

Introduction to the 2021 Survey of Income and Program Participation (SIPP)

Hsueh-Sheng Wu
CFDR Workshop Series
April 10, 2023

BGSU



Center for Family and
Demographic Research

Outline

- What is SIPP?
- History of SIPP
- Study design of 2021 SIPP
- Content of 2021 SIPP
- Advantages and disadvantages of SIPP
- Sample studies using SIPP
- Navigating through the website of 2021 SIPP
- A look at 2021 SIPP data
- Weights for 2021 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)
 - inadequate 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 19 panels. The 2021 SIPP panel is the latest one and will cover a four-year time period from January 2021 to December 2025.
- 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; (3) changing the reference period of survey questions from four months to one year (4) use Event History Calendar (EHC) to facilitate respondent recall of residences, marriage and cohabitation, school enrollment, labor force participation, use of welfare programs, and health insurance.
- In 2018, Census Bureau further adopted the overlapping panel design in which respondents of different panels are interviewed in the same year.
- These new study design has been used in subsequent SIPP panels.

Study Design of 2021 SIPP

- Target population is U.S. non-institutionalized civilians aged 15 and over
- Sampling frame was created from files of 2010 Decennial Census, Master Address File (MAF), and U.S. Postal Service's Delivery Sequence
- A two-stage stratified sampling was used to select households. The first stage selected one or more contiguous counties (Primary Sampling Unites) from the sampling frame. The second stage divided 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 2021 SIPP (Cont.)

- United States Census Bureau (April 2023) Survey of Income and Program Participation: 2021 Panel Users' Guide. Retrieved from https://www2.census.gov/programs-surveys/sipp/tech-documentation/methodology/2021_SIPP_Users_Guide_AUG22.pdf
- Overlapping panel design of SIPP from 2018 to 2025
- Data collection for 2019 SIPP faced problems including insufficient geographic coverage for some primary sample units (PSUs), and lower-than-average unit response rates nationally. Subsequently, Census Bureau decided to focus on retaining 2018 SIPP respondents instead. Therefore, SIPP 2019 contains only one wave of data.

FIGURE 1-2. OVERLAPPING PANELS ILLUSTRATION

...	2018 SIPP	2019 SIPP	2020 SIPP	2021 SIPP	2022 SIPP	2023 SIPP	2024 SIPP	2025 SIPP	...
	2018 Panel								
		2019 Panel							
			2020 Panel						
				2021 Panel					
					2022 Panel				
						2023 Panel			
							2024 Panel		
								2025 Panel	

Study Design of 2021 SIPP (Cont.)

