

**REQUEST FOR APPLICATIONS TO SUMMER WORKSHOP**

**CONDUCTING RESEARCH USING THE SURVEY OF INCOME AND PROGRAM PARTICIPATION (SIPP)**

**ADVANCED WORKSHOP ON THE SIPP SYNTHETIC BETA (SSB)**

**AUGUST 8-11, 2016**

**Lead Instructors: Lori Reeder & Holly Monti, US Census Bureau**

Co-Organizers:

Gary Benedetto, U.S. Census Bureau & H. Luke Shaefer, University of Michigan

With support from Lars Vilhuber, Cornell University

**APPLICATION DEADLINE:**

The Survey Research Center at the Institute for Social Research at the University of Michigan and the United States Census Bureau invite applications to participate in a four-day workshop, August 8-11, 2016, in Ann Arbor, Michigan. This advanced workshop will introduce participants to the use of the Survey of Income and Program Participation Synthetic Beta (SSB) and provide hands-on applications to prepare them to conduct their own SSB-based research.

The SIPP Synthetic Beta (SSB) is a Census Bureau data product that integrates person-level micro-data from a household survey (SIPP survey data) with W-2 earnings and OASDI benefits data. Census has synthesized the data to preserve the underlying covariate relationships between variables while protecting respondent confidentiality. Unlike the original, administrative data, the SSB is publicly available.

The Census Bureau offers all SSB users the opportunity to submit their programs for validation on the internal, confidential version of these data. As part of this workshop, participants will learn how to prepare programs for submission to the validation process and how to request the release of research results from Census disclosure officers. This experience will help participants not only begin their SSB research projects but also see them through to successful completion after the seminar has finished.

Lori Reeder and Holly Monti of the U.S. Census Bureau will lead the workshop, in collaboration with researchers at the University of Michigan, the U.S. Census Bureau, Cornell University, and other nationally recognized SIPP scholars. The Survey Research Center will pay travel, lodging, and meal costs for a limited number of participants.

*Aims for this workshop include the following:*

- ❑ Familiarize participants with the SIPP Synthetic Beta's (SSB) data structure and outline the synthesis process.
- ❑ Provide hands-on demonstrations for accessing SSB data via the Synthetic Data Server housed at the VirtualRDC at Cornell University.
- ❑ Provide hands-on demonstrations of analyses of SSB data related to employment, annual earnings, and OASDI program participation.
- ❑ Provide participants with an opportunity to conduct analyses for their own research projects and produce preliminary findings based on the SSB that can be continued when the participant

returns to her/his home institution. Provide directions on how to submit programs for validation and disclosure review after seminar has ended.

- ❑ Provide opportunities for participants to discuss the current status of key research and policy issues with nationally recognized experts.

Classes will meet for two and one-half hour sessions in a computer classroom each morning and afternoon. Participants will work directly with the SSB data via the VirtualRDC at Cornell University. This workshop will not offer instruction in statistics or formal research methods, but presentations will cover appropriate use of this unique dataset. Participants may want to consider attending, at their own expense, other summer courses offered in Ann Arbor by the Institute for Social Research:

<http://www.icpsr.umich.edu/icpsrweb/sumprog/> and <http://si.isr.umich.edu/>

The intent of the SSB is to enable research focusing on the relationship between lifetime earnings, OASDI, and SSI benefit receipt and SIPP outcome variables such as wealth, education, fertility and marital history. Priority will be given to projects that demonstrate creative and insightful uses of the longitudinal nature of these data and which study the relationship between SSA benefit variables and other variables on the file.

### To Apply

Applications will be accepted from faculty, postdoctoral fellows, advanced doctoral students, federal, state and local-level policy and research analysts, researchers at non-profit organizations, and others who would benefit from this workshop. Preference will be given to applicants who meet one of the following criteria:

- ❑ Emerging scholars (Assistant Professors, Postdoctoral Fellows, and Advanced Doctoral Students) working at universities and colleges that do not offer instruction in the use of SIPP.
- ❑ Members of groups historically under-represented in the social sciences. Faculty members from Historically Black Colleges and Universities (HBCU), Tribal Colleges and Universities (TCU) and Hispanic Serving Institutions (HSI) are particularly encouraged to apply.

