

SUPPORTING STATEMENT
U.S. Department of Commerce
U.S. Census Bureau
Survey of Income and Program Participation Panel
OMB Control No. 0607-1000

A. Justification

1. Necessity of Information Collection

The U.S. Census Bureau requests authorization from the Office of Management and Budget (OMB) to conduct the Survey of Income and Program Participation (SIPP). The SIPP is authorized by Title 13, United States Code, Sections 141 and 182.

This request is for the SIPP collection starting February 2021. The Census Bureau plans to conduct the SIPP using an overlapping sample design. The Census Bureau's SIPP Computer-Assisted Personal Interviewing (CAPI) will use an Event History Calendar (EHC) interviewing method and a 12-month, calendar-year reference period. This is the same approach as the 2014 and the 2018 SIPP Panels. The SIPP instrument content and post-production processing will remain similar to that of the 2018 SIPP. See Attachment A, the SIPP 2021 Instrument Items Booklet, for the interview questions.

The main objective of the SIPP is to provide accurate and comprehensive information about the income and program participation of individuals and households in the United States. The survey's mission is to provide a nationally representative sample for evaluating: 1) annual and sub-annual income dynamics; 2) movements into and out of government transfer programs; 3) family and social context of individuals and households; and 4) interactions among these items. A major use of the SIPP is to evaluate the use of, and eligibility for, government programs and to analyze the impacts of modifications to those programs. The SIPP collects detailed information on cash and non-cash income (including participation in government transfer programs) once per year. The current SIPP panel continues to reduce the cost of collection, improve accuracy, increase relevance and timeliness, reduce respondent burden, and increase accessibility.

Providing the same, or better, quality data at a reduced burden to respondents is a high priority for the Census Bureau and for the SIPP program. To accomplish this, the Census Bureau uses an EHC-based instrument to gather SIPP data. The EHC allows recording dates of events and spells of coverage, and provides measures of monthly

transitions of program receipt and coverage, labor force transitions, health insurance transitions, and others.

While the SIPP samples households, after a successful first interview, the SIPP longitudinally follows interviewed household members aged 15-years and older who move from the prior wave household. After the initial interview, future waves interview all household members who reside with those Wave 1 interviewed household members, and SIPP incorporates dependent data to help improve data quality and reduce burden. Dependent data is information collected from the prior wave interview brought forward to the current interview.

New Sample Design and Overlapping Panels

To increase the ability to respond to changing budget constraints, the SIPP modified the sample design and data collection strategy starting in 2018. We conducted the 2018 SIPP collection under the scenario of selecting a sample of 53,000 cases and following the households over a period of 4 years, with no sample replacement over the life of the Panel.

Due to a high noninterview rate in the SIPP 2018 Panel, we decided to consider an overlapping panels sample design. This ensured there would be enough data to continue to produce accurate cross-sectional and longitudinal estimates. In the SIPP 2018 Panel, of the original 53,000 designated housing units, 45,000 were eligible, and we interviewed 26,000 units (58.5%) in Wave 1. If we assumed 25% noninterview for each subsequent wave, we would expect to have 32.8% or less overall interviewed cases (including the spawn cases) by the end of the panel. This low interview rate would lead to a larger variability on the key estimates, and we may not have met the minimum Census quality standard rules for publications. Additionally, we could have faced technical issues in the weighting, requiring much more cell collapsing. For these reasons, we concluded that an overlapping panels design would provide better cross-sectional and longitudinal estimates.

Starting with the 2019 survey year, we proposed and started using a sample design scenario of overlapping panels where we would supplement sample each year, maintaining the level of interviewed cases in each subsequent year to keep the total pool of interviewed cases constant, thus mitigating sample loss. This meant there would be a new Wave 1 sample introduced each year. The new sample design has SIPP interviewing each independent sample in its entirety over the three to four-month interview period of February to June of each year. All panels interviewed during the

survey year report information for the same reference period, which is January through December of the previous year.

The Census Bureau uses Computer Audio-Recorded Interview (CARI) technology. CARI is a tool available during data collection to capture audio along with response data. With the respondent's consent, a portion of each interview is recorded unobtrusively and both the sound file and screen images are returned with the response data to Census Headquarters for evaluation. Census staff may review the recorded portions of the interview to improve questionnaire design and for quality assurance purposes.

