

U.S. Census Bureau

SIPP-EHC

2011 and 2012 Field Test Evaluation



September 2013

The Survey of Income and Program Participation

Principal Investigator

David Johnson, *Chief, Social, Economic, and Housing Statistics Division, U.S. Census Bureau*

Survey Director

Jason Fields

Contributors to the Report

Kurt Bauman	Hubert Janicki
Megan Benetsky	Lynda Laughlin
Matthew Brault	Tracy Loveless
Rebecca Chenevert	Stephen Kingsbury
Jamie Choi	Rose Kreider
Tyler Crabb	Peter Mateyka
Ralph Culver III	Brett O'Hara
Judy Eargle	Daniel Perez-Lopez
Ashley Edwards	Lori Reeder
Renee Ellis	Trudi Renwick
C. Sol Espinoza	Erik Scherpf
Stephanie Ewert	Jeremy Skog
Alison Fields	Adam Smith
Graton Gathright (Editor)	Amy Steinweg
Katherine Giefer	Martha Stinson
Al Gottschalck	Jamie Taber
Timothy Grall	Marina Vornovytsky
John Hisnanick	Christopher Wignall
Shelley Irving	Kelly Wilkin

This report is produced to inform interested parties of ongoing Census Bureau research. The results have been formally reviewed to ensure no confidential data have been disclosed.

Contents

	Executive Summary	1
1	Introduction.....	7
2	Methods and Data	10
3	Demographics	16
4	Asset Ownership.....	19
5	Child Support	24
6	Disability.....	27
7	Education	29
8	Employment and Earnings	33
9	Health Insurance.....	41
10	Household Composition.....	46
11	Housing Subsidies.....	50
12	Medicaid.....	53
13	Medicare	58
14	Nativity and Citizenship	62
15	Old-Age, Survivors, and Disability Insurance	65
16	Poverty	71
17	Residence.....	74
18	Supplemental Nutrition Assistance Program	77
19	Supplemental Security Income.....	82
20	Temporary Assistance for Needy Families.....	87
21	Unemployment Insurance.....	91
22	Transitions and Seams.....	95
	Works Cited	99

Executive Summary

The 2014 panel of the Survey of Income and Program Participation will employ a completely re-engineered survey instrument that includes an Event History Calendar (EHC) structured around a one-year reference period for each interview. The new format is designed to reduce survey costs and respondent burden without sacrificing data quality.

This report evaluates data from a two-wave field test of the re-engineered instrument (SIPP-EHC) by comparison with data from the traditional question list instrument (SIPP) and with administrative records. With very few exceptions, agreement between survey and administrative data is higher for SIPP-EHC or not statistically different between surveys. While estimates from the two surveys (SIPP-EHC and SIPP) do differ statistically in many cases, these differences are typically small and correspond to rates of agreement with administrative data that are not lower for SIPP-EHC than for SIPP. There is little evidence that key estimates from SIPP-EHC data are less accurate for periods earlier in the one-year reference period as might be expected if respondents had difficulty reporting events further in the past. However, reported transitions in program participation or other status do tend to fall disproportionately at the beginning of reference periods. It also appears that this bias in the measurement of transitions can be improved by using information from prior waves in interviewing and editing.

The SIPP-EHC, like SIPP, collects information about a variety of topics, including employment, income, participation in various government programs, health insurance coverage, and demographics. This report evaluates survey estimates for nineteen SIPP-EHC topics: assets, child support, disability, education, employment and earnings, health insurance, household composition, housing subsidies, Medicaid, Medicare, migration, nativity and citizenship, Old-Age Survivors and Disability Insurance (OASDI), poverty, Supplemental Nutrition Assistance Program (SNAP), Supplemental Security Income (SSI), Temporary Assistance for Needy Families (TANF), and unemployment insurance.

The SIPP-EHC, like SIPP, uses a panel design which collects monthly information by regularly interviewing respondents about the recent past. The SIPP-EHC is structured around one-year reference periods, the SIPP around four-month reference periods. The SIPP-EHC instrument visually represents the reference period, and interviewers are encouraged to reference landmark events to aid respondents recall. This calendar design is intended to work with the structure of autobiographical memory to improve accuracy of reporting about the longer reference period.

A longer reference period raises two possible data quality issues addressed in this report: reverse telescoping and seam bias. Reverse telescoping refers to less precise reporting about events further in the past. Seam bias is the disproportionate incidence of transitions in status at the beginning of reference periods. Transitions in status refer to the reported beginning or ending of a spell of program participation, employment, educational enrollment, health insurance coverage, or other activity.

Methods

This report evaluates estimates of key population characteristics for the nineteen topics listed above. Estimates are compared across SIPP-EHC and SIPP for calendar years 2010 and 2011. For topics for which administrative records are available, mean absolute deviations are calculated between the survey responses and administrative records and are compared across surveys and years. Mean participation rates are calculated for all programs. Where appropriate, mean values of amounts are provided. Medians are presented for asset values and employment earnings to limit the influence of outliers on the comparisons. Annual rates, number of spells, per person and number of transitions per person are also studied.

The sample for the SIPP-EHC field test is not nationally representative, but was designed to include persons likely to be participating in government programs. In this evaluation, the comparison data from 2008 SIPP for CY2010 and CY2011 is appropriately restricted to provide a valid benchmark for the SIPP-EHC field test. Attrition from 2008 SIPP could affect the comparability of the samples and person-month observations between SIPP-EHC and SIPP.

For the analysis in this report, the data from the two surveys are harmonized to allow valid comparisons. Preliminary editing of the raw SIPP-EHC survey responses is performed to create variables comparable to those in edited SIPP data. However, no imputation of missing data in SIPP-EHC was performed, so, for comparability, all imputed data in SIPP is excluded from this analysis. Consequently, differences in the amount of missing data between SIPP-EHC and SIPP could explain some observed differences or mask differences not detected in this analysis. Attrition from 2008 SIPP could affect the comparability of the sample of person-month observations from 2008 SIPP to the sample from SIPP-EHC. The differences between the surveys for child support, housing subsidy, and employment could be explained by the differences in item-nonresponse rates between the surveys for these topics. In particular, sample persons without employment may be more likely to be coded as missing in SIPP-EHC than in SIPP since the former required respondents to positively report spells of non-employment. Additional data harmonization details are discussed in each topic chapter.

This report makes use of administrative records for the following topics: employment and earnings, Medicare, Medicaid, OASDI, SNAP, SSI and TANF. These data are linked to survey data using standard Census Bureau linking processes.

Results

Table 1 presents evidence that while the survey estimates typically differ statistically, in almost all cases the SIPP-EHC is not found to be less accurate than SIPP. The table summarizes comparisons between the two surveys and between the surveys and administrative records. The columns labeled "Survey estimates differ" indicate with a check mark the topics for which a *t*-test rejects the hypothesis that the estimates in the surveys are equal. The columns labeled "SIPP-EHC at least as accurate as SIPP" indicate with a check mark the topics for which the rate of disagreement with administrative data for SIPP-EHC is lower or not statistically different from SIPP.

Table 2 presents monthly estimates for the two surveys for CY2010 and CY2011 of rates of participation in various programs, rates of ownership of specific types of assets, rates of child support receipt, and rates of conditions such as poverty or disability.

T-statistics greater than 1.96 are taken to indicate that the difference in the estimates for the indicated year the two surveys is statistically significant. Table 3 presents estimates for the two surveys for CY2010 and CY2011 of rates of participation in various programs, rates of ownership of specific types of assets, rates of child support receipt, and rates of conditions such as poverty or disability. None of the differences in estimates are large enough to indicate a problem with the SIPP-EHC data.

Table 4 presents means of reported dollar amounts related to income, assets, and program benefits. The difference in median retirement savings between SIPP-EHC and SIPP could be explained by the lower item-nonresponse rate about amount of retirement savings in SIPP-EHC. If the marginal retirement savings responses captured in SIPP-EHC are more likely to be about low balances, this would produce the observed difference. The difference in UI amounts in CY2011 is anomalous and being investigated.

The analysis in this report also looks for evidence of reverse telescoping, the possibility that respondents will report less accurately about events further in the past. There is mixed evidence of such problems in

the pattern of agreement with administrative records for monthly indicators. Rates of under-reporting of program participation show almost no evidence of this pattern (only Medicaid in CY2010 shows this pattern). Rates of over-reporting of program participation do tend to decline over the course of each reference year for some topics (Medicaid, Medicare, OASDI, and SNAP).

There is evidence that transitions in status tend to fall disproportionately on the seam between the SIPP-EHC waves, in January 2011. However, incorporating information from the prior interview in dependent interviewing appears to mitigate the problem somewhat. There is also evidence from initial investigation of reported Medicaid transitions, that data can be edited based on information from the previous interview to increase accuracy and reduce seam bias in the edited data.

Table 1: Accuracy of SIPP-EHC reporting: administrative records comparison

Topic	Variable	CY 2010		CY 2011		Reference
		Estimates differ	SIPP-EHC accurate	Estimates differ	SIPP-EHC accurate	
Employment	number of employers	✓	✓	✓	✓	Table 8.2
	mean annual earnings	✓	✓	✓	✓	Table 8.8
Medicaid	monthly participation	.	✓	✓	✓	Table 12.1
	annual participation	✓	✓	✓	✓	Table 12.2
	months of participation	✓	✓	✓	✓	Table 12.2
	number of spells			✓	✓	Table 12.3
	number of CY2011 transitions			✓	✓	Table 22.1
Medicare	monthly participation	✓	✓	.	✓	Table 13.1
	annual participation	✓	✓	✓	✓	Table 13.2
	months of participation	✓	✓	✓	✓	Table 13.2
	number of spells			✓	✓	Table 13.3
	number of CY2011 transitions			✓	✓	Table 22.1
OASDI	monthly participation	✓	✓	✓	✓	Table 15.1
	annual participation	✓	✓	✓	✓	Table 15.2
	months of participation	✓	✓	✓	✓	Table 15.2
	amounts	.	✓	✓	✓	Table 15.3
	retirement	.	✓	.	✓	Table 15.4
	disability	.	✓	.	✓	Table 15.4
	number of spells			.	✓	Table 15.5
SNAP	monthly participation	✓	✓	✓	✓	Table 18.1
	annual participation	✓	✓	✓	✓	Table 18.2
	months of participation	✓	✓	✓	✓	Table 18.2
	amounts	✓	✓	.	✓	Table 18.3
	number of spells			✓	✓	Table 18.4
	number of CY2011 transitions			✓	✓	Table 22.1
SSI	monthly participation	✓	.	✓	.	Table 19.1
	annual participation	✓	✓	✓	✓	Table 19.2
	months of participation	✓	✓	✓	✓	Table 19.2
	amounts	✓	.	✓	.	Table 19.3
	number of spells			✓	✓	Table 19.4
	number of CY2011 transitions			.	✓	Table 22.1
TANF	monthly participation	.	.	.	✓	Table 20.1
	number of CY2011 transitions			✓	✓	Table 22.1

Columns labeled “Estimates differ” indicate with a ✓ whether a *t*-test rejects the hypothesis that SIPP-EHC and SIPP estimates have equal mean values for the indicated variables. For the CY2011 transitions rows, the comparison of the survey estimates is made only for sample-persons for whom administrative records are available.

Columns labeled “SIPP-EHC accurate” indicate with a ✓ whether the rate of disagreement (mean absolute deviation) between the survey and administrative data for SIPP-EHC is lower or not statistically different than for SIPP.

For spell count rows, the comparisons reported in the CY2011 column are for both years combined.

Table 2: Summary of results: monthly participation rates

Variable	CY 2010			CY 2011			Reference
	SIPP	EHC	<i>t</i> -stat	SIPP	EHC	<i>t</i> -stat	
Child support receipt	0.23	0.14	4.50	0.22	0.15	3.58	Table 5.1
Employment	0.48	0.57	8.25	0.48	0.57	8.78	Table 8.1
School enrollment	0.15	0.14	2.75	0.15	0.17	3.15	Table 7.2
Health insurance coverage	0.42	0.37	7.13	0.42	0.39	3.70	Table 9.1
Housing subsidy	0.24	0.18	7.98	0.23	0.16	7.79	Table 11.1
Medicaid participation	0.33	0.33	0.28	0.32	0.35	3.92	Table 12.1
Medicare participation	0.17	0.15	3.09	0.17	0.17	0.08	Table 13.1
OASDI participation	0.15	0.12	6.09	0.16	0.14	2.32	Table 15.1
OASDI retirement	0.61	0.64	1.29	0.63	0.65	1.05	Table 15.4
OASDI disability	0.26	0.27	0.39	0.26	0.29	1.43	Table 15.4
Relocation	0.02	0.02	3.97	0.01	0.01	2.01	Table 17.2
SNAP participation	0.27	0.21	8.45	0.27	0.24	4.05	Table 18.1
SSI participation	0.05	0.08	6.06	0.05	0.09	7.42	Table 19.1
TANF participation	0.03	0.03	1.60	0.03	0.03	0.02	Table 20.1
UI benefits	0.03	0.03	0.15	0.02	0.02	2.64	Table 21.1

Table 3: Summary of results: annual participation rates

Variable	CY 2010			CY 2011			Reference
	SIPP	EHC	<i>t</i> -stat	SIPP	EHC	<i>t</i> -stat	
Interest-bearing assets	0.36	0.34	2.13	0.39	0.35	3.66	Table 4.1
Retirement savings	0.17	0.14	4.63	0.17	0.14	4.55	Table 4.1
Disabled	0.21	0.21	0.12	0.18	0.22	5.31	Table 6.1
Poverty	0.38	0.37	1.34	0.40	0.41	0.57	Table 16.1
Relocation	0.12	0.20	11.68	0.10	0.13	5.92	Table 17.2
UI benefits	0.05	0.05	0.45	0.04	0.03	1.33	Table 21.1

Table 4: Summary of results: amount variables

Variable	CY 2010			CY 2011			Reference
	SIPP	EHC	<i>t</i> -stat	SIPP	EHC	<i>t</i> -stat	
Median retirement savings	\$63,202	\$15,667	5.02	\$68,907	\$18,000	1.43	Table 4.2
Child support amount	\$406.55	\$375.86	0.57	\$373.78	\$396.44	0.40	Table 5.2
OASDI benefits	\$848.34	\$899.25	1.26	\$840.60	\$940.72	2.78	Table 15.3
Median earnings	\$1,944	\$1,964	1.37	\$1,947	\$2,033	4.45	Table 8.6
SNAP benefits	\$301.52	\$353.92	3.70	\$279.83	\$298.52	1.50	Table 18.3
SSI benefits	\$553.57	\$669.16	5.69	\$558.47	\$732.40	5.40	Table 19.3
TANF benefits	\$406.31	\$370.57	0.94	\$412.14	\$420.07	0.13	Table 20.3
UI benefits	\$900.43	\$935.72	0.34	\$932.41	\$623.25	3.54	Table 21.2

1. Introduction

This report presents an evaluation of the new survey instrument that has been developed for the 2014 redesign of the Survey of Income and Program Participation. The analysis compares survey measures between the field tests of the redesigned survey instrument (SIPP-EHC) and the current production instrument (SIPP) for concurrent periods. For several topics, linked administrative records are used to compare misreporting in the two surveys. Topics analyzed include employment, earnings, health insurance coverage, assets, child support receipt, an array of federal and state program participation, and characteristics of persons and households including demographics, migration, and poverty.

A focus of this report is to compare key estimates between SIPP-EHC and SIPP. Another focus is to investigate the impact on the survey estimates of the primary innovations in the redesign: a one-year reference period (changed from a four-month period in the previous design) and the incorporation of an event-history calendar for much of the survey content. In particular, the analysis investigates whether reporting about events less temporally proximate to the time of interviewing is less precise.

This introduction provides background information on the SIPP, the new SIPP redesign, and literature related to this evaluation. The next chapter discusses the methodology and data employed in the analysis presented in this report. The subsequent 19 chapters discuss the data and analysis results for particular SIPP topics. The final chapter presents an analysis of reporting of transitions in program participation, employment, and school enrollment.

The Survey of Income and Program Participation

The Survey of Income and Program Participation¹ is sponsored by the U.S. Census Bureau under the authority of Title 13, United States Code, Section 182. The purpose of the survey is to collect reports of income amounts, labor force activity, program eligibility and participation, and general demographic characteristics to permit measurement of the effectiveness of federal, state, and local government programs; 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 survey design is a continuous series of national panels. The sample is a multistage-stratified sample of the U.S. civilian non-institutionalized population, with sample size ranging from approximately 14,000 to 45,000 interviewed households, depending on the panel. The first panel of the survey, called the 1984 panel, began in October 1983. Complete panels have interviewed sampled households every four months for two to six, depending on the panel. The current panel, called the 2008 panel, began in 2008, with interviewing every four months for a total of six years.

All adult household members (age 15 year or older) are interviewed by self-response if possible; proxy interviews are conducted for household members aged less than 15 years or who are not available for interviewing. Interviews are conducted by personal visit or by telephone. Beginning with the 1996 panel, SIPP interviews have been conducted using a computer-assisted interview.

The 2014 panel will begin interviewing early in 2014. The sample for the panel is expected to include 52,000 households. Interviewing for the 2014 panel will have a one year reference period for each interview, with calendar year 2013 as the first reference period. It is anticipated that the panel will last for four one-year waves.

¹Detailed information about the survey is available at <http://www.census.gov/sipp/>

Re-engineering of SIPP

In 2006, the U.S. Census Bureau began a complete redesign of the Survey of Income and Program Participation. See National Research Council (2009) and Moore, et al. (2009) for additional background on SIPP and the redesign. The goals of the redesign are to reduce costs and respondent burden and to improve data quality and timeliness. The new survey instrument, called SIPP-EHC, is scheduled to become the production survey instrument for the SIPP program beginning in 2014. The SIPP-EHC instrument is a complete redevelopment in Blaise and C# of the previous SIPP survey instrument that was implemented in CASES. The instrument is built around the change in survey reference period for the 2014 panel from four months to one year. The SIPP-EHC incorporates an event history calendar design to help ensure that the 2014 panel will continue to collect intra-year dynamics of income, program participation, and other activities with at least the same data quality as earlier panels. Event history calendars are discussed in more detail in the next section. The instrument collects information at the monthly level. A 2009 National Academy of Sciences study describes the design of SIPP-EHC with particular emphasis on design features. See National Research Council (2009).

Five field tests of the SIPP-EHC instrument have taken place (in 2008, 2010, 2011, 2012, and 2013). The 2012 SIPP-EHC field test is a wave 2 interview of the 2011 SIPP-EHC field test sample. The reference year for waves one and two of the 2011 SIPP-EHC field tests were calendar years 2010 and 2011. The 2013 SIPP-EHC field test is a wave 3 interview of this same sample. This report evaluates the 2011 and 2012 field tests.

Previous evaluations of the SIPP-EHC

This report is the most comprehensive evaluation of the SIPP-EHC to date and the present analysis evaluates the SIPP-EHC using a broader array of administrative records than previous work. Gathright, Stinson, and Reeder (2012) evaluated the 2010 and 2011 SIPP-EHC using a more limited set of administrative records and did not compare SIPP-EHC to SIPP for topics for which no administrative records were available. Stinson, Gathright, and Skog (2012) evaluated reporting about employment and earnings in 2010 SIPP-EHC using administrative data on employment and earnings. An earlier Census Bureau working paper (SIPP-EHC Data Evaluation Workgroup, 2011) compared program participation rates for SNAP, TANF, WIC, Medicare, and Medicaid and employment rates between 2010 SIPP-EHC and benchmark samples from 2008 SIPP. Moore, et al. (2009) compared responses in SIPP and SIPP-EHC surveys for a sample that was first interviewed in the 2004 SIPP panel and then re-interviewed using an early version of the SIPP-EHC instrument.

The use of administrative records for evaluation of SIPP data quality dates back to the first SIPP panel in 1984 (Marquis and Moore, 1990). Abowd and Stinson (2011) use administrative data to estimate measurement error in SIPP annual job earnings. The methodology of the analysis in this report is related to the work of Meyer and Goerge (2011) on misreporting of SNAP participation in the Current Population Survey and the American Community Survey.

Literature on testing of event history calendars

EHC instruments are designed to work with the structure of autobiographical memory to improve the accuracy of survey reporting. An EHC is designed to support the respondent's ability to accurately recall changes over a reference period by providing visual representation of the relevant months and event cues. An EHC has three components: a visualization of the reference period, generally subdivided into smaller units of time; cues relating to themes of interest; and temporal boundaries defined by landmark events (Glasner and van der Vaart, 2009). Calendar methods are generally used to examine extended reference periods. Life course research may design instruments to cover the reference period from the respondent's birth to the time of the interview. That frame provides a very different set of challenges than designing an instrument to cover a calendar year or cohort, but an EHC instrument can be effective for covering relatively

brief reference periods by creating relevant landmarks and dividing the period into spells (Axinn, Pearce, and Ghimire, 1999). The provision of landmark events, reference points, and spells allows a respondent to define more specific periods of time within a larger reference period.

Theories from cognitive science and previous evaluations both support the merits of an EHC for improving data quality. Though many previous evaluations have been undertaken, none have been as extensive as the SIPP-EHC field tests. EHC methods are supported by the move toward understanding the cognitive underpinnings of survey design (Tourangeau 2000). A calendar, providing temporal and thematic queues, is thought to work better with the structure of autobiographical memory. Memory is relational and dependent upon contextual clues for accurate recall. Traditional questionnaire designs that attempt to segment off particular topics or periods work against the structure of memory, particularly for events which have recurred over a period of time (Belli 1998).

An EHC offers flexibility at the expense of standardization, so interviewer effects may be magnified by this format. Interviewer experience was controlled for through matched pair random assignment (interviewers were paired based on years experience, and one randomly assigned to the EHC and the other to the Q-List condition), demonstrating that a slightly stronger interviewer effect is present in EHCs (Sayles, Belli, and Serrano, 2010). Seam effects were found to be decreased in EHCs over a reference period of two years, however a “false seam” near the middle of the reference period developed (Callegaro, 2007).

The SIPP-EHC field tests have provided the opportunity to empirically test the effectiveness of EHCs through comparison with administrative records. Previous direct comparisons between EHC instruments and traditional question list (Q-List) formats have generally relied upon reinterviewing respondents from longitudinal studies. For instance, Belli, Shay and Stafford (2001) took a random sample of respondents from The Panel Study of Income Dynamics (PSID) and reinterviewed them in 1998 about the previous two years. Respondents and interviewers were randomly assigned to either a Q-List or EHC condition. The results were compared to the responses those respondents had given in the 1997 PSID interview about the events of 1996, assuming that the information collected nearer to the reference period would be more accurate than information collected a year later. Most of the measures did not statistically differ from EHC or Q-List format, but the differences which were observed all favored the EHC format.

2. Methods and Data

This chapter describes the analysis performed to evaluate the SIPP-EHC instrument. The analysis compares key estimates from SIPP-EHC to corresponding estimates from SIPP for a variety of topics. Where administrative records for a topic are available, the analysis also assesses the accuracy of reporting in SIPP-EHC relative to SIPP.

The following section describes the split-sample design that underlies this evaluation of SIPP-EHC. The subsequent section explains how the administrative records are used as standard for the surveys. The final two sections describe the specific statistics that we calculate, describe the statistical tests that we perform, and provide a guide to interpreting the tables presented in the report.

2011 SIPP-EHC and 2008 SIPP Samples

The comparison of SIPP-EHC and SIPP presented in this report is essentially a split-sample experiment. The samples for these two surveys were created at the same time using the same process. Effectively, units (addresses) were randomly assigned to either the SIPP-EHC treatment or to the SIPP treatment.

Adjustments were made to both of these samples before interviewing began. These post-sampling adjustments are easily accounted for in our analysis so that the estimates from each survey are for the same population. The present analysis re-weights the 2008 SIPP units to account for difference in state composition between the two surveys. The 2008 SIPP sample had been adjusted to make the sample state representative for the 20 largest states.

The post-sampling sample restrictions imposed for 2011 SIPP-EHC were designed to produce a low-cost field test sample that still adequately sampled participants in federal and state programs to permit an informative test of the SIPP-EHC instrument. The sample was restricted in four ways. Units were included only from (1) the unit frame, (2) the self-representing primary sampling units (PSU), (3) the low-income within PSU stratum, and (4) 20 particular states. For this evaluation, these same restrictions have been applied to the SIPP data to create a suitable benchmark sample.

Table 2.1: Persons

	Survey	AR
SIPP	12,855	10,429
SIPP-EHC	7,435	5,327
Total	20,290	15,756

Neither sample is nationally representative, but both samples permit estimation for the *same population*. Calendar years 2010 and 2011 (CY2010 and CY2011) are the reference periods for waves 1 and 2 of the 2011 and 2012 SIPP-EHC instrument. There are two elements of uncertainty introduced by the fact that the SIPP has been going for nearly 3 years longer than the SIPP-EHC. The first is a possible difference in sample compositions owing to the time elapsed between the rostering of SIPP and SIPP-EHC. The second is the possibility of a selected sampling effect produced through attrition. The many more waves and longer

time period provided more opportunities for attrition; if respondents do not attrit randomly, then this may introduce a difference in the composition of the responding sample.

Table 2.2: Person-months in sample

		Survey	AR
CY 2010	SIPP	117,543	97,198
	SIPP-EHC	83,044	63,882
CY 2011	SIPP	109,055	90,199
	SIPP-EHC	59,560	43,389
Total		369,202	294,668

The total number of sample persons included in the analysis is 20,290 (see Table 2.1), with 7,435 from SIPP-EHC and 12,885 from SIPP. The total number of sample person-months is 369,202. In CY2010, SIPP-EHC had 83,044 while SIPP had 117,543 person-months. In CY2011, SIPP-EHC had 59,560 and SIPP had 109,055 person months (see Table 2.2). The total number of sample person-years is 35,084. 26,200 (74.6 percent) of those person-years are complete. In CY2010, SIPP-EHC had 6,969 person-years (with 98.7 percent complete years) and SIPP had 12,014 (65.2 percent complete) person-years. In CY2011, SIPP-EHC had 4,996 person-years (98.6 percent complete) and SIPP had 11,105 person-years (63.1 percent complete) (see Table 2.3).

