This document was generated: November 2, 2018 at 3:59 PM

SIPP Synthetic Beta v7

(141 variables)
Last update to metadata: 2018-11-02 15:55:02 (upload date)
Document Date: 1 September 2018

Codebook prepared by: Cornell NSF-Census Research Network

Data prepared by: United States Department of Commerce. Bureau of the Census.

Principal Investigator(s): United States Department of Commerce. Bureau of the Census. , Social Security Administration. , Internal Revenue Service. , and Cornell University. Labor Dynamics Institute.

Citation

Please cite this codebook as:

Lori B. Reeder and Jordan C Stanley and Lars Vilhuber. Codebook for the SIPP Synthetic Beta 7.0 [Codebook file]. Cornell Institute for Social and Economic Research and Labor Dynamics Institute [distributor]. Cornell University, Ithaca, NY, 2018

Please cite this dataset as:

U.S. Census Bureau. SIPP Synthetic Beta: Version 7.0 [Computer file]. Washington DC; Cornell University, Synthetic Data Server [distributor], Ithaca, NY, 2018

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. Eight SIPP panels (1990, 1991, 1992, 1993, 1996, 2001, 2004, 2008) 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.

Users should be aware that time-varying variable arrays are collapsed in this codebook to a placeholder variable. Thus, *afdc_Y_M* stands in for the year and month variables *afdc_1990_1--afdc_2008_12* .

Datasets

ssb_v7_0_syntheticK_M.sas7bdat
http://www.census.gov/programs-surveys/sipp/methodology/sipp-synthetic-beta-data-product.html (SAS)
ssb_v7_0_syntheticK_M.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:

http://www.census.gov/programs-surveys/sipp/methodology/sipp-synthetic-beta-data-product.html

Access Conditions

You will need to use an NX client to logon to the Synthetic Data Server. Information about how to set-up your account and use the Synthetic Data Servers will come to you directly from the staff that maintains this server, after approval of your access by Census staff.

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 Grants SES-1042181 and BCS-0941226, and through a grant from the Alfred P. Sloan Foundation. 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

I. 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 feasability 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 .

IV. Changes since v6:

- The calculation of "Total Usual Weekly Hours" has changed. It now relies solely on SIPP information. The variable name has therefore changed from *tothours_M* to *totalhrs_Y_M* .

- The following variables are new: *date_enter_sipp, mdy_lft_sipp, panel_1styear, flag_in_tmhist_ms, first_birth_year, last_birth_year, halfsamp, varstrat, mom_personid, ser_posearn_YYYY, mbr_ssdi_ben_dof, mbr_ssdi_deny_dof, mbr_ssdi_bdof_app, mbr_ssdi_ddof_app, flag_enter_late, flag_lft_sipp, life_ins_in_scope, pension_in_scope_age, fert_in_scope, brthmn_sipp, brthyr_sipp, disab_worklimit_t, disab_nowork_t* .

- The following variable roots have changed (*oldvar_W* ind_cat

- The following variables were removed: *ind_exist, occ_exist, left_layoff_M, disab_worklimit, disab_nowork*

- The variable nomenclature based on months *_M* was changed to *_Y_M* - see more details elsewhere in this document.

Related Publications

I. U.S. Census Bureau, "Disclosure Review Board Memo: Second Request for Release of SIPP Synthetic Beta Version 6.0," U.S. Census Bureau 2015.

   Available at http://www.census.gov/content/dam/Census/programs-surveys/sipp/methodology/DRBMemoTablesVersion2SSBv6_0

II. J. M. Abowd, M. Stinson, and G. Benedetto, "Final Report to the Social Security Administration on the SIPP/SSA/IRS Public Use File Project," U.S. Census Bureau 2006. Available at https://www2.vrdc.cornell.edu/news/wp-content/papercite-data/pdf/ssafinal.pdf .

Related Studies

I. L. B. Reeder, M. Stinson, K. E. Trageser, and L. Vilhuber, "Codebook for the SIPP Synthetic Beta v5.1 [Codebook file]," {Cornell Institute for Social and Economic Research} and {Labor Dynamics Institute} [distributor]. Cornell University, Ithaca, NY, USA, DDI-C document, 2014. Available at http://www2.ncrn.cornell.edu/ced2ar-web/codebooks/ssb/v/v51 .

