United States Department of Veteran Affairs to veterans, their dependents, and their survivors.

Veterans Pension: A monthly pension administered by the United States Department of Veteran Affairs for veterans who served in the military but did not qualify for military retirement pay.

Reserve Forces Retirement: A retirement pension plan for service members who complete 20 years of service and/or who become eligible for retirement from the Reserve Forces at age 60 or in some cases before age 60 depending on active service.

Service-Connected Disabilities: A disability benefit paid to veterans disabled by injury or illness during active duty. Monthly benefits vary with the type of disability and the number of eligible dependents.

4.4.22 Workers' Compensation

The Workers' Compensation section collects person-level data for household members aged 15 or older regarding Workers' Compensation benefits received during the reference period.

The section begins by asking if the respondent received any Workers' Compensation benefits since the beginning of the reference period. If respondents replied that they have received benefits, the instrument then asks if they receive benefits now. If not, they are asked the month in which they last received benefits. Benefit amounts are collected by moving backwards chronologically, either from the interview month if they currently receive benefits or the last month they received benefits, to the start of when

receipt began. Respondents can report up to four amounts received and start months. If the respondent received more than four different monthly amounts throughout the reference period, the instrument asks the amount of benefit received during the first start month of the reference period.

Level of information the Workers' Compensation data provide

Monthly benefit amounts for each type of benefit reported are collected at the person level for all household members aged 15 or older during the reference period. Children under the age of 15 by the end of the reference period are not in universe for the Workers' Compensation section. Workers' Compensation amounts are available at the person-month level.

Additional guidance for using the Workers' Compensation data

The Workers' Compensation section collects the amount(s) received and when the amount was received for up to four amount changes during the reference period.

Other concepts related to the Workers' Compensation data

Workers' Compensation: Employer's insurance providing wage and medical benefits to an employee who is injured during the course of employment. Workers' Compensation insurance is required by almost every state. Rules and requirements may vary by state. An employer pays for Workers' Compensation and the employer is covered for work-related accidents. Workers' Compensation benefits can help pay for medical and hospital bills. They also can be provided if the employee cannot return to work. If an employee dies due to a work-related accident, Workers' Compensation insurance can also pay an insurance benefit to the family.

Workers' Compensation Payments: Plans and coverage may vary. Plans can make weekly payments in place of wages (similar to disability insurance), compensation for loss in wages, and reimbursement or payment of medical and health care expenses (similar to health insurance). Workers' Compensation benefits can be payable to families of workers killed during employment (similar to life insurance). Most states have Workers' Compensation coverage requirements; however, some states do not require Workers' Compensation coverage for specific industries such as agriculture and some small private-sector employers.

Job-Related Injury: An injury that occurred while working or because of a work-related role.

Job-Related Illness: An occupational-related disease or disability that occurred as a result of a work-related role or during employment.

4.5 Income and Poverty Recodes

Unlike the preceding content areas, data on total income and poverty status are not collected directly from the SIPP instrument, but instead are aggregated or assigned based on existing survey content. The following sections describe these aggregations.

4.5.1 Income Recodes

SIPP creates a number of income recodes at the person, family, and household levels. The income recodes aggregate reported income from the Assets, Employment and Earnings, and Programs and Income Transfer sections of the SIPP.

For detailed information on the individual components of the Income Recodes, see their respective sections of this Users' Guide.

Total Personal Income (TPTOTINC) is created from:

TPEARN_ALT	Alternate recode of earnings and profits/losses from all jobs stabilized to accommodate months of varying length
TPPRPINC	Total personal investment/property income
TPTRNINC	Total personal means-tested transfer income
TPSCININC	Total personal social insurance payments
TPOTHINC	Total personal other income

Total Personal Earnings (TPEARN_ALT) is created from:

The income types included in this variable are wage and salary income, bonus payments, commissions, overtime payments, tips, other income from self-employed businesses, self-employed business profits, and accounting for time spent away from a job without pay. TPEARN_ALT is the alternate recode of earnings and profits from all jobs; for more detail about this variable, see section "Alternative Earnings Recode (TPEARN_ALT)."

TPEARN_ALT differs from TPEARN by removing all variation in earnings due to variation in the length of a month. For example, TPEARN assigns earnings in February based on it being 28 or 29 days in length, TPEARN_ALT assigns earnings in February assuming it has the same length as all other months.[20]

Total Personal Investment & Property Income (TPPRPINC) is created from:[21]

---

[20] For more information see the user note TPEARN Historical Comparison User Note comparing RPEARN and RPEARN_ALT: <www2.census.gov/programs-surveys/sipp/tech-documentation/methodology/2018_SIPP_UsersGuide.pdf>.

[21] For additional details about the 2019 SIPP universes for variables with a TJS or TJO prefix, please refer to the user note *2019 SIPP: Universe for Jointly Owned Assets* at <https://www.census.gov/programs-surveys/sipp/tech-documentation/user-notes.html>.

TINC_BANK Income earned from interest-earning assets held at financial institutions, created from:

TOCHKINC	Total interest income earned from individually-owned checking accounts
TJSCHKINC	Share of total interest income earned from jointly-owned checking accounts [for respondents with a spouse in the household; this joint amount question is asked of only one spouse, and this response is divided by 2 and copied to both spouses' records]
TJOCHKINC	Share of total interest income earned from jointly-owned checking accounts [for respondents without a spouse in the household]
TOSAVINC	Total interest income earned from individually-owned savings accounts
TJSSAVINC	Share of total interest income earned from jointly-owned savings accounts [for respondents with a spouse in the household; this joint amount question is asked of only one spouse, and this response is divided by 2 and copied to both spouses' records]
TJOSAVINC	Share of total interest income earned from jointly-owned savings accounts [for respondents without a spouse in the household]
TJSCDINC	Share of total interest income earned from jointly-owned certificates of deposit [for respondents with a spouse in the household; this joint amount question is asked of only one spouse, and this response is divided by 2 and copied to both spouses' records]
TJOCDINC	Share of total interest income earned from jointly-owned certificates of deposit [for respondents without a spouse in the household]
TOCDINC	Total interest income earned from individually-owned certificates of deposit
TJSMMINC	Share of total interest income earned from jointly-owned money market deposit accounts and money market funds [for respondents with a spouse in the household; this joint amount question is asked of only one spouse, and this response is divided by 2 and copied to both spouses' records]

	TJOMMINC	Share of total interest income earned from jointly-owned money market deposit accounts and money market funds [for respondents without a spouse in the household]
	TOMMINC	Share of total interest income earned from individually-owned money market deposit accounts and money market funds
TINC_BOND	Income earned from other interest-earning assets, created from:	
	TJSGOVSINC	Share of total interest income earned from jointly-owned government securities [for respondents with a spouse in the household; this joint amount question is asked of only one spouse, and this response is divided by 2 and copied to both spouses' records]
	TJOGOVSINC	Share of total interest income earned from jointly-owned government securities [for respondents without a spouse in the household]
	TOGOVSINC	Total interest income earned from individually-owned government securities
	TJSMCBDINC	Share of total interest income earned from jointly-owned municipal and corporate bonds [for respondents with a spouse in the household; this joint amount question is asked of only one spouse, and this response is divided by 2 and copied to both spouses' records]
	TJOMCBDINC	Share of total interest income earned from jointly-owned municipal and corporate bonds [for respondents without a spouse in the household]
	TOMCBDINC	Total interest income earned from individually-owned municipal and corporate bonds
TINC_STMF	Income earned from stocks and mutual funds, created from:	
	TJSMFINC	Share of dividend income earned from jointly-owned mutual funds [for respondents with a spouse in the household; this joint amount question is asked of only one spouse, and this response is divided by 2 and copied to both spouses' records]
	TJOMFINC	Share of dividend income earned from jointly-owned mutual funds [for respondents without a spouse in the household]
	TOMFINC	Total dividend income earned from individually-owned mutual funds

	TJSSTINC	Share of dividend income earned from jointly-owned stocks [for respondents with a spouse in the household; this joint amount question is asked of only one spouse, and this response is divided by 2 and copied to both spouses' records]
	TJOSTINC	Share of dividend income earned from jointly-owned stocks [for respondents without a spouse in the household]
	TOSTINC	Total dividend income earned from individually-owned stocks
TINC_RENT	Net income from rental properties, created from:	
	TJSRPNETINC	Share of net income earned from jointly-owned rental property [for respondents with a spouse in the household; this joint amount question is asked of only one spouse, and this response is divided by 2 and copied to both spouses' records]
	TJORPNETINC	Share of net income earned from jointly-owned rental property [for respondents without a spouse in the household]
	TORPNETINC	Net income earned from individually-owned rental property
TINC_OTH	Net income from other assets, created from:	
	TANNINC	Amount of income received from annuities
	TTRINC	Amount of income received from trusts
	TOINVINC	Amount of income received from other financial investments

Total Means-Tested Transfer Payments (TPTRNINC) is created from:

TTANF_AMT	Amount received from monthly TANF benefit
TSSI_AMT	Amount received from monthly SSI benefit
TGA_AMT	Amount received from monthly GA benefit
TVA2AMT	Amount received from monthly Veterans' pension

Total Social Insurance Payments (TPSCININC) is created from:

TVA1AMT	Amount received from VA benefit payments for a service-connected disability
TVA3AMT	Amount received from other VA payments
TVA4AMT	Amount received for G.I. Bill benefits
TVA5AMT	Amount received from VA Insurance proceeds
TWCAMT	Amount received in workers' compensation
TUC1AMT	Amount of regular unemployment compensation received
TUC2AMT	Amount of supplemental unemployment compensation received
TUC3AMT	Amount of other unemployment compensation received
TSSSAMT	Amount received in Social Security benefit payments
TSSCAMT	Amount received in Social Security benefit payments on behalf of a child(ren)

Total Other Income (TPOTHINC) is created from:

TSUR1AMT	Amount received in survivor benefits from a pension from a company or union including income from a profit-sharing plan
TSUR2AMT	Amount received in survivor benefits from Veterans' compensation or pension
TSUR3AMT	Amount received in survivor benefits from a Federal civil service or other Federal civilian employee pension
TSUR4AMT	Amount received in survivor benefits from U.S. Government Railroad Retirement in month
TSUR5AMT	Amount received in survivor benefits from a state government pension
TSUR6AMT	Amount received in survivor benefits from a local government pension
TSUR7AMT	Amount received in survivor benefits from income from a paid-up life insurance
TSUR8AMT	Amount received in survivor benefits from military retirement pay
TSUR9AMT	Amount received in survivor benefits from Black Lung Benefits
TSUR10AMT	Amount received in survivor benefits from workers' compensation
TSUR11AMT	Amount received in survivor benefits from payments from an estate or trust
TSUR12AMT	Amount received in survivor benefits from National Guard or Reserve Forces retirement

TSUR13AMT	Amount received in survivor benefits from other survivor income
TRET1AMT	Amount received in retirement income from a pension from a company or union including income from a profit-sharing plan
TRET2AMT	Amount received in retirement income from Federal civil service or other Federal civilian employee pension
TRET3AMT	Amount received in retirement income from State government pension
TRET4AMT	Amount received in retirement income from local government pension
TRET5AMT	Amount received in retirement income from military retirement pay
TRET6AMT	Amount received in retirement income from U.S. Government Railroad Retirement
TRET7AMT	Amount received in retirement income from National Guard or reserve forces retirement
TRET8AMT	Amount received in retirement income from other retirement income
TDIS1AMT	Amount received in payments from a sickness, accident, or disability insurance policy
TDIS2AMT	Amount received in payments from employer disability payments
TDIS3AMT	Amount received in pension from a company or union including income from a profit-sharing plan
TDIS4AMT	Amount received in payments from Federal civil service or other federal civilian employee pension
TDIS5AMT	Amount received from a state government pension
TDIS6AMT	Amount received from a local government pension
TDIS7AMT	Amount received from U.S. military retirement pay
TDIS8AMT	Amount received from U.S. government railroad retirement
TDIS9AMT	Amount received in payments from Black Lung Benefits
TDIS10AMT	Amount received from other disability income
TFCCAMT	Amount of foster child care payments received
TCSAMT	Amount of child support payments received
TALIAMT	Amount of alimony payments received
TLMPAMT	Amount of lump sum payment received

TDEFERAMT	Amount received as a deferred payment or final paycheck from a job or business that ended during the reference period
TLIFEAMT	Amount received from a paid up life insurance policy or annuity
TMINC_AMT	Amount of miscellaneous income received

Level of information the Income Recode data provide

The income recodes are calculated for each month by summing income sources. For income sources reported annually, reported income is allocated evenly across the months of the reference year by dividing by 12. To avoid double counting, transfer payment amounts that are family or household based are only placed on the record of the family member identified as the benefit owner.

Additional guidance for using the Income Recode data

The SIPP instrument creates aggregate monthly income recodes at the person, family, and household levels. Additional recodes are created at the person level based on income source. Table 4-3 lists these recodes.

TABLE 4-3. INCOME RECODES

Aggregate			Description
Person	**Family**	**Household**	
TPTOTINC	TFTOTINC	THTOTINC	Total monthly income excluding Type 2 people
	TFTOTINCT2	THTOTINCT2	Total monthly income including Type 2 people
Components			**Description**
Person	**Family**	**Household**	
TPEARN_ALT			Monthly earnings: The sum of gross earnings, wages, and salary, and/or the amount of monthly income (positive or negative) from self-employment for each job and/or business recorded for the reference month. This recode is stabilized to accommodate months of varying length
TPPRPINC			Monthly investment and property income: The sum of dividend, interest, and property/rental income
TPTRNINC			Monthly means-tested transfer payments: The sum of all means-tested cash transfers (e.g., SSI, GA, TANF)
TPSCININC			Monthly social insurance payments: The sum of non-means-tested social insurance program transfers (e.g., Social Security, unemployment compensation)
TPOTHINC			Monthly income from other sources: The sum of support payments (e.g., alimony, child support),

			retirement income (e.g., pensions), and any other income not captured elsewhere

Other concepts related to the Income Recode data

Family income: Family-level income variables include the income of all household members related by birth, marriage, or adoption in a given month. In other words, primary family members, including related subfamily members, are treated as one family by the Census Bureau when calculating family-level income amounts.

Household income: Household-level income variables include the income of all household members in a given month.

Distribution of Annual Income: For income sources reported annually, reported income is allocated evenly across the months of the reference year by dividing by 12.

Reporting of Joint Income: To avoid double counting, transfer payment amounts that are family based are only placed on the record of the family member identified as the owner of the benefit. For respondents with a spouse in the household, income from jointly owned accounts are asked of only one spouse and this response is divided by 2 and copied to both spouses' records.

4.5.2 Poverty Recodes

Poverty Recodes are assigned based on data collected on household and family composition as well as the aggregate Income Recodes.

Monthly family poverty status (TFINCPOV) is one of the most frequently used variables by SIPP users, as this provides a monthly measure of poverty, allowing data users to evaluate the movement of families into and out of poverty over the course of the reference period. Additional variables at the household level and including Type 2 individuals allow data users to measure the ability of a family or household to meet their basic needs.

Level of information the Poverty Recode data provide

The monthly poverty thresholds are assigned based on the composition of a family or household in a given month, depending on the variable being used. The monthly poverty thresholds are also adjusted to reflect changes in the Consumer Price Index for all Urban Consumers (CPI-U) across months.

The “base” annual poverty thresholds from 1982 are available on the Census Bureau’s website at <www.census.gov/data/tables/time-series/demo/income-poverty/historical-poverty-thresholds.html>. These annual thresholds are converted to monthly thresholds by dividing by 12 and adjusting for inflation based on the difference in the reference month CPI-U and the 1982 annual CPI-U value.

