SIPP Synthetic Beta v6

(121 variables)
Last update to metadata: 2015-10-26 13:17:35 (upload date)
Document Date: January 14, 2015

Codebook prepared by: Cornell NSF-Census Research Network


Citation

Please cite this codebook as:
Comprehensive Extensible Data Documentation and Access Repository. Codebook for the SIPP Synthetic Beta 6.0

Please cite this dataset as:
U.S. Census Bureau. SIPP Synthetic Beta: Version 6.0 [Computer file]. Washington DC; Cornell University, Synthetic Data Server [distributor], Ithaca, NY, 2015

Abstract

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 economists and statisticians 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 publicly available due to confidentiality concerns. To overcome these concerns, Census has synthesized, or modeled, all the variables in a way that changes the record of each individual in a manner designed to preserve the underlying covariate relationships between the variables. The only variables that were not altered by the synthesis process and still contain their original values are gender and a link to the first reported marital partner in the survey. Seven SIPP panels (1990, 1991, 1992, 1993, 1996, 2001, 2004) form the basis for the SSB, with a large subset of variables available across all the panels selected for inclusion and harmonization across the years. Administrative data were added and some editing was done to correct for logical inconsistencies in the IRS/SSA earnings and benefits data.
Datasets

ssb_v6_0_synthetic1_1.sas7bdat

ssb_v6_0_synthetic1_1.dta
http://www.census.gov/programs-surveys/sipp/methodology/sipp-synthetic-beta-data-product.html (Stata)
Terms of Use

Access Levels

released
The data can only be used on the VirtualRDC Synthetic Data Server at Cornell University. While no SSB data downloads are permitted at this time, users do not have to operate behind the Census Bureau firewall to access this server.

restricted
No description given

Access Restrictions (Default)

The data can only be used on the VirtualRDC Synthetic Data Server at Cornell University. While no SSB data downloads are permitted at this time, users do not have to operate behind the Census Bureau firewall to access this server.

Access Requirements

Researchers interested in using the SSB can submit an application to the Census Bureau. The application form and instructions can be downloaded from http://www.census.gov/programs-surveys/sipp/methodology/sipp-synthetic-beta-data-product.html. Applications will be judged solely of feasibility of the proposed project (i.e., that the necessary variables are available on the SSB). Once an application has been accepted, the new user will be given an account on a server where the data can be accessed and analyzed.

Additional information:

Access Permission Requirements

The SSB files have been cleared by the Census Bureau Disclosure Review Board, SSA, and IRS for use by individuals without Census Bureau Special Sworn Status and outside of Census Bureau facilities.

Citation Requirements

We request that researchers who publish results from analyses done using these data cite the SSB as their data source and acknowledge the use of the SDS server at Cornell and the support of Census staff in running any validation programs. These citations will help ensure continued funding for the SDS server and the creation of the Gold Standard File and the SSB.

Suggested acknowledgement:

This analysis was first performed using the SIPP Synthetic Beta (SSB) on the Synthetic Data Server housed at Cornell University which is funded by NSF Grant #SES-1042181. These data are public use and may be accessed by researchers outside secure Census facilities. For more information, visit http://www.census.gov/sipp/synth_data.html. Final results for this paper were obtained from a validation analysis conducted by Census Bureau staff using the SIPP
Completed Gold Standard Files and the programs written by this author and originally run on the SSB. The validation analysis does not imply endorsement by the Census Bureau of any methods, results, opinions, or views presented in this paper.

Disclaimer

The data synthesis process employed by Census to protect the linked data from the risk of disclosing the identity of individuals is relatively new and substantially changes both the survey and administrative data. The intent of the modeling done as part of the synthesis is to preserve relationships among variables that are of interest to researchers while ensuring that personally identifiable information is not revealed to the data user. It has not been feasible to ensure accuracy by comparing every relationship among SSB variables with the corresponding relationship in the underlying confidential micro-data. Hence, we strongly urge researchers not to publish results produced from the SSB without first requesting that Census validate these results with confidential data housed in a secure environment at the Census Bureau. Census will perform this validation free of charge to researchers, as resources permit and according to the protocol established by the three agencies involved and outlined below. Without validation of results, Census, SSA, and IRS make no guarantee of the validity of the SSB for any research purpose. See http://www.census.gov/programs-surveys/sipp/methodology/sipp-synthetic-beta-data-product.html for validation conditions.

Contact

For questions regarding this data collection, please contact: sehsd.synthetic.data.use.list@census.gov

Additional Information

Related Material

1. Using SSB:

The GSF and Completed Data implicates contain personally identifiable information protected by Titles 13, 26, and 42 and cannot be accessed without Census Bureau Special Sworn Status nor outside of Census Bureau facilities. The SSB files, however, have been cleared by the Census Bureau Disclosure Review Board, SSA, and IRS for use by individuals without Census Bureau Special Sworn Status and outside of Census Bureau facilities.

Researchers interested in using the SSB can submit an application to the Census Bureau. The application form and instructions can be downloaded from http://www.census.gov/programs-surveys/sipp/methodology/sipp-synthetic-beta-data-product.html. Applications will be judged solely on feasibility of the proposed project (i.e., that the necessary variables are available on the SSB). Once an application has been accepted, the new user will be given an account on a server where the data can be accessed and analyzed. While no SSB data downloads are permitted at this time, users do not have to operate behind the Census Bureau firewall to access this server.

The SSB is designed to be analytically valid in that sense that point estimates should be unbiased and estimated variances should lead to inferences similar to those that would be drawn from an identical analysis on the Completed Data implicates. Initial tests of analytic validity of the SSB have been promising. All SSB users are invited to help further test the analytic validity of the SSB by submitting programs used to analyze the SSB to be run on the Completed Data and/or Gold Standard files. Users need only inform Census Bureau staff
of the location on the server of such programs and work with Census Bureau staff to ensure that the programs run without error. Census Bureau staff will run the programs on the confidential data and release to the user resulting output that are cleared for release by the Census Bureau Disclosure Review Board. In order to evaluate the effects of the data synthesis separate from the effect of imputing missing data, comparisons should be made between results from the SSB and the Completed Data. To evaluate the effects of missing data imputation, comparisons should be made between results from the Completed Data and the Gold Standard.

II. When analyzing the SSB, users should account for the multiple imputation aspect of the SSB by averaging statistics of interests across all sixteen implicates. Variance measures should be created following the appropriate multiple imputation formulae as described in the document Using the SIPP Synthetic Beta for Analysis.

III. Protocol for Validation of Results:

Census will validate results obtained from the SSB on the internal, confidential version of these data (Completed Gold Standard Files). Users who wish to obtain validated results should follow the protocol outlined here. The restricted access site will provide SAS and Stata analysis software and a computing environment similar to the one used to analyze the confidential Completed Gold Standard data on Census Bureau internal computers. Researchers should follow the Census Bureau programming requirements described in SSB Validation Request Guidelines to ensure that the programs will successfully transfer to internal Census computers for validation. Researchers should plan to share their results and programs from the synthetic data analysis with Census, ORES/SSA and SOI/IRS. After programs have successfully run without error on the synthetic data, researchers may request that Census run these programs on the Completed Gold Standard Files. Only programs successfully run without error on the SDS will be eligible to be run on the confidential data by Census staff. Any programs that produce errors on the Completed Gold Standard Files will be returned to users for correction. Once an analysis has been repeated on the Completed Gold Standard File, the results will be reviewed by Census staff for disclosure concerns. Researchers should familiarize themselves with standard Census disclosure rules for outside projects (See the RDC Researcher Handbook here) and should fill out the appropriate memo documenting the requested output (see RDC Disclosure Request Memo). Data products and output approved by Census staff will be released to the users, ORES/SSA, and SOI/IRS. The validation process can be accomplished in as little as one week for simple results that are generated by clean code and have no disclosure issues. However if the code does not run properly, the sample sizes are too small, or the researcher does not accurately fill out the disclosure memo, the process can take much longer. Census makes no guarantee on the length of time between submission of programs and the release of results from the confidential data. For more information about the validation process, including advice on how to make the process go smoothly and quickly, please see SSB Validation Request Guidelines.
Variable Groups - SIPP Synthetic Beta v6

*Codebook does not contain variable groups.*
<table>
<thead>
<tr>
<th>Values</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Did not receive public assistance payment</td>
</tr>
<tr>
<td>1</td>
<td>Received public assistance payment</td>
</tr>
</tbody>
</table>

Variable Name: afdcamt_YYYYM

Label: Amount of AFDC received

Concept

Type

Files

Full Description

This variable is the amount of public assistance payments (AFDC or TANF) that a respondent received in this month and year. The variable comes from the core file. I20AMT1-I20AMT4 in 1984; S20AMT in the 1990, 1991, 1992, and 1993 panels; T20AMT in the 1996, 2001, 2004, and 2008 panels.
Variable Name          birthdate
Label                  Date of Birth
Concept
Type                  numeric
Files                 F1 F2

Full Description

This variable was taken from a hierarchy of SSA sources instead of the respondent-provided value in the SIPP. Date of birth was selected from the first non-missing value in the following files: (i) SSA's Master Benefits Record (MBR) file, (ii) SSA's Supplemental Security Record (SSR) file, and (iii) the Census Bureau's Person Characteristic File (PCF) whose main input is the SSA Numident file. Thus, this variable is administrative and sometimes differs from the birth date reported in the SIPP survey itself. When missing due to the lack of a validated SSN for the SIPP respondent, date of birth was imputed using date of birth from the Census-internal version of the SIPP as a predictor variable. We chose the administrative source for two reasons. First, the administrative birth date was more consistent with the MBR and DER data and provided more accurate ages for first OASDI benefit receipt and first W-2 or self-employment earnings. Second, the differences between the administrative birth date and the birth date reported in the survey helped to increase the difficulty of re-identifying a record in the original SIPP public use data using information from a record in the synthetic data, thus improving the confidentiality protections. This variable is coded as a SAS date variable. This format gives the number of days between the date of birth and January 1, 1960. An individual born on January 1, 1959 would have birthdate=-365 and an individual born on January 1, 1961 would have birthdate=365.

Value Ranges

Value Range

Range: [-24204.5838012695, 10589.4281616211]
Variable Name: current_enroll_coll

Label: Currently Enrolled in College

Concept

Type: numeric

Files: F1 F2

Full Description

Indicates whether an individual is enrolled in college at the time of the SIPP education history topical module and has not finished his/her education. This variable can be used to differentiate between individuals who completed some college and stopped school and those who have finished some college but not yet stopped attending school. Education variables come from the following waves, by panel: Wave 3 in 1984 (determined from the year began attending college or university and last year student was at a college or university: TM8026 & TM8040); Wave 2 in the 1990-1993 panels (TM8420, TM8426, TM8440, TM8442); Wave 2 in the 1996 panel (TLASTCOL, TCOLLSTR, TVOCYR, TASSOCYR, TBACHYR); Wave 2 in the 2001-2008 panels (ELASTCOL, ECOLLSTR, EVOCYR, EASSOCYR, EBACHYR).