Variable Groups - SIPP Synthetic Beta v7

Aged Spouse Benefit
Benefits Variables
Demographic Variables
Detailed Earnings Record Variables
Disability Benefit
Disability Variables
Economic Variables
Education Variables
Fertility Variables
Geographic Variables
Health Insurance Variables
IRS/SSA Variables
Identifiers
Income Variables
Labor Force Variables
Lifespan Variables
MBR/PHUS Variables
Marital History Variables
Other Benefit
Retirement Benefit
SIPP Arrays
Summary Earnings Record Variables
Supplemental Security Record Variables
Widowed Spouse Benefit

Variable Name

panel_1stfullyear

Label

First full calendar year observed in panel

Concept

Type

numeric

Files

F1 F2

Full Description

This is the first year in the panel for which every rotation group is in scope to have all the monthly SIPP variables from January to December.

Groups

Identifiers

Variable Name

flag_valid_ssn

Label

Respondent has a valid Social Security Number

Concept

Type

numeric

Files

F1 F2

Full Description

This variable indicates whether a respondent successfully linked to a valid Social Security Number.

Values (2 total)

0	Did not link to a Social Security Number
1	Successfully linked to a Social Security Number

Groups

Variable Name

afdc_Y_M

Label

Indicator for receipt of AFDC or TANF benefits

Concept

Type

Files

Full Description

This variable indicates that a respondent received public assistance payments (AFDC or TANF) in month M of year Y. M goes from 1 to 12 and represents the calendar month. Y goes from 1 to 6 and represents the number calendar year for a given SIPP panel. The variable comes from the core file. The relevant SIPP variables are I20REC1, I20REC2, I20REC3, I20REC4 in 1984; R20 in the 1990, 1991, 1992, and 1993 panels; ER20 in the 1996, 2001, 2004, and 2008 panels.

Values (2 total)

0	Did not receive public assistance payment
1	Received public assistance payment

Groups

Variable Name

afdcamt_Y_M

Label

Amount of AFDC received

Concept

Type

numeric

Files

Full Description

This variable is the amount of public assistance payments (AFDC or TANF) that a respondent received month M of year Y. M goes from 1 to 12 and represents the calendar month. Y goes from 1 to 6 and represents the number calendar year for a given SIPP panel. The variable comes from the core file. The relevant SIPP variables are I20AMT1-I20AMT4 in 1984; S20AMT in the 1990, 1991, 1992, and 1993 panels; T20AMT in the 1996, 2001, 2004, and 2008 panels.

Groups

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.

Groups

Demographic Variables
Lifespan Variables

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)

0	Not currently enrolled in college
1	Currently enrolled in college
Sysmiss	

Groups

Education Variables

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 and 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 (2 total)

1	Currently enrolled in high school
Sysmiss	

Groups

Education Variables

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)

0	No defined benefit pension plan
1	Had defined benefit pension plan
Sysmiss	

Groups

Economic Variables

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)

0	No defined contribution pension plan
1	Had defined contribution pension plan
Sysmiss	

Groups

Economic Variables

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 death and January 1, 1960. An individual who died on January 1, 1959 would have deathdate=-365 and an individual who died on January 1, 1961 would have deathdate=365.

Groups

Demographic Variables
Lifespan Variables

Variable Name

defer_der_fica_YYYY

Label

DER: Deferred FICA

Concept

Type

numeric

Files

F1 F2

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 1990-2014. 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.

Groups

Detailed Earnings Record Variables

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 1990-2014. 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.

Groups

Detailed Earnings Record Variables

Variable Name

disab_nowork_t

Label

Disability Prevents Work - (wave a or b)

Concept

Type

numeric

Files

F1 F2

Full Description

This variable indicates that a person had a work-preventing disability at time t where t = a (earlier wave) or b (later wave). If an individual indicates a work-preventing disability in wave 2, disab_nowork_a = 1. In all panels but 1984, this question is asked again in a later wave: individuals indicating a work-preventing disability in that question have disab_nowork_b = 1. The universe is all individuals who were at least age 15, no older than age 70 by the end of the panel, and answered "yes" to having a work-limiting disability. The relevant SIPP variable for disab_nowork_a is TM8332 for the 1990-1993 panels and eprevwk for the 1996, 2001, 2004, and 2008 panels. For disab_nowork_b, the relevant SIPP variable is TM8924 in the topical module for the 1990-1993 panels and edisprev in the core of the 1996, 2001, 2004, and 2008 panels. The variable comes from wave 3 for 1990, 1991, and 1993; wave 5 for 1996, 2001, and 2004; and wave 6 for 1992 and 2008.

