

Introduction to the 2018 Survey of Income and Program Participation (SIPP)

Hsueh-Sheng Wu
CFDR Workshop Series
March 29, 2021

BGSU



Center for Family and
Demographic Research

Outline

- What is SIPP?
- History of SIPP
- Study design of 2018 SIPP
- Content of 2018 SIPP
- Research questions using 2018 SIPP
- Navigating through the website of 2018 SIPP
- 2018 SIPP data
- Replicate weights for 2018 SIPP
- Conclusions

BGSU



Center for Family and
Demographic Research

What is SIPP?

- SIPP is a series of national survey conducted by U.S. Census Bureau
- The purposes of SIPP is to provide accurate and comprehensive information on economic well-being of the American individuals and families
- SIPP focuses on:
 - the distribution of income and economic resources of individuals and families
 - the extent to which people and family participate in government assistance programs
 - family and social contexts of individuals and families
 - the interactions among the constructs above

History of SIPP

- The first SIPP survey was conducted in October 1983
- Before SIPP, researchers relied on March Supplement of Current Population Survey (CPS) to understand economic well-being of individuals and families
- CPS, designed as a survey of labor force participation of U.S. population, has its limitation to understand the economic well-being of individuals and families
 - very long recall period (i.e., a 4-8-4 design)
 - inadequately measures of irregular sources of income, government program participation, and changes in family structure
 - respondents are followed for only 16 months

History of SIPP (Cont.)

- CPS would have become too huge to manage if it had been designed to overcome these limitations of studying economic well-being. Thus, SIPP was designed as an independent survey
- To date, SIPP surveys have been conducted in 16 panels. The 2018 SIPP panel is the latest one and will cover a four-year time period from January 2018 to December 2021
- Census Bureau redesigned the 2014 SIPP study to save costs, reduce respondent burden, and still maintain data quality. These changes include (1) sampling a smaller number of households; (2) interviewing respondents one time, rather than four times, a year; and (3) changing the reference period of survey questions from four months to one year
- The SIPP panels from 2018 to 2025 will all be conducted using the study design initiated in the 2014 SIPP panel

Study Design of 2018 SIPP

- Target population is U.S. non-institutionalized civilians aged 15 and over
- Sampling frame is created from files of 2010 Decennial Census, Master Address File (MAF), and U.S. Postal Service's Delivery Sequence
- A two-stage stratified sampling is used to select households. The first stage selects one or more contiguous counties (Primary Sampling Unites) from the sampling frame. The second stage divides addresses with PSUs into two strata: One with high concentration of low-income households and the other with low concentration of low-income households. Addresses from the strata with high concentration of low-income households were oversampled