Once the appropriate monthly threshold is assigned, income-to-poverty ratios are calculated by dividing topcoded total family (or household) income by the appropriate threshold. If the resulting ratio is less than 1, then that family (or household), and all individuals in it, are in poverty for the given month.

For individuals with a valid poverty threshold in each month of the reference period, annual income-to-poverty ratios are calculated. Values are based on the annual sum of the individual's monthly total family (or household) income divided by the annual sum of the individual's monthly poverty thresholds (reflecting changing family or household composition over the reference period). The resulting annual values are then divided to calculate the individual's annual income-to-poverty ratio.

Additional guidance for using the Poverty Recode data

Poverty Recodes are available at both the monthly and person level and include thresholds as well as poverty status.

Monthly Variables:

Thresholds

RFPOV	Monthly family poverty threshold excluding Type 2 persons
RFPOVT2	Monthly family poverty threshold including Type 2 persons
RHPOV	Monthly household poverty threshold excluding Type 2 persons
RHPOVT2	Monthly household poverty threshold including Type 2 persons

Poverty Status

TFINCPOV	Monthly family income-to-poverty ratio excluding Type 2 persons
TFINCPOVT2	Monthly family income-to-poverty ratio including Type 2 persons
THINCPOV	Monthly household income-to-poverty ratio excluding Type 2 persons
THINCPOVT2	Monthly household income-to-poverty ratio including Type 2 persons

Annual Variables:

Poverty Status

TFCYINCPOV	Annual family income-to-poverty ratio excluding Type 2 persons
TFCYINCPOVT2	Annual family income-to-poverty ratio including Type 2 persons
THCYINCPOV	Annual household income-to-poverty ratio excluding Type 2 persons
THCYINCPOVT2	Annual household income-to-poverty ratio including Type 2 persons

Other concepts related to the Poverty Recode data

Family: A family is defined as all household members related by birth, marriage, or adoption in a given month. In other words, primary family members, including related subfamily members, are treated as one family by the Census Bureau when creating family assignments. When poverty status is determined at the family level, all family members have the same poverty status.

Household: A household consists of all people who occupy a housing unit, regardless of their relationships to each other. Household-level poverty status is based on the income of all household members in a given month. When poverty status is determined at the household level, all household members have the same poverty status.

Income-to-Poverty Ratios: Income-to-poverty ratios are calculated by dividing total family or household income by the appropriate threshold. Resulting values below 1 indicate a family/household is "in poverty," with values closer to 1 indicating that a family/household is closer to meeting their threshold. Values equal to or above 1 indicate a family is not in poverty.

5 Resources and Research

This section details where data users may find the data and its corresponding documentation, including links for all publicly available SIPP data.

5.1 SIPP Website

The SIPP website provides a number of resources related to SIPP, from its inception in 1984 to the present, including:

- SIPP data files for all panels and waves
- Research products using SIPP data
- Technical documentation

The SIPP webpage is located at <www.census.gov/sipp>.

5.1.1 Data

Users may directly access SIPP data via the SIPP Data webpage, located at <www.census.gov/programs-surveys/sipp/data.html>. SIPP data and documentation, as released by the Census Bureau, are not copyrighted. The data files and supporting documentation can therefore be freely copied and distributed to other users.

The SIPP Data webpage contains datasets for all panels and waves of SIPP, as well as links to supporting documentation. The data are available in STATA (2014 Panel and later), SAS, and pipe-delimited CSV (2014 and later) file formats. To reduce the data download burden, each data file format is compressed using both *zip* compression (.zip; commonly used in Windows environments) and *gzip* compression (.gz; commonly used in GNU/Linux environments). The SIPP Data webpage also contains links to user notes, variable crosswalks, and other metadata, as well as links to interactive data tools and resources.

5.1.2 Technical Documentation

The technical documentation section of the SIPP website contains a wealth of information to assist data users including an online codebook, user notes, source and accuracy statements, data dictionaries, questionnaires, and instrument specifications.

The main SIPP technical documentation is located at <www.census.gov/programs-surveys/sipp/tech-documentation.html>.

Online Codebook

The SIPP online codebook provides information and metadata for every variable of the SIPP public-use files for the 2014 SIPP Panel and later years. The SIPP online codebook is available at <www.census.gov/programs-surveys/sipp/tech-documentation/codebooks.html>. Metadata for SIPP data

prior to the 2014 Panel are available on the Data Dictionary page at <https://www.census.gov/programs-surveys/sipp/tech-documentation/data-dictionaries.html>.

User Notes

User notes for SIPP provide information on possible changes in the data or in the questionnaire and give guidance on how to account for any changes that may have occurred. User notes also include documentation and code to correct any possible inconsistencies in the data and provide any other necessary clarifications.

SIPP user notes are located at <https://www.census.gov/programs-surveys/sipp/tech-documentation/user-notes.html>.

Source and Accuracy Statements

SIPP's Source and Accuracy Statements summarize the source of data as well as information on sampling, estimation procedures, population controls, how to use the weights, possible types of error in SIPP data, and how to calculate different types of standard errors.

The Source and Accuracy Statements are located at <www.census.gov/programs-surveys/sipp/tech-documentation/source-accuracy-statements.html>.

5.2 Research Products

The SIPP website contains a number of research products analyzing SIPP data. These include P-70 reports, conference papers and presentations, table packages, working papers, and a SIPP bibliography. The website is updated on a regular basis to include the most recent and up-to-date research related to SIPP data.

SIPP research products are located at <www.census.gov/programs-surveys/sipp/library.html>.

5.2.1 P-70 Reports

The P-70 report series was designed to inform the public about the economic status and well-being of American households using data from the SIPP. The P-70 reports cover a wide range of topics, including living conditions, child care arrangements, health insurance coverage, participation in government programs, poverty, income distribution, and employment.

The Census Bureau's P-70 series is the primary source for published estimates from SIPP. These reports help data users in a number of ways:

- Published estimates may contain estimates needed for the research project at hand, thus saving users the need to generate those estimates themselves.
- Published estimates can often provide a useful cross-check for closely related estimates prepared by analysts.
- Published estimates are based on the Census Bureau's internal data files, which are often impossible for external data users to replicate because the internal files have not been subjected to topcoding and other confidentiality protection techniques.

A list of P-70 SIPP reports is available in Appendix D, and the reports are published at <www.census.gov/programs-surveys/sipp/library/publications/sipp-pubs-auto.html>.

5.2.2 Table Packages

SIPP table packages cover a variety of topics, including the dynamics of poverty, social insurance programs, and disability rates. Most table packages include summary reports of the statistics with tables and graphs to explain the main highlights from the data, other related research reports and presentations, and the complete data tables in Excel format.

SIPP table packages are located at <https://www.census.gov/programs-surveys/sipp/library/publications/tables.html>.

5.2.3 Conference Papers and Presentations

Census Bureau analysts produce a substantial amount of research drawn from SIPP data and are active participants at both national and regional conferences, including the annual meetings of the Population Association of America, the American Economic Association, and the American Sociological Association. The SIPP website makes this research available to other scholars through a series of working papers.

SIPP working papers are located at <www.census.gov/programs-surveys/sipp/library/working-papers.html>.

5.3 Webinars and Workshops

In 2019, the Census Bureau hosted a webinar series on the 2014 SIPP Panel. This seven-part series covered the following topics:

- Overview of the 2014 Panel of the SIPP
- Demographics and Residences
- Jobs
- Assets, Income, and Poverty
- Programs, Adult Well-being, and Food Security
- Health insurance, Health Care, and Disability
- Family and Fertility

Webinar materials, including audio recordings, transcripts, presentation slides, handouts, and exercises, are available from the SIPP website at <www.census.gov/programs-surveys/sipp/events/training-material.html> or the Census Academy website at <www.census.gov/data/academy/webinars/2019/sipp-series.html>.

The Census Bureau also hosts workshops to provide a hands-on introduction to SIPP and its data, including the SIPP Synthetic Beta file.[22] See <www.census.gov/programs-surveys/sipp/events/past-events.html> for more information and workshop materials.

[22] The SIPP Synthetic Beta (SSB) is a Census Bureau product that integrates person-level micro-data with administrative tax and benefit data, while maintaining confidentiality through synthetically-produced data. For more information on SSB, please visit <www.census.gov/programs-surveys/sipp/guidance/sipp-synthetic-beta-data-product.html>

6 Data Editing and Imputation

This section describes the editing and imputation procedures applied to SIPP data after the completion of each wave of interviews. Three different approaches are used for dealing with missing data in SIPP:

- Data editing (also referred to as logical imputation) for some types of item nonresponse
- Statistical (or stochastic) imputation for some types of unit nonresponse and some types of item nonresponse.
- Weighting adjustments are used for some types of noninterviews

This chapter briefly discusses the types of missing data and the goals of imputation in SIPP. It then presents an overview of the editing and imputation procedures used to deal with missing and inconsistent data. Next, this chapter provides a detailed description of each of the major steps used by the Census Bureau when creating its internal files and the files that are released for public use.

6.1 Data Editing

All respondent-reported data are edited for logical consistency. For example, say the respondent reported an age of 67 years old. However, the respondent is enrolled in 11th grade and is the biological child of a 47-year-old household member, indicating that the respondent's actual age is likely to be 17 and not 67. This is a simple example of when data editing would correct the reported age. Variables whose names start with an 'E' are edited variables.

During processing, there are also variables created that are based on the values of one or more other variables. These are recoded variables and have names starting with an 'R.' Some variables are topcoded, bottomcoded, or collapsed before they are put on the public-use file. These variables have names starting with a 'T.'

6.1.1 Types of Missing Data

As in all surveys, there are two general types of missing data in SIPP: unit nonresponse and item nonresponse. Unit nonresponse occurs when one or more of the people residing at a sample address is not interviewed and no proxy interview is obtained. Most types of unit nonresponse are dealt with through weighting adjustments.

Item nonresponse occurs when a respondent completes most of the questionnaire but does not answer one or more individual questions. Item nonresponse data in SIPP occur under the following circumstances:

- Responding sample persons refuse or are unable to provide requested information
- Interviewers fail to ask a question or incorrectly record a response
- A response is inconsistent with related responses or is incompatible with response categories

Item nonresponse data are generally imputed for all variables.

Previous SIPP Panels employed so-called Type Z imputation methods in cases of unit nonresponse. Type Z noninterviews are cases where an interview was obtained from at least one household member, but interviews were not obtained from one or more other sample persons in that household.[23] Type Z imputation methods involve imputing an entire set of data from a single donor. Starting with the 2014 SIPP Panel, Type Z imputation methods were phased out in favor of model-based imputation.

6.1.2 Goals of Imputation

There are two key problems caused by missing data:

- A lack of consistency across analyses because data users compensate for missing data in different ways, and their analyses may be based on different subsets of data
- Nonresponse is unlikely to be completely random, so estimates of population parameters are biased due to a potential non-representative sample

Because missing data are always present to some degree, analyses of survey data must be based on assumptions about patterns of missing data. When missing data are not imputed or otherwise accounted for in the model being estimated, the implicit assumption is that data are missing at random after controlling for other variables in the model. The imputation procedures used for SIPP are based on the assumption that data are missing at random within subgroups of the population. The statistical goal of imputation is to reduce the bias of survey estimates. This goal is achieved to the extent that systematic patterns of item nonresponse are correctly identified and modeled. In SIPP, the statistical goals of imputation are general, rather than specific. Instead of addressing the estimation of specific parameters, SIPP procedures are designed to provide reasonable estimates for a variety of analytical purposes.

Data editing is generally preferred over statistical imputation, and it is used whenever a missing item can be logically inferred from other data that have been provided. When information exists on the same record from which missing information can logically be inferred, that information is used to replace the missing information.

6.1.3 Assessing the Influence of Imputed Data on Analysis

Users of SIPP data interested in assessing the influence of imputed data on their analyses should consider whether SIPP imputation procedures have properties that affect their specific analytical requirements.

An evaluation of the effects of imputed data should include a review of rates of unit nonresponse and an assessment of the extent of item nonresponse. Unit nonresponse tends to increase over the life of a panel, as does the likelihood that nonresponse is not a random effect.

The effects of imputation will likely be small for items with low rates of missing data as long as rates of item nonresponse are not high among important subclasses. Lepkowski et al. (1987), using data from a large federal survey, provide a framework for evaluating the effect of imputed values on analyses. This framework can be readily adapted to SIPP analyses.

[23] This can happen because people refuse to be interviewed or they are unavailable, and a proxy is not obtained.

6.2 Imputation

SIPP uses three imputation strategies:

1. Model-Based Imputation
2. Sequential Hot-deck Imputation
3. Cold-deck Imputation

This section discusses each strategy in more detail.

6.2.1 Model-Based Imputation

Model-based imputation creates topic flags that determine whether a respondent should have answered questions about a specific content area (e.g., Social Security or TANF) if the respondent originally did not report information for that topic. The output of this prediction is a Y/N *topic flag* variable, with 'Y' indicating that there should be data related to the topic.

In addition to topic flag variables, select variables pertaining to earnings, assets, liabilities, and employment characteristics also use model-based imputation to fill in missing information. The output of these specific variables can be binary (e.g., owning an IRA retirement account), nominal (e.g., wageworker/self-employed/other), or continuous (e.g., earnings). Table 6-1 shows the group of variables imputed through model-based imputation. All other variables are imputed using hot-deck and cold-deck imputation.

TABLE 6-1. MODEL-BASED IMPUTED VARIABLES (MBIVS) IN THE 2020 SIPP

Topic-Flag-Related Variables		Non-Topic-Flag-Related Variables		
EED_SCRNR	ECREPAYANYON	EDEBT_CC	EOWN_RP	EJB5_JBORSE
EJB1_SCRNR	EVAANY	EDEBT_ED	EOWN_SAV	EJB6_INCPB
EJB2_SCRNR	ESSSANY	EDEBT_MED	EOWN_ST	EJB6_PFTLOSS
EJB3_SCRNR	ESSCANY	EDEBT_OT	EOWN_THR401	EJB6_JBORSE
EJB4_SCRNR	EWC_ANY	EIRAKEOVAL	ETHR401VAL	EJB7_INCPB
EJB5_SCRNR	EUCANY	EMED_AMT	EJB1_INCPB	EJB7_PFTLOSS
EJB6_SCRNR	EENERGY_ASST	EOWN_ANNTR	EJB1_PFTLOSS	EJB7_JBORSE
EJB7_SCRNR	ESCHOOLLUNCH	EOWN_BSI	EJB1_JBORSE	TDRAW_AMT
EJB8_SCRNR	ESCHOOLBREAK	EOWN_CD	EJB2_INCPB	
RSSI_YRYN	ELMPNOW	EOWN_CHK	EJB2_PFTLOSS	
RSNAP_YRYN	EDISABL	EOWN_GOVS	EJB2_JBORSE	
RTANF_YRYN	EDISANY	EOWN_IRAKEO	EJB3_INCPB	
RGA_YRYN	EEVERET	EOWN_LIFE	EJB3_PFTLOSS	
RWIC_YRYN	ERETANY	EOWN_MCBD	EJB3_JBORSE	
EPR_SCRNR	ESURANY	EOWN_MF	EJB4_INCPB	
EMC_SCRNR	ELIFEYN	EOWN_MM	EJB4_PFTLOSS	
EMD_SCRNR	EFINDJOB	EOWN_OINV	EJB4_JBORSE	
EML_SCRNR	EPAR_SCRNR	EOWN_PENSION	EJB5_INCPB	
EOT_SCRNR		EOWN_RE	EJB5_PFTLOSS	

The imputation process developed to handle missing data (explained below) involves estimating sequential models that predict values for an MBIV, conditional on demographic data, all other topic flags, and IRS and SSA administrative data on earnings and benefits. These data are non-monotonically missing, meaning they look like "Swiss cheese," with pockets of missing data scattered throughout. Because of this feature, we iterate our process multiple times, estimating the sequential models repeatedly and always conditioning on the most up-to-date imputations for any explanatory variable in a model. This process is commonly called Sequential Regression Multiple Imputation (SRMI) and was initially developed by Raghunathan et al. (2001).