In CY2010, SIPP-EHC had 2,596 household-years, and SIPP had 2,596 household-years. In CY2011, SIPP-EHC had 3,452 household-years, and SIPP had 1,942 household-years (see Table 2.4). The total number of sample household-years is 11,769.

Table 2.3: Person-years in sample

		Ever in year		All year	
		Survey	AR	Survey	AR
CY 2010	SIPP	12,014	9,839	7,385	6,180
	SIPP-EHC	6,969	5,327	6,879	5,315
CY 2011	SIPP	11,105	9,106	7,010	5,834
	SIPP-EHC	4,996	3,617	4,926	3,614
Total		35,084	27,889	26,200	20,943

Table 2.4: Household-years

		Survey
CY 2010	SIPP	3,779
	SIPP-EHC	2,596
CY 2011	SIPP	3,452
	SIPP-EHC	1,942
Total		11,769

Administrative records as a standard for survey reports

The Census Bureau has access to person-month level administrative records for several SIPP-EHC topics under consideration: Medicare, Medicaid, Supplemental Nutritional Assistance Program (SNAP), Temporary Aid to Needy Families (TANF), Old-Age, Survivors, and Disability Insurance (OASDI), Supplemental Security Income (SSI), and annual employment and earnings. These records provide a benchmark for survey reports on these topics. The data available for each topic are also described in the relevant chapters.

Comparison between administrative records and the responses of both survey instruments provide a test of the relative accuracy of SIPP-EHC to SIPP. Of the 12,885 sample-persons between SIPP-EHC and SIPP, 15,756 were successfully linked to administrative records, with 5,327 from SIPP-EHC and 10,429 from SIPP. Though administrative records are not without error, this report treats the administrative records as true and refers to observations where a survey response matches an administrative record as accurate. Sampling units were randomly assigned to SIPP or SIPP-EHC, so any measurement error in the administrative records should be uncorrelated with survey assignment and should not introduce bias in comparisons of misreporting between surveys. The accuracy of survey reports is measured at the person, person-year, and person-month levels. Each survey report is compared to administrative records, and then the rate of agreement between survey response and administrative record is compared across SIPP and SIPP-EHC. This difference-in-difference approach gauges the relative accuracy of SIPP-EHC and SIPP in a way that isolates differences not due to sample composition. Difference in rates of item-nonresponse between the surveys is a very important caveat for our comparisons of survey estimates, since item non-response is likely to be non-random.

Harmonization of SIPP, SIPP-EHC, and administrative data

The data from the SIPP-EHC field tests are unedited in the sense that only the raw question responses are available. Consequently, the person month status indicators and other variables used in the analysis had to be constructed from responses to multiple questions, sometimes across the interviews of multiple persons within a household.

These field test data also have none of the missing data imputed. In order to make valid comparisons with SIPP, imputed values in the SIPP data have been coded as missing for this analysis (except where otherwise indicated). For some topics, analysis variables were constructed from unedited SIPP data. For other topics, edited SIPP data were used, but imputed data were excluded. Sample person-months from either SIPP-EHC or SIPP that correspond to a nonresponding (neither self nor proxy) adult (15+) in a responding household are excluded from this analysis.

For the government program topics (SNAP, TANF, SSI, OASDI), the exclusion of missing SIPP data meant excluding all but the first observed receipt amount for those who report participation. This is because pro-

duction SIPP processing applied the first observed receipt amount to all subsequent months of participation. These “imputations” have also been excluded from the present analysis.

Description of the analysis

The analysis presented in this report compares estimates of key population characteristics between SIPP-EHC and SIPP and investigates the relative quality of microdata from the two surveys using administrative records.

For all topics covered in the report, estimates of key population characteristics are presented for each survey in each of CY2010 and CY2011. Virtually all of these estimates are means. Estimates are presented at the person-month, person-year, and person levels as indicated. Standard errors of these means were clustered at the person level.

Tests are presented of equality of means across surveys within the calendar year. In order to draw inferences about improvements in data quality across waves of the field test, tests are also presented of the statistical significance of apparent differences across years in the within-year differences across surveys. All of the statistical tests that we perform on medians and means and the difference between medians and means are 5% two-tailed t-tests using the asymptotic distribution of the t-statistics. We reject null hypotheses based on t-tests greater than 1.96. Standard errors for medians are calculated via the woodruff method using SAS proc surveymeans.

For topics with available administrative records, estimates based on administrative records are presented alongside the corresponding estimates from the survey data. The survey based measures are presented for just the linked sample for added comparability. We measure agreement between the two sources of information for a given topic by calculating the mean absolute deviation (MAD) between the reports and records. The mean absolute deviation is calculated as

$$\frac{\sum_{i=1}^{N_L} |s_{it} - a_{it}|}{N_L}$$

where s_i and a_i are the survey report and administrative record, respectively, and N_L is the number of observations in the linked sample. This measure has the advantage (over simply differencing the mean status in the two sources) of not masking offsetting disagreements across person-months.

For observations where reported status and recorded status agree, we calculate, where applicable, the mean absolute deviation between the survey-reported and administratively recorded amount of benefit or earnings.

For participation indicators, the analysis also decomposes these mean absolute deviations measures into the two types of possible errors: false negatives and false positives. We classify each person-month report in the data as false negative (FN), false positive (FP), true negative (TN), or true positive (TP) based on agreement between the survey report and the corresponding measure in the administrative records. We classify a report about participation in a particular program as false negative if the administrative records for the program indicate that the sample person participated in the program in a particular month, but the corresponding survey report disagrees. The rate of FN is the proportion of false negative reports among person-month observations with participants recorded in administrative data. The rate of FP is the proportion of false positive reports among the survey-reports of participants.

For topics with person-month indicators, the analysis investigates the possibility that reports about events further in the past may be less precise in the SIPP-EHC with its one-year reference period. Two possible issues are considered: straight-lining and reverse-telescoping. Straight-lining refers to the possibility that

respondents simply project backwards through the reference period whatever status they report in a more recent reference month. A possible reason for this is that respondents may consciously mitigate the interview burden that they would otherwise incur by being more detailed in their reporting. Reverse telescoping refers to the possibility that respondent memory of events further in the past may simply be less precise.

The possibility of straight-lining or reverse-telescoping in SIPP-EHC data is evaluated by investigating the pattern of reporting errors across calendar months. The analysis tests for equality across surveys and within calendar year of the differences in the FN (FP) rates between the last four months of each calendar year and 1) the FN (FP) rate in January and 2) the FN rate for the first four months of the calendar year. The last four months of the calendar year are used as the benchmark since the reference period in the previous design was four months long. We also test for statistically significant changes across calendar years in the difference across surveys in the difference across these two periods of the calendar year in the rate of FN (FP) reporting.

When no administrative records are available for a topic, a similar analysis is performed on the monthly status indicators in which we look for more similar estimates at the end of a year than at the beginning. We test the equality across surveys and within calendar year of differences in the mean reported status between the last four months of each calendar year and the mean reported status for 1) January of the calendar year and 2) the first four months of the calendar year.

Description of table elements

This section describes tables presented in the rest of the report.

Results for SIPP are from a sub-sample from 2008 SIPP that is comparable to the SIPP-EHC sample which is not representative of the U.S. population. Results are weighted only to adjust for over-sampling of states in 2008 SIPP relative to 2011 and 2012 SIPP-EHC.¹

Columns labeled "Survey" present statistics based on survey reports for all in-universe observations for which the relevant measure is non-missing. Columns labeled "Linked" present statistics based on survey reports for the observations in the corresponding "Survey" sample that were successfully linked to administrative records. Columns labeled "AR" present statistics based on administrative data for the observations in the linked sample. Columns labeled "MAD" present the mean absolute deviation between the survey-reported and administratively recorded measures for the linked sample.

Rows labeled "*t*-stat" present *t*-statistics from tests of equality between the SIPP and the SIPP-EHC. Rows labeled "Diff-in-diff *t*-statistic" present *t*-statistics from tests of whether the difference between SIPP and SIPP-EHC in CY2010 is equal to the difference between SIPP and SIPP-EHC in CY2011. Rows labeled "Person-month observations," "Person-year observations," or "Person observations" present counts of observations with no missing data.

Columns labeled "NIU" present the proportion of the sample person-months or person-years that are not-in-universe for the indicated variable.

Household level estimates are presented for households as constituted in December of the indicated calendar year for SIPP and as of the interview date for SIPP-EHC.

Rows labeled Test 1 and Test 2 present *t*-statistics of tests of equality of the indicated rates between the last four months of each calendar year and 1) the rate in January (Test 1) or the rate for the first four months of the

¹Source: U.S. Census Bureau, Survey of Income and Program Participation, 2008 Panel. Survey of Income and Program Participation – Event History Calendar (SIPP – EHC) Field Test internal files, 2011 Panel. For information on sampling and nonsampling error see <http://www.census.gov/sipp/source.html>

calendar year (Test 2). Rows labeled "Pooled" present the indicated rate for all calendar months combined. Columns labeled "Diff" present t-statistics for tests of difference in difference in the indicated rate across the surveys across the calendar years.

3. Demographics

This chapter presents comparisons of reporting about demographics in SIPP-EHC and SIPP. The following section describes the demographic data that are employed in the analysis discussed in the subsequent section.

Description of demographics data

The data sources for the tables in this chapter are 2008 SIPP, 2011 SIPP-EHC, and 2012 SIPP-EHC. See Chapter 2 for discussion of the SIPP-EHC and SIPP samples and weighting.

Person-year analysis variables of demographic characteristics are created for this analysis, including age, sex, marital status, race, and Hispanic origin. All sample-persons are in universe for these variable. Sample-persons age 15 years or older are in universe for the marital status variable. Questions about demographic characteristics are asked using the same wording in both years of SIPP-EHC and in the 2008 SIPP.

The 2011 and 2012 SIPP-EHC variables are created as of the interview month. The 2008 SIPP variables are created from the edited data, dropping all imputations. Because the reference year includes three interview points, person-year level data are taken from the latest wave where a report for the item is not missing. This differs from the SIPP-EHC for which data is only taken from the time of interview.

Age is broken into five age categories: under 15, 15 to 24, 25 to 44, 45 to 64, and 65 and above.

Race and origin are combined into five categories: White alone (non-Hispanic), Black alone (non-Hispanic), Asian alone (non-Hispanic), other (which includes those that choose the some other single race category and those that include more than one race but are not Hispanic), and Hispanic (of any race).

Discussion of demographics results

This section presents and discusses the analysis of the demographic variables in SIPP-EHC.

In SIPP-EHC compared to SIPP, the proportion of sample-persons reported to be under age 15 is significantly, but not substantially, lower in CY2010 (see Table 3.1). In the same calendar year, the proportion of sample-persons reported to be ages 25 to 44 is significantly, but not substantially, higher in SIPP-EHC than in SIPP.

In CY2011, the proportion of sample-persons reported to be male is significantly, but not substantially, higher in SIPP-EHC than in SIPP (see Table 3.2).

In CY2010 and CY2011, SIPP-EHC compared to SIPP reports fewer non-Hispanic whites and more Hispanics (of any race) (see Table 3.3). The proportion of sample-persons reported to be non-Hispanic black is slightly lower in SIPP-EHC than in SIPP in CY2010.

In SIPP-EHC compared to SIPP, the proportion of sample-persons reported to have a spouse absent is significantly, but not substantially, higher in both calendar years (see Table 3.4). In CY2011, the proportion of sample-persons reported to be married is significantly, but not substantially, higher in SIPP-EHC than in SIPP. SIPP-EHC compared to SIPP reports fewer sample-persons as separated or never married in CY2011.

For race, there is lower item non-response in SIPP-EHC than in SIPP for both CY2010 and CY2011 (see Table 3.5). For age, there is a statistically significant, but not substantial, difference in item non-response between the two surveys in both calendar years. For sex, there is a significant, but not substantial, difference in item non-response in CY2011.

Table 3.1: Age

		Under 15	15–24	25–44	45–64	65 and above
CY 2010	SIPP	0.25	0.16	0.26	0.22	0.11
	SIPP-EHC	0.24	0.17	0.28	0.22	0.10
	<i>t</i> -statistic	2.14	1.14	2.71	0.51	1.67
CY 2011	SIPP	0.25	0.16	0.26	0.22	0.11
	SIPP-EHC	0.24	0.15	0.26	0.23	0.12
	<i>t</i> -statistic	0.45	0.76	0.53	1.13	0.71
Diff-in-diff <i>t</i> -statistic		2.17	2.27	4.21	2.45	3.64
Person-years		35,046	35,046	35,046	35,046	35,046

Table 3.2: Sex

		Male
CY 2010	SIPP	0.47
	SIPP-EHC	0.49
	<i>t</i> -statistic	3.08
CY 2011	SIPP	0.46
	SIPP-EHC	0.48
	<i>t</i> -statistic	1.83
Diff-in-diff <i>t</i> -statistic		1.47
Person-years		35,064

Table 3.3: Race

		White	Black	Asian	Other nonhispanic	Hispanic
CY 2010	SIPP	0.28	0.29	0.05	0.03	0.35
	SIPP-EHC	0.22	0.26	0.05	0.03	0.44
	<i>t</i> -statistic	9.54	4.42	0.78	1.81	11.20
CY 2011	SIPP	0.28	0.27	0.05	0.03	0.36
	SIPP-EHC	0.22	0.26	0.06	0.03	0.43
	<i>t</i> -statistic	8.47	1.56	1.70	0.19	7.92
Diff-in-diff <i>t</i> -statistic		0.01	3.80	1.81	2.05	3.26
Person-years		34,086	34,086	34,086	34,086	34,086

4. Asset Ownership

This chapter presents comparisons of asset ownership rates and asset amounts between SIPP-EHC and SIPP. The following section describes the asset data that were employed and the subsequent section discusses the results.

Description of assets data

The data sources for the tables in this section are 2008 SIPP, 2011 SIPP-EHC, and 2012 SIPP-EHC. See Chapter 2 for discussion of the SIPP-EHC and SIPP samples and weighting.

Variables specific to asset ownership are created from the survey sources. The variables include ownership indicators for interest-bearing assets (interest-earning checking account, savings account, money market account, and certificates of deposit), stocks and mutual funds, and retirement accounts (401(k), Thrift, Keogh, and IRA accounts). They also include the total balance of retirement savings accounts, the value of home, and the value of first mortgage.

Sample-persons 15 years of age or older are in-universe for asset ownership variables. Ownership indicator variables indicate ownership of having at least one element of the corresponding asset group. For example, a sample-person with a reported stock and a missing report for mutual fund ownership is coded as an asset owner for the mutual funds asset ownership indicator. Respondents are classified as having 'no ownership' for a particular asset group if they do not own any of the assets within that group. Sample-persons with any missing report and no reported ownership in a given asset group are coded as missing. For example, a respondent who refuses to answer the question about ownership of stocks and reports having no mutual funds is classified as 'no response.' Sample-persons with no missing reports and no reported ownership of assets in an asset group are coded as non-owners for the given asset group.

Total balance of retirement savings accounts refers to the total amount in 401(k), Thrift, Keogh, and IRA accounts. The universe for this variable is anyone who reports having one of these accounts during the reference period. The value of home and the value of first mortgage are both household-level variables. The universe for the value of home is households whose primary residence is owned or being bought and whose primary residence is not a mobile home. The universe for the value of first mortgage is households who have at least one mortgage on their primary residence that is owned or being bought and whose primary residence is not a mobile home. All values are collected as of the last day of the reference period.

An amount response is considered valid if the sample-person has a report of all amounts for the items he or she was in-universe. For example, a sample-person that has a reported amount for the value of a 401(k) account but has a reported refusal to the question asking about the value of an IRA account is coded as missing. Similarly, a sample-person who has a reported refusal to the question about whether there is any debt against his or her home is coded as missing, even though the sample-person does not have a directly reported refusal to the question about the value of that debt. In the case of retirement accounts, a response of 'don't know' or 'refuse' to a question that sets the universe for the amount question or to the amount question itself results in the respondent being coded as missing. If a respondent has multiple retirement accounts and provided amounts for each, the values are added to produce the total value of retirement accounts.

For 2011 and 2012 SIPP-EHC, the reference period was 1 year. The last day of the reference period for both was December 31 of the corresponding calendar year. For SIPP, the data came from the Core and the Wealth Topical Module, administered in Waves 7 and 10 of 2008 SIPP. The reference period for SIPP was 4 months,

which means that, depending on a particular respondent's rotation group, for each calendar year the end of the reference period ranged from August to November. All imputations were dropped from SIPP data and only unedited data were used.

All amounts are collected as of the last day of the reference period, which, in the case of SIPP-EHC, always refers to December 31 of the corresponding reference year. However, in the case of 2008 SIPP, the end of the reference period could range from August to November. This means that amounts might not be directly comparable across the surveys if balances in retirement accounts, home values, and values of mortgages changed between August and December.

Assets amount questions are also placed differently in the question order in the SIPP-EHC and SIPP surveys. In SIPP, the value of retirement accounts, the value of home, and the value of first mortgage are collected in the Wealth Topical Module, which is administered at the end of the interview. This means that a respondent encounters asset content twice (once in the Core and once in the Topical Module) and gets to the Topical Module only upon successful completion of the entire Core instrument. In SIPP-EHC, a respondent is asked about the assets content only once. It is unclear what impact, if any, these differences might have on item non-response rates for assets items or on respondents report the value of a particular asset.

The number of questions a respondent needs to answer in order to be in-universe for questions about the value of an asset also differs between SIPP and SIPP-EHC. In SIPP-EHC, respondents are asked two questions about ownership of retirement accounts. The first asks whether a respondent had an IRA / Keogh account at any point during the reference period. The second asks whether a respondent had a 401(k) / Thrift account. Those who report ownership are then asked about the value as of the last day of the reference period. If a respondent no longer owned such an account, the FR is instructed to enter the value of 0.

In 2008 SIPP, global ownership questions are asked in Core SIPP. Respondents are asked three, rather than two, questions, since information about IRA and Keogh accounts is collected separately. If respondents get to the Topical Module content, they are asked whether they still had the account for which they reported ownership as of the last day of the reference period. Only those who answer in the affirmative are asked about amounts. This means that in addition to differences in the placement of the asset content in the survey, in SIPP respondents answer more questions about these assets.

For SIPP-EHC, sample-persons are in-universe for questions in the assets section if they are 15 years or older as of the time of the interview, which did not take place until the spring of the calendar year following the reference year. For the 2008 SIPP, respondents are in-universe if they are 15 years or older as of the last day of the reference period, which ranges from August to December depending on respondent's rotation group.

Discussion of assets results

This section presents and discusses the analysis of asset ownership rates and amounts in SIPP and SIPP-EHC. The table elements are described in Chapter 2, and, where appropriate, in table notes.

Some ownership rates for interest-earning assets and retirement savings accounts are lower in SIPP-EHC relative to SIPP (see Table 4.1). However, even though these differences are statistically significant, they are also quantitatively small. For example, in 2010 the ownership rates for interest-earning assets were 0.36 for SIPP and 0.34 for SIPP-EHC; for retirement savings accounts, the ownership rates were 0.17 for SIPP and 0.14 for SIPP-EHC. There is no statistically significant difference in ownership rates of stocks and mutual funds (see Table 4.1). There is no statistically significant difference in how SIPP and SIPP-EHC compare over the two calendar years.

Between SIPP-EHC and SIPP, the mean value of retirement savings does not differ statistically in either

Table 4.1: Asset ownership rates

		Interest-bearing assets	Stocks and mutual funds	Retirement savings
CY 2010	SIPP	0.36	0.05	0.17
	SIPP-EHC	0.34	0.05	0.14
	<i>t</i> -statistic	2.13	0.22	4.63
CY 2011	SIPP	0.39	0.04	0.17
	SIPP-EHC	0.35	0.04	0.14
	<i>t</i> -statistic	3.66	0.37	4.55
Diff-in-diff <i>t</i> -statistic		1.77	0.17	0.35
Person-years		21,952	22,015	21,996

CY2010 or CY2011 (see Table 4.2). The 25th percentile value is lower for SIPP-EHC in both years. The SIPP-EHC also has lower median and 75th percentiles values in CY2010.

Table 4.2: Value of retirement savings

		Mean	25th percentile	Median	75th percentile
CY 2010	SIPP	\$91,606	\$21,381	\$63,202	\$115,032
	SIPP-EHC	\$65,603	\$3,980	\$15,667	\$49,917
	<i>t</i> -statistic	1.42	3.89	5.02	3.42
CY 2011	SIPP	\$131,446	\$18,812	\$68,907	\$149,599
	SIPP-EHC	\$53,902	\$4,932	\$18,000	\$56,500
	<i>t</i> -statistic	2.26	4.01	3.52	6.66
Diff-in-diff <i>t</i> -statistic		1.32	5.53	1.43	1.43
Person-years		775	775	775	775

The mean home values are lower in SIPP-EHC than in SIPP in CY2010 and CY2011 (see Tables 4.3). SIPP-EHC has lower estimates for the median and 75th percentile home value in CY2011.

In SIPP-EHC compared to SIPP, the mean value of the first mortgage is lower in both CY2010 and CY2011 (see Table 4.4). The 25th and 75th percentile values of first mortgage are lower in SIPP-EHC than in SIPP in both years. The median value is lower in SIPP-EHC in CY2011; the median values do not differ statistically between the surveys in CY2010.

For global ownership questions, SIPP-EHC has a slightly higher nonresponse rate relative to SIPP for interest-bearing assets and retirement accounts in 2010 (see Table 4.5). However, the differences are not quantitatively significant. There is no statistically significant difference in nonresponse rates for stocks and mutual stocks. In 2011, there are no statistically significant differences for any of the assets.

Table 4.3: Home value

		Mean	25th percentile	Median	75th percentile
CY 2010	SIPP	\$197,050	\$74,654	\$125,154	\$248,710
	SIPP-EHC	\$163,561	\$66,250	\$114,038	\$196,992
	<i>t</i> -statistic	2.97	1.82	0.09	1.11
CY 2011	SIPP	\$184,189	\$69,981	\$119,295	\$223,898
	SIPP-EHC	\$147,297	\$59,250	\$99,464	\$174,333
	<i>t</i> -statistic	3.25	0.19	2.92	3.30
Diff-in-diff <i>t</i> -statistic		0.21	0.39	0.82	0.09
Person-years		2,843	2,843	2,843	2,843

Table 4.4: Value of first mortgage

		Mean	25th percentile	Median	75th percentile
CY 2010	SIPP	\$145,933	\$59,996	\$100,374	\$198,870
	SIPP-EHC	\$118,699	\$42,250	\$89,500	\$158,500
	<i>t</i> -statistic	2.88	3.51	1.38	3.11
CY 2011	SIPP	\$148,232	\$59,094	\$99,378	\$190,061
	SIPP-EHC	\$111,465	\$41,875	\$79,500	\$145,313
	<i>t</i> -statistic	2.87	3.07	2.96	2.67
Diff-in-diff <i>t</i> -statistic		0.59	0.07	0.87	0.21
Person-years		1,510	1,510	1,510	1,510

For asset amounts, SIPP-EHC has lower non-response rates relative to SIPP in both 2010 and 2011.

Table 4.5: Asset ownership and amounts item-nonresponse rates

		NIU (own- ership)	Interest- bearing assets	Stocks and mutual funds	Retirement savings	Amount of retirement savings	Value of home	Value of first mortgage
CY 2010	SIPP	0.39	0.02	0.02	0.02	0.91	0.31	0.39
	SIPP-EHC	0.32	0.03	0.02	0.02	0.53	0.20	0.30
	<i>t</i> -statistic	10.71	2.10	0.07	1.98	17.57	5.75	3.06
CY 2011	SIPP	0.39	0.02	0.02	0.02	0.90	0.31	0.37
	SIPP-EHC	0.25	0.02	0.01	0.02	0.49	0.19	0.26
	<i>t</i> -statistic	18.19	2.50	0.58	1.41	17.08	5.81	3.89
Diff-in-diff <i>t</i> -statistic		9.34	0.40	0.38	0.35	1.21	0.36	0.80
Person-years		35,084	22,444	22,444	22,446	2,856	3,887	2,317

5. Child Support

This chapter presents comparisons of reporting about child support receipt between SIPP-EHC and SIPP. The following section describes the child support data that are employed in the analysis discussed in the subsequent section.

Description of child support data

The data sources for the tables in this chapter are 2008 SIPP, 2011 SIPP-EHC and 2012 SIPP-EHC. See Chapter 2 for discussion of the SIPP-EHC and SIPP samples. This section discusses the creation of analysis variables specific to child support. There are two child support-specific variables: an indicator of receipt of child support and amount of child support for those who report receiving child support.

The universe for receipt of child support is household members who are at least 15 years old and who are the parent or legal guardian of a child under the age of 21 and the child(ren)'s other biological parent is not a household member. All household members under the age of 15 are not in universe. This definition is consistent for all three survey sources.

In all three survey sources, respondents are asked if they received any kind of financial support payments from their child's other parent. In the 2008 SIPP, imputed responses to this question are recoded to missing. After reporting receipt of child support, the respondent is then asked the amount received in child support. Only sample persons for whom child support receipt was reported are in-universe for a child support amount.

Discussion of child support results

This section presents and discusses the analysis of reporting in SIPP-EHC of child support. The table elements are described in Chapter 2, and, where appropriate, in table notes.