Values (2 total)

0	Disability does not prevent work
1	Disability prevents work

Groups

Disability Variables

Variable Name

disab_worklimit_t

Label

Work Limiting Disability - (wave a or b)

Concept

Type

numeric

Files

F1 F2

Full Description

This variable indicates that a person had a work-limiting disability at time t where t = a (earlier wave) or b (later wave). The information for wave a always comes from the disability topical module in wave 2. If an individual indicates a work-limiting disability in wave 2, disab_worklimit_a = 1. In all panels but 1984, this question is asked again in a later wave. The wave 2 topical module variable that informs disab_worklimit_a is TM8452 in the 1984 panel, TM8306 in the 1990, 1991, 1992, and 1993 panels, and ELMTVER in the 1996, 2001, 2004, and 2008 panels. Individuals asked this question are aged 16 to 67 except in 1984 where the age universe extends to age 72. The respective SIPP variables that inform disab_worklimit_b are TM8918 and TM8920 (in the topical module) for 1990, 1991, 1992, and 1993 panels and EDISABL (in the core) for 1996, 2001, 2004, and 2008 panels. The relevant variable can be found in wave 3 for 1990, 1991, and 1993; wave 5 for 1996, 2001, and 2004; and wave 6 for 1992 and 2008. 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.

Values (2 total)

0	No work-limiting disability
1	Work-limiting disability

Groups

Disability Variables

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 variables 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), 1996 and 2001 panels (EATTAIN); and 2004 and 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]

Groups

Education Variables

Variable Name

field_bach

Label

Field of Bachelors Degree

Concept

Type

numeric

Files

F1 F2

Full Description

Field of bachelors degree as reported in the education history topical module. Universe is individuals who were age 15 by beginning of their SIPP panel and who had a bachelors degree. This topical module was asked in the following waves, by panel: Wave 3 in 1984 panel (TM8038); Wave 2 in the 1990-1993 panels (TM8428, TM8436); Wave 2 in the 1996-2008 panels (EBACHFLD). Categories vary between the 1996-2008 panels and the 1984 and 1990-1993 panels.

Values (22 total)

0	Unknown
1	Agriculture/Forestry
2	Art/Architecture (1996-2008 panels); Biology (1984, 1990-1993 panels)
3	Business/Management
4	Communications (1996-2008 panels); Economics (1984, 1990-1993 panels)
5	Computer and Information Sciences (1996-2008 panels); Education (1984, 1990-1993 panels)
6	Education (1996-2008 panels); Engineering (1984, 1990-1993 panels)
7	Engineering (1996-2008 panels); English/Journalism (1984, 1990-1993 panels)
8	English/Literature (1996-2008 panels); Home Economics (1984, 1990-1993 panels)
9	Foreign Language (1996-2008 panels); Law (1984, 1990-1993 panels)
11	Health Sciences (1996-2008 panels); Liberal Arts/Humanities (including arts, architecture, music, languages, philosophy) (1984, 1990-1993 panels)
11	Liberal Arts/Humanities (1996-2008 panels); Mathematics/Statistics (1984, 1990-1993 panels)

12	Mathematics/Statistics (1996-2008 panels); Medicine (1984, 1990-1993 panels)
13	Nature Sciences (Biological and Physical) (1996-2008 panels); Nursing, Pharmacy, Health Technologies (1984, 1990-1993)
14	Philosophy/Religion/Theology (1996-2008 panels); Physical or Earth Sciences (1984, 1990-1993 panels)
15	Pre-Professional (1996-2008 panels); Police Science or Law Enforcement (1984, 1990-1993 panels)
16	Psychology
17	Social Sciences/History (1996-2008 panels); Religion/Theology (1984, 1990-1993 panels)
18	Other (1996-2008 panels); Social Sciences (1984, 1990-1993 panels)
19	Vocational or Technical Studies (1984, 1990-1993 panels)
20	Other (1984, 1990-1993 panels)
Sysmiss	

Value Ranges

Value Range

Range: [0 , 20]

Groups

Education Variables

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.