Values (3 total)

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Not currently enrolled in college</td>
</tr>
<tr>
<td>1</td>
<td>Currently enrolled in college</td>
</tr>
</tbody>
</table>

Sysmiss

Value Ranges

Value Range

Range: [0, 1]
Variable Name: current_enroll_hs

Label: Currently Enrolled in HS (or less)

Concept:

Type: numeric

Files: F1 F2

Full Description:

Indicates whether an individual is attending high school. This variable can be used to differentiate between people who do not have a high school degree (educ_5cat=1) but are still attending school and those who do not have a high school degree but are not currently attending high school. Researchers should be careful to consider the age of the individual when using current_enroll_hs and educ_5cat. Individuals who are still young children at the time of the SIPP will have educ_5cat=1 and current_enroll_hs=0 because they are still attending preschool or elementary school. Older individuals with these values for the education variables can be classified as not having finished high school. Education variables come from the following waves, by panel: Wave 2 in the 1990-1993 panels (TM8400, TM8404, TM8406); Wave 2 in the 1996 panel (TLSTSCHL, THSYR); Wave 2 in the 1996-2008 panels (ELSTSCHL, EHSYR). Note: this information is not available in the 1984 panel.

Values (3 total):

0  Not currently enrolled in high school
1  Currently enrolled in high school

Sysmiss

Value Ranges:

Value Range:

Range: [0, 1]
Variable Name        db_pension
Label                Defined Benefit Pension Plan
Concept
Type                 numeric
Files                F1 F2

Full Description

Indicator for whether SIPP respondent was enrolled in a defined benefit pension plan at the time of pension topical module in his/her SIPP panel. Waves for this topical module, by panel, are as follows: Wave 4 from the 1984, 1990, & 1992 panels; Wave 9 from the 1993 panel; Wave 7 from the 1991, 1996, 2001, & 2004 panels; Wave 3 from the 2008 panel.

Values ( 3 total)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No defined benefit pension plan</td>
</tr>
<tr>
<td>1</td>
<td>Had defined benefit pension plan</td>
</tr>
</tbody>
</table>

Sysmiss

Value Ranges

Value Range

Range: [0, 1]
Variable Name | dc_pension
---|---
Label | Defined Contribution Pension Plan
Concept |
Type | numeric
Files | F1 F2

**Full Description**

Indicator for whether the individual was enrolled in a defined contribution pension plan at the time of the pension topical module in his/her SIPP panel. Waves for this topical module, by panel, are as follows: Wave 4 from the 1984, 1990, & 1992 panels; Wave 9 from the 1993 panel; Wave 7 from the 1991, 1996, 2001, & 2004 panels; Wave 3 from the 2008 panel.

**Values (3 total)**

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No defined contribution pension plan</td>
</tr>
<tr>
<td>1</td>
<td>Had defined contribution pension plan</td>
</tr>
</tbody>
</table>

**Sysmiss**

**Value Ranges**

**Value Range**

Range: [0, 1]
Variable Name: deathdate

Label: Date of Death

Concept

Type: numeric

Files: F1 F2

Full Description

Date of death from administrative data. This variable is obtained using a hierarchy of administrative sources: (i) SSA's MBR file, (ii) SSA's SSR file, and (iii) the Census PCF with death information coming from the SSA Numident and Master Death Files. This variable is coded as a SAS date variable. This format gives the number of days between the date of birth and January 1, 1960. An individual born on January 1, 1959 would have birthdate=-365 and an individual born on January 1, 1961 would have birthdate=365.

Value Ranges

Value Range

Range: [10958, 17897]
<table>
<thead>
<tr>
<th><strong>Variable Name</strong></th>
<th>defer_der_fica_YYYY</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Label</strong></td>
<td>DER: Deferred FICA</td>
</tr>
<tr>
<td><strong>Concept</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Type</strong></td>
<td>numeric</td>
</tr>
<tr>
<td><strong>Files</strong></td>
<td>F1 F2</td>
</tr>
</tbody>
</table>

**Full Description**

Deferred earnings from jobs covered by FICA tax; summed across all employers in the DER to give a person-level total for each year. The practice of withholding deferred wages from employee pay and reporting this on W-2 forms began in 1987. However in the SSB, we restrict the time series for this variable to be 1990-2011. This decision is due to the fact that so few people had deferred wages between 1987 and 1989 that we could not reliably synthesize these variables.

**Value Ranges**

**Value Range**

Range: [0, 32112]
Variable Name: defer_der_nonfica_YYYY

Label: DER: Deferred Non-FICA

Concept:

Type: numeric

Files: F1 F2

Full Description:
Deferred earnings from jobs NOT covered by FICA tax; summed across all employers in the DER to give a person-level total for each year. The practice of withholding deferred wages from employee pay and reporting this on W-2 forms began in 1987. However in the SSB, we restrict the time series for this variable to be 1990-2011. This decision is due to the fact that so few people had deferred wages between 1987 and 1989 that we could not reliably synthesize these variables.

Value Ranges:

Value Range:

Range: [0, 13902]
<table>
<thead>
<tr>
<th>Values (2 total)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No</td>
</tr>
<tr>
<td>1</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Variable Name       disab_worklimit
Label              Work Limiting Disability - (CORE or TM)
Concept
Type               numeric
Files              F1 F2

Full Description
This variable indicates that a person had a work-limiting disability. This variable is created from a combination of reports in the core and the disability topical module. We look across all waves of the panel and at the topical module and if ever there is a report of a work-limiting disability, we set this indicator to 1. The universe for this variable is all individuals who were at least age 15 and no older than age 70 by the end of the panel. The following disability variables from the core were used: 1984, 1990-1993 (DISAB); 1996, 2001, 2004, & 2008 (EDISABL). Disability topical modules were asked in the following waves, by panel: Wave 3 in 1984 (TM8450, TM8452); Wave 3 in 1990, 1991, & 1993 (TM8918, TM8920); Wave 6 in 1992 (TM8918, TM8920); Wave 5 in 1996, 2001, & 2004 (EJOBDIF); Wave 6 in 2008 (EJOBDIF).

Values (2 total)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No</td>
</tr>
<tr>
<td>1</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Variable Name: educ_5cat

Label: Education Category

Concept:

Type: numeric

Files: F1 F2

Full Description:

This variable was created from data collected by the education history topical module and represents the highest level of education achieved at that point in time. The universe for this variable is all individuals who were age 15 at the beginning of their SIPP panel. This topical module was asked in the following waves, by panel: Wave 3 in 1984 panel (TM8024, TM8028); Wave 2 in the 1990-1993 panels (TM8400, TM8408, TM8416, TM8422); Wave 2 in the 1996 & 2001 panels (EATTAIN); in the 2004-2008 panels (EEDUCATE).

Values (5 total):

1. No high school degree
2. High school degree
3. Some college
4. College degree
5. Graduate degree

Value Ranges:

Value Range:

Range: [1, 5]
<table>
<thead>
<tr>
<th>Values</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Unknown</td>
</tr>
<tr>
<td>1</td>
<td>Agriculture/Forestry</td>
</tr>
<tr>
<td>2</td>
<td>Art/Architecture (1996-2004 panels); Biology (1990-1993 panels)</td>
</tr>
<tr>
<td>3</td>
<td>Business/Management</td>
</tr>
<tr>
<td>4</td>
<td>Communications (1996-2004 panels); Economics (1990-1993 panels)</td>
</tr>
<tr>
<td>5</td>
<td>Computer and Information Sciences (1996-2004 panels); Education (1990-1993 panels)</td>
</tr>
<tr>
<td>6</td>
<td>Education (1996-2004 panels); Engineering (1990-1993 panels)</td>
</tr>
<tr>
<td>7</td>
<td>Engineering (1996-2004 panels); English/Journalism (1990-1993 panels)</td>
</tr>
<tr>
<td>8</td>
<td>English/Literature (1996-2004 panels); Home Economics (1990-1993 panels)</td>
</tr>
<tr>
<td>9</td>
<td>Foreign Language (1996-2004 panels); Law (1990-1993 panels)</td>
</tr>
<tr>
<td>10</td>
<td>Health Sciences (1996-2004 panels); Liberal Arts/Humanities (including arts, architecture, music, languages, philosophy) (1990-1993 panels)</td>
</tr>
<tr>
<td>11</td>
<td>Liberal Arts/Humanities (1996-2004 panels); Mathematics/Statistics (1990-1993 panels)</td>
</tr>
<tr>
<td></td>
<td>Description</td>
</tr>
<tr>
<td>---</td>
<td>-------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>14</td>
<td>Philosophy/Religion/Theology (1996-2004 panels); Physical or Earth Sciences (1990-1993 panels)</td>
</tr>
<tr>
<td>15</td>
<td>Pre-Professional (1996-2004 panels); Police Science or Law Enforcement (1990-1993 panels)</td>
</tr>
<tr>
<td>16</td>
<td>Psychology</td>
</tr>
<tr>
<td>17</td>
<td>Social Sciences/History (1996-2004 panels); Religion/Theology (1990-1993 panels)</td>
</tr>
<tr>
<td>18</td>
<td>Other (1996-2004 panels); Social Sciences (1990-1993 panels)</td>
</tr>
<tr>
<td>19</td>
<td>Vocational or Technical Studies (1990-1993 panels)</td>
</tr>
<tr>
<td>20</td>
<td>Other (1990-1993 panels)</td>
</tr>
</tbody>
</table>

**Value Ranges**

**Value Range**

Range: [0, 20]
Variable Name: first_admin_birthdate

Label: Administrative birthdate of first born child

Concept:

Type: numeric

Files: F1 F2

Full Description:

This variable contains the administrative birthdate for the first biological child. The universe for this variable is all women between the ages of 15 and 65 at the time of the fertility history topical module. This variable was created by first looking for biological children on the SIPP household roster and choosing the birthdate of the oldest child (if no children were found, the value was set to missing). The total number of biological children reported on the roster (could possibly be zero) was then compared to the woman’s report in her fertility history about the number of children born to her. If the number born to her was larger, first_admin_birthdate was set to missing and imputed using the woman’s report of the first year she gave birth as a predictor variable. This process allowed us to create a fertility history that was consistent with the children reported on the roster and their administrative birthdates but still handle cases where older children or all children lived outside the household.
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>flag_deathdate_exist</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>Flag: Existence of Date of Death</td>
</tr>
<tr>
<td>Concept</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>numeric</td>
</tr>
<tr>
<td>Files</td>
<td>F1 F2</td>
</tr>
</tbody>
</table>

**Full Description**

Flag to indicate that this respondent died after being interviewed and before 2011. This flag indicates the existence of a valid value for deathdate.