Table 5.1: Child support receipt

		Monthly participation	Annual participation	Months of participation
CY 2010	SIPP	0.23	0.34	9.22
	SIPP-EHC	0.14	0.16	10.59
	<i>t</i> -statistic	4.50	6.63	3.15
CY 2011	SIPP	0.22	0.29	10.24
	SIPP-EHC	0.15	0.17	10.68
	<i>t</i> -statistic	3.58	4.61	0.96
	Diff-in-diff <i>t</i> -statistic	0.53	1.78	1.54
	Observations	31,420	2,053	492

Table 5.1 presents the rates of monthly and annual participation in receipt of child support, as well as the mean number of months in which child support was received. In both CY2010 and CY2011, reported receipt

Table 5.2: Monthly child support amounts

		Benefits
CY 2010	SIPP	\$406.55
	SIPP-EHC	\$375.86
	<i>t</i> -statistic	0.57
CY 2011	SIPP	\$373.78
	SIPP-EHC	\$396.44
	<i>t</i> -statistic	0.40
Diff-in-diff <i>t</i> -statistic		1.09
Person-months		3,087

of child support is higher in SIPP than in SIPP-EHC. In CY2010, SIPP respondents report fewer months of participation in child support receipt. The unit of observation for the monthly participation column is person-months. The unit of observation for the annual participation column is person-years for persons with a reported participation status for all months of the year. The unit of observation for the months participation column is person-years for participants with a reported participation status for all months of the year.

The mean monthly amounts of child support received, among those who received child support, are not significantly different between the SIPP and the SIPP-EHC (see Table 5.2).

The number of child support spells does not differ statistically between the SIPP and the SIPP-EHC (see Table 5.3).

Table 5.3: Spells of child support receipt for the years 2010–2011

	Spells
SIPP	1.07
SIPP-EHC	1.07
<i>t</i> -statistic	0.01
Persons	492

There is no evidence of reverse telescoping (see Chapter 2) in the pattern of month-by-month child support receipt rates in SIPP-EHC (see Table 5.4). The difference between rates of child support receipt between the two surveys do not differ statistically for September through December than for January (Test 1) or January through April (Test 2) (see the *t*-stat column values for Test 1 and Test 2 in Table 5.4).

Finally, item-nonresponse is considered in Table 5.5. Item-nonresponse of receipt and amount of receipt among recipients is much higher in the SIPP than in the SIPP-EHC. Item-nonresponse of amount reported is significantly, but not substantially, higher in the SIPP than in the SIPP-EHC.

Table 5.4: Monthly child support receipt

	CY 2010			CY 2011			Diff
	SIPP	EHC	<i>t</i> -stat	SIPP	EHC	<i>t</i> -stat	
Jan	0.22	0.14	4.06	0.22	0.15	3.32	0.44
Feb	0.23	0.14	4.16	0.22	0.15	3.11	0.73
Mar	0.24	0.14	4.56	0.21	0.15	2.93	1.28
Apr	0.24	0.14	4.60	0.22	0.15	3.17	1.10
May	0.24	0.14	4.64	0.22	0.15	3.10	1.20
Jun	0.24	0.14	4.43	0.22	0.15	3.24	0.88
Jul	0.24	0.14	4.51	0.23	0.15	3.43	0.75
Aug	0.23	0.14	4.27	0.23	0.15	3.70	0.24
Sep	0.23	0.14	4.13	0.23	0.15	3.65	0.17
Oct	0.23	0.14	4.01	0.23	0.14	3.74	0.01
Nov	0.21	0.15	3.28	0.23	0.15	3.53	0.55
Dec	0.21	0.15	3.14	0.23	0.15	3.74	0.90
Pooled	0.23	0.14	4.50	0.22	0.15	3.58	0.53
Test 1	0.23	0.64	0.47	0.79	0.28	0.84	0.84
Test 2	0.76	0.75	0.99	1.26	0.53	1.34	1.53

Table 5.5: Child support item-nonresponse rates

	NIU	Receipt	Amount for recipients
CY 2010	SIPP	0.88	0.88
	SIPP-EHC	0.91	0.00
	<i>t</i> -statistic	7.22	21.44
CY 2011	SIPP	0.88	0.86
	SIPP-EHC	0.90	0.00
	<i>t</i> -statistic	3.40	20.44
Diff-in-diff <i>t</i> -statistic	3.76	0.16	0.48
Person-months	369,317	40,040	604

6. Disability

This chapter presents comparisons of reporting of disability status in SIPP-EHC and SIPP. The following section describes the disability status data that are employed in the analysis discussed in the subsequent section.

Description of disability data

The data sources for the tables in this chapter are the 2008 SIPP, 2011 SIPP-EHC and 2012 SIPP-EHC. See Chapter 2 for discussion of the SIPP-EHC and SIPP samples and weighting.

Disability status is defined as having any of six types of difficulties: vision, hearing, ambulatory, cognitive, self-care, or independent-living. This definition was developed for the American Community Survey but has since been used in a number of federal surveys. This core definition of disability has also been the chosen standard for Department of Health and Human Services data collections as required by Section 4302 of the Affordable Care Act.

In the 2008 SIPP, the six core disability questions were asked during the Medical Expenses and Utilization of Health Care topical module, which was fielded during waves 4, 7, and 10. The wave 7 data is used to generate an estimate for 2010, while the wave 10 data is used to generate an estimate for 2011. The six disability questions in this module are asked only of people age 15 and older. No disability information was collected for children under age 15.

In the 2011 and 2012 SIPP-EHC, data are collected on several types of disability, including the six core questions. Disability status is collected as a current period measure and the survey does not attempt to capture month-to-month dynamics in a respondent's status. For children under age 5, only the vision and hearing questions are asked; for children between 6 and 14, vision, hearing, ambulatory, cognitive, and self-care questions are asked. All six questions are asked for people 15 years and older.

Table 6.1: Disability

		Disabled	Hearing	Vision	Cognitive	Ambulatory	Self-care	Ind.
CY 2010	SIPP	0.21	0.05	0.04	0.09	0.14	0.05	0.09
	SIPP-EHC	0.21	0.05	0.05	0.09	0.13	0.04	0.08
	<i>t</i> -statistic	0.12	0.83	3.19	0.27	2.25	2.67	3.63
CY 2011	SIPP	0.18	0.03	0.03	0.07	0.12	0.04	0.08
	SIPP-EHC	0.22	0.05	0.05	0.09	0.14	0.05	0.09
	<i>t</i> -statistic	5.31	3.50	5.56	4.58	2.59	2.55	1.74
Diff-in-diff <i>t</i> -statistic		5.74	2.75	2.35	4.93	5.25	5.16	5.44
Person-years		22,243	22,263	22,263	22,243	22,247	22,247	22,244

To make comparable estimates of disability status between the data sources for this analysis, only respondents

age 15 and older are considered in universe. In the SIPP, responses with imputed information are coded as missing and excluded from the calculation of estimates.

Discussion of disability results

This section presents and discusses the analysis of disability reporting in SIPP-EHC. The table elements and weighting are described in Chapter 2, and, where appropriate, in table notes.

For CY2010, there are no statistically significant differences between the SIPP and SIPP-EHC in the rates of overall disability, hearing difficulty, or cognitive difficulty (see Table 6.1). Between the two surveys, there are small differences in the estimates for vision difficulty, ambulatory difficulty, self-care difficulty, and independent-living difficulty. For CY2011, except for the independent-living category for which reported difficulty does not statistically differ between the two surveys, reported difficulty in each specific category is one or two percentage points higher in SIPP-EHC than in SIPP. The overall rate of disability in CY2011 is four percentage points higher in the SIPP-EHC (see Table 6.1).

Table 6.2: Disability item-nonresponse rates

		NIU	Disabled
CY 2010	SIPP	0.25	0.15
	SIPP-EHC	0.24	0.11
	<i>t</i> -statistic	2.49	6.73
CY 2011	SIPP	0.25	0.24
	SIPP-EHC	0.25	0.02
	<i>t</i> -statistic	0.01	43.11
Diff-in-diff <i>t</i> -statistic		3.33	26.19
Person-years		35,084	26,486

The item-nonresponse rates for SIPP-EHC are consistently lower than rates for similar items in the SIPP in both CY2010 and CY2011 (see Table 6.2). From CY2010 to CY2011, item-nonresponse for disability in SIPP-EHC decreased relative to SIPP.

7. Education

This chapter compares reported educational attainment and school enrollment in SIPP-EHC and SIPP data. The following section describes the data on educational attainment and school enrollment that are employed in the analysis discussed in the subsequent section.

The data sources for the tables in this chapter are the 2008 SIPP, 2011 SIPP-EHC, and 2012 SIPP-EHC. See Chapter 2 for discussion of the SIPP-EHC and SIPP samples and weighting.

Description of data on educational attainment

The universe for reporting about educational attainment in both SIPP and SIPP-EHC is adults age 15 and over. Both SIPP and SIPP-EHC asked the same question on educational attainment: "What is the highest level of school [sample person] has completed or the highest degree received?"

In SIPP, respondents were also asked whether sample persons attended a vocational, technical, trade, or business school, and, if so, whether they received a diploma from the school. The edited educational attainment measure was a recoded variable combining information on educational attainment and receipt of vocational diplomas. Reports of having less than an associate's degree but also a vocational certificate were recoded to a vocational certificate category on the final educational attainment measure.

The response categories for educational attainment in SIPP-EHC differed from SIPP in several ways. First, the SIPP-EHC did not include a category for vocational certificates. SIPP-EHC sample persons with some college but no degree were divided into two groups: some college credit, but less than one year; and one or more years of college, but no degree. Both SIPP and SIPP-EHC captured educational attainment during the interview month only.

For this analysis, a recoding of educational attainment in SIPP and SIPP-EHC is performed to make the data comparable across the two surveys. SIPP sample persons whose highest level of educational attainment was a vocational degree are recoded into the following categories: less than high school, high school completion, or some college but no degree. SIPP sample persons' educational attainment for December is used because attainment can change within the reference period. In SIPP-EHC, a) some college credit, but less than one year, and b) one or more years of college, but no degree are collapsed into a single response category for some college but no degree. In both SIPP and SIPP-EHC, response categories for specific grades of elementary, middle, and high school are collapsed into a single less than high school category. Any imputed data about educational attainment is recoded as missing for this analysis.

Description of data on school enrollment

SIPP measured school enrollment by asking if respondents were enrolled in each of the four reference months. Therefore, the edited data already include an indicator for school enrollment for each month. In SIPP-EHC, an indicator for monthly enrollment was created. If the value for the start and end month of an enrollment spell had a value greater than or equal to one, then a respondent was coded as enrolled for that month.

In SIPP-EHC, school enrollment information was collected for all sample persons aged 3 and over. In SIPP, only adults aged 15 or over were in-universe for school enrollment questions. Only sample-persons aged 15 or older are included in the analysis. Sample-persons with imputed values are excluded from this analysis.

Discussion of educational attainment and school enrollment results

In both CY2010 and CY2011, SIPP-EHC compared to SIPP shows higher rates of less than high school completion and lower rates of high school completion (see Table 7.1). The differences between SIPP and SIPP-EHC estimates of educational attainment are significant only for the less than high school and high school completion levels.

Table 7.1: Educational attainment

		<HS	HS	Some coll.	Assoc.	Bach.	Mast.	Prof.	Doc.
CY 2010	SIPP	0.30	0.35	0.16	0.06	0.09	0.03	0.01	0.00
	SIPP-EHC	0.34	0.30	0.17	0.06	0.09	0.03	0.01	0.01
	<i>t</i> -statistic	5.24	6.63	1.65	0.44	0.53	0.10	0.75	1.48
CY 2011	SIPP	0.28	0.36	0.16	0.06	0.09	0.03	0.01	0.00
	SIPP-EHC	0.33	0.31	0.16	0.06	0.10	0.03	0.01	0.01
	<i>t</i> -statistic	5.46	6.17	0.20	0.62	0.40	0.24	0.19	1.78
Diff-in-diff <i>t</i> -statistic		0.95	0.37	1.83	0.25	1.05	0.20	0.49	0.91
Person-years		26,091	26,091	26,091	26,091	26,091	26,091	26,091	26,091

Table 7.2 shows pooled monthly and annual enrollment rates for SIPP and SIPP-EHC respondents. In CY2010, the SIPP-EHC pooled monthly enrollment rate is lower than in SIPP, but in CY2011, the SIPP-EHC pooled monthly enrollment rate is higher than in SIPP. In CY2010, the annual enrollment rate in SIPP-EHC and SIPP are not statistically different but in CY2011 the SIPP-EHC annual enrollment rate is higher than in SIPP. These results show the enrollment rates for adults age 15 and over, which is why the rates are low compared to enrollment rates for school-age children in other sources. In CY2011, there are more person-months of enrollment in SIPP-EHC than in SIPP.

Table 7.3 shows that there are more enrollment spells in SIPP-EHC than in SIPP across the two-year period.

Table 7.4 shows monthly enrollment rates in SIPP and SIPP-EHC for CY2010 and CY2011. As some evidence of reverse-telescoping, the difference between SIPP-EHC and SIPP in the monthly school enrollment rate is smaller at the end of the calendar year (September through December) than at the beginning of each calendar year (January or January through April) in each calendar year (see Test 1 and Test 2 in Table 7.4).

In SIPP-EHC compared to SIPP, the not-in-universe rate for educational attainment is higher in CY2010 and CY2011. The cases in the SIPP-EHC data that are not-in-universe for educational attainment reflect missing values. In CY2010, item-nonresponse rates for the variables of educational attainment and school enrollment are lower in SIPP-EHC than SIPP while the item-nonresponse rates do not differ statistically between the two surveys in CY2011 (see Table 7.5).

Table 7.2: School enrollment

		Monthly enrollment	Annual enrollment	Months of enrollment
CY 2010	SIPP	0.15	0.19	8.90
	SIPP-EHC	0.14	0.18	9.15
	<i>t</i> -statistic	2.75	0.64	1.74
CY 2011	SIPP	0.15	0.18	8.87
	SIPP-EHC	0.17	0.21	9.55
	<i>t</i> -statistic	3.15	3.72	4.46
Diff-in-diff <i>t</i> -statistic		6.66	4.85	2.16
Observations		278,675	19,784	3,667

The unit of observation for the Monthly enrollment column is person-months.

The unit of observation for the Annual enrollment and Months of enrollment columns is person-years.

Table 7.3: Total number of school enrollment spells for 2010–2011

	Spells
SIPP	1.47
SIPP-EHC	1.59
<i>t</i> -statistic	5.17
Persons	3,562

Table 7.4: Month-by-month school enrollment

	CY 2010			CY 2011			Diff
	SIPP	EHC	<i>t</i> -stat	SIPP	EHC	<i>t</i> -stat	
Jan	0.16	0.15	2.62	0.16	0.19	4.69	7.69
Feb	0.17	0.15	2.95	0.16	0.19	4.49	7.78
Mar	0.17	0.15	2.48	0.16	0.19	3.78	6.55
Apr	0.17	0.15	2.35	0.16	0.19	3.56	6.16
May	0.17	0.15	2.45	0.16	0.19	3.76	6.49
Jun	0.14	0.12	2.67	0.13	0.16	3.29	5.72
Jul	0.10	0.07	6.10	0.09	0.09	0.60	5.39
Aug	0.13	0.10	4.45	0.12	0.13	1.12	4.97
Sep	0.16	0.15	1.19	0.15	0.17	2.92	4.26
Oct	0.16	0.15	1.17	0.15	0.17	2.17	3.45
Nov	0.16	0.16	0.99	0.16	0.17	1.84	2.91
Dec	0.16	0.15	0.89	0.16	0.17	1.35	2.30
Pooled	0.15	0.14	2.75	0.15	0.17	3.15	6.66
Test 1	0.94	2.33	2.16	0.98	6.98	3.97	4.15
Test 2	1.56	1.77	2.32	1.77	7.01	3.46	3.93

Table 7.5: Education item-nonresponse rates

		NIU (attainment)	Attainment	NIU (enrollment)	Enrollment
CY 2010	SIPP	0.00	0.26	0.00	0.26
	SIPP-EHC	0.02	0.24	0.00	0.23
	<i>t</i> -statistic	11.51	2.11	0.00	3.46
CY 2011	SIPP	0.00	0.25	0.00	0.25
	SIPP-EHC	0.02	0.25	0.00	0.24
	<i>t</i> -statistic	9.33	0.45	0.00	1.24
Diff-in-diff <i>t</i> -statistic		0.58	2.10	0.00	2.82
Observations		35,084	34,856	369,317	369,317

The unit of observation for the Attainment column is person-years.

The unit of observation for the Enrollment column is person-months.

8. Employment and Earnings

This chapter presents comparisons of reporting about employment rates and earning amounts in SIPP-EHC and SIPP. The following section describes the data that are used and the subsequent section discusses the results.

Description of employment data

The data sources for the tables in this chapter are 2008 SIPP, 2011 SIPP-EHC, 2012 SIPP-EHC, and administrative data from IRS Form W-2.

Variables relating to employment are created from survey sources and include a person-month indicator for reported employment and person-year counts of jobs held and businesses owned during the reference year. A spell is defined as one or more consecutive months in which a respondent reported working for an employer or owning a business. Spells may not reflect continuous employment with the same employer or business since periods of non-employment may occur within the same months when employment occurred. For this analysis employment spells are delimited by one or more consecutive months of non-employment.

Person-months in which the sample-person is age 15 years or older are in-universe for employment reporting in the surveys. Because these SIPP interviews occur every 4 months, it is possible that individuals who turn age 15 during a calendar year will have partial year in-universe and not-in-universe employment variables. Unedited SIPP-EHC data used here report only the respondent's age at the time of the interview. For this analysis, the SIPP and SIPP-EHC samples are restricted to the set of full-year respondents; that is, sample-persons with months missing in a given year due to non-interviews are excluded from the analysis sample for that year. Finally, employment variables are set to not-in-universe if their recorded age is not-in-universe (this only affects the unedited SIPP-EHC data).

Monthly employment status in SIPP and SIPP-EHC is determined by comparing the monthly calendar date of the reference year to a respondent's date(s) of reported employment at jobs or owned businesses. Job and business start dates are reported in SIPP and SIPP-EHC if the respondent answers in the affirmative to screener questions about whether or not they held employment during the reference period. Up to two jobs and two businesses may be reported simultaneously in SIPP while a total of eight jobs or businesses may be reported in SIPP-EHC. In either case, a sample-person is classified as employed in a given month if they are in-universe for employment questions and the calendar month and year lie within the bounds of the reported start and end dates of a job or business. Moreover, sample-persons are classified as employed in SIPP if they have an ongoing spell of employment that began prior to the calendar month and year of the reference year with no reported job end date. In SIPP, monthly employment status is set to 'not employed' if a sample person is in-universe for employment questions but has no reported date(s) of employment or the reported date(s) of employment lie outside the particular calendar month and year of the reference period. In SIPP-EHC, monthly employment status is set to not employed if a respondent reports no valid periods of employment *and at least one valid spell of being without a job*. In SIPP, ongoing employment spells typically have a reported in-universe job or business start date while the job or business end date is out of universe. This issue does not arise in SIPP-EHC since a start and end date (month) is reported for all employment spells.

For SIPP, an employment spell is set to missing if the sample-person is in-universe for employment questions and the jobs and owned businesses in the given reference month have imputed job or business start dates. Sample persons holding one or more jobs or businesses (or a combination of the two) during any reference month where at least one start date is not imputed are classified as employed. The same approach is used

for SIPP-EHC with one exception: monthly employment status is set to missing if a sample-person has no reported employment and no reported non-employment spells.

Annual job and own-business counts are derived quite differently in SIPP compared to SIPP-EHC. The variables EENO and EBNO characterize unique across-wave employer and own-business index variables, respectively, in edited SIPP data. Although one job or own-business may be recorded on separate job or own-business lines (e.g. job 1 versus job 2) in different waves, the employer and own business index variables are designed to be consistent across waves. Annual and own-business counts are simply the number of unique employers and own-businesses worked for during the reference year. Note that EENO and EBNO do not have corresponding allocation flags to indicate whether the data are imputed, and, as such, annual job and own-employer counts are never set to missing for sample-persons that are nonmissing due to non-interview.

In SIPP-EHC, each job line is treated as a separate employer or own-business during the reference year. It is possible that a respondent may work multiple periods for the same employer or own business, which should occupy the same job line as long as both employment spells occurred—not necessarily entirely—within the reference year. Work for employers and for own-businesses is treated separately for consistency with SIPP. Total annual job and own-business counts are the number of jobs and own-businesses reported during the reference year. Unlike 2008 SIPP, annual job and business counts may be classified as missing even without missing interviews, which occurs if a respondent reports no employment and no non-employment spells during the reference year or if the type of work cannot be classified. Attached to each job line is the variable JBORSE, which classifies the employment spell as work for either an employer, self-employment, or some other work arrangement. There are a few cases in the unedited SIPP-EHC data where JBORSE is recorded as missing.

Description of earnings data

The process of collecting earning data is structured very differently in SIPP-EHC and SIPP. In the SIPP, respondents were explicitly told that the purpose of the earnings questions was to measure monthly amounts. If respondents were not able to recall monthly totals, they were encouraged to refer to earnings records and to report each pay check in the month it was received. The survey instrument would then prompt the interviewer to ask for confirmation of the amount from the respondent. If respondents were unable to report monthly totals, but could report an hourly wage, the instrument calculated weekly and monthly earnings using usual hours, and then prompted for verification from the respondent. In the SIPP, earnings are tied to the month in which they are received, not earned; if a job ended, there could still be monthly earnings from a job that had ended the previous month, and if a job had just begun, it could be possible to have no earnings reported for the first month even if the respondent was employed in a job.

In the SIPP-EHC, respondents were asked about periods when payment for work was earned. Respondents were invited to report earnings in one of eight ways: hourly wage, weekly, biweekly, monthly, bimonthly, monthly average, annual salary, and actual annual pay. Up to two changes in earnings rate in a year could be reported for each sample person. From these reports, monthly earnings are calculated for this analysis.

Notice that there is a slight discrepancy here; in SIPP, the earnings were recorded in the month that they were received, while in SIPP-EHC they were assigned to the month in which they were earned. The exception is for extra earnings which includes tips, commission, overtime, and bonuses; these were collected at the monthly level, and a single amount was applied to all months if the respondent said they received the income every month. Note that in 2012 SIPP-EHC (CY 2011), there appears to be a discrepancy in extra earnings between the raw data and the reformatted data provided for this analysis. While most of the amounts are correct, some sample-persons are missing these types of earnings who should not be. This issue is being

investigated.

Only SIPP sample-persons who have a reported non-imputed response to monthly job earnings, monthly business earnings, and moonlighting earnings are included in the earning analysis. In SIPP, moonlighting earnings are collected quite differently than earnings from jobs or businesses, and moonlighting is not factored into the calculations of employment status. The same is true for contingent workers. In SIPP-EHC, respondents are asked to report information on every job they worked, whether a primary job or not. In the earnings calculations for SIPP-EHC, we include all sample persons who had nonmissing values for all of the relevant fields: this includes the type of pay rate reported (hourly, annual salary, etc.), the pay rate amount, whether they had a change in earnings, the week(s) of such changes, the weeks of changes and amounts for hours (when reporting an hourly wage), and the first and last week within the year that the job was held. This means more fields of nonmissing data are required from SIPP-EHC in the analysis sample.

Annual earnings in the SIPP and SIPP-EHC are calculated by summing the monthly earnings amounts described above for each of the twelve months in the reference year. We restrict the sample to those with a nonmissing monthly report for every month in the given reference period year.

The universe for monthly earnings in SIPP is anyone who held a job for pay, a business, a moonlighting job, or a contingent job during the month. If the sample-person was away from work without pay during the month, the earnings value for that job is set to zero. For SIPP-EHC, the universe for earnings includes anyone who was eligible to report a job spell during the month *and did not report that they were not working*. Earnings variables for sample-persons with missing responses for both the job and no-job questions during the month were coded as missing. Out of necessity, cases with missing data on the timing of employment/earnings changes were also classified as missing for SIPP-EHC; for example, cases in which the first week the job was held are classified as missing for earnings, even if earnings rate data are nonmissing.

Description of employment and earnings results

This section discusses the analysis of the employment and earnings variables in SIPP-EHC.

Rates of monthly and annual employment¹ are higher in SIPP-EHC than in SIPP (see Table 8.1). The difference between the surveys could be due to the higher non-response rate about employment in SIPP-EHC (see Table 8.9) if those without employment in SIPP-EHC are more likely to have missing data. Since SIPP-EHC sample-persons are coded as not employed only if there is a negative report about employment and a positive report about non-employment (a step not required of SIPP respondents), this is a plausible explanation.

The rate of disagreement between survey and administrative data about the number of distinct employers per sample-person in a year is lower in SIPP-EHC than in SIPP in both years (see Table 8.2). The mean absolute deviation between the survey-reported and the administratively-recorded number of distinct employers per sample-person annually is lower for SIPP-EHC than for SIPP in both CY2010 and CY2011. The number of businesses owned annually is somewhat lower in SIPP-EHC than in SIPP (see Table 8.3). The number of employment spells is lower in SIPP-EHC than in SIPP (see Table 8.4).

There is no evidence of reverse telescoping in reporting about employment in SIPP-EHC in the pattern of month-by-month rates of employment (see Table 8.5). The difference between the surveys in the difference between reported rates of employment at the beginning and end of the year is not statistically different (see the values for tests 1 and 2 in the CY2011 t-stat column in Table 8.5).

Mean reported monthly earnings are higher in SIPP-EHC than in SIPP in both years, more so in CY2010 (see Table 8.6). The differences appear to be due to outliers. The values at the median and 25th percentile do not

¹Rates of employment here are simply defined as the proportion of sample-persons aged 15 or older with reported employment in the specific period (month or year).

differ statistically in CY2010 and differ between the surveys by a relatively small amount in CY2011. The values at the 75th percentile differ statistically between the survey in both years, but not substantially.

Mean reported annual earnings are higher in SIPP-EHC than in SIPP in both years and by a greater amount in CY2010 (see Table 8.7). These differences also appear to be due to outliers. The values at the 25th percentile do not differ statistically in CY2010 but differ between the surveys in CY2011. The values at the median and the 75th percentile differ somewhat between the surveys in both years.

The extent of disagreement between survey and administrative data about mean annual earnings from employment does not differ statistically between the two surveys in either year. Between the two surveys, the mean absolute deviation between the survey-reported and administratively-recorded mean annual earnings does not statistically differ for either calendar year (see Table 8.8).

Allowing multiple ways of reporting in the SIPP-EHC may have led to some extreme outliers. It is believed that this is due in part to incongruity between reported pay rates and reported type of pay rate that does not correspond to the pay rate they report (e.g., reporting paid "hourly" and then reporting biweekly salary).