Groups

Demographic Variables
Fertility Variables

Variable Name

flag_deathdate_exist

Label

Flag: Existence of Date of Death

Concept

Type

numeric

Files

F1 F2

Full Description

Flag to indicate that this respondent died after being interviewed and before 2014. This flag indicates the existence of a valid value for deathdate.

Values (2 total)

0	Death date does not exist, respondent did not die during this interval
1	Death date exists, respondent died during this interval

Value Ranges

Value Range

Range: [0 , 1]

Groups

Demographic Variables
Lifespan Variables

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)

0	No additional marriage occurred with unknown date
1	An additional marriage occurred but with unknown date
Sysmiss	

Value Ranges

Value Range

Range: [0 , 1]

Groups

Marital History Variables

Variable Name

foodstp_Y_M

Label

Indicator for receipt of SNAP/Food Stamps

Concept

numeric

Type

Files

Full Description

This variable indicates that a respondent received food stamps/SNAP benefits in month M of year Y. M goes from 1 to 12 and represents the calendar month. Y goes from 1 to 6 and represents the number calendar year for a given SIPP panel. This variable comes from the core file. The relevant SIPP variables are I27REC1, I27REC2, I27REC3, I27REC4 in the 1984 panel; R27 in the 1990, 1991, 1992, and 1993 panels; and ER27 in the 1996, 2001, 2004, and 2008 panels.

Values (2 total)

0	Did not receive food stamps/SNAP benefits
1	Received food stamps/SNAP benefits

Groups

Variable Name

foreign_born

Label

Foreign Born

Concept

Type

numeric

Files

F1 F2

Full Description

Immigrant Status, born in country other than U.S. Taken from the Wave 8 topical module in the 1984 panel and Wave 2 topical module in the 1990, 1991, 1992, 1993, 1996, 2001, 2004, and 2008 panels. The relevant variables from the SIPP are TM8128, TM8166, and TM8174 in the 1984 panel; TM8730, TM8734, and TM8709 in the 1990, 1991, 1992, and 1993 panels; EPRSTATE, EBRSTATE, and RCITIZNT in the 1996 panel; EPRSTATE, EBRSTATE, and TCITIZNT for the 2001 panel; EPRSTATE, EBRSTATE, CITIZ, and EBORNUS in the 2004 and 2008 panels.

Values (2 total)

0	Born in U.S.
1	Born in country other than U.S.

Value Ranges

Value Range

Range: [0 , 1]

Groups

Demographic Variables

Variable Name

fsamt_Y_M

Label

SNAP/Food Stamps Amount Received

Concept

Type

numeric

Files

Full Description

This variable is the amount of food stamps/SNAP benefits that a respondent received in month M of year Y. M goes from 1 to 12 and represents the calendar month. Y goes from 1 to 6 and represents the number calendar year for a given SIPP panel. The relevant SIPP variables are I27AMT1, I27AMT2, I27AMT3, I27AMT4 in the 1984 panel; S27AMT in the 1990, 1991, 1992, and 1993 panels; T27AMT in the 1996, 2001, 2004, and 2008 panels.

Groups

Variable Name

hicov_Y_M

Label

Health Insurance Coverage

Concept

Type

numeric

Files

F1 F2

Full Description

A variable in this array indicates whether an individual was covered by health insurance during month M of year Y. M goes from 1 to 12 and represents the calendar month. Y goes from 1 to 6 and represents the number calendar year for a given SIPP panel. 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 SIPP variables: HIIND for the 1984, 1990, 1991, 1992, and 1993 panels; EHIMTH for the 1996, 2001, 2004, and 2008 panels.