2. Needs and Uses

The 2018 SIPP collects information about a variety of topics including demographics, household composition, education, nativity and citizenship, health insurance coverage, Medicaid, Medicare, employment and earnings, unemployment insurance, assets, child support, disability, housing subsidies, migration, Old-Age Survivors and Disability Insurance (OASDI), poverty, and participation in various government programs like Supplemental Nutrition Assistance Program (SNAP), Supplemental Security Income (SSI), and Temporary Assistance for Needy Families (TANF). The SIPP sample is nationally representative, with an oversample of low-income areas, in order to increase the ability to measure participation in government programs.

The SIPP program provides critical information necessary to understand patterns and relationships in income and program participation. It will fulfill its objectives to keep respondent burden and costs low, maintain high data quality and timeliness, and use a refined and vetted instrument and processing system. The SIPP data collection instrument maintains the improved data collection experience for respondents and interviewers, and focuses on improvements in data quality and better topic integration.

Starting in 2019, the Census Bureau and the Social Security Administration (SSA) entered into a joint agreement where both agencies support the SIPP program by contributing resources to add, process, review, and maintain additional content on marital history, parental mortality, retirement and pension, and disability. These items represent a reduced set of content that had been included in past SIPP topical modules and most recently in the 2014 SIPP SSA Supplement. This joint agreement started in September 2019 and goes until September 30, 2023.

The SIPP instrument is currently written in Blaise and C#. It incorporates an EHC design to help ensure that the SIPP will collect intra-year dynamics of income, program participation, and other activities with at least the same data quality as earlier panels. The EHC is intended to help respondents recall information in a more natural “autobiographical” manner by using life events as triggers to recall other economic events. For example, a residence change may often occur contemporaneously with a change in employment. The entire process of compiling the calendar focuses, by its nature, on consistency and sequential order of events, and attempts to correct for otherwise missing data.

Since the SIPP EHC collects information using this “autobiographical” manner for the prior year, due to the coronavirus pandemic, select questions were modified to include answer options related to the pandemic as well as adding new questions based largely on the content in the Household Pulse Survey and the Current Population Survey pertaining to the pandemic. For instance, we adjusted the question regarding being away from work part-time to include being possibly furloughed due to coronavirus pandemic business closures. We also added new questions to collect information on whether the respondent receive any stimulus payments.

Information quality, as described by the Census Bureau’s Information Quality Guidelines, is an integral part of the pre-dissemination review of information released by the Census Bureau. Information quality is essential to data collections conducted by the Census Bureau and is incorporated into the clearance process required by the Paperwork Reduction Act.

3. Use of Information Technology

The survey is administered using CAPI and CARI methodologies. The Census Bureau field representatives (FRs) collect the data from respondents using laptop computers and transmit to the Census Bureau Headquarters via high-speed modems. Automation significantly enhances our efforts to collect high quality data with skip instructions programmed into the instrument and information obtained in earlier interview segments fed back to the respondent. Response burden is minimized by incorporating design features that make it easier to collect and record respondent information. Therefore, screening questions and lead-in questions are built into the automated instrument to skip respondents out of sections of the questionnaire that are not relevant or applicable.

Review of current internet instruments and analysis from an internet field test conducted by the SIPP Methods Panel (2000) suggest that using the internet as a data-

collection mode for a long, complex demographic survey such as SIPP is currently not a feasible option for the SIPP program. The SIPP automated instrument contains many complicated skip patterns and roster related components. While the Methods Panel findings are now twenty-years old, they remain salient. The public is not expected to be tolerant of long, complex, self-administered internet data collections. Major revisions to design and to content would be necessary to shift to an online-data collection. The costs of converting a complex questionnaire such as SIPP to an online survey far outweigh the benefits even in a multi-mode environment. Further the SIPP design for collecting information about transfer programs has had some comparative advantage over other household surveys, with relatively lower levels of underreporting for most programs (Meyer et al, 2015)¹. The SIPP program continuously reviews options for simplification of data collection, options for multi-mode data collection, and ways to reduce respondent burden.

4. Efforts to Identify Duplication

The demographic data collected in the SIPP must be collected in conjunction with the labor force and program participation data in order for the information to be most useful; therefore, although we collect demographic data in conjunction with almost all Census Bureau surveys, we need to continue its present collection in the SIPP. The SIPP overlaps categories of content (Labor Force, Health Insurance, Housing, Programs, Medical Expenditures, etc.) found in other federal surveys. However, none of these other surveys include data for all of these topics and capture intrayear dynamics and monthly data. The SIPP presents unique value and the ability to calculate measures of monthly eligibility and participation for transfer programs, as well as uniquely detailed family and household dynamics. SIPP also provides a platform for calculating and evaluating composite constructs like supplemental poverty. It enables broad uses for, and integration of, administrative data. So, while other data sources can provide several of the constructs collected in SIPP, the SIPP has the unique ability to support integrated analyses of dynamics, a unique platform for the integration of administrative data, and the development of model-based estimates.