In 2011 SIPP-EHC (CY 2010), two soft checks were added to the survey instrument to verify particularly large and particularly small values reported for an annual salary and an hourly wage. As an example, if an hourly wage greater than \$250 is reported, a screen comes up for the field representative which reads "This amount seems unusually high. Are you sure this is the amount of your/his/her hourly rate?" In 2012 SIPP-EHC (CY 2011), checks like these were added for all eight ways of reporting wage and salary earnings. This likely helped lead to fewer outliers in CY2011 relative to CY2010, and more reasonable mean values.

Non-response about employment is higher in SIPP-EHC than in SIPP in both CY2010 and CY2011 (see Table 8.9). As previously described, this could be because respondents, essentially, had to affirm a report of non-employment in a second way.

Non-response about monthly earnings is also higher in SIPP-EHC than in SIPP (see Table 8.9). As previously described, SIPP-EHC requires a response to more questions in order to provide a measure of annual earnings.

Table 8.1: Reported employment

		Monthly reported employment	Annual reported employment	Months of reported employment
CY 2010	SIPP	0.48	0.54	10.71
	SIPP-EHC	0.57	0.62	10.86
	<i>t</i> -statistic	8.25	7.62	1.91
CY 2011	SIPP	0.48	0.54	10.71
	SIPP-EHC	0.57	0.63	10.94
	<i>t</i> -statistic	8.78	8.03	2.81
Diff-in-diff <i>t</i> -statistic		1.40	1.12	0.83
Observations		214,969	17,492	10,032

The unit of observation for the Monthly reported employment column is person-months. The unit of observation for the Annual reported employment and Months of reported employment columns is person-years.

Table 8.2: Annual number of employers

		Survey	Linked	AR	MAD
CY 2010	SIPP	1.28	1.27	1.23	0.41
	SIPP-EHC	1.12	1.13	1.27	0.35
	<i>t</i> -statistic	11.30	9.06	1.47	2.91
CY 2011	SIPP	1.25	1.25	1.19	0.44
	SIPP-EHC	1.14	1.15	1.22	0.33
	<i>t</i> -statistic	7.40	5.97	1.02	4.14
Diff-in-diff <i>t</i> -statistic		2.85	2.05	0.22	1.51
Person-years		9,760	7,726	7,726	7,726

Table 8.3: Number of businesses owned

		Businesses
CY 2010	SIPP	1.12
	SIPP-EHC	1.02
	<i>t</i> -statistic	4.96
CY 2011	SIPP	1.14
	SIPP-EHC	1.07
	<i>t</i> -statistic	2.14
Diff-in-diff <i>t</i> -statistic		1.51
Person-years		1,115

Table 8.4: Employment spells

		Spells
SIPP		1.11
SIPP-EHC		1.06
<i>t</i> -statistic		6.44
Persons		6,516

Table 8.5: Month-by-month employment status

	CY 2010			CY 2011			Diff
	SIPP	EHC	<i>t</i> -stat	SIPP	EHC	<i>t</i> -stat	
Jan	0.48	0.55	7.35	0.47	0.57	8.06	1.42
Feb	0.48	0.56	7.59	0.47	0.57	8.22	1.35
Mar	0.48	0.56	8.06	0.48	0.57	7.92	0.54
Apr	0.48	0.56	7.45	0.48	0.57	7.93	1.16
May	0.48	0.56	7.50	0.48	0.57	8.11	1.31
Jun	0.48	0.57	8.20	0.48	0.58	8.46	0.98
Jul	0.49	0.57	8.14	0.48	0.57	8.36	0.94
Aug	0.49	0.57	7.76	0.48	0.57	8.46	1.42
Sep	0.49	0.57	7.75	0.48	0.57	8.58	1.56
Oct	0.48	0.57	8.00	0.47	0.57	8.63	1.36
Nov	0.49	0.57	7.95	0.47	0.57	8.83	1.62
Dec	0.48	0.57	8.34	0.47	0.57	8.89	1.30
Pooled	0.48	0.57	8.25	0.48	0.57	8.78	1.40
Test 1	2.31	3.82	1.10	0.01	1.43	1.06	0.05
Test 2	2.22	3.08	0.71	0.62	1.16	1.28	0.44

Table 8.6: Monthly earnings

		Mean	25th percentile	Median	75th percentile
CY 2010	SIPP	\$2,432	\$1,195	\$1,944	\$3,077
	SIPP-EHC	\$9,261	\$1,198	\$1,964	\$3,296
	<i>t</i> -statistic	3.95	0.20	1.37	6.24
CY 2011	SIPP	\$2,462	\$1,195	\$1,947	\$3,112
	SIPP-EHC	\$3,996	\$1,303	\$2,033	\$3,296
	<i>t</i> -statistic	2.00	8.75	4.45	4.48
Diff-in-diff <i>t</i> -statistic		2.80	5.97	2.7	0.64
Person-months		89,080	89,080	89,080	89,080

Table 8.7: Annual earnings

		Mean	25th percentile	Median	75th percentile
CY 2010	SIPP	\$28,395	\$12,997	\$22,021	\$35,839
	SIPP-EHC	\$111,200	\$13,285	\$23,053	\$38,977
	<i>t</i> -statistic	3.93	1.57	4.91	10.84
CY 2011	SIPP	\$28,716	\$13,731	\$21,983	\$36,261
	SIPP-EHC	\$47,335	\$15,025	\$23,462	\$39,071
	<i>t</i> -statistic	3.91	11.31	17.62	29.69
Diff-in-diff <i>t</i> -statistic		2.81	3.32	2.43	1.16
Person-years		8,876	8,876	8,876	8,876

Table 8.8: Mean annual earnings: administrative records comparison

		Survey	Linked	AR	MAD
CY 2010	SIPP	\$28,395	\$25,618	\$28,950	\$9,132
	SIPP-EHC	\$111,200	\$117,836	\$26,188	\$98,920
	<i>t</i> -statistic	3.93	1.12	3.03	1.09
CY 2011	SIPP	\$28,716	\$25,951	\$28,854	\$10,198
	SIPP-EHC	\$47,335	\$29,019	\$26,347	\$10,171
	<i>t</i> -statistic	3.91	2.55	2.34	0.03
Diff-in-diff <i>t</i> -statistic		2.81	1.08	0.29	1.09
Person-years		8,876	6,718	6,718	6,718

Table 8.9: Employment and earnings item-nonresponse rates

		Employment	Monthly earnings
CY 2010	SIPP	0.01	0.12
	SIPP-EHC	0.23	0.49
	<i>t</i> -statistic	37.99	37.10
CY 2011	SIPP	0.01	0.13
	SIPP-EHC	0.16	0.47
	<i>t</i> -statistic	25.21	29.01
Diff-in-diff <i>t</i> -statistic		9.20	2.25
Person-months		238,734	130,003

9. Health Insurance

This chapter presents comparisons of reporting about private health insurance coverage in SIPP-EHC and SIPP. The following section describes the health insurance data that are employed in the analysis discussed in the subsequent section.

Description of private health insurance data

The data sources for the tables in this chapter are 2008 SIPP, 2011 SIPP-EHC, and 2012 SIPP-EHC. See Chapter 2 for discussion of the SIPP-EHC and SIPP samples and weighting. This section discusses the creation of analysis variables specific to private health insurance.

Variables measuring private health insurance coverage include person-month level indicators of private health insurance coverage and of type of private health insurance coverage held. All person-months are in-universe for the health insurance coverage indicator, and all person-months with a positive report of coverage and an age of 15 or older are in-universe for the type-of-coverage indicators. Indicators of coverage (and type of coverage) are coded as missing for all in-universe person-months for which no report of health insurance coverage (or coverage type) is available.

From 2008 SIPP, these variables are created from unedited person-month data. For adults, the information about health insurance coverage is obtained from the sample person's own record. Adults also could report additional household members (of any age) on their plan. These additional reports about other household members are merged onto the appropriate person-month records, reconciling multiple reports such that if any report indicated coverage, the coverage indicator variable would indicate coverage. For children, coverage information was merged from a guardian's report on behalf of the child. For these cases, months of coverage were not asked, and so it is assumed that a covered child had health insurance for all four months of the wave. Teenagers over age 15 can have reports from both themselves and a guardian, in which case this data are reconciled to indicate coverage if either source so indicated. The variable for type of health insurance is created from unedited person-level data for adults only because the questions for children did not inquire about plan type.

In the 2011 and 2012 SIPP-EHC, information was collected about any spells of coverage under several types of health insurance during the reference year. The same questions were asked for adult and child sample-persons. If coverage was indicated for a sample-person, then further questions inquired about the start and end dates of the coverage, other covered members on the plan, and plan type. As with the SIPP data, any reports of other covered household members are merged onto the appropriate person-month records, reconciling multiple reports such that if any report indicated coverage, the final health insurance variable would indicate coverage. Monthly indicators for health insurance coverage and type of insurance are created for CY2010 and CY2011.

The definition of private health insurance coverage for this analysis is designed to be comparable across SIPP-EHC and SIPP. The definition of private coverage in SIPP has always excluded non-comprehensive health plans (e.g. vision or dental only), but other details have varied across panels. When the SIPP began in 1984, the private health insurance question defined health insurance as excluding Medicaid, Medicare, CHAMPUS, and CHAMPVA. Beginning in 1988, the private health insurance question changed to include military coverage, and types of military coverage were added to the answer list for the type of health insurance question. This allowed SIPP to capture military coverage without adding a new question to the (at the time) paper survey.

SIPP-EHC was designed to capture health insurance types better than in SIPP. SIPP embedded military insurance into a direct pay health insurance question that did not probe specifically for military coverage (including its public forms). The estimates of direct pay coverage rates from SIPP were lower than expected. The SIPP-EHC added a separate set of questions about military-related health insurance coverage allowing respondents to report both private and public types of military or VA coverage. To help reduce under-reporting, the SIPP-EHC also added questions to probe for other types of health insurance coverage that, to a respondent, may not seem to fit into the other categories provided.

In order to compare the SIPP-EHC instrument data to the SIPP private health insurance coverage indicators, SIPP-EHC data are used from the questions on private health insurance, military (all types), as well as “other” coverage (when other type is not indicated to be government/public assistance). Thus, like the SIPP private health insurance variable, the SIPP-EHC variable includes all comprehensive coverage excluding Medicare and Medicaid. We have continued to refer to this variable as private health insurance although, technically speaking, it contains small percentages of public military/VA coverage. Type of private health insurance in SIPP-EHC are reconciled with the list of private types from SIPP and includes coverage through: a current employer, a former employer, a union, TRICARE, CHAMPVA, other military/VA, direct-purchase, and other.

Discussion of health insurance results

This section presents and discusses the analysis of health insurance reporting in SIPP-EHC. The table elements are described in Chapter 2, and, where appropriate, in table notes.

Table 9.1: Health-insurance coverage

		Monthly coverage	Annual coverage	Months of coverage
CY 2010	SIPP	0.42	0.51	9.96
	SIPP-EHC	0.37	0.38	11.56
	<i>t</i> -statistic	7.13	14.70	21.75
CY 2011	SIPP	0.42	0.50	10.11
	SIPP-EHC	0.39	0.41	11.44
	<i>t</i> -statistic	3.70	9.39	16.61
Diff-in-diff <i>t</i> -statistic		3.48	4.93	2.91
Observations		319,183	22,373	10,115

The unit of observation for the monthly coverage column is person-months.

The unit of observation for the annual coverage and months of coverage columns is person-years.

For monthly and annual health insurance coverage, the SIPP-EHC captures lower rates of monthly and annual health insurance coverage than SIPP in both years, with smaller differences between the two surveys

Table 9.2: Health insurance type

		Curr. emp.	Form. emp.	Union	TRICARE	CHAMPVA	Mil.	Direct Pay	Other
CY 2010	SIPP	0.76	0.08	0.02	0.01	0.00	0.01	0.10	0.02
	SIPP-EHC	0.65	0.06	0.02	0.05	0.00	0.05	0.13	0.04
	<i>t</i> -statistic	7.19	1.69	0.96	5.51	1.01	6.95	2.36	2.60
CY 2011	SIPP	0.76	0.08	0.02	0.01	0.00	0.01	0.10	0.02
	SIPP-EHC	0.69	0.06	0.02	0.03	0.00	0.04	0.11	0.04
	<i>t</i> -statistic	5.28	1.54	0.25	3.64	1.97	5.35	1.78	3.70
Diff-in-diff <i>t</i> -statistic		2.20	0.15	0.60	2.75	0.95	2.52	0.72	0.89
Person-months		94,690	94,690	94,690	94,690	94,690	94,690	94,690	94,690

The first column reports the percentage of insured persons with coverage through a current employer.

The second column reports the percentage of insured persons with coverage through a former employer.

The third column reports the percentage of insured persons with coverage through a union.

The fourth column reports the percentage of insured persons with coverage through TRICARE or CHAMPUS.

The fifth column reports the percentage of insured persons with coverage through CHAMPVA.

The sixth column reports the percentage of insured persons with coverage through the U.S. military or Department of Veterans Affairs (VA).

The seventh column reports the percentage of insured persons with coverage through a direct pay policy.

The eighth column reports the percentage of insured persons with coverage through a some other source.

in CY2011 than in CY2010 (see Table 9.1). For sample persons reported to be covered at some point during the year, SIPP-EHC captures more months of coverage than SIPP for both years, with a smaller difference between the surveys in CY2011 than in CY2010.

For adult person-months with private coverage, coverage rates by plan type differ between SIPP-EHC and SIPP for multiple types (see Table 9.2). The SIPP-EHC captures lower rates of coverage through a current employer than SIPP with a smaller difference between the surveys in CY2011 than in CY2010. As anticipated, the SIPP-EHC captures higher rates of coverage than SIPP for TRICARE and other military/VA coverage, although these differences are smaller in CY2011 than in CY2010. In addition, SIPP-EHC captures higher rates of direct-purchase than SIPP in CY2010 and higher rates of other private health insurance than SIPP in both years.

Among persons with any coverage during a calendar year, the SIPP-EHC captures fewer spells of coverage than SIPP (see Table 9.3). Because the SIPP was administered 3 times a year, there is a risk that reporting error across the waves could create false transitions in coverage status. Comparison to administrative data for public health insurance types suggests that SIPP did indeed capture spurious transitions in public coverage due to its design (see Chapter 12). The same may be true for private health insurance; the number of spells in SIPP may be inflated as an artifact of the shorter reference period in SIPP.

There is no evidence of reverse-telescoping in the pattern of reporting of health insurance coverage across the SIPP-EHC reference period. In CY2010, the difference between SIPP-EHC and SIPP in the rate of reported

Table 9.3: Total number of insurance-coverage spells for 2010–2011

	Spells
SIPP	1.13
SIPP-EHC	1.01
<i>t</i> -statistic	20.54
Persons	8,781

health insurance coverage is higher at the end of the calendar year (September through December) than at the beginning of the calendar year (January through April) (see the value for Test 2 in the CY2010 *t*-stat column in Table 9.4). In CY2011, the difference between the surveys in the difference between reported rates of coverage at the beginning and end of the year is not statistically different (see the values for tests 1 and 2 in the CY2011 *t*-stat column in Table 9.4).

The rate of item-nonresponse about health insurance coverage is lower in SIPP-EHC than in SIPP (see Table 9.5). From CY2010 to CY2011, the rate of item-nonresponse for SIPP-EHC decreases relative to SIPP. In

Table 9.4: Month-by-month health-insurance coverage

	CY 2010			CY 2011			Diff
	SIPP	EHC	<i>t</i> -stat	SIPP	EHC	<i>t</i> -stat	
Jan	0.42	0.37	6.63	0.43	0.39	3.58	2.80
Feb	0.42	0.37	6.26	0.42	0.39	4.17	1.74
Mar	0.42	0.37	6.04	0.42	0.39	3.90	1.82
Apr	0.41	0.37	5.61	0.42	0.39	3.82	1.47
May	0.42	0.37	6.20	0.42	0.39	3.98	1.91
Jun	0.42	0.36	6.65	0.42	0.39	3.44	3.00
Jul	0.43	0.36	7.38	0.42	0.39	3.56	3.62
Aug	0.43	0.37	7.35	0.42	0.39	3.61	3.54
Sep	0.43	0.37	7.13	0.42	0.39	3.37	3.61
Oct	0.43	0.37	7.03	0.42	0.39	3.39	3.51
Nov	0.43	0.37	6.89	0.41	0.39	2.81	4.02
Dec	0.43	0.37	7.41	0.41	0.39	2.45	5.02
Pooled	0.42	0.37	7.13	0.42	0.39	3.70	3.48
Test 1	1.53	1.05	1.04	2.07	1.84	0.96	1.30
Test 2	2.76	1.43	2.04	2.05	0.41	1.70	2.49

Table 9.5: Health insurance item-nonresponse rates

		NIU (coverage)	Coverage	NIU (type)	Type
CY 2010	SIPP	0.00	0.18	0.70	0.09
	SIPP-EHC	0.00	0.10	0.66	0.47
	<i>t</i> -statistic	0.00	15.47	6.33	34.67
CY 2011	SIPP	0.00	0.17	0.70	0.09
	SIPP-EHC	0.00	0.04	0.68	0.12
	<i>t</i> -statistic	0.00	30.24	2.10	3.44
Diff-in-diff <i>t</i> -statistic		0.00	12.39	4.39	27.07
Person-months		369,202	369,202	369,202	116,559

both calendar years, the rate of item-nonresponse about type of health insurance is higher in SIPP-EHC than in SIPP. From CY2010 to CY2011, the rate for SIPP-EHC decreases relative to SIPP.

10. Household Composition

This chapter presents comparisons of reporting about household composition and living arrangements of children in SIPP-EHC and SIPP. The following section describes the household composition and child living arrangement data that are employed in the analysis discussed in the subsequent section.

Description of household composition data

The data sources for the tables in this section are 2008 SIPP, 2011 SIPP-EHC, and 2012 SIPP-EHC. See Chapter 2 for discussion of the SIPP-EHC and SIPP samples and weighting. This section discusses the creation of the analysis variables for household composition and the living arrangements of children.

The variables discussed in this section are created from three survey sources (2008 SIPP, 2011 SIPP-EHC, and 2012 SIPP-EHC). SIPP-EHC unedited data are used to create the analysis variables. From SIPP-EHC, the household composition variable is created at the household level for households as constituted at the time of interview. All households are in-universe. From SIPP, the household composition variable is created from edited data, dropping all imputations, for households as constituted in December of each year. Households with an interview in December of each calendar year are in-universe.

Household composition is a household level variable indicating whether the household consists of a married couple with children, a married couple with no children, a female head with children, a male head with children, a householder with other relatives, a single person, or a non-family household. Households with an invalid or imputed householder or an invalid or imputed value for the gender of the householder are coded as missing. If the household contains some respondents that have valid, non-imputed relationship data and others for whom the relationship to householder is invalid or imputed, household composition is coded based only on those with valid, non-imputed information. If only the householder has valid, non-imputed relationship data and all other data about relationships in the household are invalid or imputed, the household is coded as a non-family household.

Description of living arrangement of children data

Children's living arrangements are classified in ten categories, living with: both biological parents; a biological mother and stepfather; a biological father and stepmother; a biological mother and an adoptive father; a biological father and an adoptive mother; two adoptive parents; two parents of some other type; mother only; father only; and without parents. Parental type (biological, step, or adoptive) is specified only when children are living with two parents. Imputed data are not used to construct these categories so if the parental type for either parent is imputed or incomplete for any child living with two parents then the child is coded as living with two parents of another type. Adults age 18 and over are out of universe for analysis of child living arrangements.

Discussion of household composition results

This section presents and discusses the analysis of the reporting of household composition variable in SIPP-EHC.

The proportion of sample-households composed of married couples with children is lower in SIPP-EHC than in SIPP in CY2010 (see Table 10.1). In the same calendar year, the proportion of male headed households with children is significantly, but not substantially, higher in SIPP-EHC than in SIPP. The single and nonfamily categories are combined for SIPP but not for SIPP-EHC, so we do not make any inference from the comparisons between the surveys for those categories.

Table 10.1: Household composition

		M-C	M-NC	FH-C	MH-C	Other	Single	Nonfamily
CY 2010	SIPP	0.19	0.17	0.11	0.02	0.18	0.00	0.34
	SIPP-EHC	0.15	0.16	0.12	0.03	0.17	0.30	0.07
	<i>t</i> -statistic	3.51	0.69	1.25	2.51	1.14	33.20	28.42
CY 2011	SIPP	0.18	0.18	0.09	0.02	0.20	0.00	0.33
	SIPP-EHC	0.17	0.17	0.10	0.03	0.19	0.27	0.07
	<i>t</i> -statistic	1.07	0.70	1.04	1.62	0.78	26.76	25.76
Diff-in-diff <i>t</i> -statistic		2.82	0.14	0.17	0.79	0.26	3.68	1.01
Household-years		11,698	11,698	11,698	11,698	11,698	11,698	11,698

The first column reports the percentage of households comprised of a married couple with children.

The second column reports the percentage of households comprised of a married couple with no children.

The third column reports the percentage of households comprised of a female head with children.

The fourth column reports the percentage of households comprised of a male head with children.

The fifth column reports the percentage of households comprised of other relatives.

The sixth column reports the percentage of one-person households.

The seventh column reports the percentage of non-family households.

For household composition, SIPP-EHC has lower item non-response than SIPP in both calendar years (see Table 10.2).

Discussion of children's living arrangements results

This section presents and discusses the analysis of the reporting of children's living arrangements in SIPP-EHC.

The proportion of children living with either the biological mother and a stepfather or the biological father and a stepmother is statistically, but not substantially, lower in SIPP-EHC than in SIPP in both calendar years (see Table 10.3). In SIPP compared to SIPP-EHC, there is a slightly higher proportion of children living with two parents of another type or no parents in CY2010 and CY2011.

Though item-nonresponse is low in both SIPP-EHC and SIPP for child living arrangements, SIPP-EHC compared to SIPP has higher non-response in CY2010 (see Table 10.4).

Table 10.2: Household composition item-nonresponse rates

		NIU	Household composition
CY 2010	SIPP	0.00	0.01
	SIPP-EHC	0.00	0.00
	<i>t</i> -statistic	0.00	5.60
CY 2011	SIPP	0.00	0.01
	SIPP-EHC	0.00	0.00
	<i>t</i> -statistic	0.00	3.31
Diff-in-diff <i>t</i> -statistic		0.00	0.27
Household-years		11,769	11,769

Table 10.3: Child living arrangements

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
CY 2010	SIPP	0.46	0.05	0.01	0.00	0.00	0.00	0.04	0.39	0.04	0.08
	SIPP-EHC	0.47	0.03	0.00	0.00	0.00	0.00	0.01	0.40	0.04	0.04
	<i>t</i> -statistic	0.77	3.73	2.26	1.35	1.73	1.69	7.53	0.68	0.44	5.27
CY 2011	SIPP	0.48	0.05	0.01	0.00	0.00	0.00	0.03	0.38	0.05	0.07
	SIPP-EHC	0.50	0.03	0.00	0.00	0.00	0.00	0.01	0.36	0.05	0.05
	<i>t</i> -statistic	1.07	3.46	2.36	0.28	0.00	0.16	5.45	1.67	0.98	3.07
Diff-in-diff <i>t</i> -statistic		0.49	0.06	0.04	0.34	1.73	1.07	1.16	2.84	1.50	1.67
Person-years		10,559	10,559	10,559	10,559	10,559	10,559	10,559	10,559	10,559	10,559

Column (1) reports the percentage of children living with both biological parents.

Column (2) reports the percentage of children living with a biological mother and a stepfather.

Column (3) reports the percentage of children living with a biological father and a stepmother.

Column (4) reports the percentage of children living with a biological mother and an adoptive father.

Column (5) reports the percentage of children living with a biological father and an adoptive mother.

Column (6) reports the percentage of children living with both adoptive parents.

Column (7) reports the percentage of children living with two parents of another type.

Column (8) reports the percentage of children living with a mother only.

Column (9) reports the percentage of children living with a father only.

Column (10) reports the percentage of children living with no parents.

Table 10.4: Child living arrangements item-nonresponse rates

		NIU	Child living arrangement
CY 2010	SIPP	0.69	0.00
	SIPP-EHC	0.72	0.01
	<i>t</i> -statistic	4.30	4.04
CY 2011	SIPP	0.69	0.00
	SIPP-EHC	0.71	0.00
	<i>t</i> -statistic	2.04	0.00
Diff-in-diff <i>t</i> -statistic		2.70	4.04
Person-years		35,084	10,577

11. Housing Subsidies

This chapter presents comparisons of reporting about housing-subsidy receipt between SIPP-EHC and SIPP. The following section describes the housing-subsidy data that are employed in the analysis of the subsequent section.

Description of housing subsidy data

The data sources for the tables in this chapter are the 2008 SIPP, 2011 SIPP-EHC, and 2012 SIPP-EHC. See Chapter 2 for discussion of the SIPP-EHC and SIPP samples and weighting. This section discusses the creation of analysis variables specific to receipt of housing subsidies.

The variable indicating receipt of housing subsidies is created from the survey sources (2008 SIPP, 2011 SIPP-EHC, and 2012 SIPP-EHC). Each survey source includes a person-month indicator of whether the sample-person lived in a unit that received a housing subsidy.

The universe for the housing-subsidy variable includes all person-months where either the sample-person lived in a renter-occupied unit, a rent-free unit, or no response was provided to the housing tenure question. The housing subsidy variable is coded as missing when a sample-person is in universe but there is no survey report available.

For the 2008 SIPP, this variable is created from edited person-month data, dropping all imputations. Information on housing-subsidy receipt for the interview address is a report by the householder. This information is merged on to the records of non-householders for all months during the interview reference period where they lived at the same address as the householder. The 2008 SIPP included several questions to identify housing-subsidy receipt and type of housing subsidy. Two variables are used to construct a single variable indicating sample-persons who lived in a unit that received any type of housing subsidy. The first variable, for all sample-persons who were in the housing-subsidy universe described above, is an indicator of whether the current unit was owned by a local housing authority or other agency. The second variable, for those who responded in the negative to the previously described public-housing question, is an indicator of whether the respondent lived in a unit that received a federal, state, or local government housing subsidy. Affirmative responses to either of these questions are combined to create the SIPP indicator of housing-subsidy receipt.