Values (3 total)

0	Respondent did not have health insurance coverage during this month
1	Respondent had health insurance coverage during this month
Sysmiss	

Value Ranges

Value Range

Range: [0 , 1]

Groups

Health Insurance Variables

Variable Name

hiemp_Y_M

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 month M of year Y. M goes from 1 to 12 and represents the calendar month. Y goes from 1 to 6 and represents the number calendar year for a given SIPP panel. Employer-provided health insurance coverage is derived from the following variables: HISRC in the 1990, 1991, 1992, and 1993 panels; EHEMPLY in the 1996, 2001, 2004, and 2008 panels. Note that this question is not asked in the 1984 panel.

Values (3 total)

0	Respondent did not have employer-provided health insurance
1	Respondent had employer-provided health insurance
Sysmiss	

Value Ranges

Value Range

Range: [0 , 1]

Groups

Health Insurance Variables

Variable Name

hispanic

Label

Hispanic

Concept

Type

numeric

Files

F1 F2

Full Description

In the 1984, 1990, 1991, 1992, and1993 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, 2001, 2004, and 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)

0	Non-hispanic
1	Hispanic

Value Ranges

Value Range

Range: [0 , 1]

Groups

Demographic Variables

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 panel (variable name HHTHEQ); Wave 4 in the 1990, 1991, 1992, and 1993 panels (HH_THEQ); Wave 3 in the 1996, 2001, and 2004 panels (THHTHEQ); and Wave 4 in 2008 panel (THHTHEQ).

Groups

Economic Variables

Variable Name

ind_cat_W

Label

Industry Category for Primary Job in Wave W

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 (21 total)

1	Agriculture, Fishing, Forestry, Hunting
2	Gas, Mining, Oil
3	Construction
4	Manufacturing
5	Wholesale Trade
6	Resale Trade
7	Transportation and Warehousing
8	Utilities
9	Information
10	Finance and Insurance
11	Real Estate, Rental of Goods

12	Professional, Scientific, and Technical Services
13	Management of companies/enterprises
14	Administrative, support, waste management services
15	Educational Services
16	Health Care and Social Assistance
17	Arts, Entertainment, and Recreation
18	Accommodation and Food Services
19	Other Services, except public domain
20	Public Administration
21	Military (specifically, last or present job was or is in the armed forces)

Value Ranges

Value Range

Range: [1 , 4]

Groups

Economic Variables

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-instutionalized 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.

Groups

Variable Name

last_admin_birthdate

Label

Administrative birthdate of last born child

Concept

Type

numeric

Files

F1 F2

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.

Groups

Demographic Variables
Fertility Variables

Variable Name

layoff_Y_M

Label

On Layoff (Without Pay)

Concept

Type

numeric

Files

F1 F2

Full Description

Indicator that the individual was on layoff without pay in month M of year Y. M goes from 1 to 12 and represents calendar month. Y goes from 1 to 6 and represents number calendar year of the given SIPP panel. Months that are outside the time frame covered by an individual's SIPP panel will always be missing and out of universe. The relevant SIPP variables are SC1098 in the 1984 panel; REASAB in the 1990, 1991, 1992, and 1993 panels; and ELAYOFF in the 1996, 2001, 2004, and 2008 panels.

Values (2 total)

0	Not on layoff without pay
1	On layoff without pay

Groups

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)

0	No
1	Yes

Groups

Variable Name

male

Label

Male

Concept

Type

numeric

Files

F1 F2

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.

Values (2 total)

0	Female
1	Male

Value Ranges

Value Range

Range: [0 , 1]

Groups

Demographic Variables

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)

0	No
1	Yes

Groups

MBR/PHUS Variables
Aged Spouse Benefit

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.

Groups

MBR/PHUS Variables

Aged Spouse Benefit

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).

Groups

MBR/PHUS Variables
Aged Spouse Benefit

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)

0	No
1	Yes

Groups

MBR/PHUS Variables

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.

Groups

MBR/PHUS Variables

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.

Groups

MBR/PHUS Variables

Variable Name

mbr_ssdi_applied_k

Label

MBR SSDI: Application submitted

Concept

Type

numeric

Files

F1 F2

Full Description

This set of variables (mbr_ssdi_applied_1mbr_ssdi_applied_4 where K=1 to 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_ssdi_applied_1 = 1, mbr_ssdi_applied_2 = 1, mbr_ssdi_applied_3 = 1, mbr_ssdi_applied_4 = 0.