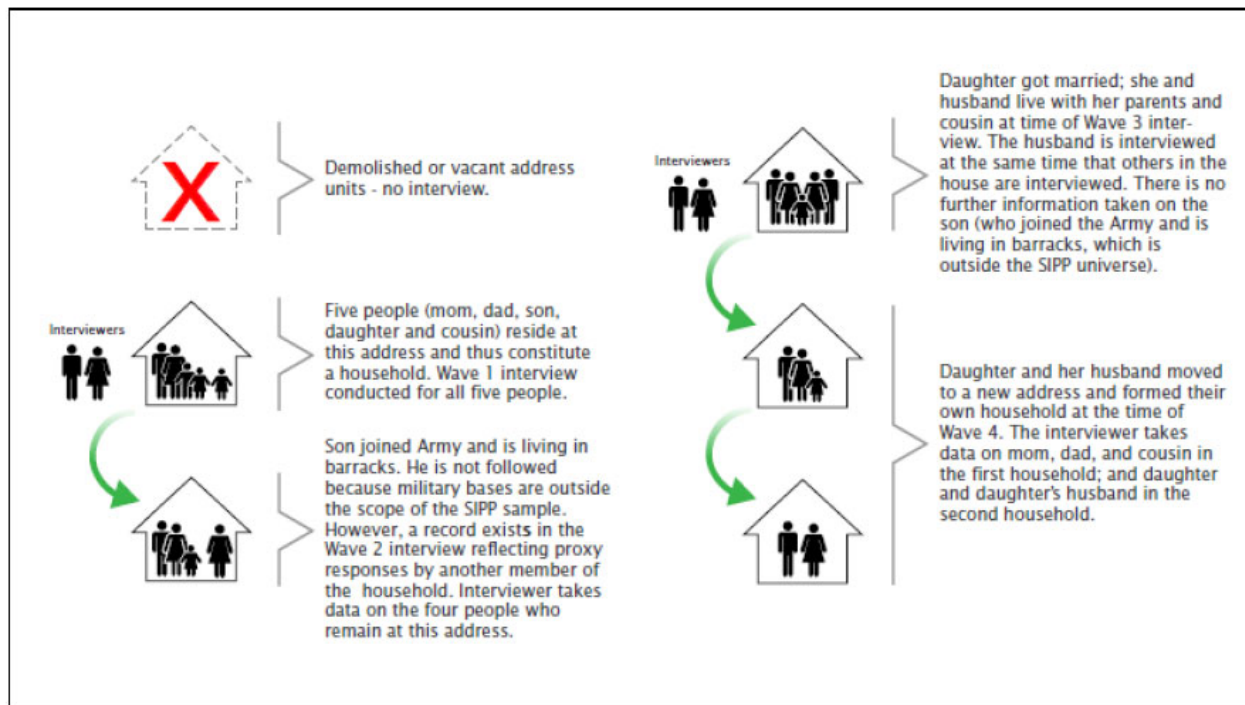
TABLE 1-2. SUMMARY OF THE 1984 TO 2021 SIPP PANELS

Panel	Date of First Interview	Date of Last Interview	Number of Wave 1 Eligible Households	Number of Waves
1984	Oct. 83	Jul. 86	20,897	9
1985	Feb. 85	Aug. 87	14,306	8
1986	Feb. 86	Apr. 88	12,425	7
1987	Feb. 87	May 89	12,527	7
1988	Feb. 88	Jan. 90	12,725	6
1989	Feb. 89	Jan. 90	12,867	3
1990	Feb. 90	Sep. 92	19,800	8
1991	Feb. 91	Sep. 93	15,626	8
1992	Feb. 92	May 95	21,577	10
1993	Feb. 93	Jan. 96	21,823	9
1996	April 96	Mar. 00	40,188	12
2001	Feb. 01	Jan. 04	50,500	9
2004	Feb. 04	Jan. 08	51,379	12
2008	Sept. 08	Dec. 13	52,031	16
2014	Feb. 14	Jun. 17	42,348	4
2018	Feb. 18	May. 21	44,870	4
2019	Mar. 19	Jul. 19	24,500	1*
2020	Feb. 20	Ongoing	22,241	Ongoing
2021	Feb. 21	Ongoing	14,624	Ongoing

Study Design of 2021 SIPP (Cont.)

- Following rules in 2021 SIPP: Respondents of the sampled households and people who later co-reside in the same households were followed over time

FIGURE 2-1. EXAMPLES OF SIPP FOLLOWING RULES



BGSU



Center for Family and Demographic Research

Source: United States Census Bureau (April 2023) Survey of Income and Program Participation: 2021 Panel Users' Guide. Retrieved from https://www2.census.gov/programs-surveys/sipp/tech-documentation/methodology/2021_SIPP_Users_Guide_AUG22.pdf

Content of 2021 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>	Economic Impact Payments (EIP)
Educational enrollment	Interest-earning assets	EITC and tax filing
Educational attainment	Other income-generating assets	Energy assistance
Family and household relationships	Other assets	General Assistance (GA)
Fertility	Retirement accounts	'Other' assistance
Language	<i>Liabilities:</i>	School meals
Marital status and marital history	Debts secured by assets	Social Security – child
Nativity, citizenship, and parent nativity	Unsecured debt	Social Security – self
Parent mortality	<i>Additional topics:</i>	Special Supplemental Nutrition Program for Women, Infants, and Children (WIC)
Residence	Rent and mortgage payments	Supplemental Nutrition Assistance Program (SNAP)
Veteran status	Utility payments	Supplemental Security Income (SSI)
<u>Employment</u>	<u>Health and Well-Being</u>	Temporary Assistance for Needy Families (TANF)
Commuting and work schedule	Adult well-being	Unemployment Compensation/Insurance
Earnings	Child care	Veterans benefits
Hours and pay changes	Child well-being	Workers' compensation
Job characteristics	Dependent care expenses	<u>Other Income</u>
Labor force participation	Disability	Annuity and life insurance retirement income
Parental leave	Food security	Disability income payments
Reasons for not working	Health care utilization and medical expenditures	Lump sum severance pay and retirement plan income
Time away without pay	Health insurance	Miscellaneous income
		Pension Income
		Retirement account withdrawals
		Support paid
		Support received
		Survivor income benefit