Administrative data from the Social Security Administration, the Internal Revenue Service, state benefit programs, housing data, and others are available to the SIPP program to facilitate the validation and evaluation, but perhaps more importantly to improve data quality in collection and in data processing. As examples, the SIPP is able to leverage administrative inputs to facilitate interviews with movers, to help guide adaptive design decisions to prioritize work in data collection, to supplement

¹ Meyer, B. D., Mok, W. K. C., & Sullivan, J. X. (2015). Household Surveys in Crisis. *Journal of Economic Perspectives*, 29(4), 199-226. doi:10.1257/jep.29.4.199

imputation models bringing additional information to bear in the assignment of missing values, and to make corrections to misreported information. No other current data source is available which provides as comprehensive a set of statistics for analysis as described above in Section 2.

5. **Minimizing Burden**

The Census Bureau uses appropriate technology to keep respondent burden to a minimum. Examples of technology used to minimize respondent burden include: use of appropriate screening and lead-in questions that serve to skip respondents out of sections of the CAPI instrument that are not relevant or applicable to them; use of flash cards to aid respondents with multiple response categories; and the arrangement of questions and sections of the CAPI instrument that facilitate the flow of administration from one topic area to another.

The SIPP instrument also generates a large amount of paradata, administrative instrument data that are not seen by the interviewers nor the respondents. We will use the paradata to evaluate the survey in several ways. First, we plan to run a number of comparisons related to interview timing – for example, how much longer are adults’ interviews compared to children’s, and how different are proxy interviews from self-reports? Additionally we intend to use paradata to evaluate differences that are observable and show effects by interviewer to improve training and instrument design.

The paradata will also provide us with metrics that allow us to evaluate the respondent burden and to produce better data collection estimates. For example, we will know the average number of questions asked during each interview, allowing us to pinpoint content areas that we could streamline or change during the research panel. We also know how many visits to a household it takes to get a completed interview, so we can use this statistic to estimate future data collection costs more precisely.

The paradata provides information we can use to improve the overall survey or identify errors in the data collection instrument. This includes reviewing item-level don’t know and refusal rates, as well as particularly time-consuming items. Field representatives have the ability to enter notes, both at the item level and at the case level. We use these notes to identify and correct a number of issues in the instrument, and regularly follow-up to get feedback from field staff and SIPP stakeholders.

In addition to these evaluation tools, the SIPP instrument will continue to utilize the CARI Interactive Data Access System (CARI System), an innovative, integrated, multifaceted monitoring system that features a configurable web-based interface for

behavior coding, quality assurance, and coaching. This system assists in coding interviews for measuring question and interviewer performance and the interaction between interviewers and respondents. By recording the household interviews, supervisors will have the ability to select recordings for the supplemental observation and coaching of interviewers. As the recordings are used during and after data collection, the Census Bureau will use the recordings to evaluate the quality of the interviews by the FRs. These recordings can be used to not only assure the quality of the SIPP data, they can be used to help update training materials for better future data collection.

Finally, for a small subset of characteristics, and for a subset of sample areas, we will have access to administrative record data. Administrative record data are used to improve modeling and imputation of missing data in addition to continued research and evaluation. The administrative record data are also used as an external data source to assess the accuracy and validity of the survey estimates.

6. Frequency of Collection

The SIPP interviews respondents annually, using the previous calendar year as the reference period. One possible consequence of the one-year reference period is the possibility of increased recall difficulties for respondents. Use of the EHC methodology of interview, however, should help to alleviate these issues by linking respondents' memories to significant life events. See earlier explanation in Section 2.

7. Special Circumstances

There are no special circumstances associated with this clearance request.

8. Consultations Outside the Agency

The SIPP program has always relied upon and valued the input of our stakeholders outside the Census Bureau. Input from partner agencies and OMB continue to be valuable and welcome guidance for the content and procedures for the SIPP.

Additionally, the Census Bureau has continued involvement with the Committee for National Statistics (CNSTAT) at the National Academies of Science, and other interested representatives from policy, research, and government. The Census Bureau has continued to work actively to assure that the SIPP stakeholders' interests and

priorities are represented, that the content and procedures to collect SIPP data are appropriate, and the duplication among surveys are minimized to the extent possible.

We published a notice in the *Federal Register* on May 19, 2020, Vol. 85 FR 29924, pages 29924-29925, inviting public comment on our plans to submit this request (Attachment J). We received one comment asking for the full OMB package materials when available and one comment about the necessity of the survey that was not relevant.