SRMI is based on Bayesian statistical theory and is a method of estimating an approximation of the posterior predictive distribution (PPD), a conditional probability distribution that describes the data generation process and captures the relationships between the variables. To impute missing values, we take a draw from this PPD; in other words, we predict a value conditional on everything else we observe or have previously imputed for that record.

This modeling method has several advantages over hot-deck imputation. We can include many more explanatory variables in the models than can be included as stratifiers in a hot-deck. For topic flag variables, this means we can condition the imputation for a given topic flag on the imputed values for every other topic flag, hopefully approximating a joint distribution of values instead of a series of independent imputations. Parent and spouse variables can also be used as regressors or conditioning variables in the models, which allows us to better preserve the relationships among the MBIVs of household members. Finally, non-SIPP data can be used to mitigate the problem of respondents with missing values being different in unobservable ways from respondents with non-missing values. The use of rich administrative data from the Social Security Administration (SSA) and Internal Revenue Service (IRS) in model-based imputation is particularly helpful in predicting values not only for missing topic flags, but for missing information regarding earnings, assets, liabilities, and employment. This is why these select non-topic flag variables use model-based imputation.

Determining the Universe

The first step prior to imputing missing values is to establish the universe of respondents that pertain to that specific variable. The universe varies depending on the screener questions for that topic. For some topics, the universe is determined strictly by age (for example, job topics, education enrollment, SSI, and unemployment compensation). For other topics, sex and presence of children are also universe determinants (for example WIC, school meals, Social Security for kids). Another group of topics is only asked of household respondents and the answers covered a group of individuals linked through family relationships to that household respondent (for example TANF, GA, and SNAP). Private health insurance is the most complicated topic with respondents allowed to choose to report coverage for themselves, for a family member, or both. The creation of the private health insurance topic flag requires a complete accounting and reconciliation of reports from all respondents within a sampling unit to create an accurate person-level indicator of coverage. For the earnings, assets, and employment variables that use SRMI, the universe often is determined by age and/or work history.

When a respondent does not have information on a given variable needed to create a topic flag, the key element in this process is determining why this information is not present. That is, the respondent may be missing information on a given variable because they were asked, but did not answer, the particular

survey question. Alternatively, they may not have any information on that variable because they were not or should not have been asked the questions about that topic (i.e., they were not in universe for the question). Only in the former case do we seek to impute the missing data. If the universe for a topic flag includes cases intentionally missing by survey design, we would potentially have too much data that appears to be missing and we would impute values for people that should not have them. Hence, we make great efforts to specify the universe exactly as it was implemented in the survey instrument.

Building an Imputation Model

The first step in the imputation model for each MBIV output variable is the identification of important related variables. These auxiliary variables fall into two categories: stratifying and regressor variables. A stratifying variable is an indicator variable that divides respondents into homogeneous groups where each group might be expected to have different models. A regressor variable is any variable that is potentially related to the output variable and could go on the right-hand side of a regression. For each group defined by the stratifying variables, the output variable is regressed on the regressor variables for the cases where the data are not missing. The resulting regression coefficients are then used to predict the values of the output variable for the missing cases. Once a particular output variable has been imputed, it can be used as a predictor in the model for another output variable, entering into the regression for the non-missing cases, and used to impute values for the missing cases. This process is how dependency between output variables is built into the models. We rely on multiple iterations of our process to prevent the order of the modeling from affecting the imputations. The final result is a set of MBIVs with no missing values for every in-universe SIPP respondent because all the originally missing values have been replaced by imputed values.

Choosing a List of Stratifying Variables

A stratifying variable should be a categorical characteristic(s) that best predicts the presence of a topic. We create an optimal set of stratifying variables and then alternative lists in case the cell sizes created by the optimal list are too small for regressions. Any sub-sample created by the stratifying variables that does not contain at least 100 observations will not be used for regressions. Instead, it will be combined with other sub-samples that are too small and divided again using the next smallest set of stratifying variables.

Choosing a List of Regressor Variables

Regressor variables are the additional variables used to predict the output variable. Regressors may be binary, categorical, or continuous. Continuous variables are particularly useful on the regressor list since too many dummy variables in the regression can cause estimation problems in logistic regressions. For example, age categories are excellent as stratifying variables, but as regressors, it is preferable to use age, age squared, and age cubed rather than a series of dummy variables defining the age category. When variables are dropped from the optimal stratifying list, we generally include them on the regressor list so that they can still provide predictive value to the modeling process even though they are no longer a stratifying variable.

Administrative Records in Model-Based Imputation

The model-based imputation makes use of six different sources of data shared with the Census Bureau by the SSA and the IRS for stratifying and regressor variables in the models. Once the MBIVs are imputed, all administrative records are stripped from the data and no administrative records are on public-use files.

Derived from W-2 forms filed with SSA by employers, the Detailed Earnings Record (DER) extract reports uncapped income-taxable earnings for each employer that filed a W-2 record after 1977. It also contains a report of earnings that were not income taxable and were deferred into accounts like 401(k) plans. We utilize the DER to create a measure of total earnings in a given year and to count the number of jobs an individual held. From the DER we also create measures of self-employed earnings and an indicator of any deferred earnings. Finally, the DER provides the work history of respondents going back to 1978, providing indicators for years worked and average earnings across years worked in our imputation models. SIPP uses the Master Beneficiary Record (MBR) and Payment History Update System (PHUS) extracts to create indicators for whether an individual was eligible for and received OASDI payments due to retirement, disability, spouse retirement or death, parent retirement or death, or some combination of reasons. These extracts contain both present benefit receipt and historical information, so we are able to tell what year an individual started receiving benefits and whether they ever stopped. The Supplemental Security Record (SSR) provides the same information about SSI benefits. These three files combined together give us a very accurate picture of who was receiving OASDI benefits, SSI benefits, or both. This information in turn is very helpful in predicting reports of OASDI and SSI receipt.

We make use of the Numident, a register of all Social Security Numbers (SSNs) ever issued in the United States, along with the MBR and SSR, as an administrative source of birth date information. If a person is receiving benefits, we utilize the birth date from the benefits files in order to create an age for the individual during the survey reference period. If the person is not receiving benefits, we use the birth date from the Numident. While this does not replace the survey reported age on the final public-use data, we do use this age derived from administrative data as an explanatory variable in our models.

Finally, we pull in the W-2 universe using data from the IRS. This helps provide additional useful variables for the imputation process, including the size of the firm where employees work, whether the employer offers a 401(k), and whether the person is the highest paid employee. In addition to the DER, it also provides another source of earnings data, if needed during SIPP production.

Some SIPP respondents do not match to administrative data. This happened because they did not consent to having their data linked or they did not provide enough information for the Census PVS system to find a Protected Identity Key (PIK) for them. The PVS system relies on matching name, date of birth, sex, and address to administrative data files to find an SSN which is then replaced with a PIK. When administrative data is missing, it is imputed during same model-based imputation process that imputes the topic flags and select non-topic flag variables.

6.2.2 Sequential Hot-Deck Imputation

The statistical imputation method used to impute most missing items in SIPP is known as a sequential hot-deck imputation. For many topics, SRMI models determine whether a respondent with missing data should have data for a topic (e.g., receipt of unemployment insurance any time during the reference year),

whereas detailed information about the topic (e.g., months of receipt or monthly amount received) is imputed using sequential hot-deck imputation. In some cases a ratio is imputed. This ratio is used to derive the value instead of imputing the value itself. This is done to preserve relationships between certain variables (e.g., asset value and income).

In a general sense, the sequential hot-deck procedure matches a record with missing data to that of a donor with similar background characteristics and uses the donor's values. This procedure differs from data editing, which replaces missing data with inferred values based on non-missing data from the same case.

The hot-deck is cross-sectional; only values from current wave responses are used in the definition of the hot-deck cells. SIPP hot-deck procedures are designed to preserve the univariate distribution of each variable subjected to imputation. However, they do not generally preserve the covariances among variables. One consequence is that imputation can introduce inconsistencies into the data. For example, if a respondent has reported program participation, but his or her income is too high for that program, it is possible that the income data have been imputed. Whenever users detect inconsistencies, it is wise to check the status flag (described in Section 3.3) to see if the inconsistent data might have been imputed. The discussion of status flags later in this chapter provides more information.

The hot-deck procedure used in SIPP for questions is sequential because the selection of replacement values is implemented one record at a time from an ordered file.

The sequential hot-deck procedure used in SIPP involves five key steps:

1. Specifying cold-deck values
2. Sorting the sample cases
3. Identifying records with no item nonresponse or topic flag set and creating hot-deck values
4. Classifying cases into subclasses of the population, referred to as imputation classes or adjustment cells, according to values on a set of classification or auxiliary variables that are non-missing for all cases (this step is omitted in the initial processing of the key demographic items: race, sex, etc.)
5. Selecting replacement values from donor cases to impute item-missing data

Specifying cold-deck or initial donor values

Cold-deck values are the values to which each cell in the hot-deck matrix is initialized. We never intend for an imputed item to receive the cold-deck value, but it is there as a last resort, if we cannot assign a value via logical imputation or the hot-deck imputation process. The cold-deck value is usually the most commonly reported value for that particular item.

Sorting the sample cases

The records in the sample file are sorted by three geographic variables prior to imputing item missing data: primary sampling unit, segment number, and serial number. The cases are sorted prior to processing and are not resorted at any other time during the imputation process. The sorting operation creates a file in which neighboring records represent geographically proximate households.

Identifying records with no item nonresponse or topic flag set and creating hot-deck values

Once the cases have been sorted, they are processed through a series of programs. During the first pass against the programs, the cold-deck values are updated with information from the current wave. Missing data are not yet imputed, the first record in the sorted file with consistent and non-missing data for a particular group of variables is identified and the values from that case replace the cold-deck values for that section in the matrix.

The values for each subsequent record with consistent and non-missing information update the previous set of consistent and non-missing values written to the matrix. The checking and updating operation continues until all records in the data file have been processed. The last values written to the matrix serve as the starting values in the subsequent sequential hot-deck procedure. In this way, cold-deck values are rarely used as replacement values in SIPP because the initial processing usually replaces all starting values with values from the current wave of data.

Classifying cases into subclasses of the population

In the next step of the imputation procedure, each respondent record or noninterview record in the sorted file is allocated to one of the imputation classes or adjustment cells according to its values on the set of classification, or auxiliary, variables.

1. The auxiliary variables are chosen for each item or set of related items on the basis of their level of correlation with the item receiving the imputation (i.e., classification variables are chosen on the basis of their ability to explain the variability of the item or set of related items); Census Bureau researchers assign different sets of classification variables to different sets of items.
2. The auxiliary variables are either dichotomous or categorical variables (e.g., sex, race); if they are continuous, they are categorized into a parsimonious number of levels (e.g., income, asset levels).
3. The level of the auxiliary variables then defines a matrix, with the number of cells in this matrix being the product of the number of levels for each auxiliary variable. For example, an imputation defined by five variables, each with three levels, has a total of 243 (or 3^5) cells. Any given item or set of related items may have imputation matrices with the numbers of cells ranging from 10, or so, to over 100 in limited circumstances, depending on the matrix. Auxiliary variables such as sex, race, and categorizations of age (with different categorizations for different items) are used frequently in the matrices, as are more specialized auxiliary variables that are relevant for particular items (such as industry and occupation category for the monthly gross pay item).

The allocation of sample cases into imputation classes (also known as subclasses or strata) according to a set of classification variables serves several purposes. Ideally, the set of classification variables should account for a large proportion of the variance in the variable being imputed and should be associated with variations in response rates. To the extent that this is accomplished, the classification procedure creates homogeneous adjustment cells containing similar cases. In this way, donors and recipients are similar under the assumption that the nonresponse mechanism within the imputation class is not related to the item being imputed; that is, an underlying assumption is made that item nonresponse data are distributed randomly within the subclass defined by the cross-classification of the auxiliary variables. The selection of classification variables may also place bounds on the range of values that can be imputed and

implicitly satisfy edit constraints. The implicit stratification created by the sort order of the file further improves the opportunity for better imputation to the extent that nearby cases are more similar to each other than cases that are farther apart in the file. This step is omitted for the imputation of the primary demographic values that are imputed before the person-level variables.

Selecting replacement values from donor cases to impute item-missing data

The selection of replacement values for missing items is restricted to donor and recipient records within each particular cell; that is, records allocated to one cell never donate information to records in another cell with missing items. As the file is processed through the set of programs the second time, the imputations are performed and the set of hot-deck values is updated once again. The records are processed sequentially, according to the sort order of the file. A missing item is given the value of the last corresponding item that is non-missing from a record in that imputation class. If the value of an item in the current record is non-missing, it replaces the previous hot-deck value for that imputation class. In this way, the hot-deck value for each imputation class is constantly being updated with the value of the last non-missing case.

The updating is performed item by item. Missing items in one record receive the current set of replacement values. Then the non-missing values in that record are used to update the hot-deck in preparation for the next record. At any point during the process, the donated values in the hot-deck likely come from many different respondents, even within imputation classes. That is why this imputation procedure does not preserve covariances among the variables being imputed.

7 Nonsampling Error, Sampling Error, and Weighting

Statistics from surveys are subject to both sampling and non-sampling errors. This chapter addresses:

- Sources of non-sampling error and the effect of non-sampling error on estimates
- Sampling error and incorporating SIPP's complex sampling design when computing variance estimates
- SIPP weighting, available weights, and using weights to produce reliable estimates

For further information on the source of the data and accuracy of the estimates, including standard errors and confidence intervals, see the SIPP Source and Accuracy Statements at <www.census.gov/programs-surveys/sipp/tech-documentation/source-accuracy-statements.html>.

7.1 Non-sampling Error

All surveys are subject to non-sampling error. SIPP contains non-sampling error common to most surveys, as well as error unique to SIPP's longitudinal design. For example, under-coverage in household surveys is due primarily to within-household omissions; the omission of entire households is less frequent. SIPP experiences some differential under-coverage of demographic subgroups; for example, the coverage ratio[24] of black males aged 15 to 64 years is lower than that for white males in the same age group. To compensate for this differential under-coverage, the Census Bureau adjusts SIPP sample weights to population control totals. However, it is not certain to what extent those adjustments eliminate biases.

Sample attrition (when respondents leave the sample) is another source of error. Attrition reduces the available sample size and to the extent that those leaving the sample are systematically different from those who remain, survey estimates could be biased.

Response errors in SIPP take on a number of forms. Many recall errors are thought to be the source of seam bias. This effect results from the respondent's tendency to project current circumstances back onto each of the months that constitute the SIPP reference period (the prior calendar year for the 2014 Panel through 2020 SIPP, and the prior 4 months for 2008 and earlier panels). This causes any changes in respondent circumstances that occurred during the reference period to appear to have happened in the first month of the reference period. The effect is a disproportionate number of changes that appear to occur between the last month of one wave and the first month of the following wave, which is the "seam" between the two waves – hence the name.