BGSU



Center for Family and Demographic Research

Source: United States Census Bureau (April 2023) Survey of Income and Program Participation: 2021 Panel Users' Guide. Retrieved from https://www2.census.gov/programs-surveys/sipp/tech-documentation/methodology/2021_SIPP_Users_Guide_AUG22.pdf

Advantages of SIPP

- Contain comprehensive data on many topical areas
- Use Event History Calendar to collect information on residences, marriage and cohabitation, school enrollment, labor force (including jobs, business, contingent work, unemployment, time not in labor force, commuting, industry, occupation, and work schedule), programs (GA, SNAP, SSI, TANF, WIC), health insurance (private, Medicare, Medicaid, military, other health insurance coverage)
- Contains information on how household members are related to each other and if they live in the same place
- Follow all original household members over time even when they move into a new household
- Generate annual and monthly estimates at both state and national levels
- Study how households and families influence individuals in various topical areas
- Examine longitudinal stability and change of households, families, and individuals of the SIPP panel and their interactions over the course of four years

BGSU

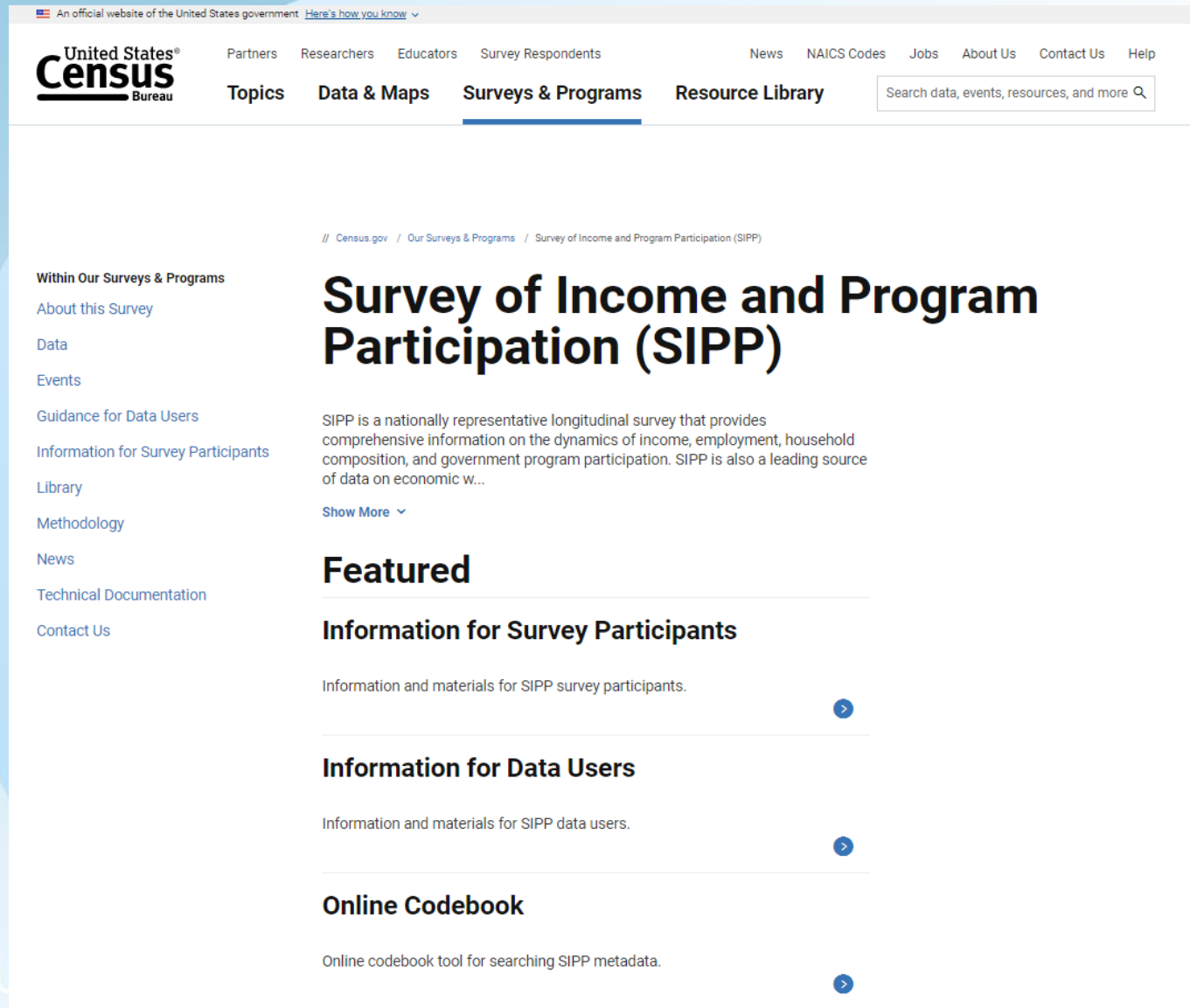
Disadvantages of SIPP

- 2021 SIPP follows respondents only for four years and is not adequate for long-term longitudinal analysis
- Respondents with very high income are top coded, making it less adequate to study respondents at the top of the income distribution
- The size of data file is big, and it is important to create data sets containing only variables needed for analysis.
- The data structure of 2021 SIPP may be difficult to grasp at first because variables can be measured at individual, family, or household levels and/or at monthly or annual level.
- 2021 SIPP provides many weight variables, and it may be challenging to choose the correct one for the analysis

Sample Studies Using SIPP

- Manning, Brown, and Stykes (2016) used data from the 2008 SIPP panel to compare the stability of same-sex and different-sex cohabiting couples' relationships. The study found that same-sex couples had lower rates of relationship stability than different-sex couples, with same-sex male couples being the least stable of all couple types examined.
- Shaefer and Cooney (2020) used data from the 2018 SIPP panel to analyze poverty dynamics in the US. The authors found that children, women, and Black and Hispanic individuals at higher risk of poverty than other groups. The study also found that poverty was often a persistent and long-term problem for many individuals and families, with relatively few able to escape poverty during the course of the panel.

Navigating through the Website of 2021 SIPP