**Values (2 total)**

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Death date does not exist, respondent did not die during this interval</td>
</tr>
<tr>
<td>1</td>
<td>Death date exists, respondent died during this interval</td>
</tr>
</tbody>
</table>

**Value Ranges**

Value Range

Range: [0, 1]
Variable Name  flag_mar4t
Label  Flag: 4 or More Marriages
Concept
Type  numeric
Files  F1 F2

Full Description
Flag for existence of a marriage for which date is unknown because it was not collected in the SIPP. The marital history topical module asks about a person's first and second marriages and then his or her most recent marriage. If any other marriages occurred after the second but before the most recent, no information about this marriage is collected. However, individuals are categorized as having 1, 2, 3, or more than 3 marriages. We create flag_mar4t to identify individuals who reported more than 3 marriages.

Values (3 total)

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No additional marriage occurred with unknown date</td>
</tr>
<tr>
<td>1</td>
<td>An additional marriage occurred but with unknown date</td>
</tr>
</tbody>
</table>

Sysmiss

Value Ranges

Value Range

Range: [0, 1]
Variable Name: foodstp_YYYYM

Label: Indicator for receipt of SNAP/Food Stamps

Concept

Type

Files

Full Description


Values (2 total)

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Did not receive food stamps/SNAP benefits</td>
</tr>
<tr>
<td>1</td>
<td>Received food stamps/SNAP benefits</td>
</tr>
</tbody>
</table>
Variable Name: foreign_born

Label: Foreign Born

Concept

Type: numeric

Files: F1 F2

Full Description


Values (2 total)

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Born in U.S.</td>
</tr>
<tr>
<td>1</td>
<td>Born in country other than U.S.</td>
</tr>
</tbody>
</table>

Value Ranges

Value Range

Range: [0, 1]
Variable Name: fsamt_YYYYM

Label: SNAP/Food Stamps Amount Received

Concept

Type

Files

Full Description

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>hicov_YYYYM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>Health Insurance Coverage</td>
</tr>
<tr>
<td>Concept</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>numeric</td>
</tr>
<tr>
<td>Files</td>
<td>F1 F2</td>
</tr>
</tbody>
</table>

**Full Description**

A variable in this array indicates whether an individual was covered by health insurance during year YYYY, month MM. Only the first 24 months of the individual's SIPP panel are in universe. A later release of these data will add months after 24 for longer SIPP panels. Months that are outside the time frame covered by an individual's SIPP panel will always be missing and out of universe. Health insurance coverage is derived from the following variables: HIIND in the 1984 & 1990-1993 panels; EHIMTH in the 1996-2008 panels.

**Values (3 total)**

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Respondent did not have health insurance coverage during this month</td>
</tr>
<tr>
<td>1</td>
<td>Respondent had health insurance coverage during this month</td>
</tr>
</tbody>
</table>

**Sysmiss**

**Value Ranges**

**Value Range**

Range: [0, 1]
Variable Name: hiemp_YYYYYM

Label: Health Insurance Coverage from Employer

Concept

Type: numeric

Files: F1 F2

Full Description

A variable in this array indicates whether an individual was covered by employer-provided health insurance during year YYYY, month MM. Only the first 24 months of the individual's SIPP panel are in universe. A later release of these data will add months after 24 for longer SIPP panels. Months that are outside the time frame covered by an individual's SIPP panel will always be missing and out of universe. Employer-provided health insurance coverage is derived from the following variables: HISRC in the 1990-1993 panels; EHEMPLY in the 1996-2008 panels. Note: this question was not included in the 1984 panel.

Values (3 total)

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Respondent did not have employer-provided health insurance</td>
</tr>
<tr>
<td>1</td>
<td>Respondent had employer-provided health insurance</td>
</tr>
</tbody>
</table>

Sysmiss

Value Ranges

Value Range

Range: [0, 1]
Variable Name: hispanic

Label: Hispanic

Concept

Type: numeric

Files: F1 F2

Full Description

In the 1984 & 1990-1993 SIPP panels, a value for ethnicity is included on each wave file. Thus, there are actually as many ethnicity variables as there are waves of the survey and some changes occur across waves as a result of data collection error. Ethnicity is chosen by creating an array of variables ethncty1-ethncty{max number of waves} and choosing the first non-missing value. Thus, ethnicity comes from the first wave in which the individual was interviewed instead of from a fixed point in the survey. Respondents are coded as Hispanic if they have an ethnicity code between 14 and 20. In the 1996-2008 panels, the longitudinally-edited version contains only one value for ethnicity across all waves (eorigin) and this value is used. Respondents are coded as Hispanic if they have an ethnicity code between 20 and 28 in 1996 and 2001, or if they have an ethnicity code of 1 in 2004 & 2008.

Values (2 total)

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Non-hispanic</td>
</tr>
<tr>
<td>1</td>
<td>Hispanic</td>
</tr>
</tbody>
</table>

Value Ranges

Range: [0, 1]
Variable Name: homeequity

Label: Home Equity

Concept:

Type: numeric

Files: F1 F2

Full Description:

Home equity value as reported in the wealth topical module, collected in the following waves, by panel: Wave 4 in the 1984 (HHTHEQ) & 1990-1993 (HH_THEQ) panels; Wave 3 in 1996, 2001, & 2004 (THHTHEQ) panels; Wave 4 in 2008 panel (THHTHEQ).

Value Ranges:

Value Range:

Range: [-194713.3972656 , 652816.126953125 ]
Variable Name          ind_4cat
Label                   Industry Category (4)
Concept
Type                    numeric
Files                   F1 F2

Full Description

Industry is a characteristic of an individual's job and hence varies over time. There are industry values reported for (potentially) two jobs in each wave of the survey. Industry is chosen by summing earnings associated with the array of variables ws1ind1-ws1ind{max number of waves} and ws2ind1-ws2ind{max number of waves} in the 1984,1990-1993 panels, and ejbind1_1-ejbind1_{(max number of waves)} and ejbind2_1-ejbind2_{(max number of waves)} in the 1996-2008 panels and choosing the industry associated with the greatest total earnings. Thus industry is the industry from which greatest earnings are derived in the survey.

Values ( 5 total)

1     Manufacturing
2     Wholesale/retail trade
3     FIRE, services, public administration, military
4     Agriculture, mining, construction, transportation, communications, and public utilities
Sysmiss

Value Ranges

Value Range

Range: [ 1 , 4 ]
Variable Name: ind_exist

Label: Flag: Industry Assigned

Concept:

Type: numeric

Files: F1 F2

Full Description:
This variable indicates whether an individual has a valid industry from a job held during the survey and governs the ind_4cat variable. If ind_exist=1, the ind_4cat will have a value between 1 and 4.

Values (2 total):

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No, last worked 1984 or earlier, or no valid industry reported</td>
</tr>
<tr>
<td>1</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Value Ranges:

Value Range:

Range: [0, 1]
Variable Name: initwgt

Label: initial SIPP weight

Concept

Type: numeric

Files: F1 F2

Full Description

INITWGT contains the base survey weight for the sample unit to which the sample person belongs. This base survey weight is the inverse of the probability of selection for the sample unit, adjusted only for unit non-response and, in rare instances, for sampled units that turned out to represent more than one separate residence. Unlike final panel and calendar year weights, INITWGT includes no adjustment for attrition of households from SIPP panels and are not adjusted at the person level to match any external controls. SSB users can utilize the weights to produce estimates that account for idiosyncratic probability of selection of sample units and for unit-nonresponse. Estimates for each SIPP panel, using these weights, become representative of the U.S. non-institutionalized population as of the beginning of the panel. For instance, using INITWGT, the distribution of calendar year 2010 administrative earnings for sample persons in the SSB from the 1984 SIPP panel would be representative of the U.S. population as of calendar year 1984.
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>last_admin_birthdate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>Administrative birthdate of last born child</td>
</tr>
<tr>
<td>Concept</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>numeric</td>
</tr>
<tr>
<td>Files</td>
<td>F1 F2</td>
</tr>
</tbody>
</table>

**Full Description**

This variable contains the administrative birthdate for the last biological child. The universe for this variable is all women between the ages of 15 and 65 at the time of the fertility history topical module. This variable was created by first looking for biological children on the SIPP household roster and choosing the birthdate of the youngest child (if no children were found, the value was set to missing). The total number of biological children reported on the roster (could possibly be zero) was then compared to the woman's report in her fertility history about the number of children born to her. If the number born to her was larger, last_admin_birthdate was set to missing and imputed using the woman's report of the last year she gave birth as a predictor variable. This process allowed us to create a fertility history that was consistent with the children reported on the roster and their administrative birthdates but still handle cases where older children or all children lived outside the household.
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>layoff_YYYYM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>On Layoff (Without Pay)</td>
</tr>
<tr>
<td>Concept</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>numeric</td>
</tr>
<tr>
<td>Files</td>
<td>F1 F2</td>
</tr>
</tbody>
</table>

**Full Description**

Indicator that individual was on layoff without pay in year YYYYY, month M. Only the first 24 months of the individual’s SIPP panel are in universe. A later release of these data will add months after 24 for longer SIPP panels. Months that are outside the time frame covered by an individual's SIPP panel will always be missing and out of universe. The variables from which layoff without pay are derived are the following: SC1098 in 1984; REASAB in the 1990-1993 panels; ELAYOFF in the 1996-2008 panels.

**Values ( 2 total)**

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No</td>
</tr>
<tr>
<td>1</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Variable Name          left_layoff_YYYYM
Label                      Left a Job Because of Layoff
Concept
Type                      numeric
Files                       F1 F2

Full Description

The variable indicates that the individual left a job due to a layoff in year YYYY, month M. Only the first 24 months of the individual's SIPP panel are in universe. A later release of these data will add months after 24 for longer SIPP panels. Months that are outside the time frame covered by an individual's SIPP panel will always be missing and out of universe. Note that these questions were not asked in the 1984 panel.

Values ( 2 total)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No</td>
</tr>
<tr>
<td>1</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Variable Name: life_ins_1

Label: Life Insurance Ownership

Concept:

Type: numeric

Files: F1 F2

Full Description:

This variable indicates that the individual owns a life insurance plan at the time of the assets topical module. This module was asked in the following waves, by panel: Wave 4 in the 1984 & 1990-1993 panels (TM8308); Wave 3 in the 1996 panel (EALLI); wave 3 in the 2001 (EALLIE); wave 3 in 2004 (EALLIE); and wave 4 in 2008 (EALLIE).