Another potential source of response error is the time-in-sample effect. This effect refers to the tendency of sample members to "learn the survey" over time. The more times a sample member is interviewed, the

[24] The coverage ratio is the estimated population before the post-stratification ratio (second stage) adjustment divided by the independent population control.

better he or she learns the questionnaire. The concern is that sample members will alter their responses to the survey questions in an effort to conceal sensitive information or to minimize the length of the interview.

Additionally, measurement error can occur when respondents misinterpret the questions.

7.1.1 Effects of Non-sampling Error on Survey Estimates

A considerable amount of research has been conducted to investigate the various sources of non-sampling error in SIPP. The results of this research are summarized in the *SIPP Quality Profile, 3rd edition*, available at <www.census.gov/library/working-papers/1998/demo/SEHSD-WP1998-11.html>. Additional findings about SIPP data quality, especially for more recent panels, can be found in the National Research Council's 2018 report *The 2014 Redesign of the Survey of Income and Program Participation: An Assessment*, and in Appendix A of the National Research Council's 2009 report *Reengineering the Survey of Income and Program Participation*. Despite the volume of methodological research, it remains difficult to quantify the combined effects of non-sampling errors on SIPP estimates. This problem is made more complex because the effects of different types of non-sampling error on survey estimates vary, depending on the estimate under consideration. However, there are some findings about non-sampling error that SIPP users should bear in mind when conducting their analyses and examining their results. Those findings include:

- Some demographic subgroups are underrepresented in SIPP because of under-coverage and nonresponse. They include young black males, metropolitan residents, renters, people who changed addresses during a panel (movers), and people who were divorced, separated, or widowed. The Census Bureau uses weighting adjustments and imputation to correct the underrepresentation. However, those procedures may not fully correct for all potential biases (SIPP Quality Profile, 3rd Ed., Chapter 8).
- SIPP estimates of the working population differ from those produced from CPS. The differences may be explained largely by substantial conceptual and operational differences in the collection of labor force data in the two surveys (SIPP Quality Profile, 3rd Ed., Chapter 10).
- SIPP estimates of the number of births compare favorably with CPS estimates. Both surveys, however, provide estimates that are low relative to records from the National Center for Health Statistics (NCHS). SIPP estimates of the number of marriages are fairly comparable with NCHS counts, but SIPP estimates of the number of divorces are consistently lower than NCHS estimates (SIPP Quality Profile, 3rd Ed., Chapter 10).
- Across all age groups, particularly children and the elderly, SIPP continues to identify more sources of family income than CPS. SIPP's greater effectiveness than CPS in capturing income from multiple sources among retired workers demonstrates an important way in which SIPP appears to provide a better tool for policy analysis (Czajka et al., 2008).
- In 2005, SIPP captured a higher share of aggregate annual benefits than CPS for Food Stamps, AFDC/TANF, OASI, and SSI, but was only marginally better for SSDI. In 1987, SIPP was on par with CPS for AFDC/TANF and SSDI. Whether because of poor recall or because respondents sometimes answer on the basis of their current situation, CPS estimates of persons who ever participated in a program sometimes line up with SIPP estimates of average monthly participants (Czajka, 2009).

- When compared to the Survey of Consumer Finances (SCF) by the Federal Reserve Board for late 1998 and early 1999, SIPP is much more effective in capturing liabilities than assets. SIPP's estimate of aggregate assets was 55 percent of the SCF estimate of $34.1 trillion, but its estimate of aggregate liabilities was 90 percent of the SCF estimate of $5.0 trillion (Czajka, 2009).
- Average monthly estimates of health insurance coverage from SIPP compare closely to estimates of health insurance coverage obtained in the National Health Interview Survey (NHIS) and CPS (Czajka, 2009).
- When examining the use of housing unit controls versus population controls, a team at the Census Bureau concluded that the weighting adjustment for within-household under-coverage when using population controls by age, sex, and race tended to be higher than the weighting adjustment for housing unit coverage when using housing unit controls, which focus on coverage of housing units (including whole households). Thus, when population-control-based weights are applied to characteristics such as household relationship, the estimate of householders (family plus nonfamily) will almost always be higher than the corresponding housing-unit-control-based weights that are applied to obtain the estimate of occupied housing units (Cresce et al., 2013).
- There seems to be evidence of potential bias due to nonresponse for some key statistics in SIPP. The Census Bureau has done nonresponse bias studies to investigate the effect of decreasing response rates, but more work needs to be done to truly quantify that bias (McMillan & Culver, 2013). These studies are available at <www.census.gov/programs-surveys/sipp/tech-documentation/nonresponse-reports.html>.

7.2 Sampling Error and Variance Estimation

This section discusses methods for obtaining the sampling error estimates derived from the SIPP panels. The sample selected for each SIPP panel is a stratified, multistage probability sample. This complex sample design must be taken into account when calculating the variances of SIPP estimates. The SIPP data files contain variables related to the sample design, which are created for the purpose of variance estimation. Several software packages are now available for computing variance estimates for a wide range of statistics based on complex sample designs, such as the PROC SURVEYMEAN procedure in SAS and the SVYSET procedure in Stata. Using the variables that specify the design, these programs can calculate appropriate variances of survey estimates. The Census Bureau also provides generalized variance functions (GVFs) that can be used to obtain approximate estimates of sampling variance for SIPP estimates. Finally, sets of replicate weights are also provided in the SIPP data files and can be used to estimate more accurate standard errors and variances for SIPP estimates. While replicate weighting methods require more computing resources, many statistical software packages have procedures that simplify the use of replicate weights for users.

A common mistake in the estimation of sampling error for survey estimates is to ignore the complex survey design and treat the sample as a simple random sample (SRS) of the population. That mistake occurs because most standard software packages for data analysis assume simple random sampling for variance estimation. When applied to SIPP estimates, SRS formulas for variances typically underestimate the true variances. This section describes how appropriate variance estimates, which take into account the complex sample design, can be obtained for SIPP estimates.

7.2.1 Direct Variance Estimation

The primary sampling unit (PSU) plays a key role in variance estimation with a multistage sample design. SIPP PSUs are mostly counties, groups of counties, or independent cities, which are sampled with probability proportional to size within strata. Self-representing (SR) PSUs are so large that they are included in the sample with certainty. Because no sampling is involved, the SR PSUs are, in fact, not PSUs but strata. Non-self-representing (NSR) PSUs are stratified, and two NSR PSUs per stratum are sampled without replacement, so that no PSU is selected more than once for the sample.

Although the SIPP PSUs are selected without replacement (as is the case with most multistage designs), for the purpose of variance estimation they are treated as if they were sampled with replacement. The with-replacement assumption greatly facilitates variance estimation since it means that variance estimates can be computed by taking into account only the PSUs and strata, without the need to consider the complexities of the subsequent stages of sample selection. This widely used simplifying assumption leads to an overestimation of variances, but the overestimation is not great.

7.2.2 Variance Units and Variance Strata

Starting with the 1990 Panel, sample member records contain data on the PSU and stratum from which a person was sampled. Software packages use this information to estimate variance. However, to avoid potential identification of sampled persons in small areas, the original PSU and stratum codes are not included in the SIPP public data files. Instead, sets of PSUs are combined across strata to produce variance units and variance strata, with two variance units in each variance stratum. Variance units and variance strata may be treated as PSUs and strata for variance estimation purposes. Their use does not give rise to any bias in the variance estimates. However, the variance estimates are somewhat less precise than those obtained from using PSUs and strata that have not been combined.

7.2.3 Replicate Weights

Starting with the 1996 Panel, the Census Bureau released replicate weight files along with the SIPP data files. Data users should use Fay's modified balanced repeated replication (BRR) method for estimating variances for the SIPP panels. The difference between the basic BRR method and Fay's method is that the BRR method uses replicate factors of 0 and 2, whereas Fay's method uses one factor, k, which is in the range (0, 1), with the other factor equal to $2 - k$. In Fay's method, the introduction of the perturbation factor $(1 - k)$ allows the use of both halves of the sample. Thus, Fay's method has the advantage that no subset of the sample units in a particular classification will be totally excluded. Figure 7-1 shows Fay's variance estimation formula used for SIPP estimates.

FIGURE 7-1. SIPP VARIANCE ESTIMATION FORMULA: 2020 SIPP

$$Var(\theta_0) = \{1/[G(1-k)^2]\} \sum_{i=1}^{G} (\theta_i - \theta_0)^2,$$

where:

G = number of replicates;

$1-k$ = perturbation factor;

i = replicate $i, i = 1\ to\ G$;

θ_i = i^{th} estimate of the parameter θ based on the observations included in the ith replicate;

θ_0 = survey estimate of the parameter θ based on the full sample.

The 1996 and 2001 SIPP Panels have 108 replicate weights available. The 2004 and 2008 Panels have 120 replicate weights, and the 2014 Panel and the 2018 through 2020 SIPP have 240. All replicate weights are calculated based on a perturbation factor of 0.5 (k = 0.5). For example, inserting the 2020 values into Figure 7-1's equation results in the variance formula of

$$Var(\theta_0) = \{1/(240 * 0.5^2)\} \sum_{i=1}^{240} (\theta_i - \theta_0)^2$$

The SIPP replicate weights are available via the SIPP Data webpage at <www.census.gov/programs-surveys/sipp/data.html>.

Data users should account for the complex survey design and sampling structure to create unbiased estimates. Below is example STATA code, R code, SAS code, and Python code that can produce an unbiased point estimate and variance estimate of mean personal monthly income for those aged 15 and older (after merging the primary data with the replicate weights):

STATA

```
svyset [pw=wpfinwgt], brrweight(repwgt1-repwgt240) fay(.5) vce(brr) mse

svy: mean tptotinc
```

R (using the *survey* library)

```
sipp.svy = svrepdesign( data = pu_rw_merged_2020, weights = ~WPFINWGT,
       repweights = matches("REPWGT[1-240]+") , type = "Fay" , rho = 0.5 )
svymean( ~TPTOTINC , sipp.svy , na.rm = TRUE )
```

SAS

```
proc surveymeans data=pu_rw_merged_2020 varmethod = BRR (Fay=0.5) mean cv sum cvsum var T;
       var TPTOTINC;
       weight WPFINWGT;
       repweights REPWGT1-REPWGT240;
run;
```

Python (using *NumPy* and *Pandas*)

```
df = pu_rw_merged_2020.loc[df.TPTOTINC.isna() != True]

point_estimate = np.nansum(df.TPTOTINC*df['WPFINWGT'])/np.nansum(df['WPFINWGT'])

rep_means   =   [np.nansum(df.TPTOTINC*df['REPWGT'+str(i)])/np.nansum(df['REPWGT'+str(i)])
       for i in range(1,241)]

variance = (1/(240*.5**2))*sum((rep_means - point_estimate)**2)

print("Point estimate:{:.2f} , Standard error:{:.2f}".format(point_estimate,variance**.5))
```

7.2.4 Approximate Variance Estimates

Although replicate weights are highly recommended to produce accurate estimates of standard errors and variances, the Census Bureau provides two forms for approximate variance estimation: GVFs and design effects (DEFF) for person and household-level estimates in multiple domains. The generalized estimates provide indications of the magnitude of the sampling error in the survey estimates. They serve as convenient ways to summarize the sampling errors for a broad variety of estimates. The GVFs for SIPP were derived by modeling the standard error behavior of groups of estimates with similar standard errors. The mathematical form of the function adopted is

$$s = (ax^2 + bx)^{1/2}$$

where s represents the standard error and x represents the value of an estimate. The parameters a and b are derived on the basis of a selected group of estimates. They are updated annually and are included in the SIPP Source and Accuracy Statements that accompany the SIPP data files for each panel, available at <www.census.gov/programs-surveys/sipp/tech-documentation/source-accuracy-statements.html>. It is essential to use the parameter estimates for a specific panel and to follow the instructions to apply necessary adjustments to obtain the correct estimates for subgroups. Besides GVFs, the Census Bureau provides design effects, which are also available in the Source and Accuracy Statements (along with more details and examples for how to use these methods).

Data users should note that the generalized variance estimates for estimating the standard errors of other statistics may not be accurate for small subgroups. Using the 1984 SIPP Panel, Bye and Gallicchio (1989) developed variance functions for participants of Old-Age, Survivors, and Disability Insurance (OASDI) and Supplemental Security Income (SSI) programs. They found that for estimated totals with less than 10 million people, the generalized standard error estimates provided by the Census Bureau were 1.20 to 1.75

times larger than those obtained from the variance functions developed specifically for that subgroup. Therefore, it is recommended to use replicate weights whenever possible.

7.2.5 Variance Estimation with Imputed Data

As discussed above, we use several imputation methods to fill in missing data in SIPP. Imputation fills in gaps in the data set, making data analysis easier, and allowing more people to be retained as panel members for longitudinal analysis. The concern, however, is that imputation fabricates data to some degree. Treating the imputed values as actual values in estimating the variance of survey estimates leads to an overstatement of the precision of the estimates. It is important to recognize this fact when sizable proportions of values are imputed.

7.3 Weighting

Person weights estimate the number of people in the target population that a person represents. In general, since population units may be sampled with different selection probabilities and since response rates and coverage rates may vary across subpopulations, different responding units represent different numbers of units in the population. The use of weights in survey analysis compensates for this differential representation, thus producing estimates that relate to the target population.

SIPP weights vary due to differential sampling rates as a result of oversampling and because response and coverage rates vary across subpopulations. For example, in the 2020 SIPP, the final person lower quartile weight is 4,200 and the upper quartile weight is 7,600 (the maximum weight is 39,000). A respondent with a final person weight of 4,200 represents 4,200 people in the U.S. population for the reference month, whereas a respondent with a weight of 7,600 represents 7,600 people.

7.3.1 Choosing a Weight, 2018 Through 2020 SIPP

The decision of which weight to use for a given analysis depends on 1) the population the results are intended to apply to, and 2) the duration of interest for that analysis. The weights in the SIPP files are constructed for sample cohorts grouped by:

- Month (e.g., the reference month weights)
- Year (e.g., the calendar year weights), and
- Multi-Year periods (e.g., the longitudinal weights that cover consecutive two-, three-, or four-year periods).

Using these weights, data users can choose to base their analyses on:

- A cross-sectional sample at a given month,
- A sample that provides continuous monthly data over a year, or
- A longitudinal sample that provides monthly data over more than one consecutive year.

Using longitudinal samples, data users can study issues such as the dynamics of program participation, lengths of poverty spells, and changes in other circumstances (e.g., household composition). The multi-year longitudinal weights include weights for the cohort of respondent who have positive calendar year weights for every year of the longitudinal period, regardless of panel. Each yearly SIPP release (excluding 2018 SIPP) includes one, two, or three sets of longitudinal weights covering a reference period of two,

three, and four earlier calendar years respectively. The number of longitudinal weights produced depends on the number of longitudinal sample cohorts available during the year of interview. For example, two longitudinal weights, FINYR2 and FINYR3, are available for the 2020 SIPP, covering the two-year reference period of 2018-2019 and the three-year reference period of 2017-2019 respectively. The Census Bureau makes nonresponse adjustments to the longitudinal weights to compensate for attrition and post-stratification adjustments to make the weighted sample totals conform to population totals for key variables.