BGSU

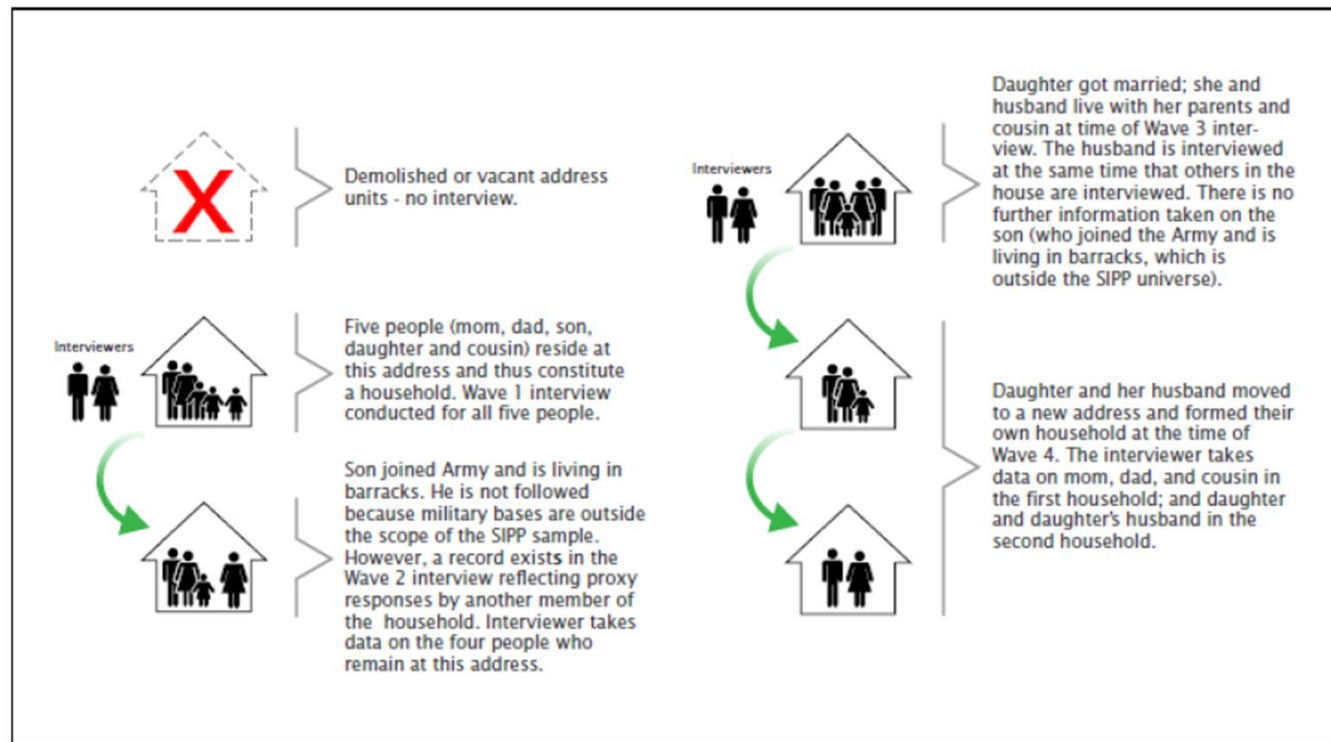


Center for Family and
Demographic Research

Study Design of 2018 SIPP (Cont.)

- Following rules in 2014 SIPP: Respondents of the sampled households and people who later co-reside in the same households were followed

FIGURE 2-1. EXAMPLES OF SIPP FOLLOWING RULES



Source: United States Census Bureau (August 2020) Survey of Income and Program Participation: 2018 Panel Users' Guide. Retrieved from https://www2.census.gov/programs-surveys/sipp/tech-documentation/methodology/2018_SIPP_Users_Guide.pdf

Content of 2018 SIPP

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

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

Source: United States Census Bureau (August 2020) Survey of Income and Program Participation: 2018 Panel Users' Guide. Retrieved from https://www2.census.gov/programs-surveys/sipp/tech-documentation/methodology/2018_SIPP_Users_Guide.pdf

Research Questions Using SIPP?

- Examine the eligibility and effectiveness of government programs, for example, studying how many people will be eligible for food stamp if there are changes to the eligibility criteria and how many people will face food insecurity without the food stamp program
- Studying how family and social contexts influence people's income and economic resources and subsequently their participation in government assistance programs, for example, Valle and Perez-Lopez (2020) examines family participation rates in Nutrition Assistance Programs.
- Studying union stability and family relations. For example, Manning, Brown, and Strykers (2016) compare the relationship stability between same-sex and different-sex cohabiting couples

BGSU



Center for Family and
Demographic Research

Navigating through the Website of 2018 SIPP

The screenshot displays the official website for the Survey of Income and Program Participation (SIPP). The browser address bar shows the URL <https://www.census.gov/programs-surveys/sipp.html>. The page features a left-hand navigation menu with links to 'About this Survey', 'Information for Survey Participants', 'Data', 'Events', 'Guidance for Data Users', 'Library', 'Methodology', 'News', 'Technical Documentation', and 'Contact Us'. The main content area is titled 'Survey of Income and Program Participation (SIPP)' and includes a descriptive paragraph: 'SIPP is the premier source of information for income and program participation. SIPP collects data and measures change for many topics including: economic well-being, family dynamics, education, assets, health insurance, childcare, and food security.' Below this text are three large buttons: 'About SIPP' (with an image of a green folder), 'Access SIPP Data' (with an image of two silhouettes), and 'SIPP Codebook' (with an image of a blue book). To the right of these buttons is a word cloud shaped like the United States, containing terms such as 'Households', 'Policy Programs', 'Health', 'Assets', 'Poverty', 'Disability', 'Event-History-Calendar', 'Living Arrangements', 'Well-Being', 'Dynamics', 'Research', 'Ownership', 'WIC', 'Health Insurance', 'Demographics', 'Income', 'SNAP', 'Wealth', 'Jobs', 'Fertility', 'Veterans', 'Disability', 'Child Care', 'Education', 'Industry', and 'Care'. Below the buttons is a 'Highlights' section titled '2018 SIPP Data', which states: '2018 SIPP Data files covering reference year January - December 2017.' To the right of the highlights is a small widget showing the 'U.S. Population' as '330,170,118' as of 'Mar 29, 2021 10:39 UTC (+4)'. At the bottom of the page, there is a taskbar with several open files, including 'pu2014w2_v07.zip', 'nw14w2_v07.zip', 'nw2018_sasdata (1).zip', 'Work History Job a...docx', 'tenure.zip', 'interview_date.zip', and 'previous_jobs.zip'. The taskbar also shows the Windows search bar and the system clock indicating '6:39 AM 3/29/2021'.