Values (2 total):

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No</td>
</tr>
<tr>
<td>1</td>
<td>Yes</td>
</tr>
<tr>
<td>Variable Name</td>
<td>male</td>
</tr>
<tr>
<td>---------------</td>
<td>------</td>
</tr>
<tr>
<td>Label</td>
<td>Male</td>
</tr>
<tr>
<td>Concept</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>numeric</td>
</tr>
<tr>
<td>Files</td>
<td>F1 F2</td>
</tr>
</tbody>
</table>

**Full Description**

In the Census-internal SIPP panels, a value for sex is included on each wave file. Thus, there are actually as many sex variables as there are waves of the survey and some changes occur across waves as a result of data collection error. Sex is selected from the array of variables sex1-sex{max number of waves} in which the wave corresponds either to the month in which marital status is first observed (for those without spouses during the course of the SIPP) or to the month in which the respondent's spouse is assigned. As with the SIPP, the SSB does not allow same-sex couples to report being married and hence gender must be chosen to be consistent with the spouse's gender instead of from a fixed point in the survey. Thus when a spouse is never assigned, an individual's gender comes from the first wave where they report being not married. For individuals who are assigned a spouse, gender comes from the first wave where they reveal their spouse. This indicator variable is set to 1 if the individual was male and was created from the original categorical sex variable for analytic convenience. This variable is unsynthesized on the SSB and is never missing so there are no imputed values in the Completed Data.

**Values (2 total)**

<table>
<thead>
<tr>
<th>Value (0)</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Male</td>
</tr>
</tbody>
</table>

**Value Ranges**

**Value Range**

Range: [0, 1]
Variable Name          mbr_agedsp_benefit
Label                  MBR: receive aged spouse benefit
Concept
Type                  numeric
Files                 F1 F2

Full Description

This variable indicates that the individual received OASDI benefits as an aged spouse. A person is entitled to aged spouse benefits if they are at least 62 years old and are married to a worker who is receiving retirement or disability benefits. Like benefits received by a person due to their own eligibility, spouse benefits are reduced if the individual elects to receive them before the full retirement age. Unmarried divorced spouses of retirement age who were married to the worker for at least 10 years are also eligible for aged spouse benefits.

Values ( 2 total)

<table>
<thead>
<tr>
<th>0</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Variable Name: mbr_agedsp_benefit_stdate
Label: MBR: startdate of benefit
Concept:
Type: numeric
Files: F1 F2

Full Description:
Date when the person first began receiving aged spouse benefits, conditional on having ever received this type of benefit.

Value Ranges

Value Range:
Range: [ 882 , 17532 ]
Variable Name: mbr_agedsp_benefit_totamt

Label: MBR: total monthly benefit

Concept:

Type: numeric

Files: F1 F2

Full Description:

Total monthly amount of benefits received at beginning of aged spouse benefit entitlement. In most cases this amount is from the same month as in MBR_agedsp_benefit_stdate. However, if data for that month were missing in the MBR extract, we searched through the monthly benefit array to find the first positive value. This amount can be a combination of payments due to multiple entitlement reasons (i.e. dual entitlement).

Value Ranges:

Value Range:

Range: [39.5497995976014, 1048.55813207881]
Variable Name: mbr_retire_benefit

Label: MBR: receive retirement benefit

Concept

Type: numeric

Files: F1 F2

Full Description

This variable indicates receipt of an OASDI retirement benefit. Individuals are eligible for retirement payments if they are at least 62 years old and have 40 quarters of coverage. Since 1978, workers earn one quarter of coverage for a set amount of FICA-covered annual earnings and can earn up to 4 quarters per year. The amount of earnings needed to receive credit for a quarter varies by year and was $1,120 in 2011. Benefits are reduced if the worker elects to receive them before full retirement age.

Values (2 total)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No</td>
</tr>
<tr>
<td>1</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Variable Name: mbr_retire_benefit_stdate
Label: MBR: retirement benefit start date
Concept:
Type: numeric
Files: F1 F2

Full Description:

This variable contains the date when the individual first began receiving retirement benefits and is stored as a SAS date variable. Benefits always begin on the first day of the month. Individuals born before 1920 were eligible to receive their first payment in the month they turned 62 unless their birthday was the first day of the month, in which case they were eligible for their first payment the month before they turned 62. Individuals born on January 1 were eligible to receive their first payment in the December before they turned 62. Individuals born in 1920 or after were eligible to receive their first payment the month after they turned 62 unless their birthday was the first or second day of the month, in which case they were eligible for their first payment in the month they turned 62.
Variable Name: mbr_retire_benefit_totamt

Label: MBR: total monthly benefit

Concept:

Type: numeric

Files: F1 F2

Full Description:

This variable provides the monthly amount of retirement benefits received by the individual. Benefit amounts are calculated based on the worker's average indexed monthly earnings (AIME). See the Social Security Annual Statistical Supplement for details on how benefit calculations are performed. This benefit amount is the total amount paid in the first month of own retirement benefits receipt. If the respondent was dually entitled in this month, this benefit amount reflects the total payment made (i.e. the sum of the amounts due to each type of benefit). For example if a person received own and aged spouse retirement benefits, this benefit amount would be the sum of those two benefits.
**Variable Name**  
mbr_sssi_applied_k

**Label**  
MBR SSDI: Application submitted

**Concept**

**Type**  
numeric

**Files**  
F1 F2

**Full Description**

This set of variables (mbr_sssi_applied_1?mbr_sssi_applied_4) indicate whether there is a corresponding record of Social Security Disability Insurance (SSDI) application in the Master Benefit Record (MBR). Details for up to four SSDI applications are maintained. If the individual applied more than four times, then details for only a subset of the applications are recorded in this data, with priority given to approved and more recent applications. The first recorded application, the last recorded application, the first recorded application during the SIPP interview period, and the last recorded application during the SIPP interview period are always kept. For example, if a person applied for SSDI three times, this will be reflected in the following way: mbr_sssi_applied_1 = 1, mbr_sssi_applied_2 = 1, mbr_sssi_applied_3 = 1, mbr_sssi_applied_4 = 0.
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>mbr_sssi_benefit_totamt_k</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>MBR SSDI: Benefit amount</td>
</tr>
<tr>
<td>Concept</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>numeric</td>
</tr>
<tr>
<td>Files</td>
<td>F1 F2</td>
</tr>
</tbody>
</table>

**Full Description**

This variable provides the monthly amount of SSDI benefits received by the individual for each of four applications. See mbr_sssi_applied_k for details on how individual applications are recorded on this file. Benefit amounts are calculated based on the worker's average indexed monthly earnings (AIME). See the Social Security Annual Statistical Supplement for details on how benefit calculations are performed. This benefit amount is the total amount paid in the first month of disability benefits receipt.
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>mbr_sssi_ceased_k</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>MBR SSDI: Benefits ceased</td>
</tr>
<tr>
<td>Concept</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>numeric</td>
</tr>
<tr>
<td>Files</td>
<td>F1 F2</td>
</tr>
</tbody>
</table>

**Full Description**

This set of variables records whether the individual ever ceased receiving previously granted benefits under each SSDI application record \( mbr_sssi_ceased_k = 1 \). See \( mbr_sssi_applied_k \) for details on how individual applications are recorded on this file.
Variable Name: mbr_smdi_ddbc_k

Label: MBR SSDI: Date of disability benefits cessation

Concept

Type: numeric

Files: F1 F2

Full Description

If the individual ever ceased receiving previously granted SSDI benefits, this variable records the date that those benefits ceased. This date is stored as a SAS date variable. See mbr_smdi_applied_k for details on how individual applications are recorded on this file.
Variable Name: mbr_sssdi_ddo_k

Label: MBR SSDI: Date of disability onset

Concept

Type: numeric

Files: F1 F2

Full Description

The date of disability onset for the associated SSDI application. This date is stored as a SAS date variable. See mbr_sssdi_applied_k for details on how individual applications are recorded on this file.
Variable Name: mbr_ssdidig_group_k

Label: MBR SSDI: Diagnosis group

Concept

Type: numeric

Files: F1 F2

Full Description:

For each approved SSDI application, this variable contains the diagnostic group for the individual's primary code for mental or physical disability used in the medical determination of the individual's eligibility for disability benefits. See mbr_ssdidapplied_k for details on how individual applications are recorded on this file. Codes:

- 01?Diseases of the blood;
- 02?Circulatory system;
- 03?Congenital anomalies;
- 04?Digestive system;
- 05?Endocrine, nutritional, and metabolic diseases;
- 06?Genitourinary system;
- 07?Infectious and parasitic diseases;
- 08?Injuries;
- 09?Mental disorders;
- 10?Mental retardation;
- 11?Musculoskeletal system;
- 12?Neoplasms;
- 13?Nervous system and sense organs;
- 14?Respiratory system;
- 15?Skin and subcutaneous tissue;
- 16?Other or unknown.
Variable Name: mbr_ssdii_doed_k

Label: MBR SSDI: Date of entitlement to disability

Concept:

Type: numeric

Files: F1 F2

Full Description:
The date at which the disabled individual became entitled to disability insurance benefits for the associated SSDI application. This date is stored as a SAS date variable. See mbr_ssdii_applied_k for details on how individual applications are recorded on this file.
Variable Name: mbr_ssdid_ssd_k

Label: MBR SSDI: Disability adjudication date

Concept

Type: numeric

Files: F1 F2

Full Description

The date of the disability-adjudication decision for the associated SSDI application. This date is stored as a SAS date variable. See mbr_ssdid_applied_k for details on how individual applications are recorded on this file.
Variable Name: mbr_ssdii_entitled_k
Label: MBR SSDI: Application entitles to disability
Concept:
Type: numeric
Files: F1 F2

Full Description:
An indicator for whether the individual was determined to be entitled to SSDI benefits under the associated application. See mbr_ssdii_applied_k for details on how individual applications are recorded on this file.
Variable Name: mbr_widowsp_benefit

Label: MBR: receive widowed spouse benefit

Concept:

Type: numeric

Files: F1 F2

Full Description:

This variable indicates that the individual received OASDI benefits as a widowed spouse. A person is entitled to widowed spouse benefits if they are at least 60 years old and are the widow(-er) of a worker who was fully insured (i.e. had the required number of quarters of coverage for his/her age). Like benefits received by a person due to their own eligibility, widow benefits are reduced if the individual elects to receive them before the full retirement age. Surviving divorced spouses who were married to the worker for at least ten years and did not remarry before age 60 are also eligible for widow benefits.