TABLE 7-1: CHOOSING APPROPRIATE WEIGHTS, 2018 THROUGH 2020 SIPP

Time Duration	Person Level	Example
Monthly	WPFINWGT	Income in July, 2019
Calendar Year	WPFINWGT in December of the reference year	Total TANF receipt in 2019
Multi-Year	FINYR# where # = 2, 3, 4 represents the number of years in the reference period	Median duration spell of unemployment from June 2017 to June 2019

7.3.2 How Weights Are Constructed

The basic components for all the different sets of weights are the same, namely:

- A base weight that reflects the probability of selection for a sample unit;
- An adjustment for subsampling within clusters;
- An adjustment for movers (in Waves 2 and beyond);
- A nonresponse adjustment to compensate for sample nonresponse;
- An adjustment for combining sample cases from multiple panels (if overlapping panels cover the reference period); and
- A post-stratification (second-stage calibration) adjustment to correct for departures from known population totals.

7.3.3 Reference Month Weights

Reference month final weights are provided on the SIPP calendar-year data files for persons. Household, family, or subfamily weights can be derived from the final person weights. One option is to use the average of the person month weights for all persons in the household, family, or subfamily. Another option is to specify a household, family, or subfamily reference person and use his or her person month weight as the household, family, or subfamily weight. This section summarizes the steps involved in constructing the final person weights for a reference month. Positive weights are computed for all interviewed persons in the SIPP universe for the specified month.

A number of factors lead to fluctuations in sample size (and therefore weights) from month to month. They include births, deaths, and immigration and emigration from the population (and therefore from the sample). In addition to those population dynamics, people move into and out of the sample as a result of the changing household composition of sample members. Since response characteristics can differ by panel and wave, when there are overlapping panels included in a calendar-year file, the previously listed weighting components up through the nonresponse adjustment are done separately by wave first to obtain a noninterview weight for each sample person.

For Wave 1 households, the noninterview weight for each sample person per month is a product of the components:

- Wave 1 base weight: This weight is the inverse of the probability of a sample household's address being selected.
- Weighting-control factor: This factor adjusts for the occasional subsampling of clusters. Clusters are occasionally subsampled in the field when they turn out to be much larger than expected
- Wave 1 nonresponse adjustment: This adjustment compensates for different rates of household noninterview within adjustment classes. 512 nonresponse adjustment classes are defined based on a cross-classification of characteristics. Those characteristics include Census Region; MSA/Place Status (MSA-central city, MSA-non-central city, other place); race of reference person (black, nonblack); household tenure (owner, renter); household size (1, 2, 3, 4+ people); and the within-primary-sampling-unit poverty stratum (poverty, non-poverty). Some adjustment classes are combined to ensure that the adjustment for each class does not exceed a factor of 6 and each class contains at least 30 unweighted sample households.

For households in Waves 2 or later, each person receives an initial weight that is the noninterview weight carried over from Wave 1. This weight is adjusted to compensate for changes in the sample between waves resulting from movers and nonresponse. The noninterview weight for each sample person per month is the product of the following:

- Wave 2+ initial weight: This weight is the product of the Wave 1 base weight, the weighting-control factor, and the Wave 1 nonresponse adjustment factor. It is for each original sample person who is a reference person or is living in group quarters for the current wave.
- Wave 2+ mover adjustment: This adjustment is made to compensate for including people who were not in the original sample but were in the SIPP universe in Wave 1 and who moved into a sample household after Wave 1. For people in housing units that contain adult members who were not part of the original sample but were in the SIPP universe at Wave 1, the weights are decreased. For example, if a third adult moves into a household occupied by two original sample persons, all three adults would receive the initial weight of the original sample persons multiplied by a factor of two-thirds.
- Wave 2+ nonresponse adjustment: The nonresponse adjustment for Waves 2 and beyond is used to compensate for household nonresponse after the first interview. The nonresponse adjustment classes are defined on the basis of sample unit characteristics and personal demographic characteristics from the most recent wave. The information used consists of household characteristics. Reference person characteristics are used to define some of the household characteristics. Tenure (owner/renter occupied), household type (female householder, no spouse

present; 65+; other), race and Hispanic origin, and education level are defined at the household level by using reference person data. Other household characteristics include size, poverty status, type of income, type of financial assets, census division, and number of imputed items. Some adjustment classes are combined to ensure that the adjustment for each class does not exceed a factor of 4 in the 2018 SIPP and each class contains at least 30 unweighted sample households.

Once the noninterview weights are obtained separately for each wave within the calendar year, they are adjusted by a factor that accounts for combining the waves from multiple panels together. For each person in a panel, this factor is a ratio of the panel's sample size relative to the total sample size of all overlapping panels.

Finally, the person weights go through a second stage calibration. This adjustment brings the sample estimates into agreement with independent monthly estimates of population totals. The characteristics used for calibration include age, race, sex, Hispanic origin, family relationship, state, and household type. A spouse equalization is also included. A raking procedure is used to ensure that the weights agree with all the control totals included for calibration.

7.3.4 Longitudinal and Calendar Year Weights

Starting with the 2014 Panel, calendar year weights are no longer provided on the longitudinal weight files since an entire reference year is covered by a single wave of the panel. The same is true for the 2018 SIPP and later calendar-year releases.

For the 2014 Panel through 2020 SIPP, final calendar year weights are available for people who had an interview covering December of the reference year. Therefore, calendar year weights are the same as the December month weights of the same year.

Starting with the 2019 SIPP, multi-year-period longitudinal weights are available for the cohort of people who receive positive calendar year weights in every year of the reference period. These weights (and corresponding replicate weights) are released in separate longitudinal files that can be merged to the calendar-year data file using the sample unit ID (SSUID) and person number (PNUM). Depending on the data available, a yearly SIPP release could include any or all of "two-", "three-", or "four-year" longitudinal weights. For example, the 2019 SIPP release includes only "two-year" longitudinal weights files covering the reference period 2017-2018, while the 2020 SIPP contains both "two-year" longitudinal weights files covering the reference period 2018-2019 and "three-year" longitudinal weights files covering the reference period 2017-2019.

The final longitudinal weights are constructed from the following three components:

- Initial weight: This weight is the household noninterview weight from the first year of the reference period. If persons from more than one panel are receiving the longitudinal weight, this noninterview weight is adjusted by a factor to account for combining cases from overlapping panels.
- Nonresponse adjustment factors: These factors account for noninterviewed eligible sample persons not already accounted for in the noninterview adjustment component of the initial weight. The adjustment classes are similar to those used in the Wave 2+ nonresponse adjustment factors.

- Second-stage calibration factors: These factors are determined by a process similar to that used for the reference weights. The control totals used for the longitudinal weights are the population estimates for December of the last year in the reference period. Therefore, for a given yearly SIPP data release, the calendar year weights and all multi-year longitudinal weights will use the same control totals.

Appendices

A. Acronyms

ACS	=	American Community Survey
ADL	=	Activities of Daily Living
AFDC	=	Aid to Families with Dependent Children
ASA	=	American Statistical Association
BLS	=	Bureau of Labor Statistics
BW	=	Base weight
CAI	=	Computer-assisted interviewing
CAPI	=	Computer-assisted personal interviewing
CHI	=	Contact History Instrument
CMSA	=	Consolidated Metropolitan Statistical Area
CNSTAT	=	Committee on National Statistics
CPI-U	=	Consumer Price Index for all Urban Consumers
CPS	=	Current Population Survey
CV	=	Coefficient of Variation
DADS	=	Data Access and Dissemination System
DCF	=	Duplication Control Factor
DER	=	Detailed Earnings Record
DES	=	Data Extraction System
EBT	=	Electronic Benefit Transfer
EDs	=	Enumeration Districts
EHC	=	Event History Calendar
EITC	=	Earned Income Tax Credit
FERS	=	Federal Employees Retirement System
FIPS	=	Federal Information Processing Standards

GA	=	General Assistance
GED	=	General Education Development
GVFs	=	Generalized Variance Functions
HUs	=	Housing Units
ICPSR	=	Inter-university Consortium for Political and Social Research
IRA	=	Individual Retirement Account
IRS	=	Internal Revenue Service
ISDP	=	Income Survey Development Program
LQs	=	Living Quarters
MAF	=	Master Address File
MBIV	=	Model-Based Imputed Variable
MSA	=	Metropolitan Statistical Area
NAF	=	Noninterview adjustment factor
NCF	=	New-construction noninterview adjustment factor
NCHS	=	National Center for Health Statistics
NHIS	=	National Health Interview Survey
NLS	=	National Longitudinal Surveys
NOI	=	Neighborhood Observation Instrument
NSR PSUs	=	Non-self-representing PSUs
OASDI	=	Old-Age, Survivors, and Disability Insurance
OMB	=	Office of Management and Budget
PRWORA	=	Personal Responsibility and Work Opportunity Reconciliation Act
PSUs	=	Primary Sampling Units
RIP	=	Respondent Identification Policy
SCF	=	Survey of Consumer Finances
SIPP	=	Survey of Income and Program Participation

SNAP	=	Supplemental Nutrition Assistance Program
SPD	=	Survey of Program Dynamics
SRMI	=	Sequential Regression Multiple Imputation
SRS	=	Simple Random Sample
SSA	=	Social Security Administration
SSCA	=	Second-stage Calibration Adjustment
SSI	=	Supplemental Security Income
TANF	=	Temporary Assistance for Needy Families
UC	=	Unemployment Compensation
UI	=	Unemployment Insurance
USDA	=	United States Department of Agriculture
WIC	=	Special Supplemental Nutrition Program for Women, Infants, and Children

B. Glossary

A

Address Unit - This collection unit is a person or group of persons living at the same address at the time of the interview. The address unit may consist of one person living by himself or herself, a group of unrelated individuals, or one or more families.

Allocation Flag - See *Status Flag.*

B

Bottomcoding - Practice of recoding continuous variables to protect against the possibility that a user might recognize the identity of a SIPP respondent with an extreme value. Variables below a minimum value are recoded to that minimum value or to a mean of responses less than that value.

C

CAI (Computer-assisted Interviewing) - A method of interviewing in which a computer is used as the data collection instrument.

CAPI (Computer-assisted Personal Interviewing) - A method of interviewing in which field representatives use a laptop computer to collect data during in-person interviews. In SIPP, the field representatives also periodically use the laptop computers during telephone interviews conducted from their homes.

Cold-deck Matrix - The matrix of starting values that constitutes the first step in the hot-deck imputation procedure. The matrix values can be determined from information external to the current file being processed or can be determined from reported information from the current file.

Continuation Flags - Person-level variables that indicate whether spells ending in the last month of the reference period (1) ended in the last month of the reference year, (2) continued into the interview year but ended before the interview month, or (3) was ongoing as of the time of interview.

Core Content - Questions asked at every SIPP interview in the 1984-2008 Panels. These questions cover demographic characteristics, work experience, earnings, program participation, transfer income, and asset income.

Core Wave Files - Files for the 1984-2008 Panels that contain the core data from one wave of interviews.

Coverage Unit - A set of people covered by a health insurance policy or an assistance program. A coverage unit may include one person, a subset of people in the household, or all household members. Coverage units are also known as program units.

Cross-sectional - Pertaining to data collected for a single time period from a representative sample. In SIPP hot-deck imputation procedures, cross-sectional refers to current-wave data.

Current Population Survey (CPS) - A labor force survey sponsored jointly by the Census Bureau and the Bureau of Labor Statistics that is used to compute the government's official monthly unemployment statistics along with other estimates of labor force characteristics.

D

Data Dictionary - Contains information about the file structure and the names, locations, and contents of all variables in a microdata file.

Data Editing - The use of related information to replace missing or inconsistent data in the survey.

E

F

Family - Two or more people who are living together and are related by birth, marriage, or adoption.

Following Rules - SIPP rules that guide which original sample members continue to be interviewed should they move.

Full Panel Files - Files containing all data for every person who was a member of a SIPP panel at any time during the life of that panel.

G

General Income - Any type of income except earnings and asset income.

H

Hot-deck Matrix - The matrix used in all but the first stage of hot-deck imputation. As cold-deck values are replaced with information from the current wave, the resulting array of cells constitutes the hot- deck matrix.

Hot-deck Procedure - The statistical method used to impute items missing from the core questionnaire and topical modules. This procedure replaces missing item data in a wave with non-missing values from similar interviewed cases. The imputation method can be a purely cross-sectional procedure of locating donors from the current file on the basis of characteristics reported in this wave, or it can be a longitudinal procedure of locating donors from the prior wave on the basis of characteristics reported at that earlier time for items missing in the current wave.

Household - People living in a housing unit at the time of the interview. SIPP infers households from the interviews conducted at each address.

Household-level Non-interviews - See *Household Nonresponse*.

Household-level Variables - Variables that have the same value across all people in the household for all months of the reference period.

Household Nonresponse - Nonresponse that occurs when the interviewer either cannot locate a household or cannot interview any of its adult members. See Type A, Type B, Type C, and Type D non-interviews.

Household Reference Person - See *Reference Person*.

Household Respondent - The first eligible adult household member interviewed

Housing Unit - Living quarters with its own entrance and cooking facilities

I

Imputation - The most common method for handling missing data in SIPP. Imputation replaces missing values with statistical estimates that are based on the best relevant information available.

Imputation Flag - See *Status Flag*.

In Scope - Being part of the survey universe.

Interview Month - The month during which the interview takes place.

Item Nonresponse - A source of missing data that occurs when a respondent does not answer one or more questions, even though most of the questionnaire is completed.

J

K

L

Logical Imputation - See *Data Editing*.

Longitudinal - Pertaining to data collected at different times over an extended period from a representative sample. In SIPP hot-deck imputation procedures, longitudinal refers to previous-wave data.

M

Merged Households - Households created either when two separate sampling units, each containing original sample members, are merged together, perhaps because of a marriage, or when a household splits into two new households and later the households recombine.

Microdata Files - Data files containing information at the person, family, or household level. For SIPP, they include the core wave files, topical module files, and full panel files.

Missing Item Data - Data that are missing for one or more individual questions or variables, but the observation has sufficient reported information to be classified as interviewed.

Missing Waves - Waves in which a respondent has no data, although data are present for other waves.

Monthly Interview Status Variables - Variables that indicate whether a person was in sample in a particular month, and whether a person was in sample in the interview month. They are known as the PP-MIS variables.

Mover - An original sample person who moves during the life of the panel.

N

National Guard Retirement - An Army National Guard retirement plan provided for employees who have 20 years of service in the National Guard. By age 60, this retirement pension may combine with any additional income or retirement.

National Longitudinal Survey (NLS) - Collects data on current labor force and employment status, work history, and characteristics of the current or last job.

Non-self-representing (NSR) Primary Sampling Units (PSUs) - Smaller PSUs that must be grouped with similar PSUs from the same region in order to form strata for sampling. This level of geography is not available on the public-use files.

Noninstitutional Group Quarters - Noninstitutional living quarters, such as rooming and boarding houses, college dormitories, convents, and monasteries. These do not constitute households and are often treated differently from households.

O

Original Sample Members (OSPs) - All people who were interviewed in the first wave of the panel and any children subsequently born to or adopted by them.

Oversampling - Sampling that involves selecting certain groups or units with higher probabilities than others, resulting in the oversampled group having greater representation than occurs in the population from which it was drawn.

P

P-70 reports - Primary source for published estimates from the SIPP. These reports can be obtained from the SIPP website.

Panel - Refers both to a new sample that is introduced periodically in the SIPP and to the full collection of information for that sample. For example, the 2018 Panel refers to both the sample introduced in 2018 and the 4 waves of interviews conducted with that sample.

Panel Non-respondents - Persons for whom an interview is missing for a wave.

Person-level Non-interviews - This type of non-interview occurs when data are collected for at least one member of a household, but are missing for one or more other sample persons within that household. See *Type Z non-interview*.