Values (2 total)

0	Did not apply for SSDI
1	Applied for SSDI

Groups

MBR/PHUS Variables
Disability Benefit

Variable Name

mbr_ssdi_benefit_totamt_k

Label

MBR SSDI: Benefit amount

Concept

Type

numeric

Files

F1 F2

Full Description

This variable provides the monthly amount of SSDI benefits received by the individual for each of four applications. See mbr_ssdi_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.

Groups

MBR/PHUS Variables
Disability Benefit

Variable Name

mbr_ssdi_ceased_k

Label

MBR SSDI: Benefits ceased

Concept

Type

numeric

Files

F1 F2

Full Description

This set of variables records whether the individual ever ceased receiving previously granted benefits under each SSDI application record (mbr_ssdi_ceased_k = 1). See mbr_ssdi_applied_k for details on how individual applications are recorded on this file.

Values (2 total)

0	Did not cease receiving benefit
1	Ceased receiving benefit

Groups

MBR/PHUS Variables
Disability Benefit

Variable Name

mbr_ssdi_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_ssdi_applied_k for details on how individual applications are recorded on this file.

Groups

MBR/PHUS Variables
Disability Benefit

Variable Name

mbr_ssdi_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_ssdi_applied_k for details on how individual applications are recorded on this file.

Groups

MBR/PHUS Variables
Disability Benefit

Variable Name	mbr_ssdi_dig_group_k
Label	MBR SSDI: Diagnosis group
Concept	
Type	numeric
Files	F1 F2

Full Description

This variable contains the diagnostic group for the SSDI recipient's primary code for mental or physical disability used in the medical determination of the individual's eligibility for disability benefits.

Values (16 total)

1	Diseases of the blood
2	Circulatory system
3	Congenital anomalies
4	Digestive system
5	Endocrine, nutritional, and metabolic diseases
6	Genitourinary system
7	Infectious and parasitic diseases
8	Injuries
9	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

Groups

MBR/PHUS Variables

Disability Benefit

Variable Name

mbr_ssdi_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_ssdi_applied_k for details on how individual applications are recorded on this file.

Groups

MBR/PHUS Variables
Disability Benefit

Variable Name

mbr_ssdi_dsd_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_ssdi_applied_k for details on how individual applications are recorded on this file.

Groups

MBR/PHUS Variables
Disability Benefit

Variable Name

mbr_ssdi_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_ssdi_applied_k for details on how individual applications are recorded on this file.

Groups

MBR/PHUS Variables
Disability Benefit

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)

0	No
1	Yes

Groups

MBR/PHUS Variables
Widowed Spouse Benefit

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.

Groups

MBR/PHUS Variables
Widowed Spouse Benefit

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.

Groups

MBR/PHUS Variables
Widowed Spouse Benefit

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)

0	Never married
1	First marriage occurred
Sysmiss	

Value Ranges

Value Range

Range: [0 , 1]

Groups

Marital History Variables

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)

0	First Marriage did not end over course of survey
1	First marriage ended in widowhood
2	First marriage ended in divorce/separation
Sysmiss	

Value Ranges

Value Range

Range: [0 , 2]

Groups

Marital History Variables

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)

0	No second marriage
1	Second marriage occurred
Sysmiss	

Value Ranges

Value Range

Range: [0 , 1]

Groups

Marital History Variables

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)

0	Second marriage did not end over course of survey
1	Second marriage ended in widowhood
2	Second marriage ended in divorce/separation
Sysmiss	

Value Ranges

Value Range

Range: [0 , 2]

Groups

Marital History Variables

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)

0	No third marriage
1	Third marriage occurred
Sysmiss	

Value Ranges

Value Range

Range: [0 , 1]

Groups

Marital History Variables

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)

0	Third marriage did not end over course of survey
1	Third marriage ended in widowhood
2	Third marriage ended in divorce/separation
Sysmiss	

Value Ranges

Value Range

Range: [0 , 2]

Groups

Marital History Variables

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 marriges, 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)

0	No fourth marriage
1	Fourth marriage occurred
Sysmiss	

Value Ranges

Value Range

Range: [0 , 1]

Groups

Marital History Variables

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)

0	Fourth marriage did not end over course of survey
1	Fourth marriage ended in widowhood
2	Fourth marriage ended in divorce/separation
Sysmiss	

Value Ranges

Value Range

Range: [0 , 2]

Groups

Marital History Variables

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.