Values (2 total):

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No</td>
</tr>
<tr>
<td>1</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Variable Name: mbr_widowsp_benefit_stdate

Label: MBR: widowed spouse benefit start date

Concept

Type: numeric

Files: F1 F2

Full Description:

This variable contains the date when the individual first began receiving widowed spouse benefits and is stored as a SAS date variable. Benefits always begin on the first day of the month. Individuals were eligible to receive their first payment in the month they turned 60 unless their birthday was the first day of the month, in which case they were eligible for their first payment the month before they turned 60. Individuals born on January 1 were eligible to receive their first payment in the December before they turned 60. Start dates range from January 1962 to December 2012.
Variable Name: mbr_widowsp_benefit_totamt

Label: MBR: total monthly benefit

Concept:

Type: numeric

Files: F1 F2

Full Description:

This variable provides the monthly amount of widowed spouse benefits received by the individual. Benefit amounts are calculated based on the worker's average indexed monthly earnings (AIME). See the Social Security Annual Statistical Supplement for details on how benefit calculations are performed. This benefit amount is the total amount paid in the first month of widowed spouse benefits receipt. If the respondent was dually entitled in this month, this benefit amount reflects the total payment made (i.e. the sum of the amounts due to each type of benefit). For example if a person received own and widowed spouse retirement benefits, this benefit amount would be the sum of those two benefits.
Variable Name: mh1

Label: Flag: Marital History Event 1

Concept

Type: numeric

Files: F1 F2

Full Description

This variable is the first indicator for whether a respondent was married. Individuals who are never married will have mh1=0. Individuals who were married one or more times will have mh1=1.

Values (3 total)

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Never married</td>
</tr>
<tr>
<td>1</td>
<td>First marriage occurred</td>
</tr>
<tr>
<td></td>
<td>Sysmiss</td>
</tr>
</tbody>
</table>

Value Ranges

Value Range

Range: [0, 1]
**Variable Name**  
mh2

**Label**  
Flag: Marital History Event 2

**Concept**

**Type**  
numeric

**Files**  
F1 F2

**Full Description**

This variable indicates whether a respondent's first marriage ended in widowhood or divorce. This variable is only in universe for individuals with mh1=1.

**Values (4 total)**

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>First Marriage did not end over course of survey</td>
</tr>
<tr>
<td>1</td>
<td>First marriage ended in widowhood</td>
</tr>
<tr>
<td>2</td>
<td>First marriage ended in divorce/separation</td>
</tr>
</tbody>
</table>

**Sysmiss**

**Value Ranges**

**Value Range**

Range: [0, 2]
Variable Name  
mh3

Label  
Flag: Marital History Event 3

Concept

Type  
numeric

Files  
F1 F2

Full Description

This variable indicates whether a respondent entered into a second marriage. This variable is only in universe if mh2 = 1 or 2.

Values (3 total)

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No second marriage</td>
</tr>
<tr>
<td>1</td>
<td>Second marriage occurred</td>
</tr>
</tbody>
</table>

Sysmiss

Value Ranges

Value Range

Range: [0, 1]
Variable Name  

mh4

Label  

Flag: Marital History Event 4

Concept

Type  

numeric

Files  

F1 F2

Full Description

This variable indicates whether a respondent's second marriage ended in widowhood or divorce. Only respondents with mh3=1 are in universe for this variable.

Values ( 4 total)

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Second marriage did not end over course of survey</td>
</tr>
<tr>
<td>1</td>
<td>Second marriage ended in widowhood</td>
</tr>
<tr>
<td>2</td>
<td>Second marriage ended in divorce/separation</td>
</tr>
</tbody>
</table>

Sysmiss

Value Ranges

Value Range

Range: [ 0 , 2 ]
Variable Name: mh5

Label: Flag: Marital History Event 5

Concept

Type: numeric

Files: F1 F2

Full Description

This variable indicates whether a respondent entered into a third marriage. Only individuals with mh4= 1 or 2 are in universe for mh5.

Values (3 total)

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No third marriage</td>
</tr>
<tr>
<td>1</td>
<td>Third marriage occurred</td>
</tr>
</tbody>
</table>

Sysmiss

Value Ranges

Value Range

Range: [0, 1]
Variable Name: mh6

Label: Flag: Marital History Event 6

Concept

Type: numeric

Files: F1 F2

Full Description

This variable indicates whether a respondent's third marriage ended in widowhood or divorce. Only respondents with mh5=1 are in universe.

Values (4 total)

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Third marriage did not end over course of survey</td>
</tr>
<tr>
<td>1</td>
<td>Third marriage ended in widowhood</td>
</tr>
<tr>
<td>2</td>
<td>Third marriage ended in divorce/separation</td>
</tr>
</tbody>
</table>

Value Ranges

Range: [0, 2]
Variable Name  
\( mh7 \)

Label  
Flag: Marital History Event 7

Concept

Type  
numeric

Files  
F1 F2

Full Description

This variable indicates whether respondent entered into a fourth marriage and is in universe only if \( mh6=1 \) or 2. The marital history topical module only collects information about up to three marriages, and hence in cases where \( mh7=1 \), the fourth marriage occurred over the course of the SIPP panel, after the marital history topical module had been conducted.

Values (3 total)

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No fourth marriage</td>
</tr>
<tr>
<td>1</td>
<td>Fourth marriage occurred</td>
</tr>
</tbody>
</table>

Sysmiss

Value Ranges

Value Range

Range: \([0, 1]\)
Variable Name: mh8

Label: Flag: Marital History Event 8

Concept

Type: numeric

Files: F1 F2

Full Description

This variable indicates whether respondent's fourth marriage ended in widowhood or divorce and is only in universe if mh7=1.

Values (4 total)

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Fourth marriage did not end over course of survey</td>
</tr>
<tr>
<td>1</td>
<td>Fourth marriage ended in widowhood</td>
</tr>
<tr>
<td>2</td>
<td>Fourth marriage ended in divorce/separation</td>
</tr>
</tbody>
</table>

Sysmiss

Value Ranges

Value Range

Range: [0, 2]
Variable Name  mh_date1
Label  Date of Marital History Event 1
Concept
Type  numeric
Files  F1 F2

Full Description
Date of first marriage in SAS date value format. In universe if mh1=1.

Value Ranges

Value Range

Range: [-12368.9569092027 , 17471 ]
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>mh_date2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>Date of Marital History Event 2</td>
</tr>
<tr>
<td>Concept</td>
<td>numeric</td>
</tr>
<tr>
<td>Type</td>
<td>numeric</td>
</tr>
<tr>
<td>Files</td>
<td>F1 F2</td>
</tr>
</tbody>
</table>

**Full Description**

Date that first marriage ended in divorce or widowhood. Stored as a SAS date value.

**Value Ranges**

<table>
<thead>
<tr>
<th>Value Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range: [-10318.8063402205, 16071]</td>
</tr>
</tbody>
</table>
Variable Name: mh_date3

Label: Date of Marital History Event 3

Concept

Type: numeric

Files: F1 F2

Full Description

Date of beginning of second marriage. Stored as a SAS date value.

Value Ranges

Value Range

Range: [-10266.2469875312, 17501]
<table>
<thead>
<tr>
<th><strong>Variable Name</strong></th>
<th><strong>mh_date4</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Label</strong></td>
<td>Date of Marital History Event 4</td>
</tr>
<tr>
<td><strong>Concept</strong></td>
<td>numeric</td>
</tr>
<tr>
<td><strong>Type</strong></td>
<td>numeric</td>
</tr>
<tr>
<td><strong>Files</strong></td>
<td>F1 F2</td>
</tr>
</tbody>
</table>

**Full Description**

Date of ending of second marriage. Stored as a SAS date value.

**Value Ranges**

**Value Range**

Range: [-7991.34877841126, 17501]
Variable Name: mh_date5
Label: Date of Marital History Event 5
Concept:
Type: numeric
Files: F1 F2

Full Description:
Date of beginning of third marriage. Stored as a SAS date value.

Value Ranges:
Value Range:
Range: [-5511.43090506919, 17471]
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>mh_date6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>Date of Marital History Event 6</td>
</tr>
<tr>
<td>Concept</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>numeric</td>
</tr>
<tr>
<td>Files</td>
<td>F1 F2</td>
</tr>
<tr>
<td>Full Description</td>
<td>Date of ending of third marriage.</td>
</tr>
<tr>
<td>Value Ranges</td>
<td></td>
</tr>
<tr>
<td>Value Range</td>
<td>Range: ([-1416.79843489124, 17471])</td>
</tr>
</tbody>
</table>
Variable Name: mh_date7
Label: Date of Marital History Event 7
Concept:
Type: numeric
Files: F1 F2

Full Description:
Date of beginning of fourth marriage. Stored as a SAS date value.

Value Ranges:

Value Range:
Range: [11323, 15582.7638571104]
Variable Name  

mh_date8

Label  

Date of Marital History Event 8

Concept

Type

numeric

Files

F1 F2

Full Description

Date of fourth marriage ending. Stored as a SAS date value.

Value Ranges

Value Range

Range: [ 11436.2674992709 , 16010 ]
Variable Name  nonhouswealth
Label  Non-Housing Financial Wealth
Concept
Type  numeric
Files  F1 F2

Full Description


Value Ranges

Value Range

Range: [-29979.16015625, 2370592.9140625]
Variable Name  
obs_first_sipp_mar_num

Label  
Ordinal Number of First Observed Marriage

Concept

Type  
numeric

Files  
F1 F2

Full Description

This variable tells which of the marriages described in the marital history variables (mh1-mh8) is the linked marriage. For example, if obs_first_sipp_mar_num=1 then the first marriage is the one observed in the SIPP and the partner from that marriage is the linked spouse.

Values (5 total)

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Linked marriage is first reported marriage described by mh1, mh2, mh_date1, and mh_date2</td>
</tr>
<tr>
<td>2</td>
<td>Linked marriage is first reported marriage described by mh3, mh4, mh_date3, and mh_date4</td>
</tr>
<tr>
<td>3</td>
<td>Linked marriage is first reported marriage described by mh5, mh6, mh_date5, and mh_date6</td>
</tr>
<tr>
<td>4</td>
<td>Linked marriage is first reported marriage described by mh7, mh8, mh_date7, and mh_date8</td>
</tr>
<tr>
<td>Sysmiss</td>
<td>When there is no linked marriage</td>
</tr>
</tbody>
</table>

Value Ranges

Value Range

Range: [1, 4]
Variable Name  occ_3cat
Label  SIPP-reported occupation - three categories
Concept
Type  numeric
Files  F1 F2

Full Description

Occupation is a characteristic of an individual's job and hence varies over time. There are occupation values reported for (potentially) two jobs in each wave of the survey. Occupation is chosen by summing earnings associated with the array of variables ws1occ1-ws1occ{max number of waves} and ws2occ1-ws2occ{max number of waves} in the 1984, 1990-1993 panels, and tjbocc1_1-tjbocc1_{max number of waves} and tjbocc2_1-tjbocc2_{max number of waves} in the 1996-2008 panels and choosing the occupation associated with the greatest total earnings. Thus occupation is the occupation from which greatest earnings are derived in the survey.