For the 2011 and 2012 SIPP-EHC, housing-subsidy receipt for the interview address is as reported by the householder. This information is merged on to the records of all household members for the months during the reference period during which they lived at the same address as the householder. The tenure and housing-subsidy status of the residential unit for the months any sample-person did not live with the householder was collected. In SIPP-EHC, only one question was used to determine overall housing-subsidy receipt. An affirmative response to the question about receiving a federal, state, or local government housing subsidy is coded as receipt of housing subsidy.

Discussion of housing subsidy results

In both CY2010 and CY2011, the monthly and annual rates of reported housing-subsidy receipt are larger in SIPP than in SIPP-EHC (see Table 11.1). Among sample-persons who received a housing subsidy, the reported average months of housing-subsidy receipt are greater in SIPP-EHC than in SIPP for both calendar years. The number of housing-subsidy receipt spells in CY2010 through CY2011 is greater for SIPP than SIPP-EHC (see Table 11.2).

Table 11.1: Housing subsidies

		Monthly subsidy	Annual subsidy	Months of subsidy
CY 2010	SIPP	0.24	0.29	10.47
	SIPP-EHC	0.18	0.19	11.17
	<i>t</i> -statistic	7.98	10.09	5.29
CY 2011	SIPP	0.23	0.28	10.63
	SIPP-EHC	0.16	0.17	11.55
	<i>t</i> -statistic	7.79	9.96	7.28
Diff-in-diff <i>t</i> -statistic		0.61	0.32	1.22
Observations		218,085	14,203	3,356

The unit of observation for the monthly subsidy column is person-months. The unit of observation for the annual subsidy and months of subsidy columns is person-years.

Table 11.2: Total number of housing subsidy spells for 2010–2011

	Spells
SIPP	1.08
SIPP-EHC	1.00
<i>t</i> -statistic	11.20
Persons	3,177

In CY2011, there is some evidence of reverse telescoping (see Chapter 2) in the pattern of month-by-month housing subsidy receipt rates in SIPP-EHC (see Table 11.3). The difference between SIPP-EHC and SIPP in the rate of reported housing subsidy receipt is slightly lower at the end of the calendar year (September through December) than at the beginning of the calendar year (January through April) (see the value for Test 2 in the CY2011 *t*-stat column in Table 11.3).

The proportion of the full sample in-universe for the housing subsidy questions is greater in SIPP-EHC than in SIPP (see Table 11.4). This is consistent with the finding that in SIPP-EHC compared to SIPP there are more sample-persons that live in renter-occupied units. In CY2010 and CY2011, the rate of item-nonresponse about housing subsidy receipt is slightly higher in SIPP-EHC than in SIPP.

Table 11.3: Month-by-month housing subsidies

	CY 2010			CY 2011			Diff
	SIPP	EHC	<i>t</i> -stat	SIPP	EHC	<i>t</i> -stat	
Jan	0.24	0.17	8.04	0.23	0.16	7.50	0.04
Feb	0.24	0.17	8.11	0.23	0.16	7.98	0.50
Mar	0.23	0.17	7.25	0.23	0.16	8.13	1.59
Apr	0.23	0.18	7.04	0.23	0.16	7.88	1.47
May	0.24	0.18	7.32	0.24	0.16	8.53	1.93
Jun	0.24	0.18	7.84	0.24	0.16	8.30	1.18
Jul	0.24	0.18	7.90	0.23	0.16	7.63	0.40
Aug	0.24	0.17	8.17	0.23	0.16	7.45	0.05
Sep	0.24	0.17	7.88	0.23	0.16	7.06	0.19
Oct	0.23	0.17	7.29	0.22	0.16	6.43	0.26
Nov	0.23	0.18	7.05	0.22	0.16	6.33	0.20
Dec	0.23	0.18	6.73	0.22	0.16	6.26	0.08
Pooled	0.24	0.18	7.98	0.23	0.16	7.79	0.61
Test 1	1.30	0.34	1.32	0.94	1.61	1.52	0.11
Test 2	0.35	0.70	0.62	1.88	2.21	2.62	1.28

Table 11.4: Housing subsidy item-nonresponse rates

	NIU	Housing subsidy
CY 2010 SIPP	0.39	0.03
CY 2010 SIPP-EHC	0.37	0.04
<i>t</i> -statistic	3.26	3.70
CY 2011 SIPP	0.39	0.03
CY 2011 SIPP-EHC	0.37	0.04
<i>t</i> -statistic	3.27	2.61
Diff-in-diff <i>t</i> -statistic	0.57	0.62
Person-months	369,202	225,604

12. Medicaid

This chapter presents comparisons of reporting about Medicaid coverage between SIPP-EHC and SIPP, including comparison between the surveys on agreement with administrative data. The following section describes the Medicaid data that were employed in the analysis of Medicaid reporting that is discussed in the subsequent section. Note that in the context of this section, Medicaid refers to any means-tested public health insurance program, including Medicaid, State Children's Health Insurance Program (SCHIP), and other programs.

Description of Medicaid data

The data sources for the tables in this chapter are 2008 SIPP, 2011 SIPP-EHC, 2012 SIPP-EHC, and administrative data for Medicaid (Medicaid Statistical Information System, MSIS). See Chapter 2 for discussion of the SIPP-EHC and SIPP samples and weighting. This section discusses the creation of two person-month level indicators of Medicaid coverage, one from the survey data and one from the administrative data.

All person-months are in-universe for the Medicaid coverage indicator variables. The Medicaid coverage indicator created from the survey (or administrative) data is coded as missing for person-months for which no survey report (or administrative record) of Medicaid coverage status is available.

For 2008 SIPP, these variables are created from unedited person-month data. For persons aged 15 or older, information about Medicaid coverage is obtained from the sample person's own record. For persons aged 19 or younger, the information is merged from guardians' reports on behalf of a child. Persons aged 15 to 19 can have reports from both themselves and a guardian, in which case this data is reconciled to indicate coverage if either person reported coverage.

In the 2011 and 2012 SIPP-EHC, all individuals were asked if they had been covered by Medicaid at any time since the beginning of the reference year. Individuals under age 15 were asked the same questions as adults, except through a proxy respondent. If they answered in the affirmative, they were asked further questions to determine the start and end dates of the coverage. The EHC had an additional question on coverage by "other" health insurance types. If a respondent indicated coverage here, they were asked about months and type of coverage. If respondents chose the coverage type choice of "Government/public assistance," then these coverage reports were treated as Medicaid. Using these unedited responses, monthly Medicaid coverage indicators were created for CY2010 and CY2011.

The administrative data version of this variable is created from the MSIS, a database of Medicaid enrollment records collected by the states and compiled by the Centers for Medicare and Medicaid Services. The data include records submitted from the first quarter of 2010 to the fourth quarter of 2011.

Person-month indicators of monthly administrative Medicaid enrollment are created to indicate for each month whether, according to the MSIS data, the sample-person was enrolled in Medicaid or SCHIP for at least one day of the month. In order to capture only comprehensive Medicaid or other public coverage, certain categories of coverage are excluded. As a result, observations are included where the individual was eligible for full Medicaid or SCHIP benefits, was eligible for pregnancy related services, was eligible for Medicaid benefits under the Deficit Reduction Act of 2005, was eligible based on the Money Follows the Person Demonstration, or was eligible under the Psychiatric Residential Treatment Facilities Demonstration Grant Program. The Medicaid variables exclude those eligible for restricted benefits due to alien status, dual eligibility for Medicare, or other reasons. It also excludes those with benefit restrictions as well as

those eligible for Medicaid benefits using Health Opportunity Accounts and those with only SCHIP dental coverage. Additionally, those whose SCHIP status is not known are excluded.

The administrative data have a number of limitations. Quarterly data are submitted by states over time. Not all states have submitted information for all quarters of 2010, and many states have not submitted any information for 2011. Those states that have submitted information may often retrospectively add cases and correct errors, and the data used here will not capture those additions and changes. Additionally, not all states submit information on SCHIP enrollment, and SCHIP is included in our SIPP and SIPP-EHC measures of Medicaid. MSIS also does not include other public programs, which may be captured by the SIPP and SIPP-EHC questions. Another possible limitation is that not all administrative data observations have identifying variables that allow them to be matched with SIPP-EHC and SIPP data; in the 2009 data 8 percent of the observations had missing or invalid social security numbers.

Discussion of Medicaid results

This section presents and discusses the analysis of Medicaid reporting in SIPP-EHC. The table elements and weighting are described in Chapter 2, and, where appropriate, in table notes.

Medicaid coverage is documented to be under-reported within surveys similar to the SIPP, and thus increased capture of Medicaid coverage is desirable. For monthly Medicaid coverage, the SIPP-EHC captures higher rates of Medicaid coverage than SIPP in CY2011 while for CY2010 the rates between SIPP and SIPP-EHC do not differ statistically (see Table 12.1). Across SIPP and SIPP-EHC, the mean absolute deviation between survey-reported and administratively recorded rates of monthly Medicaid does not differ statistically for CY2010 while it is lower for SIPP-EHC in CY2011.

Table 12.1: Medicaid monthly participation

		Survey	Linked	AR	MAD
CY 2010	SIPP	0.33	0.32	0.33	0.14
	SIPP-EHC	0.33	0.34	0.32	0.14
	<i>t</i> -statistic	0.28	2.05	1.69	0.51
CY 2011	SIPP	0.32	0.33	0.33	0.16
	SIPP-EHC	0.35	0.38	0.33	0.13
	<i>t</i> -statistic	3.92	4.95	0.02	3.56
Diff-in-diff <i>t</i> -statistic		4.55	4.21	1.73	3.27
Person-months		349,336	241,817	241,817	241,817

Between SIPP-EHC and SIPP, the rate of disagreement between survey and administrative data about annual Medicaid participation does not differ statistically (see Table 12.2). There is no statistically significant difference between SIPP-EHC and SIPP in either CY2010 or CY2011 in the mean absolute deviation between the survey-reported and administratively recorded annual Medicaid participation for the sample person-months that link to the administrative records. Differences in the rate of survey-reported annual

Medicaid participation between SIPP-EHC and SIPP appear to be due to sample composition differences (see Chapter 2).

Table 12.2: Medicaid participation

		Annual participation				Months of participation			
		Survey	Linked	AR	MAD	Survey	Linked	AR	MAD
CY 2010	SIPP	0.42	0.41	0.36	0.13	9.04	9.69	10.94	2.04
	SIPP-EHC	0.34	0.35	0.36	0.14	11.49	11.47	10.82	1.05
	<i>t</i> -statistic	9.08	6.31	0.27	0.91	30.50	18.63	1.25	10.28
CY 2011	SIPP	0.41	0.43	0.37	0.15	9.21	10.13	11.12	1.86
	SIPP-EHC	0.38	0.42	0.39	0.14	11.27	11.53	11.10	0.97
	<i>t</i> -statistic	3.36	0.43	1.23	1.02	22.73	11.39	0.12	6.69
Diff-in-diff <i>t</i> -statistic		5.77	4.36	1.66	1.60	3.68	2.63	0.70	0.67
Person-years		24,226	14,813	14,813	14,813	9,336	4,565	4,565	4,565

In SIPP-EHC compared to SIPP, there is less disagreement between survey and administrative data about the number of months of Medicaid participation (see Table 12.2). The mean absolute deviation between the survey-reported and administratively recorded number of participation months is lower for SIPP-EHC than for SIPP in both calendar years.

In SIPP-EHC compared to SIPP, there is less disagreement between survey and administrative data about the number of Medicaid spells per person (see Table 12.3). The mean absolute deviation between the survey and administrative data is lower for SIPP-EHC than for SIPP, and the difference is statistically significant for CY2010 through CY2011.

Table 12.3: Medicaid spells

	Survey	Linked	AR	MAD
SIPP	1.25	1.25	1.08	0.26
SIPP-EHC	1.01	1.01	1.11	0.11
<i>t</i> -statistic	34.37	26.08	2.67	12.20
Persons	8,327	5,063	5,063	5,063

Based on the pooled person-month observations, the rate of false negative reporting about Medicaid coverage is lower in SIPP-EHC than in SIPP (see the row labeled "Pooled" in Table 12.4). The rate of false negative

reporting about Medicaid coverage in SIPP-EHC decreased relative to SIPP across the calendar years. The rate of false positive reporting about Medicaid coverage in CY2011 is not statistically different between the two surveys (see the row labeled "Pooled" in Table 12.4). This represents an improvement in SIPP-EHC relative to SIPP across the calendar years.

For CY2011, neither reverse telescoping nor straight-lining (see Chapter 2) appear to be a feature of reporting of Medicaid coverage in SIPP-EHC. There is no statistically significant difference between the false negative rate for September through December of CY2011 and 1) the false negative rate in January of CY2011 (Test 1), and 2) the false negative rate for January through April of CY2011 (Test 2). The pattern of false negative rates in SIPP-EHC reflects a decrease in apparent reverse telescoping in SIPP-EHC relative to SIPP across the calendar years.

There is some evidence of reverse telescoping or straight-lining in SIPP-EHC relative to SIPP in the patterns of false positive reporting about Medicaid coverage. In both years, the decline in false negative reporting over the course of the calendar year is greater in SIPP-EHC than for SIPP. The differences between these differences across the surveys are statistically significant.

Table 12.4: Medicaid reporting errors

	False negatives						Diff	False positives						Diff
	CY 2010			CY 2011				CY 2010			CY 2011			
	SIPP	EHC	t-stat	SIPP	EHC	t-stat		SIPP	EHC	t-stat	SIPP	EHC	t-stat	
Jan	0.22	0.19	2.92	0.25	0.12	8.58	5.20	0.17	0.21	3.32	0.14	0.14	0.45	2.60
Feb	0.22	0.18	3.27	0.25	0.13	8.39	4.84	0.17	0.21	3.58	0.14	0.14	0.05	3.13
Mar	0.21	0.18	2.96	0.25	0.12	8.72	5.47	0.17	0.21	3.56	0.15	0.14	0.09	3.20
Apr	0.22	0.18	3.26	0.26	0.13	8.31	4.87	0.18	0.20	2.22	0.15	0.14	0.79	2.64
May	0.21	0.17	2.97	0.24	0.13	7.97	4.78	0.17	0.21	3.60	0.15	0.14	0.27	3.39
Jun	0.21	0.17	3.73	0.24	0.12	7.86	4.08	0.17	0.20	2.99	0.15	0.14	0.08	2.69
Jul	0.22	0.17	4.38	0.23	0.12	7.52	3.44	0.17	0.19	2.36	0.14	0.14	0.23	2.26
Aug	0.22	0.16	4.51	0.23	0.12	7.49	3.39	0.17	0.18	1.54	0.14	0.14	0.34	1.64
Sep	0.23	0.16	5.64	0.23	0.12	7.55	2.61	0.18	0.19	0.66	0.15	0.14	0.92	1.39
Oct	0.24	0.17	6.04	0.22	0.11	6.48	1.55	0.17	0.17	0.19	0.14	0.11	2.35	2.18
Nov	0.24	0.16	5.86	0.22	0.12	6.12	1.42	0.16	0.17	0.69	0.13	0.11	1.95	2.27
Dec	0.24	0.16	6.42	0.22	0.13	5.60	0.63	0.16	0.17	0.96	0.13	0.12	0.97	1.69
Pooled	0.22	0.17	5.20	0.24	0.12	9.28	4.74	0.17	0.19	2.48	0.14	0.14	0.71	3.07
Test 1	1.72	3.46	3.35	2.32	0.63	1.42	2.98	0.42	5.31	3.21	0.19	3.84	2.43	0.75
Test 2	2.33	2.88	3.56	2.90	1.11	1.63	3.37	0.63	5.69	3.37	1.14	4.21	1.95	1.17

All person-months are in universe for analysis of reporting of Medicaid coverage. For the monthly Medicaid coverage indicator, the SIPP-EHC has a higher percentage of missing data than SIPP in CY2010 and CY2011. There is a decrease in the difference of the item non-response rates between SIPP-EHC and SIPP across the calendar years (see Table 12.5).

Table 12.5: Medicaid item-nonresponse rates

		NIU	Participation
CY 2010	SIPP	0.00	0.03
	SIPP-EHC	0.00	0.11
	<i>t</i> -statistic	0.00	19.84
CY 2011	SIPP	0.00	0.03
	SIPP-EHC	0.00	0.05
	<i>t</i> -statistic	0.00	3.94
Diff-in-diff <i>t</i> -statistic		0.00	13.55
Person-months		369,317	369,317

13. Medicare

This chapter presents comparisons of reporting about Medicare enrollment in SIPP-EHC, SIPP, and administrative records from Medicare. The following section describes the Medicare data that were employed in the analysis of Medicare reporting that is discussed in the subsequent section.

Description of Medicare data

The data sources for the tables in this chapter are 2008 SIPP, 2011 SIPP-EHC, 2012 SIPP-EHC, and administrative records from Medicare. See Chapter 2 for discussion of the SIPP-EHC and SIPP samples and weighting. This section discusses the creation of Medicare-specific analysis variables from administrative data.

From each data source, a person-month level indicator of Medicare coverage is created for all sample persons aged 15 or older. The Medicare coverage indicator is coded as missing for in-universe person-months for which no survey report or administrative record of Medicare coverage status is available.

From 2008 SIPP, the Medicare coverage indicator is created from unedited person-level data which includes no imputed data.

In the 2011 and 2012 SIPP-EHC, individuals were asked if they had enrolled in the Medicare program anytime since the beginning of the reference year. If they answered in the affirmative, they were asked further questions to determine enrollment start and end. Using these unedited responses, monthly Medicare coverage indicators are created for CY2010 and CY2011.

The administrative-data version of this variable was created from the Active Medicare Enrollment Database (MEDB) maintained by the Centers for Medicare and Medicaid Services. The MEDB records are issued annually and detail beginning and ending dates of enrollment. Medicare Part A enrollment dates are used to create the coverage indicator.

Discussion of Medicare results

This section presents and discusses the analysis of reporting of Medicare in SIPP-EHC. The table elements are described in Chapter 2, and, where appropriate, in table notes.

The rate of disagreement between survey and administrative data about monthly Medicare coverage does not differ statistically across SIPP-EHC and SIPP (see Table 13.1). In the two surveys, the mean absolute deviation of the survey-reported from the administratively recorded monthly Medicare coverage for the sample person-months that link to the administrative records is not statistically different for either CY2010 or CY2011.

The rate of disagreement between survey and administrative data about annual Medicare coverage does not differ statistically between SIPP-EHC and SIPP (see Table 13.2). The mean absolute deviation between survey-reported versus administratively recorded annual coverage is not statistically different between SIPP-EHC and SIPP in either CY2010 or CY2011.

In SIPP-EHC compared to SIPP, there is less disagreement between survey and administrative data about the number of months of Medicare coverage (see Table 13.2). The mean absolute deviation between the survey-reported and administratively recorded months of coverage is lower for SIPP-EHC than for SIPP in both years. SIPP-EHC appears to perform better than SIPP at capturing the full number Medicare-covered months recorded in administrative records during a calendar year.

Table 13.1: Medicare monthly participation

		Survey	Linked	AR	MAD
CY 2010	SIPP	0.17	0.19	0.20	0.04
	SIPP-EHC	0.15	0.17	0.18	0.04
	<i>t</i> -statistic	3.09	2.05	2.57	0.22
CY 2011	SIPP	0.17	0.19	0.21	0.04
	SIPP-EHC	0.17	0.20	0.19	0.04
	<i>t</i> -statistic	0.08	1.08	1.44	0.92
Diff-in-diff <i>t</i> -statistic		3.99	4.38	1.47	0.71
Person-months		270,527	222,204	222,204	222,204

In SIPP-EHC compared to SIPP, there is less disagreement between survey and administrative data about the number of spells of Medicare coverage for CY2010 through CY2011 (see Table 13.3). The mean absolute deviation between the survey-reported and administratively recorded number of spells of Medicare receipt is lower for SIPP-EHC than for SIPP.

Based on the pooled person-month observations, the rate of false negative reporting about Medicare coverage is lower in SIPP-EHC than in SIPP in CY2011 but it is not statistically different between the two surveys in CY2010 (see Table 13.4). There is no evidence of reverse telescoping in SIPP-EHC (see Chapter 2) in the pattern of month-by-month false negative rates for Medicare reporting.

There is no statistical difference in the rates of false negative reporting between January and the pooled coverage rates for September through December (Test 1), and January through April and September through April (Test 2) in CY2010 and CY2011 for both surveys. The rate of false positive reporting about Medicare coverage is higher in SIPP-EHC than in SIPP in CY2011 but it is not statistically different between the two surveys in CY2010.

There is some evidence of reverse telescoping in the pattern of month-by-month false positive rates for Medicare reporting in SIPP-EHC. The rate of false positive reporting for September through December is lower than the rate for January (Test 1) in both CY2010 and CY2011 and is lower than the rate for January through April (Test 2) in CY2011.

The proportion of person-months in-universe for the Medicare coverage variable was higher in SIPP-EHC than SIPP in CY2010 but not statistically different between the two surveys in CY2011. The rate of item-nonresponse regarding Medicare coverage is higher in SIPP-EHC than in SIPP in CY2010 while the corresponding rate does not differ statistically between the two surveys in CY2011 (see Table 13.5). From CY2010 to CY2011, the rate of item-nonresponse about Medicare coverage in SIPP-EHC decreased relative to SIPP.

Table 13.2: Medicare participation

		Annual participation				Months of participation			
		Survey	Linked	AR	MAD	Survey	Linked	AR	MAD
CY 2010	SIPP	0.22	0.24	0.22	0.04	10.12	10.76	11.83	1.15
	SIPP-EHC	0.15	0.17	0.18	0.04	11.70	11.74	11.80	0.26
	<i>t</i> -statistic	9.08	7.73	4.54	0.29	14.21	9.43	0.59	9.25
CY 2011	SIPP	0.22	0.24	0.23	0.04	10.35	10.88	11.81	1.11
	SIPP-EHC	0.18	0.21	0.20	0.04	11.61	11.73	11.77	0.38
	<i>t</i> -statistic	4.70	3.43	2.89	0.30	10.86	8.18	0.67	6.55
Diff-in-diff <i>t</i> -statistic		5.01	4.91	1.87	0.03	2.21	0.91	0.13	1.21
Person-years		19,026	15,576	15,576	15,576	3,796	3,112	3,112	3,112

Table 13.3: Medicare spells

	Survey	Linked	AR	MAD
SIPP	1.19	1.19	1.00	0.19
SIPP-EHC	1.00	1.00	1.00	0.00
<i>t</i> -statistic	18.16	15.80	1.00	15.75
Persons	2,788	2,211	2,211	2,211

Table 13.4: Medicare reporting errors

	False negatives							False positives						
	CY 2010			CY 2011			Diff	CY 2010			CY 2011			Diff
	SIPP	EHC	<i>t</i> -stat	SIPP	EHC	<i>t</i> -stat		SIPP	EHC	<i>t</i> -stat	SIPP	EHC	<i>t</i> -stat	
Jan	0.13	0.13	0.22	0.14	0.07	4.61	3.74	0.08	0.10	1.26	0.07	0.11	2.91	1.53
Feb	0.14	0.13	0.60	0.14	0.09	3.67	2.65	0.08	0.09	0.99	0.08	0.11	2.36	1.33
Mar	0.14	0.13	0.50	0.15	0.08	4.27	3.21	0.07	0.09	1.67	0.07	0.11	2.45	0.87
Apr	0.13	0.13	0.23	0.15	0.08	4.51	3.56	0.07	0.09	1.74	0.07	0.11	2.49	0.82
May	0.12	0.13	0.34	0.15	0.08	4.21	3.77	0.07	0.09	1.29	0.07	0.11	2.72	1.40
Jun	0.12	0.13	0.64	0.14	0.08	3.68	3.60	0.07	0.09	1.54	0.07	0.11	2.86	1.40
Jul	0.12	0.13	0.42	0.13	0.08	2.69	2.56	0.07	0.09	1.38	0.07	0.10	2.68	1.34
Aug	0.12	0.13	0.27	0.13	0.09	2.97	2.70	0.07	0.08	0.83	0.06	0.10	2.53	1.65
Sep	0.13	0.12	0.56	0.13	0.09	3.09	2.09	0.07	0.08	0.90	0.06	0.10	2.27	1.33
Oct	0.14	0.12	0.79	0.14	0.08	3.48	2.24	0.08	0.08	0.62	0.07	0.09	1.82	1.13
Nov	0.14	0.12	0.88	0.14	0.08	3.61	2.29	0.08	0.08	0.60	0.07	0.09	1.97	1.26
Dec	0.14	0.12	1.11	0.13	0.08	3.01	1.56	0.08	0.08	0.44	0.07	0.09	1.86	1.28
Pooled	0.13	0.13	0.30	0.14	0.08	4.14	3.44	0.07	0.09	1.22	0.07	0.10	2.60	1.46
Test 1	0.40	0.95	0.81	0.70	1.52	1.43	1.47	0.44	2.12	0.92	0.04	2.39	1.49	0.38
Test 2	0.03	1.31	0.63	1.21	0.69	1.39	1.34	0.15	1.83	1.23	0.70	3.11	1.24	0.05

Because the SIPP was administered 3 times a year, there was a risk that reporting error across the waves could create false transitions in coverage status. The comparison of survey to administrative data in Table 13.3 suggests that SIPP captured gaps in coverage not present in the administrative records. In SIPP-EHC compared to SIPP, there is less disagreement between survey and administrative data about the number of spells.

Table 13.5: Medicare item-nonresponse rates

	NIU	Participation
CY 2010		
SIPP	0.26	0.01
SIPP-EHC	0.23	0.09
<i>t</i> -statistic	3.85	20.43
CY 2011		
SIPP	0.25	0.01
SIPP-EHC	0.24	0.01
<i>t</i> -statistic	1.22	2.85
Diff-in-diff <i>t</i> -statistic	3.40	20.34
Person-months	369,202	278,795

14. Nativity and Citizenship

This chapter presents comparisons of reporting about nativity and citizenship between SIPP-EHC and SIPP. The following section describes the nativity and citizenship data that were employed in the analysis discussed in the subsequent section.