The screenshot shows the United States Census Bureau website. At the top, there is a navigation bar with the following links: Partners, Researchers, Educators, Survey Respondents, News, NAICS Codes, Jobs, About Us, Contact Us, and Help. The main navigation menu includes Topics, Data & Maps, Surveys & Programs (which is highlighted), and Resource Library. A search bar is located on the right side of the navigation menu.

The main content area features a breadcrumb trail: // Census.gov / Our Surveys & Programs / Survey of Income and Program Participation (SIPP). The title of the page is "Survey of Income and Program Participation (SIPP)". Below the title, there is a brief description: "SIPP is a nationally representative longitudinal survey that provides comprehensive information on the dynamics of income, employment, household composition, and government program participation. SIPP is also a leading source of data on economic w...". A "Show More" link is provided below the description.

On the left side of the page, there is a sidebar menu titled "Within Our Surveys & Programs" with the following links: About this Survey, Data, Events, Guidance for Data Users, Information for Survey Participants, Library, Methodology, News, Technical Documentation, and Contact Us.

The "Featured" section contains three items:

- Information for Survey Participants**: Information and materials for SIPP survey participants. (Link icon: >)
- Information for Data Users**: Information and materials for SIPP data users. (Link icon: >)
- Online Codebook**: Online codebook tool for searching SIPP metadata. (Link icon: >)

BGSU



Center for Family and Demographic Research

Navigating through the Website of 2021 SIPP

The screenshot shows the '2021 SIPP Data' page on the United States Census Bureau website. The top navigation bar includes links for Partners, Researchers, Educators, Survey Respondents, News, NAICS Codes, Jobs, About Us, Contact Us, and Help. The main navigation menu features Topics, Data & Maps, Surveys & Programs (which is highlighted), and Resource Library. A search bar is located on the right side of the navigation bar.

On the left side of the page, there is a 'Within Data' section with links for Data Tools and Datasets.

The main content area features a breadcrumb trail: // Census.gov / Survey of Income and Program Participation Data / Survey of Income and Program Participation Datasets / 2021 SIPP Data / 2021 SIPP Data. Below this is the main heading '2021 SIPP Data' and social media sharing options for Facebook, Twitter, and LinkedIn.

The text below the heading states: 'The 2021 SIPP continued to face data collection complications due to the COVID-19 pandemic and a general lower-than-average unit response rate nationally. For technical documentation and more information about SIPP data quality, please visit the Technical Documentation page. For questions about SIPP data, visit the Contact Us page.'