Values (4 total)

1 Managerial and professional specialty occupations
2 Technical, sales, and administrative support occupations
3 Other

Sysmiss

Value Ranges

Value Range

Range: [1, 3]
Variable Name: occ_exist

Label: SIPP Occupation - existence

Concept

Type: numeric

Files: F1 F2

Full Description

This variable indicates whether a person has a valid occupation from a job held during the survey.

Values (2 total)

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No, last worked 1984 or earlier, or no valid industry reported</td>
</tr>
<tr>
<td>1</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Value Ranges

Range: [0, 1]
Variable Name          own_home
Label                  Own a Home
Concept
Type                   numeric
Files                  F1 F2

Full Description

This variable indicates whether an individual owns a home at the time of the wealth topical module in the person's SIPP panel. The wealth topical module is conducted in the following waves by panel: wave 4 in the 1984 & 1990-1993 panels (TM8530, TM8608). In panels 1996-2004, tenure is available for each month of the survey, for this variable (own_home), tenure is defined as tenure during 12th month in the survey (ETENURE), and the universe for the real estate/wealth topical module is defined in wave 3 (EHREUNV). In 2008, ETENURE is also from the 12th month of the survey (ETENURE), and the universe for the real estate/wealth topical module is defined in wave 4 (EHREUNV).

Values (2 total)

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Do not own a home</td>
</tr>
<tr>
<td>1</td>
<td>Own a home</td>
</tr>
</tbody>
</table>

Value Ranges

Value Range

Range: [0, 1]
Variable Name: own_kids-ever
Label: Number of Children Ever Born
Concept:
Type: numeric
Files: F1 F2

Full Description:
This variable contains the number of children ever born to a person (i.e. count of biological children). This is taken from the wave two fertility history topical module (TM8752 and TM8754 for 1984, 1990-1993 panels; tfrch1 and tmomch1 for 1996-2008 panels) and is in universe for men and women ages 15-64.

Value Ranges

Value Range:
Range: [0, 20]
Variable Name: panel
Label: SIPP Panel Year
Concept:
Type: numeric
Files: F1 F2

Full Description:
indicates panel of source record

Values (9 total):
1984
1990
1991
1992
1993
1996
2001
2004
2008

Value Ranges:
Value Range:
Range: [1984, 2008]
Variable Name: pension_in_scope_empl

Label: In-Scope for Pension (Level II)

Concept:

Type: numeric

Files: F1 F2

Full Description:

This variable indicates that an individual was in scope for the pension questions because he or she was employed at the time of the pension topical module. The pension topical module is conducted in the following waves by panel: wave 4 in the 1984 panel (TM8324); wave 4 in the 1990 & 1992 panels (TM8324); wave 7 in 1991 panel (TM8324); wave 9 in the 1993 panel (TM6000); wave 7 in the 1996-2004 panels (EARPUNV, RMNJBBS); wave 3 in the 2008 panel (EARPUNV, RMNJBBS).

Values (2 total):

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Pension was not in scope</td>
</tr>
<tr>
<td>1</td>
<td>Pension was in scope</td>
</tr>
</tbody>
</table>

Value Ranges:

Value Range

Range: [0, 1]
Variable Name: personid

Label: SIPP Gold Standard Person ID

Concept:

Type: numeric

Files: F1 F2

Full Description:

Personid uniquely identifies individuals within each SSB implicate. Personid does not link records across implicate files or to the Gold Standard and Completed Data files.

Value Ranges:

Value Range:

Range: $[1, 1000]$
Variable Name                  phus_agedsp_benefit_stdate

Label                         PHUS: startdate of aged spouse benefits

Concept

Type                          numeric

Files                        F1 F2

Full Description

This startdate variable tells when the first aged spouse benefit payment was recorded in the Payment History Update System, the administrative database maintained by SSA to track actual payments made to beneficiaries. This startdate can differ from the MBR startdate which records only eligibility and not actual payments. The PHUS began in 1984 and hence the earliest possible start date is January 1984. The latest possible start date is December 2012.
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>phus_agedsp_benefit_totamt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>PHUS: total monthly benefit</td>
</tr>
<tr>
<td>Concept</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>numeric</td>
</tr>
<tr>
<td>Files</td>
<td>F1 F2</td>
</tr>
</tbody>
</table>

**Full Description**

Total monthly benefit payment as recorded in the Payment History Update System, the administrative database maintained by SSA to track actual payments made to beneficiaries. This benefit amount can differ from the MBR total benefit which records only eligibility and not actual payments. This benefit amount is the total amount paid in the first month of aged spouse benefits receipt. If the respondent was dually entitled in this month, this benefit amount reflects the total payment made (i.e. the sum of the amounts due to each type of benefit). For example if a person received own and aged spouse retirement benefits, this benefit amount would be the sum of those two benefits.
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>phus_retire_benefit_stdate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>PHUS: startdate of retirement benefits</td>
</tr>
<tr>
<td>Concept</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>numeric</td>
</tr>
<tr>
<td>Files</td>
<td>F1 F2</td>
</tr>
</tbody>
</table>

**Full Description**

This startdate variable tells when the first own retirement benefit payment was recorded in the Payment History Update System, the administrative database maintained by SSA to track actual payments made to beneficiaries. This startdate can differ from the MBR startdate which records only eligibility and not actual payments. The PHUS began in 1984 and hence the earliest possible start date is January 1984. The latest possible start date is December 2012.
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>phus_retire_benefit_totamt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>PHUS: total monthly benefit</td>
</tr>
<tr>
<td>Concept</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>numeric</td>
</tr>
<tr>
<td>Files</td>
<td>F1 F2</td>
</tr>
</tbody>
</table>

**Full Description**

Total monthly benefit payment as recorded in the Payment History Update System, the administrative database maintained by SSA to track actual payments made to beneficiaries. This benefit amount can differ from the MBR total benefit which records only eligibility and not actual payments. This benefit amount is the total amount paid in the first month of own retirement benefits receipt. If the respondent was dually entitled in this month, this benefit amount reflects the total payment made (i.e. the sum of the amounts due to each type of benefit). For example if a person received own and aged spouse retirement benefits, this benefit amount would be the sum of those two benefits.
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>phus_sssi_benefit_stdate_k</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>PHUS: SSDI benefit start date 1</td>
</tr>
<tr>
<td>Concept</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>numeric</td>
</tr>
<tr>
<td>Files</td>
<td>F1 F2</td>
</tr>
<tr>
<td>Variable Name</td>
<td>phus_ssdii_benefit_totamt_k</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>------------------------------------------------</td>
</tr>
<tr>
<td>Label</td>
<td>PHUS: SSDI total monthly benefit 1</td>
</tr>
<tr>
<td>Concept</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>numeric</td>
</tr>
<tr>
<td>Files</td>
<td>F1 F2</td>
</tr>
<tr>
<td>Variable Name</td>
<td>phus_widowsp_benefit_stdate</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>Label</td>
<td>PHUS: startdate of widowed spouse benefits</td>
</tr>
<tr>
<td>Concept</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>numeric</td>
</tr>
<tr>
<td>Files</td>
<td>F1 F2</td>
</tr>
</tbody>
</table>

**Full Description**

This startdate variable tells when the first widowed spouse benefit payment was recorded in the Payment History Update System, the administrative database maintained by SSA to track actual payments made to beneficiaries. This startdate can differ from the MBR startdate which records only eligibility and not actual payments. The PHUS began in 1984 and hence the earliest possible start date is January 1984. The latest possible start date is December 2012.
Variable Name: phus_widowsp_benefit_totamt

Label: PHUS: total monthly benefit

Concept:

Type: numeric

Files: F1 F2

Full Description:

Total monthly benefit payment as recorded in the Payment History Update System, the administrative database maintained by SSA to track actual payments made to beneficiaries. This benefit amount can differ from the MBR total benefit which records only eligibility and not actual payments. This benefit amount is the total amount paid in the first month of widowed spouse benefits receipt. If the respondent was dually entitled in this month, this benefit amount reflects the total payment made (i.e. the sum of the amounts due to each type of benefit). For example if a person received own and widowed spouse retirement benefits, this benefit amount would be the sum of those two benefits.
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>pos_phus_agedsp_benefit_totamt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>PHUS: indicator of positive benefit amount</td>
</tr>
<tr>
<td>Concept</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>numeric</td>
</tr>
<tr>
<td>Files</td>
<td>F1 F2</td>
</tr>
</tbody>
</table>

**Full Description**

Indicator that PHUS recorded a positive aged spouse benefit amount (and consequently had a non-missing PHUS startdate) at some point after the MBR eligibility start date.
Variable Name  pos_phus_retire_benefit_totamt

Label  PHUS: indicator of positive benefit amount

Concept

Type  numeric

Files  F1 F2

Full Description

Indicator that PHUS recorded a positive retirement benefit amount (and consequently had a non-missing PHUS startdate) at some point after the MBR eligibility start date.
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>pos_phus_ssd_benefit_totamt_k</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>PHUS: indicator of positive SSDI benefit amount</td>
</tr>
<tr>
<td>Concept</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>numeric</td>
</tr>
<tr>
<td>Files</td>
<td>F1 F2</td>
</tr>
</tbody>
</table>
Variable Name: pos_phus_widowsp_benefit_totamt

Label: PHUS: indicator of positive benefit amount

Concept

Type: numeric

Files: F1 F2

Full Description

Indicator that PHUS recorded a positive widowed spouse benefit amount (and consequently had a non-missing PHUS startdate) at some point after the MBR eligibility start date.
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>race</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>Race</td>
</tr>
<tr>
<td>Concept</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>numeric</td>
</tr>
<tr>
<td>Files</td>
<td>F1 F2</td>
</tr>
</tbody>
</table>

**Full Description**

In the Census-internal SIPP panels, a value for race is included on each wave file. Thus, there are actually as many race variables as there are waves of the survey and some changes occur across waves as a result of data collection error. Race is chosen by creating an array of variables race1-race(max number of waves) and choosing the first non-missing value. Thus race comes from the first wave in which the individual was interviewed instead of from a fixed point in the survey.

