

**SUPPORTING STATEMENT**  
**U.S. Department of Commerce**  
**U.S. Census Bureau**  
**2018 Survey of Income and Program Participation Panel**  
**OMB Control No. 0607-XXXX**

**B. Collection of Information Employing Statistical Methods**

**1. Universe and Respondent Selection**

The SIPP respondent universe is the civilian, noninstitutionalized population based on the 2010 decennial census, which contains approximately 303.5 million individuals. The SIPP uses a multistage stratified sample of the U.S. civilian noninstitutionalized population. The first stage involves the division of the United States into groups of counties called the PSUs, which are assembled into homogeneous groups called strata. Two PSUs are then selected from each stratum. The second stage involves selection of units within the selected PSUs.

Within each PSU, living quarters (LQ) are systematically selected from lists of addresses prepared for the 2010 Decennial Census. Other sampling techniques are used to represent new construction and group quarters. Low-income households were over sampled from the lists of addresses prepared for the census. The SIPP sample comes from the 2010 sample redesign. The 2018 SIPP Panel sample is currently planned for approximately 53,000 Wave 1 designated LQs, it will yield approximately 42,400 Wave 1 designated occupied LQs at the time of interview in 2018, of which approximately 31,800 will be interviewed. If not fully funded for a sample of 53,000 households, approximately 35,000 Wave 1 designated LQs will be sampled, and will yield approximately 28,000 Wave 1 designated occupied LQs at the time of interview in 2018, of which approximately 20,000 will be interviewed.

Each household contains an average of 2.1 eligible respondents (aged 15 years and older); therefore, the planned 2018 SIPP Panel sample for approximately 53,000 Wave 1 designated LQs, it should contain approximately 66,800 survey respondents in Wave 1. If only 35,000 Wave 1 designated LQs are sampled, they should contain approximately 42,000 survey respondents. The expected response rate for this panel is 75-80%.

**2018 SIPP Sample Design**

The SIPP respondent universe is the civilian, noninstitutionalized population based on the 2010 decennial census, which contains approximately 303.5 million individuals. The SIPP uses a multistage stratified sample of the U.S. civilian noninstitutionalized population. The first stage involves the definition and division of the United States into groups of counties called the PSUs, which are assembled into homogeneous groups called strata. The second stage involves selection of units within the PSU.

The 2018 SIPP Panel is the second sample for the SIPP to be fielded from the 2010 redesign. There are 820 selected PSUs in the 2010 redesigned SIPP. The selected PSUs in the 2010 SIPP sample design cover both urban and rural areas of the United States. PSU definitions and address lists, are all based on the 2010 decennial census. PSUs are formed from one or more contiguous counties. Larger populated PSUs (>100,000 HUs) are identified as self-representing (SR) PSUs, while the remaining PSUs are identified as non-self representing (NSR). SR PSUs are in the SIPP sample with certainty while the NSR PSUs are stratified and selected with a probability proportionate to their size. During the stratification process, NSR PSUs are grouped together within the same state to form strata. During the PSU selection process, two NSR PSUs are selected from each stratum. There are 252 SR PSUs and 434 NSR PSUs in sample for the 2018 SIPP.

Within each selected PSU, living quarters (LQ) are systematically selected from lists of addresses prepared for the 2010 Decennial Census. The universe of addresses within the sample PSU is divided into two strata, one with a higher concentration of low-income households and the other with a lower concentration of low income households. In the sampling strata, low income is defined based on the poverty thresholds at the national level for families of certain size and age compositions. For example, the low-income household cutoff for one person with no related individual is \$18,700, and the low income household cutoff with three people is \$28,600. Cut-offs are adjusted by the Consumer Price Index, and applied to the sample frame by household size and also by age of householder for one and two person households. Addresses are sorted by geographic and demographic variables and a systematic selection of units is taken from each stratum. A 28% higher sampling rate is used in the stratum with the higher concentration of low-income households, thereby resulting in an oversample of low-income households. Oversampling occurs to the extent that the rise in the variance for the estimate for persons 55 and over is not increased by more than 5%.

The frame for the SIPP is the Master Address File (MAF), which is maintained by the U.S. Census Bureau and is the source of addresses for the American Community Survey, other demographic surveys, and the decennial census. The MAF is updated using the U.S. Postal Service's Delivery Sequence File and various automated, clerical, and field operations.