Applicants should submit their proposal via email to [SIPP-SSB-Workshop2016@umich.edu](mailto:SIPP-SSB-Workshop2016@umich.edu). Proposals must be received by 5 PM Eastern Time on Friday, **May 27, 2016**. Items 1 through 3 of this application should be submitted as a single PDF file that includes the following elements, in the order listed below:

1. Cover sheet indicating you are applying for the SIPP SSB Workshop, with your name and institutional affiliation with mailing address, email address, and telephone number
2. Curriculum vita or resume
3. Brief summary of your current research or analysis activities and how these relate to your proposed SIPP SSB project (about 2 pages)
4. As a separate pdf document, please submit the Census Bureau Application to use the SSB, including a brief description of the project to be undertaken and list of variables to be used. Applications can be found at <https://www.census.gov/programs-surveys/sipp/guidance/sipp-synthetic-beta-data-product.html>. A codebook for the SSB can be found at <https://www2.ncrn.cornell.edu/ced2ar-web/codebooks/ssb/v/v602>. Please note that the SSB does not contain all variables released in the public use SIPP but rather contains a subset of SIPP variables linked to administrative data.

**Direct Questions to:** [SIPP-SSB-Workshop2016@umich.edu](mailto:SIPP-SSB-Workshop2016@umich.edu)

The deadline for receipt of applications is 5 PM Eastern Time on **May 27, 2016**. Selected applicants will be notified no later than Friday, June 10, 2016.

*This workshop is part of the NSF-Census Research Network project of the Institute for Social Research at the University of Michigan. It is funded by National Science Foundation Grant No. SES 1131500.*

## *Survey of Income and Program Participation Synthetic Beta*

More at: <https://www.census.gov/programs-surveys/sipp/guidance/sipp-synthetic-beta-data-product.html>

The SIPP Synthetic Beta (SSB) is a Census Bureau product that integrates person-level micro-data from a household survey with administrative tax and benefit data. These data link respondents from the Survey of Income and Program Participation (SIPP) to Social Security Administration (SSA)/Internal Revenue Service (IRS) Form W-2 records and SSA records of retirement and disability benefit receipt, and were produced by Census Bureau staff in collaboration with researchers at Cornell University, the SSA and the IRS. The purpose of the SSB is to provide access to linked data that are usually not publically available. Census synthesizes, or models, all the variables in a way that changes the record of each individual so as to preserve the underlying covariate relationships between the variables. Only gender and a link to the first reported marital partner are not altered by the synthesis process and still contain their original values.

Nine SIPP panels (1984, 1990, 1991, 1992, 1993, 1996, 2001, 2004, and 2008) form the basis for the SSB, with a subset of SIPP variables available across all the panels selected for inclusion and harmonization of variable definitions across the years covered by the panels. Administrative data are added and some editing is done to correct for logical inconsistencies in the IRS and Social Security earnings and benefits data. Thus, the SSB is a particularly appealing data set for new SIPP users because little data preparation is needed. A complete list of variables included in SSB version 6.0, along with details about the harmonization and editing, is available in the codebook at: <https://www2.ncrn.cornell.edu/ced2ar-web/all>

The application to use the SSB can be found directly here:

[https://www.census.gov/content/dam/Census/programs-surveys/sipp/methodology/SSBApplication\\_template06252012.pdf](https://www.census.gov/content/dam/Census/programs-surveys/sipp/methodology/SSBApplication_template06252012.pdf)

Materials on the SSB can also be found on the VirtualRDC website at:

<http://www2.vrdc.cornell.edu/news/data/sipp-synthetic-beta-file/>

The SSB 6.0 includes several improvements over earlier versions: first, two additional SIPP panels (1984 and 2008); second, the administrative W-2 earnings records extend through 2011; and third, OASDI and SSI benefit information are available through 2012. Several SIPP variables were added to version 6.0 of the SSB, including monthly indicators of receipt and amounts of AFDC or TANF benefits, workers' compensation benefits, veterans' compensation or pension benefits, and food stamp or SNAP benefits. Two additional SIPP point-in-time variables, a sample base weight and a life insurance policy indicator, are available. Also, the fertility history variables (e.g., `first_admin_birthdate` and `last_admin_birthdate`) have been improved by reconciling children's administrative birth dates with those birth dates reported by the children's mother. Finally, it includes measures from SSDI and SSI applications such as date of application, when and whether applications were approved or denied, and a diagnosis group classifying type of disability.

As part of the synthesis process, missing data were multiply-imputed. The resulting confidential data sets are called the Completed Gold Standard Files and contain all original, non-missing, confidential values and imputed values in place of missing data. The goal of the SSB is to produce results that are qualitatively the same as results from the Completed Gold Standard Files.

For workshop participants who attend the training and complete SSB analyses, Census will validate SSB results on the internal, confidential version of these data (**the Completed Gold Standard Files**) and release them to the researcher if they meet disclosure requirements.

*The Synthetic Data Server is funded through NSF grant SES-1042181 and a grant from the Sloan Foundation. You can find more information at: <http://www2.vrdc.cornell.edu/news/data/sipp-synthetic-beta-file/>.*