**Values ( 3 total)**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>White</td>
</tr>
<tr>
<td>2</td>
<td>Black</td>
</tr>
<tr>
<td>3</td>
<td>Other</td>
</tr>
</tbody>
</table>

**Value Ranges**

**Value Range**

Range: [1, 3]
Variable Name: rot

Label: SIPP Survey Rotation Group

Concept:

Type: numeric

Files: F1 F2

Full Description:

Each SIPP panel is divided up into four rotation groups and one rotation group is interviewed each month. This variable classifies each SIPP respondent into a particular rotation group and enables researchers to tell which month a particular respondent was asked the topical module questions. For example, if a person is in the first rotation group of the 1990 panel, his/her wave 1 interview covered January 1990-April 1990 and the wave 2 interview covered May 1990-August 1990. The wave 2 topical modules would have been asked in September 1990 at the time of the wave 2 interview. Schedules for interviewing rotation groups by panel can be found in the SIPP Users’ Guide.

Value Ranges:

Value Range:

Range: [1, 4]
Variable Name: sipp_panel_beg_date

Label: SIPP Panel Begin Date

Concept:

Type: numeric

Files: F1 F2

Full Description:

This variable is stored as a SAS date and gives the month and year of the first SIPP reference period for each individual (reference periods always begin on the first day of the month). The start date differs even within panel due to the fact that only one rotation group (1/4 of the sample) is interviewed each month.
Variable Name: sipp_panel_end_date

Label: SIPP Panel End Date

Concept

Type: numeric

Files: F1 F2

Full Description

This variable is stored as a SAS date and gives the month and year of the last month of the last SIPP reference period for each individual (reference periods always end on the last day of the month). The end date differs even within panel due to the fact that only one rotation group (1/4 of the sample) is interviewed each month.

Value Ranges

Value Range

Range: [ 11809 , 17501 ]
Variable Name: spouse_personid

Label: Personid of spouse

Concept:
Personid of linked spouse, defined as the first person to whom the SIPP respondent was married during the time period covered by the SIPP panel. Individuals could enter the panel already married and then each would be linked to the other. Individuals could also get married during the course of the panel. If this was the first observed marriage for each member of the couple, they were linked together. Individuals could also get divorced during the course of the panel and then remarry. In many cases, this later marriage caused a new individual to join the panel. This new SIPP respondent would only be linked to his or her spouse if the spouse (and original SIPP sample member) had not already been observed married to someone else. If the original SIPP sample member had been previously linked by marriage to another SIPP sample member, this original link was maintained in spouse_personid. However, the marital history reflects the ending of this marriage and the occurrence of the next marriage for the original SIPP sample member. Likewise, the new SIPP sample member who joins through marriage will have that marriage date recorded in his or her marital history but will have a blank spouse_personid. In summary, this variable captures only one marriage partner and does not provide a history of marriage partners even if this history is (partially) observed in the SIPP. The link between SIPP respondents and their spouses has not been perturbed in any way in the SSB. The same individuals will be linked as married partners in the Gold Standard, the Completed Data, and the SSB. In order to strengthen confidentiality protection, however, spouse_personid in the SSB does not link records across implicate(s) or to the Gold Standard and Completed Data files, i.e. within each synthetic implicate it is a random number that can only be used for matching spouses within that implicate.

Value Ranges:

Value Range:
Range: [1167, 444444]
Variable Name: ssr_ssi_appl_dt

Label: SSR: Application date

Concept:

Type: numeric

Files: F1 F2

Full Description:

This variable contains the date when the individual applied for SSI benefits and is stored as a SAS date variable.
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>ssr_ssi_applied</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>SSR: Applied for SSI benefits</td>
</tr>
<tr>
<td>Concept</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>numeric</td>
</tr>
<tr>
<td>Files</td>
<td>F1 F2</td>
</tr>
</tbody>
</table>

**Full Description**

An indicator for whether the individual has a recorded application for Supplemental Security Income (SSI).
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>ssr_ssi_benefit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>SSR: Received SSI benefits</td>
</tr>
<tr>
<td>Concept</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>numeric</td>
</tr>
<tr>
<td>Files</td>
<td>F1 F2</td>
</tr>
<tr>
<td>Full Description</td>
<td>An indicator for whether the applicant ever received SSI benefits.</td>
</tr>
<tr>
<td>Variable Name</td>
<td>ssr_ssi_benefit_fed_totamt</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>Label</td>
<td>SSR: Total federal benefit amount</td>
</tr>
<tr>
<td>Concept</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>numeric</td>
</tr>
<tr>
<td>Files</td>
<td>F1 F2</td>
</tr>
<tr>
<td>Full Description</td>
<td>The total SSI benefit amount from federal sources.</td>
</tr>
</tbody>
</table>
Variable Name  ssr_ssi_ceased
Label  SSR: Benefits ceased
Concept
Type  numeric
Files  F1 F2
Variable Name: ssr_ssi_dig_group

Label: SSR: Diagnosis code

Concept

Type: numeric

Files: F1 F2

Full Description

This variable contains the diagnostic group for the SSI recipient's primary code for mental or physical disability used in the medical determination of the individual's eligibility for disability benefits. Codes: 01?Diseases of the blood; 02?Circulatory system; 03?Congenital anomalies; 04?Digestive system; 05?Endocrine, nutritional, and metabolic diseases; 06?Genitourinary system; 07?Infectious and parasitic diseases; 08?Injuries; 09?Mental disorders; 10?Mental retardation; 11?Musculoskeletal system; 12?Neoplasms; 13?Nervous system and sense organs; 14?Respiratory system; 15?Skin and subcutaneous tissue; 16?Other or unknown.
Variable Name        ssr_ssi_first_pmt_dt
Label                SSR: First payment date
Concept              
Type                 numeric
Files                F1 F2

Full Description

The date of first recorded payment of SSI benefits. This variable is saved as a SAS date.
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>ssr_ssi_last_pmt_dt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>SSR: Last payment date</td>
</tr>
<tr>
<td>Concept</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>numeric</td>
</tr>
<tr>
<td>Files</td>
<td>F1 F2</td>
</tr>
<tr>
<td>Variable Name</td>
<td>ssr_ssi_type</td>
</tr>
<tr>
<td>----------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>Label</td>
<td>SSR: Type of benefit</td>
</tr>
<tr>
<td>Concept</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>numeric</td>
</tr>
<tr>
<td>Files</td>
<td>F1 F2</td>
</tr>
</tbody>
</table>

**Full Description**

Type of SSI benefit applied for or received by the individual: 1?Aged individual; 2?Disabled or blind individual; 3?Disabled or blind child.
**Variable Name**  
state

**Label**  
State of Residence: FIPS code (modified)

**Concept**

**Type**  
numeric

**Files**  
F1 F2

**Full Description**

State of residence. FIPS State Code for state of residence first recorded in the SIPP. For married couples, we take the state value for both partners at the same point in the survey when we first observed the marriage. For individuals who never have an observed marriage during the panel, we take their first ever reported state value. *All panels prior to 2004 group some states together and give only one code for the group. For these panels, the individual FIPS code will not appear for states contained in a group.

**Values (56 total)**

1. Alabama
2. Alaska *see description
3. Arizona
4. Arkansas
5. California
6. Colorado
7. Connecticut
8. Delaware
9. DC
10. Florida
11. Georgia
12. Hawaii

<table>
<thead>
<tr>
<th></th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>Idaho *see description</td>
</tr>
<tr>
<td>17</td>
<td>Illinois</td>
</tr>
<tr>
<td>18</td>
<td>Indiana</td>
</tr>
<tr>
<td>19</td>
<td>Iowa *see description</td>
</tr>
<tr>
<td>20</td>
<td>Kansas</td>
</tr>
<tr>
<td>21</td>
<td>Kentucky</td>
</tr>
<tr>
<td>22</td>
<td>Louisiana</td>
</tr>
<tr>
<td>23</td>
<td>Maine *see description</td>
</tr>
<tr>
<td>24</td>
<td>Maryland</td>
</tr>
<tr>
<td>25</td>
<td>Massachusetts</td>
</tr>
<tr>
<td>26</td>
<td>Michigan</td>
</tr>
<tr>
<td>27</td>
<td>Minnesota</td>
</tr>
<tr>
<td>28</td>
<td>Mississippi *see description</td>
</tr>
<tr>
<td>29</td>
<td>Missouri</td>
</tr>
<tr>
<td>30</td>
<td>Montana *see description</td>
</tr>
<tr>
<td>31</td>
<td>Nebraska</td>
</tr>
<tr>
<td>32</td>
<td>Nevada</td>
</tr>
<tr>
<td>33</td>
<td>New Hampshire</td>
</tr>
<tr>
<td>34</td>
<td>New Jersey</td>
</tr>
<tr>
<td>35</td>
<td>New Mexico *see description</td>
</tr>
<tr>
<td>36</td>
<td>New York</td>
</tr>
<tr>
<td>37</td>
<td>North Carolina</td>
</tr>
<tr>
<td>38</td>
<td>North Dakota *see description</td>
</tr>
<tr>
<td>39</td>
<td>Ohio</td>
</tr>
<tr>
<td>40</td>
<td>Oklahoma</td>
</tr>
<tr>
<td>41</td>
<td>Oregon</td>
</tr>
<tr>
<td>42</td>
<td>Pennsylvania</td>
</tr>
<tr>
<td>Page</td>
<td>States</td>
</tr>
<tr>
<td>------</td>
<td>--------</td>
</tr>
<tr>
<td>44</td>
<td>Rhode Island</td>
</tr>
<tr>
<td>45</td>
<td>South Carolina</td>
</tr>
<tr>
<td>46</td>
<td>South Dakota *see description</td>
</tr>
<tr>
<td>47</td>
<td>Tennessee</td>
</tr>
<tr>
<td>48</td>
<td>Texas</td>
</tr>
<tr>
<td>49</td>
<td>Utah</td>
</tr>
<tr>
<td>50</td>
<td>Vermont *see description</td>
</tr>
<tr>
<td>51</td>
<td>Virginia</td>
</tr>
<tr>
<td>53</td>
<td>Washington</td>
</tr>
<tr>
<td>54</td>
<td>West Virginia *see description</td>
</tr>
<tr>
<td>55</td>
<td>Wisconsin</td>
</tr>
<tr>
<td>56</td>
<td>Wyoming *see description</td>
</tr>
<tr>
<td>90</td>
<td>(1984 panel only) Idaho, New Mexico, South Dakota, Wyoming *see description</td>
</tr>
<tr>
<td>91</td>
<td>(1984 panel only) Mississippi, West Virginia *see description</td>
</tr>
</tbody>
</table>
Variable Name: time_arrive_usa

Label: Time Period of Arrival to US (Foreign Born)