Person-level Variables - Variables that have the same value across all months of the reference period for a respondent.

Person-month Files - Microdata files containing a record for each person in a wave, for each month of the reference period the person was in the sample.

Person-month-level Variables - Variables that may vary month-to-month for a respondent.

Person Nonresponse - Nonresponse that occurs when at least one person in the household is interviewed, while at least one other person is not. See *Type Z non-interview*.

Pointer - Variables that are used to link people within a household. An example are parent pointers, which link children to their parent(s). On a child's record, you will find EPNPAR1 and EPNPAR2. These variables which contain the PNUM(s) of any parents living in the household.

Primary Family - Family containing the household reference person and related individuals.

Primary Individual - A household reference person who lives alone or lives with only nonrelatives.

Primary Sample Members - See *Original Sample Members*.

Primary Sampling Units (PSUs) - Geographic units based on Census data and used in developing the SIPP sample. This level of geography is not available on the public-use files.

Program Units - The group of individuals which constitutes one case, as defined by a particular benefit program. In SIPP, program units apply to health insurance and transfer programs and are identified for programs in which a case can consist of more than one person.

Proxy Interviews - Interviews taken on behalf of a sample member who is unable to answer, including those under age 15 at the time of the interview.

Public-Use Microdata Files - Data files that have been prepared by the Census Bureau for public-use. These files have already been processed to impute missing data, to edit data for confidentiality, and to provide weights. Microdata files are available from the SIPP website.

Q

R

Recodes - Variables created from other variables. These variables start with "R."

Reference Months - The months that constitute the reference period for a wave.

Reference Period - The period of time for which most interview questions are in reference to. For the 2014 Panel through 2020 SIPP, the reference period is the prior calendar year (January to December). For the 1984 to 2008 Panels, the reference period was the four calendar months preceding the month of interview.

Reference Person - An owner or renter of record who is at least 15 years old and who can reasonably be expected to answer questions about the household in general and about other household members should they be unavailable for interview. All people in the household are listed according to their relationship to the reference person in the interview month. If multiple respondents are listed as the owner or renter of record, the reference person is recorded as the first person listed as the owner or renter.

Related Subfamily - A married couple and dependents or parent-child family related to the reference person but not including him or her. An example would be the reference person's daughter and son-in-law. Related subfamilies can only be identified in the interview month, since a reference person is not identified in the monthly data.

Respondent - Someone in the household for whom SIPP collects information

S

Sample Attrition - Loss of sample members. Sample attrition rates decline over time, but total attrition numbers increase.

Seam Effect - The tendency of respondents to report a disproportionate number of changes as occurring at the "seam" between the end one wave and the beginning of another.

Secondary Families - Two or more people living in the same household who are related to each other but not to the household reference person.

Secondary Individual - An individual who is neither a household reference person nor a relative of any other people in the household.

Secondary Sample Members - People living with original sample members.

Self-representing (SR) Primary Sampling Units (PSUs) - Larger PSUs that do not have to be combined with other PSUs in order to form strata for sampling. This level of geography is not available on the public-use files.

Self-response - A respondent who provides information about him/herself.

Sequential Hot-deck Procedure - See *Hot-deck Procedure.*

Source and Accuracy Statement - A statement included with the technical documentation that accompanies public-use files; it contains detailed information about weights on the files, when and how to make adjustments to the weights, and how to use generalized variance procedures to compute standard errors for some common types of estimates. It also includes cautions for users about sources of non-sampling error.

Spell - A period of consecutive months spent in an activity, such as in a job or receiving program benefits.

Spell-level Variables - Variables that have the same value across all months of a spell for a respondent, but values may vary across spells.

Status Flag - A status flag is associated with each item subject to statistical imputation and indicates whether information has been imputed, and if so, the type of procedure used. These variables start with 'A.'

Survey of Program Dynamics (SPD) - An offshoot of SIPP that began re-contacting members of the 1992 and 1993 Panels, with data collection to continue through 2001 in order to collect 10 years of data.

T

Technical Documentation - Information that accompanies microdata files and that includes a description of file contents, a glossary, codes, a data dictionary, a source and accuracy statement, and a copy of the core questions for the panel in question.

Time-in-Sample Effect - Tendency of sample members to "learn" the survey over time, possibly resulting in altered responses.

Topcoding - Practice of recoding continuous variables to protect against the possibility that a user might recognize the identity of a SIPP respondent with an extreme value. Variables exceeding a maximum value are recoded to that maximum value or to a mean of responses in excess of that value.

Topical Content - For the 1984-2008 Panels, questions that are not repeated in every wave. They cover a wide range of topics and can occur once or more than once in a panel. The questions are grouped into modules by topic.

Topical Module Files - For the 1984-2008 Panels, files containing all topical module data from the wave in question.

Topical Modules - For the 1984-2008 Panels, collections of questions asked periodically, but not at every interview, about various topics that might be outside the range of the core content.

Topical Module Imputation Procedure - For the 1984-2008 Panels, missing data in topical modules were imputed using the same hot-deck procedure used to impute missing data in the core questionnaire.

Type A Non-interview - Households that are occupied by people eligible for interview but for which no interview is obtained.

Type B Non-interview - A household non-interview that occurs when the address unit is vacant or in some way unfit for residence.

Type C Non-interview - In Wave 1, a household non-interview that occurs when the housing unit has been demolished or converted to some other use; in subsequent waves, a household non-interview that

occurs when all sample members in a household are outside the scope of the survey, for example, deceased, living abroad, living in institutions, or living in armed forces barracks.

Type D Non-interview - Households or people who have moved to an unknown address, or who have moved more than 100 miles from the nearest field representative and for whom no telephone interview is conducted. This type of non-interview applies only to Wave 2 and beyond.

Type Z Imputation - Procedures used in the 1984-2008 Panels to impute missing data for Type Z non-interviews and for situations when a person was in sample early in the wave but not in sample by the month of interview.

Type Z Non-interview - An eligible person in an interviewed household from whom the field representative could not get an interview or for whom the interviewer could not obtain a proxy interview. A non-interview also occurs when a person who was part of the household for a portion of the reference period moves and is no longer a household member on the date of the interview. If the person is an original sample member, an effort will be made to locate and follow the person.

Type 2 Person - A person who lived with a respondent during the reference year but not at time of interview. A Type 2 person does not get a person record in the public-use microdata, and can have been at any address a respondent lived at during the reference year, not just interview address.

U

Under-coverage - Underrepresentation of demographic subgroups within the surveyed population.

Unrelated Subfamily - A family, that is, a group of two or more related individuals, living at a sample address unit that does not contain the reference person or anyone related to the reference person. Unrelated subfamilies can only be identified in the interview month, since a reference person is not identified in the monthly data.

User Notes - Issued periodically by the Census Bureau, these contain updated information for specific microdata files and are available on the SIPP website.

Usual Place of Residence - Place where a person normally lives and sleeps; specific living quarters held for the person, to which he or she is free to return at any time.

V

Variable Metadata - Provide a complete characterization of a variable's content. Variable metadata are available on the SIPP website.

W

Wave - One round of interviewing, which takes four months to complete; one fourth of the panel.

Wave Files - See *Core Wave Files*.

Weights - Estimates of the number of units in the target population that a given unit represents.

X

Y

Z

C. Topical Module Descriptions (1984 to 2008 Panels)

This section describes the topical modules that were administered before the 2014 redesign. At the end of each description, the panels and waves in which the topical module was administered is included in brackets. For example, "86-7," denotes that the topical module was fielded in Wave 7 of the 1986 Panel.

Adult Well-Being. See *Extended Measures of Well-Being*. [91-6, 92-3, 93-9, 96-8, 01-8, 04-5, 08-6, 08-9]

Annual Income and Retirement Accounts. Obtains respondent estimates of calendar-year business income and personal retirement plans. The module asked about businesses owned by respondents, gross income and expenses for such businesses, net income to such businesses, retirement accounts, including IRA, Keogh, and 401(k), and respondent participation in those retirement plans. [84-9, 85-5, 85-8, 86-5, 87-5, 88-5, 90-5, 90-8, 91-5, 91-8, 92-5, 92-8,93-5, 93-8, 96-4, 96-7, 96-10,01- 4, 01-7, 04-4, 04-7, 08-5, 08-8]

Assets, Liabilities, and Eligibility. Collected information about the value of assets and debt on assets and expands data gathered in the core questions. The intent of this topical module was to derive a comprehensive measure of household net worth and to collect information used to determine eligibility for federal and local assistance programs. To that end, the topical module included selected additional questions needed to determine program eligibility. Some of the assets included were savings accounts, stocks, mutual funds, and bonds. Data on unsecured liabilities such as loans, credit cards, and medical bills were also gathered. Real estate and vehicle costs, mortgages owned, and rental properties, as well as value of business and other financial investments were collected. Assets and liabilities that were held jointly were identified to prevent double counting. [84-4, 84-7, 85-3, 85-7, 86-4, 86-7, 87-4, 90-4, 91-7, 92-4, 93-7, 96-3, 96-6, 96-9, 96-12, 01-3, 01-6, 01-9, 04-3, 04-6, 08-4, 08-7, 08-10]

Basic Needs. See *Extended Measures of Well-Being*. [93-9]

Child Care. Collected information about all child care arrangements, for all children under 15, from mothers, single fathers, or guardians, regardless of labor force status. Those with children under age 15 were asked about the type of child care arrangements, who provides the care, the number of hours of care per week, where the care is provided, and the cost of the care. The module asked whether a relative or nonrelative cared for the child, and if the child was in school. Before the 1993 Panel, the module collected information about only one to two child care arrangements from mothers, single fathers, or guardians who were either working, in school, or looking for a job during the 4-month reference period. [84-5, 85-6, 86-3, 86-6, 87-3, 87-6, 88-3, 88-6, 89-3, 90-3, 91-3, 92-6, 92-9, 93-3, 93-6, 96-4, 96-10, 01-4, 04-4, 04-8,08-05, 08-08]

Child Support Agreements. Helps determine whether money received as child support affects participation in government programs and whether lack of support from one parent causes the other parent to need government assistance. The module collected information about characteristics of child support agreements, the annual amount and frequency of payments, and provisions for health care costs. Additional questions covered custodial arrangements, contact with public agencies for assistance in collection

of child support, frequency of contact with the absent parent, current place of residence of the absent parent, and reasons for non-award of child support. Questions about paternity establishment status were also asked about children of women with non-written agreements and all never married women. [85-6, 86-3, 86-6, 87-3, 87-6, 88-3, 88-6, 89-3, 90-3, 90-6, 91-3, 92-6, 92-9, 93-3, 93-6, 93-9, 96-5, 96-11, 01-5, 01-8, 04-5, 08-06]

Child Support Paid. Serves as a counterpart to the Child Support Agreements module. It sought information about support for children of the respondent who were under 21 years old and who lived with another parent or guardian at any time during the module's reference period of 4 months. [96-3, 96-6, 96-9, 96-12, 01-3, 01-6, 01-09, 04-3, 04-6, 08-4, 08-7, 08-10]

Children's Well-Being. Asked the designated parent or guardian about the health of children in the household, care of the child by nonfamily members, activities the family did with the children (such as reading and outings), lessons and activities outside of school, rules for children's TV viewing, and the respondent's opinion about the quality of their neighborhood. The module obtained information about children in three age groups: under 6 years old, aged 6-11, and aged 12-17 for as many as seven children in each category. Certain questions targeted fathers or stepfathers who are not designated parents; other questions address whether the child attended a public or private school. Content of this module varies across different panels and waves; data users should check the documentation for exact content. [92-9, 93-6, 93-9, 96-6, 96-12, 01-7, 04-3, 04-8, 08-4, 08-10]

Education and Training History. Collected information about respondent's highest level of school completed or degree received, courses or programs studied, and dates of receipt of high school and post-secondary degrees or diplomas. The module determined if the respondent attended a public or a private high school. Job-related-training questions addressed training designed to help find or develop skills for a new job as well as to improve skills at the current or most recent job. People 15 years of age and older were asked whether they had received job training; if they had, they were asked about the duration of the training, how it was used, how it was paid for, and if it was federally sponsored. (Variations are also asked as Education and Work History [84-3] and Education and Training [84-6].) [86-2, 87-2, 88-2, 89-2, 90-2, 91-2, 92-2, 93-2, 96-2, 01-2, 04-2, 08-2]

Employer-Provided Health Benefits. Collected data on the availability of health care benefits from employers and the demographics of workers with and without employer-provided health coverage. The module asked whether the plan restricted the respondent to specified doctors, if family members were covered, and whether any family members had pre-existing conditions not covered by the plan. The module also asked about long-term health care options. [96-5, 01-5, 04-5, 08-6]

Employment History. Identified patterns of employment, length of employment at certain jobs, and reasons for any periods of unemployment subsequent to the respondent's first job. Number of times away from work, total amount of time out of work and if this interruption was because of care giving to a child, an elder or disabled. Beginning with the 1996 Panel, specific questions that addressed type of work done, job duties, and the industry in which the respondent worked were moved into the core content; previously, such questions had been part of this module. [86-2, 87-2, 88-2, 89-2, 90-2, 91-2, 92-1, 93-1, 96-1, 01-1, 04-1, 08-1]

Extended Measures of Well-Being. Assessed the standard of living of SIPP households. Three types of questions addressed the physical conditions in which the respondents live. Respondents' ability to meet specified basic needs during the reference period, and respondents' subjective assessments of the quality of their living situations. Included under the first category are questions about the presence and condition of specified consumer durable goods in the home (e.g., clothes washers, refrigerators, air conditioners) and the physical condition of the home itself (e.g., condition of the roof and walls, state of the home's electrical wiring and plumbing), while another series of questions concerned conditions in the respondent's neighborhood, such as safety, cleanliness, and traffic. The second group of questions concerned whether members of the respondent's household had sufficient food to eat during the 4-month reference period and whether they were able to pay rent and other bills or to obtain medical care when needed. Respondents were also asked about the sources of help available when the respondent is in need (e.g., family, friends, or community). Respondents rated their satisfaction with the quality of different aspects of their living conditions. Included were items such as the quality of the furnishings, convenience of the home to shopping, and the general state of repair of their home. (Some of the basic needs questions were asked as a Basic Needs module [93-9].) From 1996 forward, this module was renamed Adult Well-Being to distinguish it from the topical module focusing on children's well-being. [91-6, 92-3, 93-9]

Family Background. Asked of people between ages 25 and 64. Obtained family characteristics at the time of the respondent's 16th birthday, including how many brothers and sisters the person had, with whom the person lived, the highest grade of school completed by the parents, and the occupations of the parents. [86-2, 87-2, 88-2]

Fertility History. Asked only of females 15 years of age and older and males 18 and older. Men were asked about the number of children they have fathered, and women were asked about their birth histories. Interviewers asked women who have had children when their first and last children were born, along with questions about their employment status during pregnancy and prior to the birth of their first child, circumstances of any absence from work before and after the first birth, and the maternity leave policies of their employers. Post birth employment was also covered. [84-8, 85-4, 86- 2, 87-2, 88-2, 89-2, 90-2, 91-2, 92-2, 93-2, 96-2, 01-2, 04-2, 08-2]