SURVEY OF INCOME AND PROGRAM PARTICIPATION (SIPP)

About this Survey

Information for Survey Participants

Data

Events

Guidance for Data Users

Library

Methodology

News

Technical Documentation

Contact Us

< Back to Our Surveys & Programs

Survey of Income and Program Participation (SIPP)

SIPP is the premier source of information for income and program participation. SIPP collects data and measures change for many topics including: economic well-being, family dynamics, education, assets, health insurance, childcare, and food security.

[About SIPP](#)

[Access SIPP Data](#)

[SIPP Codebook](#)

Highlights

2018 SIPP Data

2018 SIPP Data files covering reference year January - December 2017.

Mar 29, 2021 10:39 UTC (+4)

United States | World

U.S. Population

330,170,118

Is this page helpful?

Yes No

Navigating through the Website of 2018 SIPP

The screenshot displays the United States Census Bureau website for the Survey of Income and Program Participation (SIPP) datasets. The browser address bar shows the URL: <https://www.census.gov/programs-surveys/sipp/data/datasets.html>. The website header includes the Census Bureau logo, a search bar, and navigation tabs: BROWSE BY TOPIC, EXPLORE DATA, LIBRARY, SURVEYS/ PROGRAMS, INFORMATION FOR..., FIND A CODE, and ABOUT US. The breadcrumb trail indicates the path: Census.gov > Our Surveys & Programs > Survey of Income and Program Participation (SIPP) > Survey of Income and Program Participation Data > Survey of Income and Program Participation Datasets.

The main content area features the title "Survey of Income and Program Participation Datasets" with social media links for Facebook, Twitter, and LinkedIn. Below the title, it states: "Find SIPP data by panel and wave from 1984 - 2018 in the tabbed lists below. For SIPP data related questions, visit the [Contact Us](#) page."

A horizontal tabbed list shows years: 2018 (selected), 2014, 2008, 2004, 2002, 2001, 1996, 1993, and a MORE dropdown arrow.

The "2018" tab is active, displaying the text: "Beginning with 2018 SIPP, data are collected as yearly overlapping panels, with new panels beginning each year. 2018 SIPP includes Wave 1 of the 2018 Panel only."

On the left sidebar, under "SURVEY OF INCOME AND PROGRAM PARTICIPATION (SIPP)", the "Datasets" link is highlighted. Below it, "Data Tools" is listed, and a link to "Back to Survey of Income and Program Participation Data" is provided. An image labeled "Information for Respondents" is also visible.

On the right, the "Related Information" section includes links for "DATA", "Survey of Program Dynamics (SPD) Data", "1984-2008 Topical Modules List", and "Synthetic SIPP Data".

At the bottom, the "Related Data" section highlights "2018 SIPP Data" with the description: "2018 SIPP Data files covering reference year January - December 2017." A feedback box in the bottom right corner asks "Is this page helpful?" with "Yes" and "No" buttons.

BGSU



Center for Family and Demographic Research

Navigating through the Website of 2018 SIPP

The screenshot shows the 'SIPP Codebook' page on the Census.gov website. The page header includes the United States Census Bureau logo and a breadcrumb trail: // Census.gov / Our Surveys & Programs / Survey of Income and Program Participation / SIPP Codebook. The main title 'SIPP Codebook' is prominently displayed. Below the title, a status bar indicates 'Showing 5534 of 5534 Variables' with a 'Reset' button. To the right, there are dropdown menus for 'Sort by Variable' (set to 'Ascending') and a 'Mini Codebook' button. A search bar prompts 'Browse by Variable Name or Topic Sea...' with a magnifying glass icon. On the right, a 'FILTER BY' section contains three buttons: 'Survey Year', 'Topic/Subtopic', and 'Variable Name'. Below these, a table lists variables with columns for 'VARIABLE', 'TOPIC/SUBTOPIC', and 'DESCRIPTION'. The table shows variables like SHHADID, SPANEL, SSUID, SWAVE, PNUM, MONTHCODE, and GHLFSAM. At the bottom right, a small feedback box asks 'Is this page helpful?' with 'Yes' and 'No' options. The Windows taskbar is visible at the very bottom.