Type: numeric

Files: F1 F2

Values (11 total):

1  Before 1959
2  1960-1964
3  1965-1969
4  1970-1974
5  1975-1979
6  1980-1981
7  1982-1984
8  1985-1993
9  1994-1999
10 2000-2004

Sysmiss: Structurally missing, out of scope for question (foreign_born=0)
Value Ranges

Value Range

Range: \([1, 10]\)
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>total_der_fica_YYYY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>DER: Total earnings from FICA-covered jobs</td>
</tr>
<tr>
<td>Concept</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>numeric</td>
</tr>
<tr>
<td>Files</td>
<td>F1 F2</td>
</tr>
</tbody>
</table>

**Full Description**

Total earnings from all FICA-covered jobs with W-2 or Schedule C (self-employment) filings. These earnings are the sum of amounts from Box 1 (Total Wages, Tips, and Bonuses) and Box 12 (earnings deferred to a 401(k) type account). This array extends from 1978-2011.
Variable Name: total_der_nonfica_YYYY

Label: DER: Total earnings from all non-FICA jobs

Concept:

Type: numeric

Files: F1 F2

Full Description:

Total earnings from all non-FICA-covered jobs with W-2 or Schedule C (self-employment) filings. These earnings are the sum of amounts from Box 1 (Total Wages, Tips, and Bonuses) and Box 12 (earnings deferred to a 401(k) type account). This array extends from 1978-2011.
Variable Name: totearn_YYYYM

Label: Total Earnings

Concept:

Type: numeric

Files: F1 F2

Full Description:

This variable contains total monthly person-level earnings as reported in the SIPP for year YYYY and month MM. This includes both job and business (self-employment) earnings and is top-coded. Only the first 24 months of the individual's SIPP panel are in universe. A later release of these data will add months after 24 for longer SIPP panels. Months that are outside the time frame covered by an individual's SIPP panel will always be missing and out of universe.

Values (1 total):

Sysmiss
<table>
<thead>
<tr>
<th><strong>Variable Name</strong></th>
<th>totearn_ser_YYYY</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Label</strong></td>
<td>SER: Capped Earnings from all FICA-covered jobs</td>
</tr>
<tr>
<td><strong>Concept</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Type</strong></td>
<td>numeric</td>
</tr>
<tr>
<td><strong>Files</strong></td>
<td>F1 F2</td>
</tr>
<tr>
<td><strong>Full Description</strong></td>
<td>Person-level annual earnings that were taxed by FICA; these variables include earnings only up to the FICA taxable maximum and cover the years 1951-2011. These earnings are the inputs for calculating the OASDI benefit a person and his or her spouse will receive upon retirement or disability.</td>
</tr>
<tr>
<td><strong>Value Ranges</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Value Range</strong></td>
<td>Range: [ 0 , 94192 ]</td>
</tr>
<tr>
<td>Variable Name</td>
<td>tohours_YYYYM</td>
</tr>
<tr>
<td>------------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>Label</td>
<td>Total Usual Weekly Hours Worked at All Jobs</td>
</tr>
<tr>
<td>Concept</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>numeric</td>
</tr>
<tr>
<td>Files</td>
<td>F1 F2</td>
</tr>
</tbody>
</table>

**Full Description**

Total number of usual weekly hours worked at all regular employment jobs in year YYYYY and month MM. Only the first 24 months of the individual's SIPP panel are in universe. A later release of these data will add months after 24 for longer SIPP panels. Months that are outside the time frame covered by an individual's SIPP panel will always be missing and out of universe.

**Values (1 total)**

<p>| Sysmiss |</p>
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>totinc_YYYYM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>Total Personal Income</td>
</tr>
<tr>
<td>Concept</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>numeric</td>
</tr>
<tr>
<td>Files</td>
<td>F1 F2</td>
</tr>
</tbody>
</table>

**Full Description**

Total personal income summed from all sources in year YYYY and month MM. Only the first 24 months of the individual's SIPP panel are in universe. A later release of these data will add months after 24 for longer SIPP panels. Months that are outside the time frame covered by an individual's SIPP panel will always be missing and out of universe.

**Values (1 total)**

<p>| Sysmiss |</p>
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>totnetworth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>Total Net Worth</td>
</tr>
<tr>
<td>Concept</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>numeric</td>
</tr>
<tr>
<td>Files</td>
<td>F1 F2</td>
</tr>
</tbody>
</table>

**Full Description**

Total net worth as reported in the wealth topical module, collected in the following waves, by panel: Wave 4 in the 1984 (HHTNW) & 1990-1993 (HH_TNW) panels; Wave 3 in 1996, 2001, & 2004 (THHTNW) panels; Wave 4 in 2008 panel (THHTNW).

**Value Ranges**

<table>
<thead>
<tr>
<th>Value Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range: [-200331.076019287, 2700581.59277344]</td>
</tr>
</tbody>
</table>
Variable Name: vcamt_YYYYM

Label: Amount of veterans compensation or pension benefit

Concept:

Type:

Files:

Full Description:

This variable is the amount of veterans compensation or pension benefits that a respondent received in this month and year. The variable comes from the core file. I08AMT1-I08AMT4 in 1984; S08AMT in the 1990, 1991, 1992, and 1993 panels; T08AMT in the 1996, 2001, 2004, and 2008 panels.
Variable Name  vetrecip_YYYYM

Label  Indicator for receipt of veterans compensation or pension benefits

Concept

Type

Files

Full Description

This variable indicates that a respondent received veterans compensation or pension benefits in this month and year. The variable comes from the core file. I08REC1-I08REC4 in 1984; R08 in the 1990, 1991, 1992, and 1993 panels; ER08 in the 1996, 2001, 2004, and 2008 panels.

Values (2 total)

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Did not receive veterans compensation or veterans benefits</td>
</tr>
<tr>
<td>1</td>
<td>Received veterans compensation or veterans benefits</td>
</tr>
</tbody>
</table>
Variable Name: wcamt_YYYYM

Label: Amount of Workers Compensation Received

Concept

Type

Files

Full Description

Variable Name  wkcomp_YYYYM

Label  Indicator for receipt of workers compensation

Concept

Type

Files

Full Description


Values (2 total)

0  Did not receive workers compensation benefits

1  Received workers compensation benefits
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>wksjob_YYYYM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Label</td>
<td>Weeks at a Job</td>
</tr>
<tr>
<td>Concept</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>numeric</td>
</tr>
<tr>
<td>Files</td>
<td>F1 F2</td>
</tr>
</tbody>
</table>

**Full Description**

Total number of weeks that the respondent held a job in year YYYY and month MM. Only the first 24 months of the individual's SIPP panel are in universe. A later release of these data will add months after 24 for longer SIPP panels. Months that are outside the time frame covered by an individual's SIPP panel will always be missing and out of universe.

**Values (1 total)**

Sysmiss
Variable Name  wkswp_YYYYM
Label  Weeks With Pay
Concept
Type  numeric
Files  F1 F2

Full Description
Total number of weeks worked with pay in year YYYY and month MM. Only the first 24 months of the individual's SIPP panel are in universe. A later release of these data will add months after 24 for longer SIPP panels. Months that are outside the time frame covered by an individual's SIPP panel will always be missing and out of universe.

Values (1 total)
Sysmiss
Variable Name: wqc_yrtot_YYYY

Label: SER: Annual Total Covered Quarters of Work

Concept:

Type: numeric

Files: F1 F2

Full Description:

Indicates the total number of quarters of FICA-covered work in year YYYY, where YYYY=1951-2011. In 2011, an individual received credit for one quarter of coverage for every $ he/she earned at a FICA-covered job in the year. In 1978, $250 was required in order to be credited one quarter of coverage. This number increases automatically each year and is tied to increases in average wages. A maximum of 4 quarters per year is possible. Quarters of coverage are used to calculate eligibility for OASDI benefits. For example, people who reached age 62 in 2011 were eligible to retire if they had at least 40 quarters of coverage. The number of quarters of coverage required to be fully insured has changed over time. See http://www.socialsecurity.gov/history/reports/crsleghist2.html for a summary of rules. Prior to 1978, a quarter of coverage was granted for every quarter of the calendar year in which the worker had at least $50 in earnings at a FICA-covered job. Hence the portion of this array from 1951-1977 provides more detail about whether a worker was employed throughout the year. However caution should be used when looking at these earlier years as far fewer jobs were FICA-covered and hence zero quarters of coverage does not mean the person as not employed but rather that the person was not employed by a job that required FICA taxes or made the person eligible for OASDI benefits.

Values (5 total):

<table>
<thead>
<tr>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
</tbody>
</table>

Value Ranges:
Value Range

Range: $[0, 4]$
Variable Name: year_bach

Label: Year of Bachelors Degree

Concept:

Type: numeric

Files: F1 F2

Full Description:

This variable contains the year the respondent finished his/her bachelor's degree. This information comes from the education history topical module, conducted in wave 2 of each panel. Individuals must have educ_5cat=4 to be in scope for this variable.

Value Ranges:

Value Range:

Range: [ 1926 , 2009 ]
Variable Name: year_beg_posths

Label: Year Began Post-HS Education

Concept

Type: numeric

Files: F1 F2

Full Description

This variable gives the year that post-high school education began and comes from the wave 2 Education and Training History topical module (variable TM8420 for 1984, 1990-1993 panels; variable tcollstr for 1996-2008 panels). This variable is in scope if educ_5cat is 3 or greater.

Values (1 total)

Sysmiss - Structurally missing (educ_5cat=1 or educ_5cat=2)

Value Ranges

Value Range

Range: [1922, 2009]
Variable Name: year_end_hs

Label: Year Ended HS (or less) Education

Concept:

Type: numeric

Files: F1 F2

Full Description:

This variable gives the year that high school education ended and comes from the wave 2 Education and Training History topical module (variable TM8404 and TM8412 for 1984, 1990-1993 panels; variable tistschl and thsyr for 1996-2008 panels). This variable is in scope for all levels of educ_5cat.

Value Ranges:

Value Range:

Range: [ 1914, 2009 ]
Variable Name        year_end_posths
Label                Year Ended Post-HS Education
Concept
Type                 numeric
Files                F1 F2

Full Description

This variable gives the year that post-high school education ended and comes from the wave 2 Education and Training History topical module (variables TM8426 and TM8440 for 1984, 1990-1993 panels; variables tlastcol, tvocyr, tassocyr, tbachyr, and tadvncyr for 1996-2008 panels). This variable is in scope if educ_5cat is 3 or greater.

Values (1 total)

   Sysmiss - Structurally missing (educ_5cat=1 or educ_5cat=2)

Value Ranges

Value Range

Range: [1926, 2009]