Functional Limitations and Disability. Provides data that can be used to evaluate links between types of disability, the family financial situation, and program participation. This module was asked in three variations: overall, adult, and children. Adults were asked the standard Activities of Daily Living (ADL) and Instrumental Activities of Daily Living (IADL) battery of questions. Questions addressed physical and mental conditions affecting the respondent, the use of mobility aids, vision and hearing impairments, speech difficulties, lifting and aerobic difficulties, and the ability to function independently within the home. For those under age 22, the questions were modified, referring to age-appropriate activities (e.g., questions about work activities are recast to ask about analogous school activities). Questions about children also addressed the use of special education services. For those under age 15, the interviewer asked the questions of the designated parent or guardian. [90-3, 90-6, 91-3, 92-6, 93-3 for overall module; 92-9, 93-6, 96-5, 96-11, 01-5, 01-8, 04-5, 08-6 for separate children and adults modules]

Health and Disability. Gathered data for all sample members about their general health, functional limitations (using the standard ADL battery of questions), work disability, and the need for personal

assistance. Respondents were asked about any hospital stays during the reference period, other periods of illness, other health facilities used, and their health insurance coverage. Information on children was collected from a designated parent or guardian. (Variations were also asked as Functional Activities, Disability Status of Children, and Disability Questions.) [84-3 for Health and Disability; 88-6, 89-3 for Functional Activities; 85-6, 86-3, 87-6, 88-3, 88-6, 89-3 for Disability Status of Children; 96-4 for Disability Questions]

Health Status and Utilization of Health Care Services. Asked about hospital stays, including any in psychiatric institutions; other illnesses or injuries that left the respondent bedridden for at least most of 1 day; doctor visits and frequency of visits, dental visits and frequency of visits; where the respondent seeks health advice (doctor's office, clinic, hospital); and health insurance coverage. (Also asked as Utilization of Health Care Services.) [85-6, 86-3, 87-6, 88-3, 88-6, 89-3, 90-3, 90-6, 91-3, 92-6, 92-9, 93-3, 93-6, 96-3, 96-6, 96-9, 96-12, 01-3, 01-6, 01-9, 04-3, 04-6, 08-4, 08-7, 08-10]

Home Health Care. Asked about the type and sources of help given to respondents who needed help with their personal care, household activities, and basic errands because of a health condition. Respondents were asked if caregivers were relatives or nonrelatives, and whether or not the caregivers were household members. This module also asked about members of the household who might have given such care, on a nonprofessional level, to a person outside the household. Questions determined the relationship of the caregiver and recipient(s) and the kind of care given. Starting with the 2001 Panel, this topical module was referred to as Informal Caregiving. [88-6, 89-3, 96-7]

Household Relationships. Collected information about relationships among household members. The SIPP core questions gathered extensive information about household composition for each month of the panel. This information allowed for the identification of families and subfamilies and details each household member's relationship to the household reference person. As extensive as this information is, it does not cover the interrelationships of all household members. For example, the SIPP core data provides no information about the relationships between members of two different unrelated (to the household reference person) subfamilies residing in the same household. This topical module fills that gap by providing complete information about how each member of the household is related to every other member of the household. Relationships are specified in detail; for example, a brother is a full brother, half-brother, stepbrother, or adoptive brother. In-law relationships are also identified. [84-8, 85-4, 86-2, 87-2, 88-2, 89-2, 90-2, 91-2, 92-2, 93-2, 96-2, 01-2, 04-2, 08-2]

Housing Costs, Conditions, and Energy Usage. Collected information on mortgage payments, real estate taxes, fire insurance, principal owned, when the mortgage was obtained, and interest rates; rent; type of fuel used and heating facilities; appliances; and vehicles.[25] Questions on value of home and automobile are used in conjunction with assets and liabilities reported in the Assets and Liabilities Topical Module to calculate each individual's net worth. This topical module also helped to fulfill a need for

[25] Subsequent to the 1984 Panel, questions on energy usage were combined into a separate module. Vehicles and housing values were retained in a module entitled Real Estate and Vehicles.

information concerning energy usage and conservation. The information can be used in analysis of the requirements of individuals and households who participate in energy assistance programs. [84-4]

Informal Caregiving. See Home Health Care. Starting with the 2001 Panel, this topical module was no longer referred to Home Health Care. Asked about the type and sources of help given to respondents who needed help with their personal care, household activities, and basic errands because of a health condition. Respondents were asked if caregivers are relatives or nonrelatives, and whether or not the caregivers are household members. This module also asked about members of the household who might have given such care, on a nonprofessional level, to a person outside the household. Data from this module determine the relationship of the caregiver and recipient(s) and the kind of care given. [01-7, 04-7, 08-9]

Job Offers. Asked about any job offers received by respondents who were looking for work or who were on layoff during the reference period. If the respondent was offered a job and did not accept it, obtained the reason for rejecting the job and the amount of money that was offered. [85-6, 86-3]

Long-Term Care. Focused on health-related conditions that might cause a person to need help around the home. Specific questions addressed the ability of people in the household to manage their personal care, housework, meal preparation, and basic errands outside the home. The module ascertains whether or not individuals providing such assistance are household members. Additional questions asked about community services and the financial burden of acquiring assistance. The module also asked about the activities of respondents who themselves provided such assistance on a nonprofessional basis to individuals outside the household. (Also asked as Home Health Care.) [85- 6, 86-3, 87-6, 88-3, 88-6, 89-3]

Marital History. Asked questions of all respondents aged 15 and older who have ever been married. The date of the present marriage was determined; for those married more than once, the module recorded the dates of their first two marriages and their last marriage, if married more than twice. If appropriate, respondents were asked when their previous marriages ended and whether they were widowed or divorced at the end of their marriages. [84-8, 85-4, 86-2, 87-2, 88-2, 89-2, 90-2, 91-2, 92-2, 93-2, 96-2, 01-2, 04-2, 08-2]

Medical Expenses. Collected data about out-of-pocket medical expenses, health services, doctor visits, prescription drugs, insurance reimbursement, and health and physical conditions that might affect the respondent's ability to work. The reasons for and length of any hospitalizations are determined, and respondents were asked about the types of medical professionals who delivered care. Most questions applied to both children and adults. [87-7, 88-4, 89-4, 90-7, 91-4, 92-7, 93-4, 93-7, 96-3, 96-6, 96-9, 96-12, 01-3, 01-6, 01-9, 04-3, 04-6, 08-4, 08-7, 08-10]

Migration History. Asked respondents aged 15 and older where they were born, where they have lived, and how long they have lived in those places. Respondents born in a foreign country were asked about their citizenship status and when they came to the United States to stay. [84-8, 85-4, 86- 2, 87-2, 88-2, 89-2, 90-2, 91-2, 92-2, 93-2, 96-2, 01-2, 04-2, 08-02]

Property Income and Taxes. Collected information on rental income received during the calendar year and on interest earned and/or dividends from assets such as savings accounts, money market deposit accounts, interest-earning checking accounts, bonds, or stocks. They were also asked about federal and

state income tax liabilities and certain other tax information such as type of return, use of selected schedules (for example, Schedule A, Itemized Deductions; Schedule B, Interest or Dividends; or Form 4835, Farm Rental Income), and number of exemptions. The tax questions were asked in order to develop better estimates of the distribution of after-tax income and to help build better micro simulation models of the tax and transfer system. [84-6]

Real Estate, Shelter Costs, Dependent Care, and Vehicles. Gathered information about housing tenure and financing, other real estate ownership, and automobile ownership. Homeowners were asked a series of questions that allow the estimation of net real estate equity. Questions about vehicles address ownership, type of vehicle (e.g., car, truck, or motorcycle), value, and amount owed. These questions allow determination of program eligibility. Starting in the 1996 Panel, these questions were included in the Assets and Liabilities Topical. [84-7, 85-3, 85-7, 86-4, 86-7, 87-4, 87-7, 88-4, 90-4, 90-7, 91-, 91-7, 92-4, 92-7, 93-4, 93-7, 96-3, 96-6, 96-9, 96-12, 01-3, 01-6, 01-9, 04-3, 04-6, 08-4, 08-7, 08-10]

Reasons for Not Working /Reservation Wage. Ascertains the reasons that persons are not in the labor force and the conditions under which persons might want to join the labor force. The reservation wage questions asked about the pay rate that a person would require in order to begin working (Ryscavage, 1996). Questions were also asked about job search and, if people have been offered but did not accept a job, the reason they refused it. This module was discontinued with the 1985 Panel. [84-5]

Recipiency History. Obtained a profile of a respondent's pattern of participation in certain government programs prior to the beginning of the SIPP panel. Specific questions addressed the first time a respondent participated in a particular program, the length of participation, and the number of times the respondent has been in the program. [86-2, 87-2, 88-2, 89-2, 90-2, 91-2, 92-1, 93-1, 96-1, 01-1, 04-1, 08-1]

Retirement Expectations and Pension Plan Coverage. Obtained information about the respondent's pension plan coverage for the most important current job or business, and information from persons currently receiving retirement benefits from a former job or business. Respondents were asked about their coverage and vesting in pension plans, types of plans, the reasons they were not included by or did not participate in plans, then-current contributions and amounts of money in their accounts if applicable, and how the money in their own plans was invested. Other questions concerned loans from pension accounts and treatment of lump sums received from prior job pension plans. Respondents currently receiving pension income were asked about the types of pension they received, provisions for cost-of-living adjustments, and health benefits. Respondents were also asked Industry and Occupation data about the job or business from which their pensions are received. (Also asked as Pension Plan Coverage [84-7].) [84-4, 85-7, 86-4, 86-7, 87-4, 90-4, 91-7, 92-4, 93-9, 96-7, 01-7, 04-7, 08-3, 08-11]

School Enrollment and Financing. Sought information about basic educational attainment, enrollment in public and private schools, and whether those in government programs differed from others in terms of financing their education and their sources of educational assistance. Asked of people aged 15 and older, the module included questions to pinpoint the grade level of people enrolled in a general, technical, or business school; their pattern of full- or part-time enrollment; amount of tuition and fees; costs of room and board; and books and supplies. Specific sources of educational assistance, such as the G.I. Bill or employer assistance, were also determined. (Also asked as Education Financing and

Enrollment.) [84-9, 85-5, 85-8, 86-5, 87-5, 88-5, 90-5, 90-8, 91-5, 91-8, 92-5, 92-8, 93-5, 93-8, 96-5, 01-5, 04-5]

Selected Financial Assets. Focused on the value of such assets as savings bonds, checking accounts, retirement accounts, life insurance, and the number of years respondents had held certain assets. These were also used in program eligibility determination. Starting in the 1996, these questions were included in the Assets and Liabilities topical module. [87-7, 88-4, 90-7, 91-4, 92-7, 93-4, 96-3, 96-6, 96-9, 96-12, 01-3, 01-6, 01-9, 04-3, 04-6, 08-4, 08-7, 08-10]

Shelter Costs and Energy Usage. Collected information on rent or mortgages, real estate taxes, and insurance; energy costs; and motor vehicles. The information is also pertinent to the determination of eligibility for a number of federal assistance programs. (Also asked as Housing Costs, Conditions, and Energy Usage.) [84-4, 86-6, 87-3]

Support for Non-household Members. Provided information about respondents' routine payments supporting people who were not current household members. Includes both child support payments for own children under 21 years of age and payments made to (or for) people who were not children of the respondents, for example, an elderly parent in a nursing home or an adult child living away from home and in an entry-level job. Questions about child support included number of children supported, type and year of agreement, annual amount and method of payment, health care provisions and custodial arrangements, and amount of contact with the absent children. Questions about support for other persons outside the household included their relationship to the respondent, living arrangement, and annual amount of support paid. [84-5, 84-8, 85-4, 85-6, 86-3, 86-6, 87-3, 87-6, 88-3, 88-6, 89-3, 90-3, 90-6, 91-3, 92-6, 92-9, 93-3, 93-6, 93-9, 96-5, 96-11, 01-5, 01-8, 04-5, 08-6]

Taxes. Included questions about exemptions, calendar-year wages and salaries, income from businesses, itemized deductions, and earned income credits. Respondents were asked about federal and state income tax liabilities, exemptions, amounts owed for federal and property taxes, and amounts from a variety of tax schedules. To help ensure accuracy, interviewers encouraged respondents to refer to income tax returns and other records. Generally, this module was administered at least twice per panel in the spring – when respondents were likely to be preparing their tax returns for the prior year. (Also asked as Earnings and Benefits, and Property Income and Taxes.) [84-6, 84-9, 85-5, 85-8, 86-5, 87-5, 88-5, 90-5, 90-8, 91-5, 91-8, 92-5, 92-8, 93-5, 93-8, 96-4, 96-7, 96-10, 01-4, 01-7, 04-4, 04-7, 08-5, 08-8]

Tax Rebate. Collected information about tax rebates. [08-1, 08-2]

Time Spent Outside Work Force. Collected information about work history and reasons for not working. Asked of people 21 or older, this short module addressed up to four periods of 6 months or longer in which the respondent did not work at a paid job or business. [90-6]. This modulewas discontinued after the 1990 Panel, but selected items from it were collected in the Employment History Topical Module in the 1996+ Panels.

Welfare Reform. Sought information about eligibility for and recipiency of public assistance. Specific questions addressed benefits, assistance that supported a respondent seeking work or acquiring training, requirements for receiving benefits (such as job hunting, drug testing, etc.), job subsidies, transportation

assistance, health care, and food assistance. This data was collected to measure the effect of the Welfare Reform Legislation in 1996. This module also gathered information about electronic transfer of benefits and denial of benefits to the respondent. [96-8, 01-8, 04-8, 08-3]

Work Disability History. Asked a series of questions about chronic health conditions that may affect the amount or type of work a respondent can do. Included are any such physical, mental, or other health conditions that interfere with the respondent's ability to work for at least 3 months. Questions were asked about when the limiting condition first became an issue, whether the person was working at the time, whether the condition resulted from an accident or injury, and if so, where the accident or injury occurred. Shorter-term conditions (including pregnancy) were not included as limiting conditions. [86-2, 87-2, 88-2, 89-2, 90-2, 91-2, 92-2, 93-2, 96-2, 01-2, 04-2, 08-2]

Work-Related Expenses. Asked about work-related expenses for each employer the respondent had during the reference period. Questions addressed various costs of working, such as union dues, licenses, special tools, and uniforms. Mode of transportation and mileage driven to and from work was determined, along with any parking or mass transit fees. (Also asked as Work-Related Expenses and Child Support Paid and Child Care Poverty.) [84-5, 84-8, 85-4, 86-6, 87-3, 96-3, 96-6, 96-9, 96-12, 01-3, 01-6, 01-9, 04-3, 04-6, 08-4, 08-7, 08-10]

Work Schedule. Collected information about the number of hours and days worked during a typical week in the fourth reference month. Questions about whether or not the respondent worked only at home on any days were included. [87-6, 88-3, 88-6, 89-3, 90-3, 91-3, 92-6, 92-9, 93-3, 93-6,93-9, 96-4, 96-10, 01-4, 04-4, 08-5, 08-8]