United States Census Bureau

// Census.gov / Our Surveys & Programs / Survey of Income and Program Participation / SIPP Codebook

SIPP Codebook

Showing 5534 of 5534 Variables [Reset](#)

Sort by Variable [Ascending](#) [Mini Codebook](#) [Download](#) [Help](#) [Email](#)

Browse by Variable Name or Topic Sea... [Q](#)

FILTER BY: [Survey Year](#) [Topic/Subtopic](#) [Variable Name](#)

Items per page: 20 1 – 20 of 5534 [<](#) [>](#)

VARIABLE	TOPIC/SUBTOPIC	DESCRIPTION
+ SHHADID	ID Variables ID Variables	Household address ID. Used to differentiate households spawned from an original sample household. Print
+ SPANEL	ID Variables ID Variables	Panel year Print
+ SSUID	ID Variables ID Variables	Sample unit identifier. This identifier is created by scrambling together PSU, Sequence #1, Sequence #2, and the Frame Indicator for a case. It may be used in matching sample units from different waves. Print
+ SWAVE	ID Variables ID Variables	Wave number of interview Print
+ PNUM	ID Variables ID Variables	Person number Print
+ MONTHCODE	ID Variables ID Variables	Value of reference month Print
+ GHLFSAM	ID Variables ID Variables	Half sample code. A code used to divide the sample into "half sample" replicates that are used for variance estimation. Print
+ ...	ID Variables	Variance pseudo stratum code. Strata formed for half sample variance estimation

Is this page helpful? [Yes](#) [No](#)

6:52 AM 3/29/2021

2018 SIPP Data

- Currently, only data from Wave 1 of SIPP 2018 are available
- The primary data file has 763,186 records, 4,942 variables. One of these variables is the final person weight variable (i.e., wpinwgt)
 - When SIPP data from a particular month are analyzed, the final person weight variable from that particular month should be used
 - When SIPP data from a calendar are analyzed, the final person weight variable from December of that year should be used
 - When the whole panel of SIPP data from analyzed, the final person weight variables from December of each year are combined and averaged to weight the result
- The replicate weights file has 240 replicate weight variables

2018 SIPP Data

Table 1. Select Key ID, Demographic, and Weight Variables

<u>Variable Name</u>	<u>Variable Description</u>
MONTHCODE	Value of reference month
SPANEL	Panel year
SWAVE	Wave number of interview
SSUID	Sample unit identifier. This identifier is created by scrambling together PSU, Sequence #1, Sequence #2, and the Frame Indicator for a case. It may be used in matching sample units across years.
SHHADID	Household address ID. Used to differentiate households spawned from an original sample household.
PNUM	Person number in a household
TAGE	Age as of last birthday
ESEX	Sex of this person
EMS	Marital Status
ERELRP	Household Relationships
WPFINWGT	Final person weight
REPWGT1-REPWGT240	Replicate Weights

2018 SIPP Data

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
Example	monthcode	spanel	swave	ssuid	shhadid	pnum	tage	esex	ems										erelrpe	
1	1	2018	1	11413646518	11	101	35	2: Female	4. Divorced										Child	
1	1	2018	1	11413646518	11	102	52	2: Female	1. Married, spouse present										1. Householder with relatives	
1	1	2018	1	11413646518	11	103	48	1: Male	1. Married, spouse present										3. Opposite-sex husband/wife/spouse	
1	1	2018	1	11413646518	11	104	17	2: Female	6. Never married										Child	
2	1	2018	1	28511493318	11	101	67	1	1. Married, spouse present										1. Householder with relatives	
2	1	2018	1	28511493318	11	102	63	2	1. Married, spouse present										1. Householder with relatives	
2	1	2018	1	28511493318	11	103	7	1	:										8. Grandchild	

BGSU



Center for **Family and Demographic** Research

Replicate Weights for 2018 SIPP

Stata:

- `svyset [pweight= wpfinwgt], brrweight(REPWGT1-
REPWGT240) fay(.5) vce(brr) mse`

BGSU



Center for Family and
Demographic Research

Conclusions

- 2018 SIPP offers a unique opportunity to describe and examine how economic well-being change across different family and social contexts
- The format of data is person-month. Thus, it would be better if users are familiar with how to use data in long format
- The sizes of the primary data and replicate weight data files of the Wave I of 2018 SIPP are very big. Users should use computers with enough RAMs or select necessary variables or observations when analyzing 2018 SIPP data
- 2018 SIPP over-samples low-income households. Thus, users should remember weighting their analyses
- SIPP website provides latest information about data, codebooks, user guides, notes, and corrections. Thus, SIPP website is the first place you should look before starting your SIPP analyses
- If you have any questions about 2018 SIPP, please feel free to contact me at wuh@bgsu.edu.