Under the heading '2021 SIPP Data Technical Documentation', there is a list of documents with download icons and file sizes:

- 2021 SIPP Release Notes [[< 1.0 MB](#)]
- 2021 SIPP Users' Guide [[< 1.0 MB](#)]
- 2021 SIPP Instrument Specifications [[< 1.0 MB](#)]
- 2021 SIPP Crosswalks - 2020 to 2021 [[< 1.0 MB](#)]
- 2021 SIPP Crosswalks - 2021 to 2020 [[< 1.0 MB](#)]
- 2014-2021 SIPP Variable Name Change Summary [[< 1.0 MB](#)]

At the bottom of the page, there is a section for 'Online Codebook' with the text 'Online codebook tool for searching SIPP metadata.' and a right-pointing arrow icon.

BGSU



Center for Family and Demographic Research

Navigating through the Website of 2021 SIPP

United States
Census
BUREAU OF ECONOMIC ANALYSIS

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

SIPP Codebook

Showing 3319 of 3319 Variables [Reset](#) Sort by Variable Ascending

Browse by Variable Name or Topic Search FILTER BY:

	VARIABLE	TOPIC/SUBTOPIC	DESCRIPTION
+	SHHADID	ID Variables ID Variables	Household address ID. Used to differentiate households spawned from an original sample household.
+	SPANEL	ID Variables ID Variables	Panel year
+	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.
+	SWAVE	ID Variables ID Variables	Wave number of interview
+	PNUM	ID Variables ID Variables	Person number
+	MONTHCODE	ID Variables ID Variables	Value of reference month
+	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.
+	GVARSTR	ID Variables ID Variables	Variance pseudo stratum code. Strata formed for half sample variance estimation
+	WPFINWGT	ID Variables ID Variables	Final person weight
+	EAF1	Demographics Period Served in Military	September 2001 to present
+	EAF2	Demographics Period Served in Military	August 1990 to August 2001
+	EAF3	Demographics Period Served in Military	May 1975 to July 1990
+	EAF4	Demographics Period Served in Military	Vietnam Era
+	EAF5	Demographics Period Served in Military	February 1955 to July 1964
+	EAF6	Demographics Period Served in Military	Korean War
+	EAF7	Demographics Period Served in Military	January 1947 to June 1950
+	EAF8	Demographics Period Served in Military	World War II

BGSU

A Look at 2021 SIPP Data

- The primary data file has 670,678 records, 5,214 variables
- For estimating frequencies or proportions of variables, use final person weight variable (i.e., wpfinwgt)
 - When SIPP data from a particular month are analyzed, use the final person weight variable from that particular month
 - When SIPP data from a calendar are analyzed, use the final person weight variable from December of that year
 - When the whole panel of SIPP data from analyzed, combine and average the final person weight variables from December of each year
- For accurately estimating the standard errors of the estimates, the replicate weight variables should be used. SIPP provides different sets of replicate weights for calendar years, multi-years, and SIPP panels

BGSU



Center for Family and Demographic Research

A Look at 2021 SIPP Data (Cont.)

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

A Look at 2021 SIPP Data (Cont.)

A	E	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
<u>Example</u>	<u>monthcode</u>	<u>spanel</u>	<u>swave</u>	<u>ssuid</u>	<u>shhadid</u>	<u>pnum</u>	<u>tage</u>	<u>esex</u>										<u>ems</u>		<u>erelrpe</u>
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

Weights for 2021 SIPP

Use personal weights in Stata:

```
tab1 ems if spanel ==2021 [iweight=wpfinwgt]
```

```
tab1 ems if monthcode==12 [iweight=wpfinwgt]
```

```
tab1 ems if spanel ==2021 & monthcode==12  
[iweight=wpfinwgt]
```

Use replicate weights in Stata:

```
svyset [pweight= wpfinwgt], brrweight(repwgt1-repwgt240)
```

```
fay(.5) vce(brr) mse
```

```
svy: mlogit ems tage
```

BGSU

Conclusions

- 2021 SIPP provide a unique opportunity to estimate the composition of a household cross-sectionally and examine how changes in composition over time may affect individuals' economic well-being.
- The 2021 SIPP data is in person-month format. As a result, it is preferable if users are familiar with how to use data in long format.
- The 2021 SIPP data files are large in size. When analyzing data, users should use computers with sufficient RAM or select only the variables or observations that are required.
- Low-income households are over-represented in the 2021 SIPP. As a result, users should remember to weight their analyses using final personal weight variables or replicate weight variables.
- SIPP's website contains webinar videos and up-to-date information on data, codebooks, user guides, notes, and corrections. As a result, the SIPP website is the first place you should look before beginning your SIPP analysis.

- If you have any questions about using 2021 SIPP, please feel free to contact me at wuh@bgsu.edu.