Description of nativity and citizenship data

The data sources for the tables in this chapter are the 2008 SIPP, 2011 SIPP-EHC, and 2012 SIPP-EHC. See Chapter 2 for discussion of the SIPP-EHC and SIPP samples and weighting. This section discusses the creation of the analysis variables specific to naturalization and citizenship.

The variables discussed in this section were created from three survey sources (2008 SIPP, 2011 SIPP-EHC, and 2012 SIPP-EHC). For each of the survey sources, nativity and citizenship variables were created at the person-year level for persons not missing due to non-interview in December of the given year. From SIPP, the variables were created from edited data, dropping all imputations. The variables are coded as missing for person records that are in universe but have no available report.

The original nativity and citizenship variables in all three SIPP sources include a person-wave report of whether the respondent was born in the U.S. and a follow-up person-level report of U.S. citizenship status and citizenship type. These questions are used to create variables for nativity, citizenship, and naturalization.

The analysis in this chapter is based on two variables constructed from the original data of nativity and citizenship from the survey sources. The nativity and citizenship variable is a person-year report of whether the respondent was a (1) U.S. born citizen, (2) foreign born (naturalized) citizen, or (3) foreign born non-citizen. Respondents born in the U.S. states or a U.S. Island Area territory and those born outside of the U.S. but to U.S. citizen parents are coded as U.S. born citizens. All respondents are in universe for this variable.

The second variable is a recoded indicator identifying how foreign born respondents became a citizen for the universe of only respondents who are naturalized citizens. The variable includes the following types of naturalization: (1) general, (2) military service, or (3) international adoption.

Discussion of nativity and citizenship results

This section presents and discusses the analysis of reporting in SIPP-EHC of nativity and citizenship as well as naturalization type. The table elements and weighting are described in Chapter 2, and, where appropriate, in table notes.

The proportion of sample-persons reported to be U.S. born citizens is lower in SIPP-EHC than in SIPP across both CY2010 and CY2011 (see Table 14.1). The SIPP-EHC and SIPP proportions of sample-persons reported to be foreign born citizens do not differ statistically in CY2010 while the proportion is higher in SIPP-EHC than in SIPP in CY2011. There is a higher rate of reported non-citizenship in SIPP-EHC than in SIPP in CY2010 only. The rates reported of naturalization do not differ statistically between SIPP-EHC and SIPP (see Table 14.2).

The not-in-universe rates are notably higher in SIPP than in SIPP-EHC (see Table 14.3). This is largely because only the cases with nativity and citizenship SIPP data from December of each calendar year are included. Respondents who were not present at the interview address in December were coded out of universe, even if they were present during another month of the calendar year.

Table 14.1: U.S. nativity and citizenship

		U.S. born citizen	Foreign born citizen	Foreign born noncitizen
CY 2010	SIPP	0.75	0.11	0.14
	SIPP-EHC	0.72	0.10	0.18
	<i>t</i> -statistic	4.67	0.84	6.28
CY 2011	SIPP	0.75	0.10	0.15
	SIPP-EHC	0.72	0.12	0.15
	<i>t</i> -statistic	3.24	2.84	1.46
Diff-in-diff <i>t</i> -statistic		1.54	5.61	6.42
Person-years		30,104	30,104	30,104

Table 14.2: Type of U.S. naturalization

		General	Military service	Adoption
CY 2010	SIPP	0.99	0.01	0.00
	SIPP-EHC	0.99	0.00	0.01
	<i>t</i> -statistic	1.32	0.25	1.99
CY 2011	SIPP	0.99	0.01	0.00
	SIPP-EHC	0.99	0.01	0.00
	<i>t</i> -statistic	0.45	0.09	1.00
Diff-in-diff <i>t</i> -statistic		1.00	0.41	1.65
Person-years		2,958	2,958	2,958

Though the item-specific nonresponse rates are low in both SIPP and SIPP-EHC, the nonresponse rate for naturalization is higher in SIPP compared to SIPP-EHC in CY2011 (see Table 14.3).

Table 14.3: Nativity and citizenship item-nonresponse rates

		NIU	Nativity and citizenship	Naturalization
CY 2010	SIPP	0.19	0.01	0.03
	SIPP-EHC	0.00	0.01	0.03
	<i>t</i> -statistic	52.07	0.48	0.41
CY 2011	SIPP	0.20	0.01	0.04
	SIPP-EHC	0.00	0.01	0.02
	<i>t</i> -statistic	51.42	1.54	2.18
Diff-in-diff <i>t</i> -statistic		1.59	1.83	2.66
Person-years		35,084	30,370	3,051

15. Old-Age, Survivors, and Disability Insurance

This chapter presents comparisons of reporting about OASDI receipt in SIPP-EHC, SIPP, and OASDI administrative records. The following section describes the OASDI data that are employed in the analysis discussed in the subsequent section.

Description of OASDI data

The data sources for the tables in this chapter are 2008 SIPP, 2011 SIPP-EHC, 2012 SIPP-EHC, and administrative records for OASDI. See Chapter 2 for discussion of the SIPP-EHC and SIPP samples and weighting. This section discusses the creation of OASDI-specific analysis variables from the survey and administrative data.

Two sets of OASDI variables are created, one from the survey data and one from the administrative data. Each set includes a person-month level indicator for reported participation in OASDI, a person-month level report of OASDI benefit amount, and indicator variables for receipt of different types of OASDI benefits: retirement, disability, spousal, widow, child, and other.

All person-months are in-universe for the OASDI participation indicator variables. The OASDI participation indicator is coded as missing for person-months for which no survey report or administrative record of OASDI participation status is available. Person-months for which OASDI participation is indicated are in-universe for the benefit-amount measure and benefit-type indicators. The OASDI benefit-amount and benefit-type variables are coded as missing when they are in universe but there is no report or record available. These variables are also coded as missing when the OASDI participation indicator for the person-month and source is missing.

From 2008 SIPP, these variables are created from edited person-month data, dropping all imputations. For adults, the information about OASDI receipt is obtained from the sample person's own record. For children, the information is merged on from adults' reports of receipt on behalf of a child. Reports about OASDI receipt in 2008 SIPP identify up to two types of OASDI benefits in a given month.

In the 2011 and 2012 SIPP-EHC, individuals were asked if they had received any OASDI payments since the beginning of the reference year. If they answered in the affirmative, they were asked further questions to determine the start and end dates of the benefit receipt. Similar questions were asked about benefit amounts, collecting first the current amount received and then probing for changes in amounts and the dates those changes happened. Using these unedited responses, monthly OASDI participation indicators and amounts are created for CY2010 and CY2011. The survey reports also include up to two reasons for receiving the OASDI benefit which are used to create monthly benefit type indicators.

The administrative-data versions of these variables are created from two Social Security Administration (SSA) databases that record benefit eligibility and payments: the Master Benefit Record (MBR) and the Payment History Update System (PHUS). The MBR records monthly eligibility and also records dates and reasons for the beginning and ending of eligibility. The MBR also indicates amounts that individuals were eligible to receive based on the calculation of the Primary Insurance Amount from a person's own work history or the work history of his or her spouse.

It is possible for the MBR to record OASDI eligibility for an individual in a month in which the individual received no OASDI benefit payment. For this reason, SSA maintains the PHUS to record actual payments sent to recipients.

The OASDI participation indicator based on administrative data is created based on the PHUS. The OASDI benefit amount based on administrative data is also created from the PHUS¹.

The MBR data extract used by the Census Bureau records up to four different types of OASDI benefit for each OASDI beneficiary: type of own benefit at time of initial receipt, type of spouse benefit at time of initial receipt, current own benefit, and current spouse benefit. The administrative benefit types and the accompanying dates of initial and current receipt are used to create a type-of-benefit history for each survey reference year that allows up to two benefit types each month. For example, if a person began receiving disability benefits in 2000 and then switched to retirement benefits upon reaching full retirement age in June 2010, this person would be coded as having disability benefits from January 2010 to May 2010 and retirement benefits from June 2010 to December 2011. Any spouse benefits would be coded as a second type of benefit in the months in which they were received.

Discussion of OASDI results

This section presents and discusses the analysis of reporting of OASDI in SIPP-EHC. The table elements are described in Chapter 2, and, where appropriate, in table notes.

The rate of disagreement between survey and administrative data about monthly OASDI participation does not differ statistically across SIPP-EHC and SIPP (see Table 15.1). There is no statistically significant difference between SIPP-EHC and SIPP in either CY2010 or CY2011 in the mean absolute deviation between the survey-reported and administratively recorded monthly OASDI participation for the sample person-months that link to the administrative records. Differences in the rate of survey-reported monthly OASDI participation between SIPP-EHC and SIPP appear to be due to sample composition differences (see Chapter 2).

Table 15.1: OASDI monthly participation

		Survey	Linked	AR	MAD
CY 2010	SIPP	0.15	0.17	0.17	0.04
	SIPP-EHC	0.12	0.14	0.15	0.04
	<i>t</i> -statistic	6.09	5.38	2.85	0.08
CY 2011	SIPP	0.16	0.18	0.17	0.06
	SIPP-EHC	0.14	0.17	0.16	0.06
	<i>t</i> -statistic	2.32	1.20	0.03	0.01
Diff-in-diff <i>t</i> -statistic		4.40	5.08	4.16	0.06
Person-months		352,087	283,642	283,642	283,642

Between SIPP-EHC and SIPP, the rate of disagreement between survey and administrative data about annual OASDI participation does not differ statistically (see Table 15.2). There is no statistically significant difference

¹Medicare premium is excluded from these amounts as it is deducted from the check before it is mailed or direct deposited. Both SIPP and SIPP-EHC surveys collect Medicare premiums, but the amounts are collected separately from the main OASDI amounts reported.

between SIPP-EHC and SIPP in either CY2010 or CY2011 in the mean absolute deviation between the survey-reported and administratively recorded annual OASDI participation for the sample person-months that link to the administrative records. Differences in the rate of survey-reported annual OASDI participation between SIPP-EHC and SIPP appear to be due to sample composition differences (see Chapter 2).

Table 15.2: OASDI annual participation and months of participation

		Annual participation				Months of participation			
		Survey	Linked	AR	MAD	Survey	Linked	AR	MAD
CY 2010	SIPP	0.19	0.22	0.20	0.05	10.67	11.27	11.66	0.55
	SIPP-EHC	0.12	0.14	0.16	0.04	11.57	11.56	11.62	0.33
	<i>t</i> -statistic	10.82	9.75	5.11	0.88	8.61	2.78	0.56	2.45
CY 2011	SIPP	0.19	0.21	0.19	0.06	10.85	11.40	11.69	0.45
	SIPP-EHC	0.15	0.17	0.17	0.06	11.68	11.76	11.75	0.26
	<i>t</i> -statistic	6.21	4.99	1.80	0.90	8.52	4.23	0.74	2.42
Diff-in-diff <i>t</i> -statistic		4.99	5.14	4.26	0.19	0.55	0.64	0.93	0.22
Person-years		24,609	19,882	19,882	19,882	4,188	3,249	3,249	3,249

In SIPP-EHC compared to SIPP, there is less disagreement between survey and administrative data about the number of months of OASDI participation (see Table 15.2). The mean absolute deviation between the survey-reported and administratively recorded number of participation months is lower for SIPP-EHC than for SIPP in both calendar years.

Table 15.3: OASDI monthly benefits

		Survey	Linked	AR	MAD
CY 2010	SIPP	\$848.34	\$872.85	\$826.16	\$176.79
	SIPP-EHC	\$899.25	\$915.45	\$914.31	\$132.34
	<i>t</i> -statistic	1.26	1.16	2.79	1.41
CY 2011	SIPP	\$840.60	\$880.63	\$830.67	\$145.81
	SIPP-EHC	\$940.72	\$995.00	\$902.19	\$175.31
	<i>t</i> -statistic	2.78	2.53	2.28	0.80
Diff-in-diff <i>t</i> -statistic		1.03	1.42	0.51	1.54
Person-months		21,047	16,252	16,252	16,252

For monthly OASDI benefit amount, the rate of disagreement between survey and administrative data does not differ statistically between SIPP-EHC and SIPP (see Table 15.3). Between the two surveys, the mean absolute deviation between the survey-reported OASDI benefit amount and the administratively-recorded amount is not statistically different for either CY2010 or CY2011.

Table 15.4: OASDI benefit types

		Retirement				Disability			
		Survey	Linked	AR	MAD	Survey	Linked	AR	MAD
CY 2010	SIPP	0.61	0.66	0.59	0.16	0.26	0.25	0.19	0.11
	SIPP-EHC	0.64	0.67	0.65	0.16	0.27	0.24	0.20	0.09
	<i>t</i> -statistic	1.29	0.20	2.49	0.30	0.39	0.39	0.45	1.31
CY 2011	SIPP	0.63	0.68	0.65	0.15	0.26	0.25	0.19	0.11
	SIPP-EHC	0.65	0.71	0.70	0.14	0.29	0.26	0.20	0.12
	<i>t</i> -statistic	1.05	1.10	2.00	0.07	1.43	0.47	0.40	0.79
Diff-in-diff <i>t</i> -statistic		0.21	1.05	0.51	0.20	1.19	0.98	0.02	1.89
Person-months		50,781	40,233	40,233	40,233	50,781	40,233	40,233	40,233

For CY2011, the rate of disagreement between survey and administrative data about whether retirement (disability) benefits are part of the reported/recorded OASDI benefits in a given person-month does not differ statistically between SIPP-EHC and SIPP (see Table 15.4). Between the two surveys, the mean absolute deviation between the survey-reported and administratively recorded receipt of OASDI benefits is not statistically different for either calendar year.

Table 15.5: OASDI participation spells

	Survey	Linked	AR	MAD
SIPP	1.08	1.06	1.02	0.06
SIPP-EHC	1.01	1.01	1.03	0.04
<i>t</i> -statistic	9.45	6.64	0.97	2.66
Persons	3,131	2,326	2,326	2,326

There is less disagreement in SIPP-EHC and SIPP between survey and administrative data about the number of spells of OASDI participation in CY2010 through CY2011 (see Table 15.5). The mean absolute deviation between the survey-reported and administratively recorded number of spells of OASDI receipt is lower for SIPP-EHC than for SIPP.

Table 15.6: OASDI reporting errors

	False negatives							False positives						
	CY 2010			CY 2011			Diff	CY 2010			CY 2011			Diff
	SIPP	EHC	<i>t</i> -stat	SIPP	EHC	<i>t</i> -stat		SIPP	EHC	<i>t</i> -stat	SIPP	EHC	<i>t</i> -stat	
Jan	0.11	0.18	3.89	0.14	0.16	1.10	2.18	0.12	0.10	1.16	0.17	0.16	0.79	0.14
Feb	0.12	0.18	3.26	0.14	0.16	1.31	1.45	0.12	0.10	1.39	0.17	0.16	0.89	0.23
Mar	0.12	0.18	3.57	0.13	0.16	1.59	1.50	0.12	0.10	1.56	0.17	0.15	1.20	0.08
Apr	0.13	0.18	3.09	0.13	0.16	1.73	1.00	0.12	0.09	2.01	0.18	0.15	1.33	0.31
May	0.13	0.19	3.25	0.12	0.16	2.22	0.72	0.12	0.10	1.97	0.18	0.15	1.83	0.23
Jun	0.13	0.19	3.77	0.12	0.17	2.39	0.98	0.13	0.10	1.93	0.18	0.15	1.81	0.23
Jul	0.13	0.19	3.62	0.13	0.16	2.04	1.15	0.12	0.10	1.15	0.17	0.15	1.17	0.24
Aug	0.12	0.18	3.62	0.13	0.16	1.66	1.45	0.12	0.10	1.69	0.17	0.15	1.20	0.18
Sep	0.12	0.18	3.27	0.14	0.16	1.12	1.69	0.12	0.10	1.71	0.17	0.14	1.53	0.11
Oct	0.13	0.18	3.02	0.14	0.16	0.98	1.62	0.12	0.09	1.82	0.17	0.14	1.66	0.20
Nov	0.12	0.17	3.00	0.14	0.16	0.87	1.69	0.12	0.09	2.30	0.17	0.14	1.69	0.17
Dec	0.13	0.18	2.91	0.14	0.16	1.13	1.36	0.13	0.09	2.96	0.18	0.14	1.91	0.46
Pooled	0.13	0.18	3.60	0.13	0.16	1.63	1.56	0.12	0.10	1.95	0.17	0.15	1.50	0.05
Test 1	1.42	0.28	1.28	0.63	0.61	0.11	0.74	0.75	1.81	1.49	0.07	2.28	1.46	0.08
Test 2	0.36	0.85	0.77	0.84	0.06	0.69	0.01	0.52	1.39	1.07	0.15	2.48	1.25	0.16

See Chapter 2 for definitions of false negative and false positive reporting. The monthly *t*-statistics are from tests for equality of the rate of false positives (or false negatives) between the SIPP and the SIPP-EHC. Test 1 presents the *t*-statistic from a test of equality of the error rate for January and the error rate for September through December. Test 2 presents the *t*-statistic from a test of equality of the error rate for January through April and the error rate for September through December.

Based on the pooled person-month observations, the rate of false negative reporting about OASDI participation is higher in SIPP-EHC than in SIPP in CY2010 but not statistically different between the two surveys in CY2011 (see the row labeled “pooled” in Table 15.6).

There is little evidence of reverse telescoping (see Chapter 2) in the month-by-month pattern of OASDI reporting errors in SIPP-EHC. In CY2011, the false positive rate is lower for September through December than for January (Test 1) or January through April (Test 2). There is no statistically significant difference in the false positive rates between the same two periods in CY2010. There is no evidence in CY2010 or CY2011 of declining false negative rates between the beginning and end of each year.

The rate of item-nonresponse about OASDI receipt is lower in SIPP than in SIPP-EHC in CY2010 (see Table 15.7). The difference in the rates between SIPP relative to SIPP-EHC decreased across the calendar years to a negligible but still statistically significant difference in CY2011. In both calendar years, the item-nonresponse rate for benefit type is higher in SIPP-EHC than in SIPP. For the benefit amount for participants, the item-nonresponse rate is lower in SIPP-EHC than in SIPP in CY2010 and CY2011.

Table 15.7: OASDI item-nonresponse

		NIU	Participation	Benefit type	Benefit for participants
CY 2010	SIPP	0.00	0.03	0.15	0.80
	SIPP-EHC	0.00	0.11	0.54	0.16
	<i>t</i> -statistic	0.00	21.23	26.49	42.64
CY 2011	SIPP	0.00	0.03	0.15	0.80
	SIPP-EHC	0.00	0.03	0.25	0.18
	<i>t</i> -statistic	0.00	2.12	5.97	37.97
Diff-in-diff <i>t</i> -statistic		0.00	17.35	16.56	1.23
Person-months		369,202	369,202	69,264	52,149

See "Description of OASDI data" section for details of classification of data as out-of-universe and missing.

16. Poverty

This chapter presents comparisons of family income and poverty status between SIPP-EHC and SIPP. The following section describes the assignment of individuals into family units, as well as the process for summing income and assigning poverty status within these family units.

Description of income and poverty data

The data sources for the tables in this chapter are the 2008 SIPP, 2011 SIPP-EHC and 2012 SIPP-EHC. See Chapter 2 for discussion of the SIPP-EHC and SIPP samples and weighting.

Since poverty status is calculated at a family level based on the comparison of a family's total income to its poverty threshold, this chapter covers 1) the assignment of family units, 2) the calculation of total family income, and 3) the assignment of poverty status across the 2008 SIPP, 2011 SIPP-EHC and 2012 SIPP-EHC samples.

In the 2011 and 2012 SIPP-EHC, data is collected on the relationship of each individual to the head of household. This includes those in the household during the interview month as well as any individuals who were in the household at any point in the reference period who have some missing data. In the editing process, this data is used to create monthly primary- and sub-family assignments within a household, taking into account when people transition into and out of the household and how their relationships change over time.

This assignment of individuals into family units is modified in this analysis since there is no edited monthly data on household composition or on how the relationship of individuals in the household to the householder change over time. Given this limitation in both the 2011 and 2012 SIPP-EHC, this analysis focuses on household and family composition in December of the reference year, as this is the closest period to the interview date, providing increased confidence that family composition and relationships are consistent with that reported in the interview month. In addition, this analysis is limited to primary families, defined as individuals who are directly related to the householder as a spouse, child, grandchild, parent, sibling, or other relative. Individuals in primary families in the month of December constitute approximately 95 percent of individuals in the 2008 SIPP, 2011 SIPP-EHC, and 2012 SIPP-EHC samples. Individuals who are not related to the household respondent, be they foster children, roommates, boarders, or other non-relatives, are out of scope for this income and poverty analysis.

Once individuals are placed into primary family units for the month of December, monthly income from employment earnings; Supplemental Security Income; unemployment insurance; Old-Age, Survivors, and Disability Insurance; and Temporary Assistance for Needy Families are summed for everyone in the primary family. In order for individuals in a primary family to have valid total family income for a given month, each individual within the primary family must have valid, non-missing, non-imputed reports for each of the five income sources used in the total family income calculation. If individuals within a family did not meet this criteria, that family's earnings in December are coded as missing.

In order to evaluate whether total family earnings in the month of December place a family above or below the poverty line, annual poverty thresholds for each primary family are calculated based on the size of the family, the age of the householder, and the number of children in the family. These annual thresholds are then divided by 12 and adjusted for monthly CPI rates to assign December poverty rates for the respective sample reference year. If an individual was in a primary family but the householder did not report an age,

the poverty threshold was not assigned, and each individual in that family was assigned a missing value for poverty status.

Poverty rates are calculated for all primary families with valid total family income and poverty thresholds for the month of December. If the family's total monthly income in December was below its poverty threshold for that month, that family, and all of the individuals within it, are in poverty for the month of December in that respective calendar year.

Discussion of income and poverty results

This section presents and discusses the analysis of income and poverty status in the SIPP-EHC. The table elements and weighting are described in Chapter 2, and, where appropriate, in table notes.

There is no statistical difference in poverty rates between the SIPP-EHC and SIPP surveys in either the 2010 or 2011 calendar years (see Table 16.1). However, for family income, there are significant differences in CY2011, with total family income lower in the SIPP-EHC than in the SIPP. Changes in income differences across survey designs across calendar years are not statistically significant.

Table 16.1: Poverty status and family income

		Poverty	Family income
CY 2010	SIPP	0.38	\$2,752.15
	SIPP-EHC	0.37	\$11,058.02
	<i>t</i> -statistic	1.34	1.06
CY 2011	SIPP	0.40	\$2,826.17
	SIPP-EHC	0.41	\$2,473.39
	<i>t</i> -statistic	0.57	3.00
Diff-in-diff <i>t</i> -statistic		1.57	1.10
Observations		14,743	6,308

The unit of observation for the Poverty column is person-years.

The unit of observation for the Family income column is family-years.

When looking at the distribution of earnings in the CY2010 SIPP-EHC, we find that reported family earnings are highly skewed to the right, and that median family income values are likely a more appropriate measure in this instance (See Chapter 8).

When looking at response rates as shown in Table 16.2, there are significant differences in the percent of the population not in universe for poverty, with lower rates in the SIPP-EHC than the SIPP. The population not in universe in the SIPP-EHC is composed entirely of individuals not in primary families. However, for the SIPP, this population includes both individuals not in primary families (consistent with the SIPP-EHC at around 5 percent of individuals) or not in households that were in the SIPP survey at some point in the reference period but not present in the month of December. Therefore, the majority of the difference is driven by overall differences in the data-collection schedule across the survey designs rather than changes

in the composition of families. See Chapter 2 for additional details on differences in data collection across samples.

Table 16.2: Poverty item non-response rates

		NIU (poverty)	Poverty	NIU (income)	Family income
CY 2010	SIPP	0.23	0.42	0.19	0.41
	SIPP-EHC	0.07	0.57	0.00	0.57
	<i>t</i> -statistic	33.66	18.06	50.68	19.17
CY 2011	SIPP	0.24	0.41	0.20	0.40
	SIPP-EHC	0.05	0.60	0.00	0.60
	<i>t</i> -statistic	37.60	21.27	50.48	21.84
Diff-in-diff <i>t</i> -statistic		4.24	4.30	1.99	3.93
Observations		35,084	28,713	33,215	28,713

The unit of observation for the Poverty columns is person-years.

The unit of observation for the Family income columns is family-years.

Data on family income nonresponse is reported in Table 16.2. Because this analysis is limited to primary families, there are no families that are out of universe in either the SIPP or SIPP-EHC.

The SIPP-EHC has higher rates of missing data for family income and poverty status in both CY2010 and CY2011. This is largely driven by differences in missing rates for employment earnings across the two survey designs, although there are a number of other statistical differences in missing rates across the two surveys for the income sources used in this analysis, as discussed in corresponding chapters. All individuals with missing family income are assigned missing values for family poverty, and individuals with valid family income who are in families with missing data on householder age are also assigned missing poverty status values.

17. Residence

This chapter presents comparisons of reporting about residence changes and residence tenure in SIPP-EHC and SIPP. The following section describes the residence data that were employed in the analysis of residence reporting that is discussed in the subsequent section.

Description of residence data

The data sources for the tables in this chapter are the 2008 SIPP, 2011 SIPP-EHC, and 2012 SIPP-EHC. See Chapter 2 for discussion of the SIPP-EHC and SIPP samples and weighting. This section discusses the creation of analysis variables specific to residence.

A person-month indicator of whether the respondent changed residences that month and a three-category person-month indicator of tenure status of residence. The tenure categories are owner-occupied unit, renter-occupied unit, and rent-free occupied unit.

All person-months are in universe for the residence change variable. For tenure status, all person-months are in universe except months where the respondent lived in group quarters or lived abroad. The residence change and tenure variables are coded as missing when a respondent is in universe but there is no survey report available.

The 2008 SIPP collected data on up to one move per person per four-month reference period. If a sample person had moved multiple times during the reference period, address information was only collected for the move to the interview address. Sample persons who moved between waves were asked to report the month of move. This information is used to create the residence change variable.