The addresses selected for the 2018 SIPP sample will not be eligible to be selected for another Census Bureau demographic survey (CPS, SCHIP, NCVS, CE, and AHS) before 5 years after the last SIPP interview.

### **2018 SIPP Weighting**

Each year, weights are created for every month and the calendar year at the person-level. Starting in the second year, panel weights at the person-level are produced along with monthly and calendar year weights. Each weight is calculated as the product of three components: the base weight, the household noninterview adjustment factor, and the second stage adjustment factor. The noninterview adjustment is calculated for each noninterview cell based on the following formula: Adjustment factor =  $\frac{\text{sum (interviewed weights + noninterview weights)}}{\text{sum (interviewed weights)}}$ , for each cell. These factors are applied to the base weights. There were 512 noninterview household cells in Wave 1. The second stage weights are calculated as a ratio adjustment of the sum of noninterview weights to the population controls and applied to each cell in five dimensions. The second stage weighting procedure consists of raking (with 4000 maximum iteration and 250 tolerance level), cell collapsing, and husband/wife equalization. This procedure is done for each reference month separately.

Each monthly weight for the SIPP 2018 Panel is produced based on the SIPP survey universe corresponding to that month. Therefore, the controls (benchmark population estimates) for second stage raking for each monthly weight are those for the corresponding reference month. Each calendar year weight is produced based on the SIPP survey universe in December of that year. Therefore, the controls for second stage raking for the calendar year weights are those for December of that calendar year. Meanwhile, the panel weight is based on December of the first year, so the controls for second stage raking for the panel weights are those for December 2017.

## **2. Procedures for Collecting Information**

In sample households, all people 15 years old and over will be interviewed using regular proxy-respondent rules. Each household member 15 years old or older who is present at the time of interview should respond for himself/herself. If a 15+ person is physically or mentally incompetent, a proxy respondent will be selected. Also, a proxy respondent for a person absent at the time of the interview will be selected. Any knowledgeable household member who is 15 years or older may serve as a proxy.

Instead of one large sample that we would interview over a five-month interview period, we propose to draw five smaller, independent samples, interviewing and closing out one of these samples each month of the interview period, which would run from February through June. We would draw these independent samples from the same geographies, so that staffing needs would be consistent in level and location from month to month. FRs (Field Representatives) would get case assignments each month and would close them all out of the end of that month. In other words, we would treat each of the five smaller samples independently during data collection, but combine them for data processing and file release as the full wave SIPP data file.

If the 2018 SIPP Panel sample is started with 53,000 households, approximately 31,800 households are expected to have completed interviews for the initial wave. We estimate that each household contains 2.1 people aged 15 and above, yielding approximately 66,800 person-level interviews per wave in this panel. Interviews take approximately 60 minutes per adult on average, consequently the total annual burden for 2018 SIPP interviews will be 66,800 hours per year in FY 2018, 2019, 2020, and 2021.

If the 2018 SIPP Panel sample is started with 35,000 households, approximately 20,000 households are expected to have completed interviews for the initial wave. We estimate that each household contains 2.1 people aged 15 and above, yielding approximately 42,000 person-level interviews per wave in this panel. Interviews take approximately 60 minutes per adult on average, consequently the total annual burden for 2018 SIPP interviews will be 42,000 hours per year in FY 2018, 2019, 2020, and 2021.

At the most conservative, 35,000 household sample size, we expect the minimum detectable differences between the 2018 SIPP Panel and the 2014 SIPP Panel monthly participation rates to be approximately 0.9 for TANF and SSI, 1.9 for Food Stamps and WIC, and 2.2 for Medicaid at the 10% level of significance.

### **3. Methods to Maximize Response**

In all SIPP Panels, we make special efforts to minimize non-interviews. In each wave, every household in the active sample receives an advance letter that explains the purpose of the survey and why the household's cooperation is important. In the 2018 SIPP Panel each household will be given one of four brochures (Attachments E-H) that contains specific information about the SIPP and how it specifically relates to them in one of the following four areas: (1) Health (insurance, expenditures, etc.); (2) Families with children (child care, child well-being, etc.); (3) Labor force (retirement,

employment, unemployment, etc.), and (4) Wealth (income, retirement accounts, etc.). Each household will also receive an eye-catching glossy “SIPP in the News” postcard (Attachment I) which has highlights of when SIPP has been used in major news sources such as the Washington Post. SIPP also distributes a SIPP Fact Sheet (Attachment D) which gives examples of how SIPP data are used in various government programs. For Type A refusal households, standard procedures include additional visits to the household by another FR or if needed, a Supervisory Field Representative (SFR) to convert the household response.