D. Publications in the P-70 Series

TABLE D-1. PUBLICATIONS IN THE SIPP P-70 SERIES

Publication Number	Year of Publication	Title
P70-1	1984	Economic Characteristics of Households in the U.S. Third Quarter, 1983
P70-2	1985	Economic Characteristics of Households in the U.S. Fourth Quarter, 1983
P70-3	1985	Economic Characteristics of Households in the U.S. First Quarter, 1984
P70-4	1985	Economic Characteristics of Households in the U.S. Second Quarter, 1984
P70-5	1985	Economic Characteristics of Households in the U.S. Third Quarter, 1984
P70-6	1986	Economic Characteristics of Households in the U.S. Fourth Quarter, 1984
P70-7	1986	Household Wealth and Asset Ownership, 1984
P70-8	1986	Disability, Functional Limitations, and Health Insurance Coverage: 1984-1985
P70-9	1987	Who's Minding the Kids? Child Care Arrangements: Winter 1984-1985
P70-10	1987	Male-Female Differences in Work Experience, Occupation, and Earnings: 1984
P70-11	1987	What's It Worth? Educational Background and Economic Status: Spring, 1984
P70-12	1987	Pensions: Workers Coverage and Retirement Income, 1984
P70-13	1988	Who's Helping Out? Support Network Among American Families
P70-14	1989	Characteristics of Persons Receiving Benefits from Major Assistance Programs
P70-15-RD-1	1989	Transitions in Income and Poverty Status: 1984-1985
P70-16-RD-2	1989	Spells of Job Search and Layoff...and Their Outcomes
P70-17	1990	Health Insurance Coverage, 1986-1988
P70-18	1990	Transitions in Income and Poverty Status: 1985-1986
P70-19	1990	The Need for Personal Assistance with Everyday Activities: Recipients and Caregivers
P70-20	1990	Who's Minding the Kids? Child Care Arrangements: Winter 1986-1987
P70-21	1990	What's It Worth? Educational Background and Economic Status: Spring 1987
P70-22	1990	Household Wealth and Asset Ownership: 1988
P70-23	1991	Family Disruption and Economic Hardship: The Short-Run Picture for Children
P70-24	1991	Transitions in Income and Poverty Status: 1987-1988
P70-25	1991	Pensions: Worker Coverage and Retirement Benefits, 1987
P70-26	1992	Extended Measures of Well-Being: 1984
P70-27	1992	Job Creation During Late 1980's: Dynamic Aspects of Employment Growth
P70-28	1992	Who's Helping Out? Support Network Among American Families: 1988
P70-29	1992	Health Insurance Coverage: 1987-1990
P70-30	1992	Who's Minding the Kids? Child Care Arrangements: Fall 1988
P70-31	1992	Characteristics of Recipients and the Dynamics of Program Participation: 1987-1988
P70-32	1990	What's It Worth? Educational Background and Economic Status: Spring 1990
P70-33	1993	Americans with Disabilities: 1991-1992

P70-34	1994	Household Wealth and Asset Ownership: 1991
P70-35	1994	Monitoring the Economic Health of American Households: Average Monthly Estimates of Income, Labor Force Activity, Program Participation and Health Insurance, First Quarter 1984 to Third Quarter 1991
P70-36	1994	Who's Minding the Kids? Child Care Arrangements: Fall 1991
P70-37	1994	Dynamics of Economic Well-Being: Health Insurance 1990-1992
P70-38	1994	The Diverse Living Arrangements of Children: Summer 1991
P70-39	1994	Dollars for Scholars: Postsecondary Costs and Financing 1990-1991
P70-40	1994	Dynamics of Economic Well-Being: Labor Force and Income: 1990-1992
P70-41	1995	Dynamics of Economic Well-Being: Program Participation: 1990-1992
P70-42	1995	Dynamics of Economic Well-Being: Poverty: 1990-1992
P70-43	1995	Dynamics of Economic Well-Being: Health Insurance: 1991-1993
P70-44	1995	The Effect of Health Insurance Coverage on Doctor and Hospital Visits: 1990-1992
P70-45	1995	Dynamics of Economic Well-Being: Poverty: 1991-1993
P70-46	1995	Dynamics of Economic Well-Being: Program Participation: 1991-1993
P70-47	1995	Asset Ownership of Households: 1993
P70-48	1995	Dynamics of Economic Well-Being: Labor Force: 1991-1993
P70-49	1995	Dynamics of Economic Well-Being: Income: 1991-1992
P70-50	1995	Beyond Poverty Extended Measures of Well-Being: 1992
P70-51	1995	What's It Worth? Field of Training and Economic Status: 1993
P70-52	1995	What Does it Cost to Mind Our Preschoolers?
P70-53	1996	Who's Minding Our Preschoolers?
P70-54	1996	Dynamics of Economic Well-Being: Health Insurance 1992-1993 Who Loses Coverage and for How Long?
P70-55	1996	Dynamics of Economic Well-Being: Poverty: 1992-1993 Who Stays Poor? Who Doesn't?
P70-56	1996	Dynamics of Economic Well-Being: Income 1992-1993 Moving Up and Down the Income Ladder
P70-57	1996	Dynamics of Economic Well-Being: Labor Force 1992-1993 - A Perspective on Low-Wage Workers
P70-58	1996	Dynamics of Economic Well-Being: Program Participation 1992-1993 Who Gets Assistance?
P70-59	1997	My Daddy Takes Care of Me! Fathers as Care Providers
P70-60	1999	Financing the Future: Postsecondary Students, Costs, and Financial Aid
P70-61	1997	Americans with Disabilities: 1994-95
P70-62	1997	Who's Minding Our Preschoolers? Fall 1994 Update
P70-63	1998	Dynamics of Economic Well Being: Poverty 1993-94
P70-64	1998	Dynamics of Economic Well-Being: Health Insurance 1993-1995 Who loses Coverage and For How Long?
P70-65	1998	Dynamics of Economic Well-Being: Income 1993-1994 Moving Up and Down the Income Ladder

P70-66	1998	Seasonality of Moves and Duration of Residence
P70-67	1999	Extended Measures of Well-Being: Meeting Basic Needs 1995
P70-68	2001	A Child's Day: Home, School, and Play (Selected Indicators of Child Well-Being) 1994
P70-69	1999	Dynamics of Economic Well-Being: Program Participation 1993-1994 Who Gets Assistance?
P70-70	2000	Who's Minding the Kids? Child Care Arrangements Fall 1995
P70-71	2001	Household Net Worth and Asset Ownership 1995
P70-72	2001	What's it Worth? Field of Training Economic Status, 1996
P70-73	2001	Americans With Disabilities: 1997
P70-74	2001	Living Arrangements of Children 1996
P70-75	2001	Did You Know? Homes Account for 44 Percent of All Wealth
P70-76	2001	Reasons People Do Not Work
P70-77	2001	Dynamics of Economic Well Being: Program Participation Who Gets Assistance? 1993-1995
P70-78	2001	Home-Based Workers in the United States 1997
P70-79	2001	Maternity Leave and Employment Patterns: 1961- 1995
P70-80	2002	Number, Timing, and Duration of Marriages and Divorces: 1996
P70-81	2002	Employment-Based Health Insurance: 1997
P70-82	2002	Fertility and Program Participation in the United States: 1996
P70-83	2002	Financing the Future - Post secondary Students Costs and Financial Aid: 1996-1997
P70-84	2002	Who's Helping Out? Financial Support Networks Among American Households: 1997
P70-85	2002	Work and Work-Related Activities of Mothers Receiving Temporary Assistance to Needy Families: 1996, 1998, and 2000
P70-86	2002	Who's Minding the Kids? Child Care Arrangements: Spring 1997
P70-87	2003	Extended Measures of Well Being: Living Conditions in the United States: 1998
P70-88	2003	Net Worth and Asset Ownership of Households: 1998 and 2000
P70-89	2003	A Child's Day: 2000 (Selected Indicators of Child Well Being)
P70-90	2003	Supplemental Security Income and Its Noninstitutional Recipients: July 1997 and 1999
P70-91	2003	Dynamics of Economic Well Being: Poverty 1996-1999
P70-92	2003	Dynamics of Economic Well Being: Health Insurance 1996-1999
P70-93	2003	Dynamics of Economic Well Being: Spells of Unemployment 1996-1999
P70-94	2004	Dynamics of Economic Well Being: Program Participation 1996 to 1999, Who Gets Assistance?
P70-95	2004	Dynamics of Economic Well Being: Movements in the U.S. Income Distribution 1996-1999
P70-96	2004	Dynamics of Economic Well Being: Labor Force Turnover, 1996-1999
P70-97	2005	Number, Timing, and Duration of Marriages and Divorces: 2001
P70-98	2005	What It's Worth: Field of Training and Economic Status in 2001
P70-99	2005	Support Providers: 2002
P70-100	2005	Dynamics of Economic Well Being: Moving Up and Down the Income Ladder 1998 to 1999

P70-101	2005	Who's Minding the Kids? Child Care Arrangements: Winter 2002
P70-102	2005	Participation of Mothers in Government Assistance Programs: 2001
P70-103	2005	Maternity Leave and Employment Patterns of First Time Mothers: 1961-2000
P70-104	2005	Living Arrangements of Children: 2001
P70-105	2006	Dynamics of Economic Well Being: Spells of Unemployment 2001- 2003
P70-106	2006	Health Status, Health Insurance, and Health Services Utilization: 2001
P70-107	2006	Americans with Disabilities: 2002
P70-108	2006	Dynamics of Economic Well Being: Participation in Government Programs 2001 Through 2003 Who Gets Assistance?
P70-109	2007	A Child's Day: 2003 (Selected Indicators of Child Well Being)
P70-110	2007	Extended Measures of Well Being: Living Conditions in the United States 2003
P70-111	2007	Reasons People Do Not Work: 2004
P70-112	2007	Dynamics of Economic Well Being: Fluctuations in the U.S. Income Distribution, 2001-2003
P70-113	2008	Maternity Leave and Employment Patterns of First Time Mothers: 1961-2003
P70-114	2008	Living Arrangements of Children: 2004
P70-115	2008	Net Worth and Asset Ownership of Households: 2002
P70-116	2008	Participation of Mothers in Government Assistance Programs: 2004
P70-117	2008	Americans with Disabilities: 2005
P70-118	2009	A Child’s Day: 2006 (Selected Indicators of Child Well-Being)
P70-119	2010	Economic Characteristics of Households in the United States: Third Quarter 2008
P70-120	2010	Dynamics of Economic Well-Being: Spells of Unemployment, 2004-2007
P70-121	2010	Who’s Minding the Kids? Child Care Arrangements: Spring 2005/Summer 2006
P70-122	2010	Seasonality of Moves and the Duration and Tenure of Residence: 2004
P70-123	2011	Dynamics of Economic Well-Being: Poverty, 2004-2006
P70-124	2011	Dynamics of Economic Well-Being: Fluctuations in the U.S. Income Distribution, 2004-2007
P70-125	2011	Number, Timing, and Duration of Marriages and Divorces: 2009
P70-126	2011	Living Arrangements of Children: 2009
P70-127	2011	Comparing Program Participation of TANF and Non-TANF Families Before and During a Time of Recession
P70-128	2011	Maternity Leave and Employment Patterns of First-Time Mothers: 1961-2008
P70-129	2012	What It’s Worth: Field of Training and Economics Status in 2009
P70-130	2012	Dynamics of Economic Well-Being: Participation in Government Programs, 2004-2007 and 2009 Who Gets Assistance?
P70-131	2012	Americans with Disabilities: 2010
P70-132	2012	Home-Based Workers in the United States: 2010
P70-133	2012	Health Status, Health Insurance, and Medical Services Utilization: 2010
P70-134	2013	Employment-Based Health Insurance: 2010
P70-135	2013	Who’s Minding the Kids? Child Care Arrangements: Spring 2011
P70-136	2013	Extended Measures of Well-Being: Living Conditions in the United States: 2011

P70-137	2014	Dynamics of Economic Well-Being: Poverty 2009-2011
P70-138	2014	Measuring Alternative Educational Credentials: 2012
P70-139	2014	A Child's Day: Living Arrangements, Nativity, and Family Transitions: 2011 (Selected Indicators of Child Well-Being)
P70-140	2015	Desire to Move and Residential Mobility: 2010-2011
P70-141	2015	Dynamics of Economic Well-Being: Participation in Government Programs, 2009-2012: Who Gets Assistance?
P70-142	2017	Dynamics of Economic Well-Being: Fluctuations in the U.S. Income Distribution: 2009-2012
P70-143	2017	Improvements to Measuring Net Worth of Households: 2013
P70-144	2017	Participation Rates and Monthly Payments from Selected Social Insurance Programs
P70-145	2017	Monthly and Average Monthly Poverty Rates by Selected Demographic Characteristics: 2013
P70-146	2017	Multiple Partner Fertility Research Brief
P70-147	2017	Fertility Research Brief
P70-148	2017	Demographics and Living Arrangements: 2013
P70-149	2017	Participation Rates in Other Assistance Programs: 2013
P70-150	2017	Common Pay Patterns and Extra Earnings: 2013
P70-151	2017	The Wealth of Veterans
P70-152	2018	Americans with Disabilities: 2014
P70-153	2018	Health Status and Medical Services Utilization: 2013
P70-155	2018	Net Worth of Households: 2014
P70-156	2018	Monthly and Average Monthly Poverty Rates by Selected Demographic Characteristics: 2014
P70-157	2018	Participation Rates & Monthly Payments from Social Insurance: 2014
P70-158	2018	Support Providers: 2013
P70-159	2018	A Child's Day: Parental Interaction, School Engagement, and Extracurricular Activities: 2014
P70-161	2020	Family Participation Rates in Nutrition Assistance Programs: 2015
P70-162	2019	Men's Fertility & Fatherhood: 2014
P70-163	2019	Multiple Jobholders in the United States: 2013
P70-164	2019	Net Worth of Households: 2015
P70-165	2019	Monthly and Average Monthly Poverty Rates by Selected Demographic Characteristics: 2015
P70-166	2019	Net Worth of Households: 2016
P70-167	2021	Number, Timing, and Duration of Marriages and Divorces: 2016
P70-168	2020	Health Insurance Dynamics in the Survey of Income and Program Participation: 2013-2014
P70-169	2020	Transitions in Parental Presence Among Children: 2017
P70-170	2020	The Wealth of Households: 2017
P70-171	2021	A Profile of Supplemental Security Income Recipients: 2017

P70-172	2021	Dynamics of Economic Well-Being: Poverty, 2013-2016

References

Bye, B., & Gallicchio, S. (1989). Two Notes on Sampling Variance. SIPP Working Paper 79.

Citro, C. F. & Kalton, G (1993). The future of the Survey of Income and Program Participation (SIPP). Washington, D.C.: National Academy Press.

Cresce, A. R., Cheng, Y., & Grieves, C. (2013). Household Estimates Conundrum: Effort to Develop More Consistent Household Estimates Across Surveys.

Czajka, J. L., Mabli, J., & Cody, S. (2008). Sample Loss and Survey Bias in Estimates of Social Security Beneficiaries: A Tale of Two Surveys. Mathematica Policy Research.

Lepowski, J. M., Landis. R. L., & Stehouwer, S. A. (1987). Strategies for the analysis of imputed data from a sample survey. Medical Care 25(8). 705-716.

McMillan, S., & III, R. C. (2013). Nonresponse Bias Analysis for Wave 1 2008 Survey of Income and Program Participation (ALYS-12). U.S. Census Bureau Internal Memorandum from Killion to Fields.

Moore, J. (2008). Seam Bias in the 2004 SIPP Panel: Much Improved, but Much Bias Still Remains. Washington, D.C.: U.S. Census Bureau.

Participation, P. O. (2009). Reengineering the Survey of Income and Program Participation. The National Academies.

Raghunathan, T., Lepkowski, J. M., van Hoewyk, J., & Solenberger, P. (2001). A multivariate technique for multiplying imputing missing values using a sequence of regression models. Survey Methodology, 85-96.

Ryscavage, P. (1996). An Evaluation and Analysis of Reservation Wage Data from SIPP. The Survey of Income and Program Participation, US Department of Commerce, Bureau of the Census.

U.S. Census Bureau. (1998). SIPP Quality Profile, 3rd. Ed.