A notice also appeared in the Federal Register on October 30, 2020, Volume 85, page 68839, asking for comments on this collection request. The notice inadvertently referred to the 2018 SIPP. The correct reference is the 2021 SIPP.

9. Paying Respondents

The SIPP does not have an incentive program therefore no payments are provided to respondents.

10. Assurance of Confidentiality

The U.S. Census Bureau is required by law to protect all respondent information. The Census Bureau is not permitted to publicly release responses in any way that could identify an individual or household. We are conducting this survey under the authority of Title 13, United States Code, Sections 141 and 182. Federal law protects respondent privacy and keeps all answers confidential under Title 13, United States Code, Section 9. Per the Federal Cybersecurity Enhancement Act of 2015, all respondent data are protected from cybersecurity risks through screening of the systems that transmit the data.

SIPP respondents will be informed of the confidentiality of their responses and that this is a voluntary survey by an annual letter from the Director of the Census Bureau that will be sent to all participants in the survey in advance of the interview (Attachments B and C). After completion of the SIPP survey, a thank you letter is sent to all responding households (Attachment D).

11. Justification for Sensitive Questions

The sources of income and assets are among the kinds of data collected and possibly considered of a sensitive nature. The Census Bureau takes the position that the

collection of these types of data is necessary for the analysis of important policy and program issues, and consequently has structured the questions to lessen their sensitivity.

12. Estimates of Annualized Respondent Hour and Cost Burden

Based on our experience with prior SIPP collection information, the burden estimates for the future SIPP are as follows:

12a. SIPP ESTIMATED ANNUALIZED BURDEN HOURS SUMMARY

Table 1: For a 53,000 Household Sample

	Expected Number of Respondents	Number of Responses per Respondent	Average Burden Per Response (in Minutes)	Total Burden Hours
Interview	70,560	1	63	74,088
Total	70,560	1	63	74,088

For a 53,000 household sample, we estimate that each household contains 2.1 people aged 15 and above, yielding approximately 33,600 person-level interviews in 2021. We estimate that completing the SIPP interview take approximately 63 minutes per adult on average, consequently the total annual burden for SIPP interviews will be 74,088 hours per year.

12b. 2018 SIPP PANEL ESTIMATED ANNUALIZED BURDEN COSTS

Table 2: For a 53,000 Household Sample

	Total Burden Hours	Hourly Wage Rate ²	Total Respondent Costs

² For individuals, the wage rate is \$29.37 per hour (for June 2020). This is based on the average hourly earnings for all employees on private nonfarm payrolls as reported by the Bureau of Labor Statistics (<https://www.bls.gov/news.release/empsit.t19.htm>).

Respondents	70,560	\$29.37	\$2,072,347
Total	70,560	\$29.37	\$2,072,347

13. Estimate of Cost Burden

There are no direct costs to respondents participating in the survey other than the time involved in answering the survey questions.

14. Cost to Federal Government

The data collection costs of the SIPP are approximately \$34,450,000 for 53,000 households (\$650 case for a 4-month data collection period) each year. That amount is included in the estimate of total costs to the federal government of the Census Bureau's current programs supplied to OMB. Items included in the cost of data collection are: instrument review, printing of materials, hiring and training for field representatives, data collection, interview monitoring, respondent engagement, initial data review, overhead, and support staff for the Field Division.

15. Reason for Change in Burden

The SIPP is being submitted as a revised collection; therefore, change in burden will occur. For this collection, we used the 2018 SIPP paradata to estimate the total burden hours for the existing SIPP questions. Additionally, for the added SSA content and the coronavirus pandemic modifications and additions, we used mock interview times to estimate the additional burden.

16. Project Schedule

The SIPP Advance Letters will be mailed prior to interviewing with data collection occurring from early February to June of 2021. We will release public-use data products on a schedule to be determined.

We will evaluate the survey using both collected data and paradata. As each subject-matter area is evaluating its content, we expect that in addition to comparing the results from prior SIPP collection periods, where possible they will also compare the results to

those from other surveys, such as the American Community Survey (ACS) or the Current Population Survey (CPS). While we expect some variation in estimates from the different surveys due to sample size, survey universe, etc., we will be able to tell whether SIPP's results are broadly in-line with those from other surveys.

17. Display of OMB Approval Information

The OMB control number is displayed in the advance letter that will be sent to eligible households before each interview. We request not to display the expiration date so we can reuse materials.

18. Exceptions to the Certification

There are no exceptions to the certification.