For 2008 SIPP the three-category monthly tenure variable is created from edited person-month data, dropping all imputations. Information on housing tenure of the current residence was obtained from the householder. This tenure information is attached to the records of all household members for all months during the interview reference period where they lived at the same address as the householder.

In the 2011 and 2012 SIPP-EHC, the beginning and end month of residence spells for each respondent were collected for up to five residences lived at for at least one month in the reference period. In 2011 SIPP-EHC, all respondents were asked to report the year and month of move to the address lived at during the first month (January) of the reference period. In 2012 SIPP-EHC, a wave 2 survey, only new respondents to the SIPP-EHC sample were asked this question. Restrictions were also set on following movers in the 2012 SIPP-EHC. Sample persons who moved to an address more than 50 miles away from any Field Representative's area between wave 1 and wave 2 were dropped from the sample. This decision may have artificially lowered the mover rates in the 2012 SIPP-EHC.

Collecting residence information for up to five addresses in the reference year represents a significant change from the 2008 SIPP. Identifying movers in SIPP-EHC rests primarily on respondent reports, and respondents may be asked to recall address information for a previous residence.

In the 2011 and 2012 SIPP-EHC, tenure status for the interview address was reported by the householder. This information was attached to the records of all household members for months during the interview reference period where they lived at the same address as the householder. Additionally, respondents were asked to report the tenure status of their unit for months they did not live with the householder.

Discussion of residence results

This section presents and discusses the analysis of reporting in SIPP-EHC of residence data. The table elements and weighting are described in Chapter 2, and, where appropriate, in table notes.

Rates of reporting living in an owner-occupied unit are lower, and reported rates of living in a renter-occupied unit are higher for SIPP-EHC than SIPP across both CY2010 and CY2011 (see Table 17.1). There are no differences in the reported rates of occupying a unit free of rent (“Other”) for the two surveys in CY2010 or CY2011.

Table 17.1: Housing tenure

		Owned	Rented	Other
CY 2010	SIPP	0.39	0.58	0.02
	SIPP-EHC	0.36	0.62	0.02
	<i>t</i> -statistic	3.95	4.24	1.11
CY 2011	SIPP	0.39	0.58	0.03
	SIPP-EHC	0.37	0.61	0.03
	<i>t</i> -statistic	3.22	3.19	0.00
Diff-in-diff <i>t</i> -statistic		0.42	0.86	1.04
Person-months		365,181	365,181	365,181

Rates of reported monthly and annual residence change in SIPP-EHC and SIPP differ for CY2010 and CY2011 (see Table 17.2). In CY2010, the rate of reported monthly and annual moves is slightly higher in SIPP-EHC compared to SIPP. For CY2011, the rate of monthly moves is slightly higher in SIPP compared to SIPP-EHC.

The annual number of moves per person is higher in SIPP-EHC than in SIPP for CY2010. The difference in annual number of moves between the two surveys is not statistically significant in CY2011.

Item non-response for the residence change and tenure variables, while low for both surveys, is statistically higher in SIPP-EHC compared to SIPP in both calendar years (see Table 17.3). From CY2010 to CY2011, the SIPP-EHC to SIPP differences in the not in universe rates were higher while the differences in the item non-response rates between the two surveys does not differ statistically across the two calendar years.

Table 17.2: Change of residence

		Monthly rate	Annual rate	Annual number of moves
CY 2010	SIPP	0.02	0.12	1.15
	SIPP-EHC	0.02	0.20	1.19
	<i>t</i> -statistic	3.97	11.68	2.11
CY 2011	SIPP	0.01	0.10	1.09
	SIPP-EHC	0.01	0.13	1.07
	<i>t</i> -statistic	2.01	5.92	1.40
Diff-in-diff <i>t</i> -statistic		4.59	4.41	2.63
Observations		366,135	25,975	3,616

The unit of observation for the Monthly rate is person-months. The unit of observation for the Annual rate and Annual number of moves columns is person-years.

Table 17.3: Residence item-nonresponse rates

		NIU (change)	Change	NIU (tenure)	Tenure
CY 2010	SIPP	0.00	0.00	0.00	0.00
	SIPP-EHC	0.00	0.02	0.01	0.02
	<i>t</i> -statistic	3.45	11.47	8.54	12.22
CY 2011	SIPP	0.00	0.00	0.00	0.00
	SIPP-EHC	0.01	0.02	0.01	0.02
	<i>t</i> -statistic	7.09	9.57	8.35	9.89
Diff-in-diff <i>t</i> -statistic		5.77	0.56	2.07	1.07
Person-months		369,202	368,683	369,202	367,977

18. Supplemental Nutrition Assistance Program

This chapter presents comparisons of reporting about SNAP receipt in SIPP-EHC, SIPP and SNAP administrative records. The following section describes the SNAP data that were employed in the analysis of SNAP reporting that is discussed in the subsequent section.

Description of SNAP data

The data sources for the tables in this chapter are 2008 SIPP, 2011 SIPP-EHC, 2012 SIPP-EHC, and New York State SNAP administrative records. See Chapter 2 for discussion of the SIPP-EHC and SIPP samples and weighting. This section discusses the creation of SNAP-specific analysis variables from the survey and administrative data.

Two sets of SNAP variables are created, one from the survey data and one from the administrative data. Each set includes a person-month-level indicator for reported participation in SNAP and a person-month-level report of SNAP benefit amount.

All person-months are in-universe for the SNAP participation indicator variables. The SNAP participation indicator is coded as missing for person-months for which no report (survey) or record (administrative data) of SNAP participation status was available. Person-months for which SNAP participation was indicated are in-universe for the benefit-amount measure. The SNAP benefit-amount variables are coded as missing when they are in universe but there is no report or record available; these variables are also coded as missing when the SNAP participation indicator for the person-month and source is missing.

From 2008 SIPP, these variables are created from edited person-month data, dropping all imputations. For all sample persons, the information about SNAP participation was obtained from the sample person's own record. SNAP benefit amounts data were collected at the SNAP unit-level and were obtained from the owner of the reported SNAP benefit.

In the 2011 and 2012 SIPP-EHC, adults were asked if they or someone in their family received any SNAP payments since the beginning of the reference year. If they answered in the affirmative, they were asked further questions to determine the start and end dates of the benefit receipt. Reported SNAP participation is attributed to additional specified household members. For each spell of receipt, questions were asked about the benefit amounts. Using these unedited responses, monthly SNAP participation indicators and amounts were created for CY2010 and CY2011.

The administrative-data versions of these variables are created from a database of monthly SNAP benefit payments in New York State. The database contains a record for each individual in the SNAP case unit, and individuals for whom a record exists are considered to have participated in the program for that month. The benefit amount recorded in the database is the monthly payment made to the SNAP case unit and is not prorated by the number of individuals who belong to the case unit. Since administrative SNAP records are available only for New York State, the linked SIPP and SIPP-EHC samples are restricted to individuals residing in New York State.

Discussion of SNAP results

This section presents and discusses the analysis of reporting of SNAP in SIPP-EHC. The table elements and weighting are described in Chapter 2, and, where appropriate, in table notes.

The rate of disagreement between survey and administrative data about monthly SNAP participation does not differ statistically between SIPP-EHC and SIPP (see Table 18.1). There is no statistically significant

difference between SIPP-EHC and SIPP in either CY2010 or CY2011 in the mean absolute deviation between the survey-reported and administratively recorded monthly SNAP participation.

Table 18.1: SNAP monthly participation

		Survey	Linked	AR	MAD
CY 2010	SIPP	0.27	0.34	0.32	0.12
	SIPP-EHC	0.21	0.33	0.33	0.11
	<i>t</i> -statistic	8.45	0.64	0.64	1.28
CY 2011	SIPP	0.27	0.38	0.35	0.12
	SIPP-EHC	0.24	0.35	0.38	0.10
	<i>t</i> -statistic	4.05	1.22	0.91	1.53
Diff-in-diff <i>t</i> -statistic		4.04	0.95	0.58	0.38
Person-months		361,187	37,079	37,079	37,079

Across SIPP-EHC and SIPP, the rate of disagreement between survey and administrative data about annual SNAP participation does not differ statistically (see Table 18.2). There is no statistically significant difference between SIPP-EHC and SIPP in CY2010 in the mean absolute deviation between the survey-reported and administratively recorded annual SNAP participation.

Across SIPP-EHC and SIPP, the rate of disagreement between survey and administrative data about annual SNAP participation does not differ statistically (see Table 18.2). There is no statistically significant difference between SIPP-EHC and SIPP in CY2010 in the mean absolute deviation between the survey-reported and administratively recorded months of SNAP participation.

Regarding the monthly amount of SNAP benefit, the rate of disagreement between survey and administrative data does not differ statistically between SIPP-EHC and SIPP. The mean absolute deviation between the survey-reported and administratively recorded months of SNAP participation does not differ statistically between SIPP-EHC and SIPP (see Table 18.3).

The rate of disagreement between survey and administrative data about the number of spells of SNAP participation does not differ statistically between SIPP and SIPP-EHC. The mean absolute deviation between the survey-reported and administratively-recorded does not statistically differ between SIPP and SIPP-EHC (see Table 18.4).

Based on the pooled person-month observations, rate of false negative SNAP reporting is not statistically different between SIPP-EHC and SIPP in either CY2010 or CY2011 (see Table 18.5). The rate of false positive SNAP reporting is lower in SIPP-EHC than in SIPP in both calendar years.

There is little evidence of reverse telescoping (see Chapter 2) in the pattern of SNAP reporting errors in SIPP-EHC (see Table 18.5). In CY2010, the false positive rate is lower in the last four months of the year than in the first four months of the year (Test 2). There is no statistically significant difference in the false positive rates between the same two periods in CY2011. There is no statistically significant difference between

Table 18.2: SNAP annual participation and months of participation

		Annual participation				Months of participation			
		Survey	Linked	AR	MAD	Survey	Linked	AR	MAD
CY 2010	SIPP	0.32	0.43	0.39	0.11	10.01	10.62	10.57	1.32
	SIPP-EHC	0.24	0.34	0.38	0.10	10.67	11.63	11.04	0.96
	<i>t</i> -statistic	10.58	3.41	0.21	1.14	6.55	5.09	2.00	1.66
CY 2011	SIPP	0.32	0.43	0.40	0.12	10.12	11.32	11.12	1.03
	SIPP-EHC	0.26	0.36	0.42	0.10	11.19	11.45	11.18	0.74
	<i>t</i> -statistic	7.19	2.11	0.90	1.09	10.79	0.61	0.30	1.45
Diff-in-diff <i>t</i> -statistic		2.72	1.01	1.50	0.13	3.03	3.14	1.36	0.21
Person-years		25,233	2,574	2,574	2,574	7,368	875	875	875

SIPP-EHC and SIPP in the difference between the false negative rates between the beginning and the end of either year.

The rate of item non-response about SNAP participation is slightly higher in SIPP-EHC than in SIPP for CY2010; there is no statistical difference for CY2011 (see Table 18.6 and 2). The item non-response about SNAP benefit amounts was lower SIPP-EHC than in SIPP for both years.

Table 18.3: SNAP monthly benefits

		Survey	Linked	AR	MAD
CY 2010	SIPP	\$301.52	\$271.43	\$305.69	\$57.85
	SIPP-EHC	\$353.92	\$351.49	\$355.70	\$57.07
	<i>t</i> -statistic	3.70	2.05	1.40	0.05
CY 2011	SIPP	\$279.83	\$265.77	\$300.34	\$51.60
	SIPP-EHC	\$298.56	\$351.82	\$353.63	\$105.45
	<i>t</i> -statistic	1.50	1.91	1.78	1.42
Diff-in-diff <i>t</i> -statistic		2.13	0.11	0.10	1.33
Person-months		20,458	2,724	2,724	2,724

Table 18.4: SNAP participation spells

	Survey	Linked	AR	MAD
SIPP	1.11	1.10	1.32	0.33
SIPP-EHC	1.03	1.12	1.62	0.58
<i>t</i> -statistic	8.87	0.21	1.87	1.76
Persons	3,408	205	205	205

Table 18.5: SNAP reporting errors

	False negatives							False positives						
	CY 2010			CY 2011			Diff	CY 2010			CY 2011			Diff
	SIPP	EHC	<i>t</i> -stat	SIPP	EHC	<i>t</i> -stat		SIPP	EHC	<i>t</i> -stat	SIPP	EHC	<i>t</i> -stat	
Jan	0.18	0.17	0.30	0.17	0.19	0.58	0.77	0.04	0.02	2.47	0.04	0.01	5.03	1.93
Feb	0.19	0.16	0.77	0.15	0.18	0.98	1.49	0.04	0.03	1.60	0.03	0.01	3.78	1.47
Mar	0.18	0.17	0.40	0.17	0.17	0.10	0.40	0.05	0.03	2.88	0.03	0.01	4.88	0.83
Apr	0.18	0.17	0.28	0.11	0.16	1.43	1.45	0.04	0.03	0.99	0.03	0.01	3.87	2.11
May	0.16	0.17	0.50	0.12	0.17	1.58	0.99	0.04	0.02	2.48	0.03	0.01	3.76	1.02
Jun	0.16	0.17	0.32	0.11	0.18	2.02	1.49	0.04	0.02	2.60	0.03	0.01	4.14	1.01
Jul	0.13	0.15	0.86	0.11	0.14	0.87	0.07	0.03	0.02	2.46	0.03	0.01	4.00	1.22
Aug	0.13	0.15	0.60	0.12	0.15	1.08	0.48	0.03	0.02	2.15	0.03	0.01	4.52	1.66
Sep	0.16	0.17	0.43	0.12	0.17	1.57	1.08	0.03	0.02	2.10	0.03	0.01	4.75	1.98
Oct	0.15	0.16	0.52	0.14	0.16	0.66	0.18	0.03	0.02	2.20	0.03	0.01	4.81	2.28
Nov	0.14	0.15	0.43	0.13	0.20	1.86	1.42	0.04	0.02	3.23	0.04	0.02	3.63	0.37
Dec	0.17	0.16	0.25	0.12	0.15	1.00	1.13	0.03	0.01	3.67	0.03	0.01	4.13	0.35
Pooled	0.16	0.16	0.14	0.13	0.17	1.32	1.19	0.04	0.02	2.84	0.03	0.01	5.09	1.76
Test 1	1.07	0.31	0.65	2.09	1.04	0.76	0.11	0.75	1.13	0.11	1.41	0.32	1.02	0.77
Test 2	1.38	0.38	0.88	1.31	0.36	0.65	0.14	1.80	3.40	0.70	0.48	0.37	0.21	0.60

See Chapter 2 for definitions of false negative and false positive reporting. The monthly *t*-statistics are from tests for equality of the rate of false positives (or false negatives) between the SIPP and the SIPP-EHC. Test 1 presents the *t*-statistic from a test of equality of the error rate for January and the error rate for September through December. Test 2 presents the *t*-statistic from a test of equality of the error rate for January through April and the error rate for September through December.

Table 18.6: Item-nonresponse rates

	NIU	Participation	Amount for participants
CY 2010			
SIPP	0.00	0.02	0.74
SIPP-EHC	0.00	0.03	0.23
<i>t</i> -statistic	0.00	5.80	8.93
CY 2011			
SIPP	0.00	0.02	0.71
SIPP-EHC	0.00	0.02	0.11
<i>t</i> -statistic	0.00	1.55	18.99
Diff-in-diff <i>t</i> -statistic	0.00	5.36	1.49
Person-months	369,202	369,202	9,794

19. Supplemental Security Income

This chapter presents comparisons of reporting about SSI receipt in SIPP-EHC, SIPP and SSI administrative records. The following section describes the SSI data that were employed in the analysis of SSI reporting that is discussed in the subsequent section.

Description of SSI data

The data sources for the tables in this chapter are 2008 SIPP, 2011 SIPP-EHC, 2012 SIPP-EHC, and SSI administrative records. See Chapter 2 for discussion of the SIPP-EHC and SIPP samples and weighting. This section discusses the creation of SSI-specific analysis variables from survey and administrative data.

Two sets of SSI variables are created, one from the survey data and one from the administrative data. Each set includes a person-month-level indicator for reported participation in SSI and a person-month-level report of SSI benefit amount.

All person-months are in-universe for the SSI participation indicator variables. The SSI participation indicator is coded as missing for person-months for which no report (survey) or record (administrative data) of SSI participation status was available. Person-months for which SSI participation is indicated are in-universe for the benefit-amount measure. The SSI benefit-amount variable is coded as missing when the person-month is in universe but there is no report or record available; these variables are also coded as missing for two types of cases: first, when the sample person participates but there is no amount reported; second, when the SSI participation indicator for the person-month is missing.

From 2008 SIPP, these variables are created from edited person-month data, dropping all imputations. For adults, the information about SSI receipt is obtained from the sample person's own record. For children, the information is merged on from adults' reports of receipt on behalf of a child.

In the 2011 and 2012 SIPP-EHC, individuals were asked if they had received any SSI payments since the beginning of the reference year. If they answered in the affirmative, they were asked further questions to determine the start and end dates of the benefit receipt. Similar questions were asked about benefit amounts, collecting first the current amount received and then probing for changes in amounts and the dates those changes happened. Using these unedited responses, monthly SSI participation indicators and amounts are created for CY2010 and CY2011.

In the 2011 SIPP-EHC, the information about SSI receipt for adults is obtained from the sample person's own record. For children, information from the child's own record and is reconciled with any report from a guardian of receipt on behalf of a child. Reconciled reports indicate SSI participation if either report so indicates. In the 2012 SIPP-EHC, the information about SSI receipt was always collected on the sample person's own record.

From the administrative data on SSI, comparable measures are constructed at the person-month level using the Supplemental Security Record (SSR) from the SSA. While nationwide records are available on payment of federal SSI benefits, records of payment of state SSI benefits are available only for state which elected to have their state SSI programs administered by SSA. Participation indicators are constructed for all for the all states, and these understate participation to the extent that beneficiaries in states for which we have no state SSI records receive state SSI without also receiving federal SSI. Measures of total benefit amount and source of SSI benefits are calculated only for states with SSA-administered state SSI programs.

Discussion of SSI results

In SIPP-EHC compared to SIPP, there is very slightly more disagreement between survey and administrative data about monthly SSI participation (see Table 19.1). The mean absolute deviation between the survey-reported and administratively recorded monthly SSI participation is higher for SIPP-EHC than for SIPP in each calendar year.

Table 19.1: SSI monthly participation

		Survey	Linked	AR	MAD
CY 2010	SIPP	0.05	0.06	0.07	0.04
	SIPP-EHC	0.08	0.09	0.07	0.04
	<i>t</i> -statistic	6.06	5.85	0.30	2.50
CY 2011	SIPP	0.05	0.06	0.07	0.04
	SIPP-EHC	0.09	0.10	0.07	0.05
	<i>t</i> -statistic	7.42	7.23	1.24	2.18
Diff-in-diff <i>t</i> -statistic		2.94	3.16	1.66	0.06
Person-months		347,281	279,722	279,722	279,722

Between SIPP-EHC and SIPP, the rate of disagreement between survey and administrative data about annual SSI participation does not differ statistically (see Table 19.2). There is no statistically significant difference between SIPP-EHC and SIPP in either CY2010 or CY2011 in the mean absolute deviation between the survey-reported and administratively recorded annual SSI participation. The statistically significant differences in the rate of survey-reported annual SSI participation between SIPP-EHC and SIPP appear to be due to sample composition differences (see Chapter 2).

In SIPP-EHC compared to SIPP, there is less disagreement between survey and administrative data about the number of months of SSI participation (see Table 19.2). The mean absolute deviation between the survey-reported and administratively recorded numbers is lower for SIPP-EHC than for SIPP in both years.

In SIPP-EHC compared to SIPP, there is more disagreement between survey and administrative data about the monthly amount of received SSI benefit (see Table 19.3). The mean absolute deviation between the survey-reported and administratively-recorded SSI benefit amount is higher for SIPP-EHC than for SIPP for CY2010 and CY2011. The statistically significant differences in the reported amount of the monthly SSI benefit between SIPP-EHC and SIPP appear to be due to sample composition differences (see Chapter 2). In SIPP-EHC compared to SIPP, there is less disagreement between survey and administrative data about the number of spells of SSI participation in CY2010 through CY2011. The mean absolute deviation between the survey-reported and administratively recorded number of spells of SSI receipt is lower for SIPP-EHC than for SIPP.

Based on the pooled sample, the rate of false negative reporting about SSI participation is lower in SIPP-EHC than in SIPP (see Table 19.5). The rate of false positive reporting about SSI participation is higher in

Table 19.2: SSI annual participation

		Annual participation				Months of participation			
		Survey	Linked	AR	MAD	Survey	Linked	AR	MAD
CY 2010	SIPP	0.06	0.07	0.06	0.04	8.90	10.35	11.65	1.56
	SIPP-EHC	0.07	0.08	0.07	0.04	12.00	12.00	11.71	0.29
	<i>t</i> -statistic	3.23	3.46	2.04	1.52	15.14	7.82	0.46	5.69
CY 2011	SIPP	0.06	0.07	0.07	0.04	9.42	10.33	11.52	1.33
	SIPP-EHC	0.09	0.10	0.08	0.04	12.00	12.00	11.78	0.22
	<i>t</i> -statistic	4.96	5.11	1.75	1.59	13.24	8.12	1.74	5.75
Diff-in-diff <i>t</i> -statistic		2.54	2.64	0.01	0.24	2.00	0.08	1.01	0.54
Person-years		24,058	19,358	19,358	19,358	1,633	1,000	1,000	1,000

Table 19.3: SSI benefits

		Survey	Linked	AR	MAD
CY 2010	SIPP	\$553.57	\$550.49	\$560.83	\$91.53
	SIPP-EHC	\$669.16	\$622.71	\$556.32	\$142.67
	<i>t</i> -statistic	5.69	3.47	0.17	2.53
CY 2011	SIPP	\$558.47	\$545.50	\$563.48	\$88.94
	SIPP-EHC	\$732.40	\$618.66	\$574.47	\$142.76
	<i>t</i> -statistic	5.40	3.36	0.39	2.67
Diff-in-diff <i>t</i> -statistic		1.72	0.04	0.57	0.11
Person-months		12,842	8,010	8,010	8,010

SIPP-EHC than in SIPP.

There is no evidence of reverse telescoping or straight-lining (see Chapter 2) in the patterns of SSI participation reporting errors in SIPP-EHC (see Table 19.5). For CY2010 and CY2011, there is no statistically significant difference between the false negative or false positive rates for January (Test 1), or January through April (Test 2), and the rates for September through December.

The rate of item-nonresponse about SSI receipt is higher in SIPP-EHC than in SIPP (see Table 19.6). The rate of item-nonresponse about SSI receipt in SIPP-EHC decreased relative to SIPP in CY2011 relative to CY2010. The rate of item-nonresponse about SSI amounts for participants is lower in SIPP-EHC than in SIPP. See Chapter 2 for an explanation of how the high SSI amount item-nonresponse in SIPP is an artifact of how the data were edited.

Table 19.4: SSI participation spells

	Survey	Linked	AR	MAD
SIPP	1.11	1.12	1.08	0.17
SIPP-EHC	1.00	1.00	1.09	0.09
<i>t</i> -statistic	8.52	6.46	0.34	2.18
Persons	955	457	457	457

Table 19.5: SSI reporting errors

	False negatives							False positives						
	CY 2010			CY 2011			Diff	CY 2010			CY 2011			Diff
	SIPP	EHC	<i>t</i> -stat	SIPP	EHC	<i>t</i> -stat		SIPP	EHC	<i>t</i> -stat	SIPP	EHC	<i>t</i> -stat	
Jan	0.26	0.17	3.41	0.28	0.11	6.11	2.57	0.22	0.34	3.98	0.23	0.31	2.66	1.10
Feb	0.26	0.17	3.49	0.29	0.11	6.35	2.79	0.20	0.34	4.60	0.24	0.31	2.24	2.02
Mar	0.26	0.17	3.58	0.28	0.11	6.36	2.65	0.22	0.33	3.98	0.22	0.30	2.52	1.28
Apr	0.25	0.17	3.10	0.27	0.11	6.01	2.77	0.23	0.33	3.56	0.23	0.31	2.58	0.83
May	0.26	0.16	3.65	0.26	0.11	5.74	1.97	0.22	0.34	3.79	0.22	0.31	3.05	0.62
Jun	0.27	0.16	3.71	0.27	0.11	5.77	1.91	0.23	0.34	3.50	0.22	0.31	3.10	0.30
Jul	0.27	0.16	3.97	0.26	0.11	5.58	1.52	0.20	0.33	4.38	0.21	0.31	3.37	0.79
Aug	0.27	0.17	3.67	0.27	0.11	5.63	1.91	0.21	0.32	3.86	0.23	0.31	2.50	1.11
Sep	0.27	0.17	3.81	0.27	0.11	5.56	1.68	0.24	0.33	2.92	0.24	0.31	2.34	0.49
Oct	0.27	0.17	3.34	0.27	0.11	5.56	2.15	0.23	0.33	3.17	0.25	0.32	2.35	0.65
Nov	0.27	0.17	3.41	0.28	0.11	5.98	2.51	0.26	0.32	2.14	0.24	0.32	2.38	0.31
Dec	0.28	0.18	3.81	0.27	0.11	6.08	2.11	0.25	0.33	2.53	0.24	0.32	2.60	0.14
Pooled	0.27	0.17	3.89	0.27	0.11	6.35	2.64	0.23	0.33	3.90	0.23	0.31	2.85	0.88
Test 1	0.59	0.56	0.29	0.46	0.44	0.35	0.41	1.37	1.05	1.70	0.42	0.17	0.32	0.89
Test 2	0.84	0.85	0.36	0.52	0.14	0.54	0.58	1.72	1.09	2.02	0.60	1.26	0.09	1.31

See Chapter ?? for definitions of false negative and false positive reporting. The monthly *t*-statistics are from tests for equality of the rate of false positives (or false negatives) between the SIPP and the SIPP-EHC. Test 1 presents the *t*-statistic from a test of equality of the error rate for January and the error rate for September through December. Test 2 presents the *t*-statistic from a test of equality of the error rate for January through April and the error rate for September through December.