In the 2014 SIPP panel experiment, we evaluated conditional post-paid incentives and the development of model-based incentives. The 2014 SIPP incentive experiment showed that a \$40 conditional incentive was effective in increasing the response rate by approximately 3 percentage points in all waves. Therefore, we do plan to continue utilizing model-based incentives and distributing \$40 incentives in the 2018 SIPP Panel. A report that summarizes the 2014 SIPP panel incentive experiment is forthcoming.

We plan to divide the panel into two incentive groups and a control group to monitor the effectiveness of the incentives over time. Group 1 is the control group; households in this group will not be eligible for an incentive in any wave of the 2018 panel. For Wave 1, incentives will be assigned randomly to households in incentive Group 2 and will be eligible to receive an incentive in Wave 1 and later waves. In subsequent waves, households in Group 2 will receive incentives of \$40 based on a propensity model that considers the effectiveness of the incentive for generating an interview. This assignment plan for incentives will help to increase the response rate among households where the absence of the incentive would lead to differential attrition. The use of incentives in this model-based framework will also lower costs since we would not focus incentives on households that would be over-represented in the absence of an incentive. Additionally, we will continue to evaluate the incentive modeling and develop specifications that incorporate ongoing work to utilize responsive and adaptive design to prioritize cases in interviewers’ workloads.

The third group of respondents (Group 3) will receive a \$40 incentive in Wave 1 and all subsequent waves. All incentives are conditional on completing the interview. We will inform households in both the advance letter and the introduction to the survey of their eligibility for an incentive.

Inputs for the propensity model for Group 2 will come from the Wave 1 responses, Wave 1 contact data, sample frame data, data linkable to the sample frame, and results from comparing Groups 1 and 2 with Group 3. We will evaluate what characteristics of

households seem to make them more or less likely to complete interviews, how they contribute to the eventual sample representation, and how likely they are to respond to incentives. Using this knowledge, we can design a model for Waves 2+ that optimizes our distribution of incentives to Group 2 in a way that maximizes the return on each dollar spent, reduces non-response bias, and improves sample representativeness.

For all waves, we will distribute the incentives centrally, from our National Processing Center (NPC). This centralized distribution eliminates any discretion on the part of the field representatives, ensuring that only eligible households are given (or promised) incentives. This control is necessary to ensure the success of the propensity modeling experiment. We plan to mail the debit cards containing the incentives on a weekly basis. That is, as we receive completed interviews from eligible households, we will send a list of these households to the NPC, and they will mail letters containing the PIN information and then the debit cards. Splitting the mailings this way allows us to avoid the additional expense of sending the debit cards via a signature-required service such as FedEx.

#### **4. Contacts for Statistical Aspects and Data Collection**

The Census Bureau will collect and process these data. Within the Census Bureau, please consult the following individuals for further information:

##### **Sample Design**

Tracy Mattingly      Lead Scientist (SIPP)  
Demographic Statistical Methods Division  
301-763-6445

Mahdi S Sundukchi      SIPP Survey Design Lead  
Demographic Statistical Methods Division  
301-763-4228

##### **Data Content**

Matthew Marlay      Assistant Survey Director (SIPP)  
Associate Director for Demographic Programs  
301-763-5083

##### **Data Collection and Tabulation**

Jason Fields                      SIPP Survey Director  
Associate Director for Demographic Programs  
301-763-2465

### **Attachments**

- A. Wave 1 SIPP (2018) Instrument Items Booklet
- B. SIPP-105(L1)2018–Advance Letter
- C. SIPP-105(L3)2018–Advance Letter with \$40 incentive
- D. SIPP Fact Sheet
- E. SIPP Brochure: Health Insurance
- F. SIPP Brochure: Families with Children
- G. SIPP Brochure: Labor Force and Employment
- H. SIPP Brochure: Wealth
- I. SIPP in the News postcard
- J. SIPP-106(L1)2018 Thank You Letter
- K. SIPP-106(L2)2018 Thank You Letter (Incentive)
- L. Incentives FCSM Evaluation
- M. Federal Register Notice