Table 19.6: SSI item-nonresponse rates

		NIU	Participation	Amount for participants
CY 2010	SIPP	0.00	0.04	0.79
	SIPP-EHC	0.00	0.12	0.30
	<i>t</i> -statistic	0.00	19.88	13.53
CY 2011	SIPP	0.00	0.04	0.78
	SIPP-EHC	0.00	0.06	0.26
	<i>t</i> -statistic	0.00	6.04	15.99
Diff-in-diff <i>t</i> -statistic		0.00	11.37	0.88
Person-months		369,317	369,317	11,851

20. Temporary Assistance for Needy Families

This chapter presents comparisons of reporting about TANF receipt in SIPP-EHC, SIPP and administrative records for TANF. The following section describes the TANF data that were employed in the analysis discussed in the subsequent section.

Description of TANF data

The data sources for the tables in this chapter are 2008 SIPP, 2011 SIPP-EHC, 2012 SIPP-EHC, and from TANF New York State TANF administrative records. See Chapter 2 for discussion of the SIPP-EHC and SIPP samples and weighting. This section discusses the creation of TANF-specific from survey and administrative data.

Person-month-level indicators of reported participation in TANF are created from the survey sources and from administrative data. Data from the survey sources are used to create a person-month-level report of TANF benefit amount, as well as person-year-level indicators of annual TANF participation, months of participation, and spells of participation.

All person-months are in-universe for the TANF participation indicator variables. The TANF participation indicator is coded as missing for person-months for which no report (survey) or record (administrative data) of TANF participation status was available. Person-months for which TANF participation was indicated are in-universe for the benefit-amount measure. The TANF benefit-amount variable is coded as missing when it is in universe but there is no report available; this variable is also coded as missing when the TANF participation indicator for the person-month is missing.

From 2008 SIPP, these variables were created from edited person-month data, dropping all imputations. For all sample persons, the information about TANF participation was obtained from the sample person's own record. TANF benefit amounts data was collected at the family-level and was obtained from the head of the family's record.

In the 2011 and 2012 SIPP-EHC, adults were asked if they or someone in their family received any TANF benefits since the beginning of the reference year. If they answered in the affirmative, they were asked further questions to determine the start and end dates of the benefit receipt. Coverage is attributed to any additional household members covered as indicated in the report. For each spell of receipt, questions were asked about the benefit amounts. Using these unedited responses, monthly TANF participation indicators and amounts were created for CY2010 and CY2011.

The administrative-data versions of these variables are created from a database of monthly TANF benefit payments in New York State. The database contains a record for each individual in the TANF case unit, and individuals for whom a record exists are considered to have participated in the program for that month. The benefit amount recorded in the database is the monthly payment made to the TANF case unit and is not prorated by the number of individuals who belong to the case unit. Since administrative TANF records are available only for New York State, the linked SIPP and SIPP-EHC samples are restricted to individuals residing in New York State. Due to small cell sizes in the linked sample, comparisons to the administrative records could only be made for monthly participation rates.

Discussion of TANF results

This section presents and discusses the analysis of reporting in SIPP-EHC of TANF. The table elements and weighting are described in Chapter 2, and, where appropriate, in table notes.

In SIPP-EHC compared to SIPP, there is more disagreement between survey and administrative data about the monthly TANF participation (see Table 20.1) in CY2010. For CY2010, the mean absolute deviation between the survey-reported and administratively recorded monthly TANF participation is higher in SIPP-EHC than in SIPP. In CY2011, there is no statistically significant difference between SIPP-EHC and SIPP in the mean absolute deviation between the survey-reported and administratively recorded monthly TANF participation (see Table 20.1). The decrease in mean absolute deviation in SIPP-EHC relative to SIPP across the calendar years is statistically significant.

Table 20.1: TANF monthly participation

		Survey	Linked	AR	MAD
CY 2010	SIPP	0.03	0.03	0.03	0.03
	SIPP-EHC	0.03	0.03	0.04	0.05
	<i>t</i> -statistic	1.60	0.68	1.22	2.79
CY 2011	SIPP	0.03	0.03	0.03	0.03
	SIPP-EHC	0.03	0.04	0.03	0.03
	<i>t</i> -statistic	0.02	0.91	0.32	0.29
Diff-in-diff <i>t</i> -statistic		1.48	0.46	0.87	2.43
Person-months		367,140	37,357	37,357	37,357

In CY2011, the rate of annual TANF participation is higher in SIPP-EHC than in SIPP; in CY2010, there is no statistically significant difference between the rate of annual TANF participation in SIPP-EHC and the corresponding rate in SIPP (see Table 20.2). The number of months of reported TANF participation is higher in SIPP-EHC than in SIPP in CY2010 and CY2011 (see Table 20.2).

Table 20.2: TANF annual participation

		Annual participation	Months of participation
CY 2010	SIPP	0.04	7.44
	SIPP-EHC	0.04	9.87
	<i>t</i> -statistic	0.62	7.39
CY 2011	SIPP	0.04	7.87
	SIPP-EHC	0.03	10.90
	<i>t</i> -statistic	2.45	9.71
Diff-in-diff <i>t</i> -statistic		1.82	1.51
Person-years		25,939	1,047

There is no statistically significant difference between SIPP-EHC and SIPP in either CY2010 or CY2011 in the reported TANF benefit amounts (see Table 20.3).

Table 20.3: TANF benefits

		Amount
CY 2010	SIPP	\$406.31
	SIPP-EHC	\$370.57
	<i>t</i> -statistic	0.94
CY 2011	SIPP	\$412.14
	SIPP-EHC	\$420.07
	<i>t</i> -statistic	0.13
Diff-in-diff <i>t</i> -statistic		0.75
Person-months		2,103

The number of spells of TANF participation in CY2010 through CY2011 is lower for SIPP-EHC than for SIPP (see Table 20.4).

Table 20.4: Participation spells

		Spells of participation
SIPP		1.10
SIPP-EHC		1.03
<i>t</i> -statistic		3.19
Persons		736

There is no evidence of reverse telescoping in the pattern of month-by-month TANF participation rates in SIPP-EHC. For CY2011, there is no statistical difference across the surveys in the difference between the TANF participation rate for September through December and the TANF participation rate for 1) January (Test 1) and for 2) January through April (Test 2).

For CY2010, the difference between the surveys in the difference between participation rates at the beginning and end of the year is statistically significant, but the differences between the surveys are greater at the end of the year, contrary to the reverse telescoping hypothesis (see Table 20.5).

In CY2010 and CY2011, the rate of item-nonresponse about TANF participation is higher in SIPP-EHC than in SIPP (see Table 20.6). Across both calendar years, the rate of item-nonresponse about TANF benefit

Table 20.5: Month-by-month TANF participation

	CY 2010			CY 2011			Diff
	SIPP	EHC	<i>t</i> -stat	SIPP	EHC	<i>t</i> -stat	
Jan	0.03	0.03	1.61	0.03	0.03	0.19	1.62
Feb	0.03	0.03	0.21	0.03	0.03	0.50	0.67
Mar	0.03	0.03	0.01	0.03	0.03	0.36	0.34
Apr	0.03	0.03	1.09	0.03	0.03	0.28	0.67
May	0.03	0.03	1.51	0.03	0.03	0.33	1.64
Jun	0.03	0.03	1.69	0.03	0.03	0.49	1.95
Jul	0.03	0.03	1.77	0.03	0.03	0.49	2.03
Aug	0.03	0.04	1.80	0.03	0.03	0.03	1.57
Sep	0.03	0.04	2.27	0.03	0.03	0.17	1.83
Oct	0.03	0.04	2.82	0.03	0.03	0.03	2.51
Nov	0.03	0.04	3.17	0.03	0.03	0.15	3.02
Dec	0.03	0.04	3.02	0.03	0.03	0.12	2.62
Pooled	0.03	0.03	1.60	0.03	0.03	0.02	1.48
Test 1	3.25	6.61	6.40	0.73	0.75	0.25	4.34
Test 2	1.57	6.50	5.02	0.93	0.48	0.53	3.65

amounts is lower in SIPP-EHC than in SIPP.

Table 20.6: TANF item-nonresponse rates

	NIU	Participation	Benefits
CY 2010	SIPP	0.00	0.69
	SIPP-EHC	0.00	0.47
	<i>t</i> -statistic	0.00	4.96
CY 2011	SIPP	0.00	0.75
	SIPP-EHC	0.00	0.49
	<i>t</i> -statistic	0.00	2.12
Diff-in-diff <i>t</i> -statistic	0.00	1.81	0.53
Person-months	369,202	369,202	5,250

21. Unemployment Insurance

This chapter presents comparisons of reporting about Unemployment Insurance (UI) receipt in SIPP-EHC and SIPP. The following section describes the UI data that were employed in the analysis of UI reporting that is discussed in the subsequent section.

Description of unemployment insurance data

The data sources for the tables in this chapter are 2008 SIPP, 2011 SIPP-EHC and 2012 SIPP-EHC. See Chapter 2 for discussion of the SIPP-EHC and SIPP samples and weighting. This section discusses the creation of UI-specific analysis variables.

The variables specific to UI include a person-month-level indicator for reported receipt of UI and a person-month-level report of UI benefit amount. The universe for the UI participation indicator is restricted to those 15 or older at time of interview. Person-months for which UI participation is indicated are in-universe for the UI benefit amount measure.

This participation indicator is coded as missing for person-months for which no report of UI participation status is available. The UI amount received variable is coded as missing when receipt is reported in a given person-month, but no amount report is available, or when the UI participation indicator for the person-month is missing.

The surveys collected both participation and amount received for three types of UI benefits: regular, state UI; supplemental UI; and other UI. For the purpose of this analysis, participation in any of the three types is coded as participation in UI. Benefit amounts for the three types are summed to create the measure of benefit amount for this analysis.

From 2008 SIPP, these variables were created from edited person-month data, dropping all imputations.

In the 2011 and 2012 SIPP-EHC, individuals were asked if they had received any UI benefits since the beginning of the reference year. If they answered in the affirmative, they were asked further questions to determine the start and end dates of the benefit receipt. Similar questions were asked about benefit amounts, collecting first the current or most recent amount received and then probing for changes in amounts and the months those changes occurred. Using these unedited responses, monthly UI participation indicators and benefit amounts measures are created for 2010 and 2011.

Discussion of unemployment insurance results

This section presents and discusses the analysis of reporting of UI in SIPP-EHC. The table elements and weighting are described in Chapter 2, and, where appropriate, in table notes.

There is no statistically significant difference between rates of reported monthly UI participation in SIPP-EHC and SIPP in CY2010 (see Table 21.1). In CY2011, there is a statistically significant but negative difference between the monthly rates of participation (see Table 21.1) as participation is higher in SIPP. There is no statistically significant difference between annual rates of UI participation in SIPP-EHC and SIPP in CY2010 and CY2011 (see Table 21.1).

There is no statistically significant difference between months of UI participation reported in SIPP-EHC and SIPP in CY2011 (see Table 21.1). In CY2010, there is a statistically significant difference between the reported months of participation, with SIPP-EHC person-years reporting more months of receipt than SIPP person-years, on average (see Table 21.1).

Table 21.1: Unemployment insurance receipt

		Monthly participation	Annual participation	Months of participation
CY 2010	SIPP	0.03	0.05	6.28
	SIPP-EHC	0.03	0.05	7.17
	<i>t</i> -statistic	0.15	0.45	2.41
CY 2011	SIPP	0.02	0.04	6.33
	SIPP-EHC	0.02	0.03	6.02
	<i>t</i> -statistic	2.64	1.33	0.65
Diff-in-diff <i>t</i> -statistic		2.19	1.48	1.98
Observations		270,652	19,008	796

The unit of observation for the Monthly participation column is person-months.

The unit of observation for the Annual participation column is person-years for persons with a reported participation status for all months of the year.

The unit of observation for the Months participation column is person-years for participants with a reported participation status for all months of the year.

The monthly amount of UI benefit reported in CY2011 is lower in SIPP-EHC than in SIPP; however, there was no statistically significant difference between the two surveys in CY2010 (see Table 21.2). The surprising difference in UI benefit amount between the surveys in CY2011 is still being investigated.

Across 2010 and 2011 together, fewer spells of UI receipt are reported in SIPP-EHC than in SIPP (see Table 21.3).

Table 21.2: Unemployment insurance monthly benefit amounts

		Benefits
CY 2010	SIPP	\$900.43
	SIPP-EHC	\$935.72
	<i>t</i> -statistic	0.34
CY 2011	SIPP	\$932.41
	SIPP-EHC	\$623.25
	<i>t</i> -statistic	3.54
Diff-in-diff <i>t</i> -statistic		2.61
Person-months		5,773

Table 21.3: Spells of unemployment insurance benefit receipt for the years 2010–2011

	Spells
SIPP	1.14
SIPP-EHC	1.09
<i>t</i> -statistic	2.06
Persons	931

There is no evidence of reverse-telescoping or straight-lining (see Chapter 2) in the patterns of reporting about UI participation in SIPP-EHC (see Table 21.4). There is no statistically significant difference between SIPP-EHC and SIPP in the difference between the average participation rates in January (Test 1) or January through April (Test 2) and the rates for September through December.

Table 21.4: Month-by-month unemployment insurance benefits receipt

	CY 2010			CY 2011			Diff
	SIPP	EHC	<i>t</i> -stat	SIPP	EHC	<i>t</i> -stat	
Jan	0.03	0.03	0.69	0.02	0.02	2.32	1.18
Feb	0.03	0.03	0.67	0.02	0.02	2.34	1.16
Mar	0.03	0.03	0.92	0.02	0.02	1.12	0.03
Apr	0.03	0.03	0.00	0.02	0.02	0.89	0.62
May	0.03	0.03	0.58	0.02	0.02	1.43	1.49
Jun	0.03	0.03	0.81	0.02	0.01	2.72	2.58
Jul	0.03	0.03	1.00	0.02	0.01	3.58	3.44
Aug	0.03	0.03	0.82	0.02	0.02	2.73	2.69
Sep	0.03	0.03	0.96	0.02	0.02	2.58	2.76
Oct	0.03	0.03	0.21	0.02	0.02	2.63	1.84
Nov	0.03	0.03	0.01	0.02	0.02	2.00	1.51
Dec	0.03	0.03	0.01	0.02	0.02	2.22	1.65
Pooled	0.03	0.03	0.15	0.02	0.02	2.64	2.19
Test 1	1.98	0.77	0.90	0.63	0.76	0.02	0.63
Test 2	1.93	0.72	0.86	0.95	0.21	0.86	1.15

The rate of item-nonresponse for UI receipt is higher in SIPP-EHC than in SIPP (see Table 19.6). Relative to CY2010, the difference between SIPP-EHC and SIPP in the rate of item-nonresponse for UI receipt decreases in CY2011. The rate of item-nonresponse for UI amount for participants is higher in SIPP-EHC than in SIPP in CY2010. In CY2011, the corresponding rate is not statistically different between the two surveys (see Table 19.6).

Table 21.5: Unemployment-insurance benefit item-nonresponse rates

		NIU	Receipt	Amount for participants
CY 2010	SIPP	0.26	0.01	0.09
	SIPP-EHC	0.23	0.11	0.49
	<i>t</i> -statistic	3.46	23.45	1.98
CY 2011	SIPP	0.25	0.00	0.10
	SIPP-EHC	0.24	0.01	0.50
	<i>t</i> -statistic	1.22	2.13	1.09
Diff-in-diff <i>t</i> -statistic		2.85	21.60	0.01
Person-months		369,202	278,579	247

22. Transitions and Seams

This chapter presents analysis of reporting in SIPP-EHC of changes over time in the activities and status of sample-persons. The incidence and timing of reported transitions is evaluated for employment, program participation, health insurance coverage, and school enrollment. The identification of changes in the status of sample-persons is important for evaluating many topics, such as the determinants and seasonality of program entry and exit.

For the purposes of the analysis in this section, transition rates are assigned to the first month with a change in status. This month may not immediately follow a month with a different reported status if there are intervening months with missing interviews or item non-response. The analysis of administrative data is based on 12 months of records per year for linked sample-persons, regardless of any missing months of survey data.

This section also contains a brief analysis of seam bias. Both SIPP and SIPP-EHC consist of multiple abutting series of waves, each of which covers a reference period spanning several months. If respondents project current or recent status backwards to the beginning of the reference period, then transitions will fall disproportionately on the seam months (the first month of each reference period).

The reference period for each SIPP wave is 4-months long. Therefore, if transitions are distributed uniformly across the year, then approximately 25 percent of transitions should occur during the seam month. The SIPP reference period is rotated, so every calendar month is on the seam for approximately one-fourth of respondents. The SIPP-EHC reference period is 12-months long. Therefore, if transitions are distributed uniformly across the year, then approximately 8.3 percent of transitions should occur during the seam month. In all cases, the SIPP-EHC reference period is the calendar year, and so the seam month is January. An excessive amount of transitions during the seam months is evidence of seam bias, indicating biased measurement of the timing of transitions.

The following analysis focuses on calendar year 2011. For SIPP-EHC, there is limited information on status in December 2009. Therefore, it is not possible to observe transitions in January 2010 in these SIPP-EHC data.

The following analysis of seam bias reflects an innovation in the SIPP-EHC that allows for improved capture of the timing of transitions around the seam. The SIPP-EHC interview collects status reports for the months between the reference period and the interview date. For example, a person interviewed for the 2010 reference period in March 2011 will report status for January 2010 through the first three months of 2011. This information is incorporated into the interviewing in the subsequent wave to prompt respondents to time transitions more accurately.

This overlap in the periods for which data are collected in successive waves produces two reported statuses for the months around the seam (in the case of the preceding example, for the months of January, February, and March 2011). However, only reports of positive status (employment, enrollment, participation, coverage) are incorporated into interviewing in the subsequent wave. Reports of negative status (non-coverage, non-employment, non-enrollment, non-participation) are not fed back to the instrument. For the following analysis, the SIPP-EHC data are modified in an illustrative case (Medicaid) to favor wave 1 SIPP-EHC reports of non-coverage over wave 2 reports of coverage. This reconciliation could be made for all topics in SIPP-EHC. At this time, only the Medicaid participation variable has been adjusted and, therefore, the analysis of seam bias in SIPP-EHC focuses on Medicaid.

SIPP-EHC captures the incidence of intra-year transitions well relative to SIPP (see Table 22.1.) For Medicaid, Medicare, OASDI, SNAP, and TANF, the rate of disagreement (mean absolute deviation) between survey and administrative data about number of CY2011 transitions for each sample-person is lower for SIPP-EHC than for SIPP. For the only remaining administrative-data topic, SSI, the rate of disagreement is not statistically different between the surveys.

Table 22.1: Number of transitions in CY2011: adrec topics

		Linked	AR	MAD
Medicaid	SIPP	0.27	0.07	0.28
	SIPP-EHC	0.11	0.07	0.14
	<i>t</i> -statistic	21.66	0.11	16.77
Medicare	SIPP	0.06	0.01	0.06
	SIPP-EHC	0.03	0.01	0.03
	<i>t</i> -statistic	8.87	0.01	9.00
OASDI	SIPP	0.05	0.02	0.06
	SIPP-EHC	0.04	0.02	0.05
	<i>t</i> -statistic	1.67	0.16	2.52
SNAP	SIPP	0.14	0.03	0.16
	SIPP-EHC	0.10	0.02	0.12
	<i>t</i> -statistic	5.82	1.70	5.16
SSI	SIPP	0.03	0.01	0.04
	SIPP-EHC	0.03	0.01	0.04
	<i>t</i> -statistic	1.58	0.20	1.07
TANF	SIPP	0.04	0.00	0.04
	SIPP-EHC	0.02	0.00	0.03
	<i>t</i> -statistic	3.80	1.40	3.28
N		16,101	16,101	16,101

Table 22.2 reports mean number of transitions per person for five additional topics for which administrative data are not available. Mean number of transitions per person does not differ statistically between the surveys for rent subsidies and unemployment insurance. Mean number of transitions is higher in SIPP-EHC than in SIPP for school enrollment and employment. Mean number of transitions is lower in SIPP-EHC than in SIPP for health insurance.

Seam bias is present in both the SIPP and SIPP-EHC data. For SIPP and SIPP-EHC, Table 22.3 compares the rate of transitions being reported on seam-months for CY2011 for several topics. For each survey, the

Table 22.2: Number of transitions in CY2011: non-adrec topics

	SIPP	SIPP-EHC	<i>t</i> -stat
School Enrollment	0.16	0.20	4.18
Health Insurance	0.18	0.11	9.06
Rent Subsidy	0.06	0.06	1.34
Unemployment Insurance	0.04	0.04	0.64
Employment	0.10	0.14	5.11
N	16,101	16,101	16,101

on-seam transition rates are compared to the rates that would be observed if the transitions were uniformly distributed across months: 24% for SIPP and 8.3% for SIPP-EHC. For both surveys, the rates are higher than the rates expected under uniform distribution of transitions across months for every topic.

Table 22.3: On seam transitions in CY2011: non-adrec topics

	SIPP	<i>t</i> -stat	SIPP-EHC	<i>t</i> -stat
Employment	0.47	13.43	0.50	20.50
Health Insurance	0.93	104.09	0.72	33.73
Medicaid	0.95	160.99	0.65	27.44
Medicare	0.97	99.75	0.76	18.66
OASDI	0.89	42.20	0.84	30.42
SNAP	0.81	52.61	0.72	31.96
SSI	0.93	48.14	0.92	38.43
Rent Subsidy	0.91	56.11	0.88	43.06
School Enrollment	0.51	19.73	0.21	8.24
TANF	0.89	38.62	0.77	17.40
UI	0.65	16.16	0.51	11.56
N	16,101	16,101	16,101	16,101

Table 22.4 presents the distribution of Medicaid transitions reported in SIPP-EHC across the months of CY2011. The first column gives the distribution of transitions without reconciling wave 1 and wave 2 information for months of overlap favoring the wave 2 reports. The second column favors reports from wave 1 for the overlapping months. In the latter distribution, about 23 percent of the transitions are moved from January to subsequent months, mostly to February, March, or April. The third column reports Z-statistics for tests of equality between the rates in each month. Basically, this reconciliation of multiple reports moves

transitions on to Medicaid from January to the first month following the interview date. More involved imputation techniques could be employed to generate a distribution of transitions over months that reflects the distribution from administrative records.

Table 22.4: Month-by-month Medicaid transitions in SIPP-EHC

	Original	Reconciled	<i>t</i> -stat
January	64.3%	41.3%	7.63
February	4.7%	14.6%	-5.49
March	3.8%	13.9%	-5.84
April	4.7%	9.1%	-2.83
May	2.3%	2.1%	0.15
June	4.9%	4.6%	0.22
July	2.1%	2.0%	0.14
August	2.3%	2.1%	0.15
September	3.4%	3.2%	0.18
October	2.1%	2.0%	0.14
November	2.5%	2.3%	0.15
December	3.0%	2.9%	0.17

Works Cited

- Abowd, J. and Harrison Stinson, M. (2011). Estimating measurement error in sipp annual job earnings: a comparison of census bureau survey and ssa administrative data. *US Census Bureau Center for Economic Studies Paper No. CES-WP-11-20*.
- Axinn, W. G., Pearce, L. D., and Ghimire, D. (1999). Innovations in life history calendar applications. *Social Science Research*, 28(3):243–264.
- Belli, R. F. (1998). The structure of autobiographical memory and the event history calendar: Potential improvements in the quality of retrospective reports in surveys. *Memory*, 6(4):383–406.
- Belli, R. F., Shay, W. L., and Stafford, F. P. (2001). Event history calendars and question list surveys: A direct comparison of interviewing methods. *Public Opinion Quarterly*, 65(1):45–74.
- Callegaro, M. (2007). *Seam Effect Changes due to Modifications in Question Wording and Data Collection Strategies A Comparison of Conventional Questionnaire and Event History Calendar Seam Effects in the PSID*. PhD thesis, University of Nebraska.
- Citro, C. F., Scholz, J. K., et al. (2009). *Reengineering the Survey of Income and Program Participation*. National Academies Press.
- Gathright, G., Stinson, M., and Reeder, L. (2012). An evaluation of reporting accuracy in sipp-ehc field test data using 2008 sipp and administrative records. *Federal Committee on Statistical Methodology 2012 Research Conference Papers*.
- Glasner, T. and Van der Vaart, W. (2009). Applications of calendar instruments in social surveys: a review. *Quality and Quantity*, 43(3):333–349.
- Kominski, R. (1990). The sipp event history calendar: Aiding respondents in the dating of longitudinal events. *Proceedings of the Section on Survey Research Methods. American Statistical Association, Alexandria*.
- Marquis, K. H. and Moore, J. C. (1990). Measurement errors in sipp program reports. *U.S. Census Bureau*, page 01.
- Meyer, B. and Goerge, R. (2011). Errors in survey reporting and imputation and their effects on estimates of food stamp program participation. *U.S. Census Bureau Center for Economic Studies Paper No. CES-WP-11-14*.
- Moore, J., Fields, J., Benedetto, G., Stinson, M., Chan, A., and Maples, J. (2009). The 2008 survey of income and program participation event history calendar field test: Study design and initial results. *Survey Methodology Research Report*, (2009-09).
- Sayles, H., Belli, R. F., and Serrano, E. (2010). Interviewer variance between event history calendar and conventional questionnaire interviews. *Public Opinion Quarterly*, 74(1):140–153.
- SIPP-EHC Data Evaluation Workgroup (2011). An initial evaluation of the 2010 field test of re-engineered sipp. *SEHSD Working Papers*, (2011-9).
- Stinson, M., Gathright, G., and Skog, J. (2012). Comparing job characteristics from the 2010 sipp-ehc field test to the census bureau business register. *Federal Committee on Statistical Methodology 2012 Research Conference Papers*.

Tourangeau, R., Rips, L. J., and Rasinski, K. (2000). *The psychology of survey response*. Cambridge University Press.

U. S. Census Bureau (2012). Sipp website.