Groups

Marital History Variables

Variable Name

mh_date2

Label

Date of Marital History Event 2

Concept

Type

numeric

Files

F1 F2

Full Description

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

Groups

Marital History Variables

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.

Groups

Marital History Variables

Variable Name

mh_date4

Label

Date of Marital History Event 4

Concept

Type

numeric

Files

F1 F2

Full Description

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

Groups

Marital History Variables

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.

Groups

Marital History Variables

Variable Name

mh_date6

Label

Date of Marital History Event 6

Concept

Type

numeric

Files

F1 F2

Full Description

Date of ending of third marriage.

Groups

Marital History Variables

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.

Groups

Marital History Variables

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.

Groups

Marital History Variables

Variable Name

nonhouswealth

Label

Non-Housing Financial Wealth

Concept

Type

numeric

Files

F1 F2

Full Description

Non-housing wealth = total wealth minus home equity, collected in the following waves, by panel: Wave 4 in the 1984 (HHTWLTH, HHTHEQ), 1990-1993 (HH_TWLTH, HH_THEQ); Wave 3 in 1996, 2001, & 2004 (THHTWLTH, THHTHEQ) panels; Wave 4 in 2008 panel (THHTWLTH, THHTHEQ).

Groups

Economic Variables

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)

1	Linked marriage is first reported marriage described by mh1, mh2, mh_date1, and mh_date2
2	Linked marriage is first reported marriage described by mh3, mh4, mh_date3, and mh_date4
3	Linked marriage is first reported marriage described by mh5, mh6, mh_date5, and mh_date6
4	Linked marriage is first reported marriage described by mh7, mh8, mh_date7, and mh_date8
Sysmiss	When there is no linked marriage

Value Ranges

Value Range

Range: [1 , 4]

Groups

Marital History Variables

Variable Name

occ_cat_W

Label

Occupation Category for Primary Job in Wave W

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 (24 total)

1	Management
2	Business and financial operations
3	Business and financial operations
4	Architecture and engineering
5	Life, physical, and social science
6	Community and social service
7	Legal
8	Education, training, and libraries
9	Arts, design, entertainment, sports, and media
10	Healthcare practitioner and technical
11	Healthcare support

12	Protective service
13	Food prep and service
14	Building and grounds cleaning and maintenance
15	Personal care and service
16	Sales
17	Office and administrative support
18	Farming, fishing, and forestry
19	Construction and extraction
20	Installation, maintenance, and repairs
21	Production
22	Transportation
23	Material moving
24	Military

Value Ranges

Value Range

Range: [1 , 3]

Groups

Economic Variables

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 and 1990-1993 panels (TM8530, TM8608). In the 1996, 2001, and 2004 panels, 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)

0	Do not own a home
1	Own a home

Value Ranges

Value Range

Range: [0 , 1]

Groups

Economic Variables

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; tfrchl and tmomchl for 1996-2008 panels) and is in universe for men and women ages 15-64.

Value Ranges

Value Range

Range: [0 , 20]

Groups

Demographic Variables
Fertility Variables

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]

Groups

Identifiers

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)

0	Pension was not in scope
1	Pension was in scope

Value Ranges

Value Range

Range: [0 , 1]

Groups

Economic Variables

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 implicates or to the Gold Standard and Completed Data files.

Groups

Identifiers

Variable Name

phus_agedsp_benefit_stdate

Label

PHUS: startdate of aged spouse benefits

Concept

Type

numeric

Files

F1 F2

Full Description

This start date 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 start date can differ from the MBR start date 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 2014.

Groups

MBR/PHUS Variables
Aged Spouse Benefit

Variable Name

phus_agedsp_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 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.

Groups

MBR/PHUS Variables
Aged Spouse Benefit

Variable Name

phus_retire_benefit_stdate

Label

PHUS: startdate of retirement benefits

Concept

Type

numeric

Files

F1 F2

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 eligiblity 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 2014.

Groups

MBR/PHUS Variables

Variable Name

phus_retire_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 eligiblity 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.

Groups

MBR/PHUS Variables

Variable Name

phus_ssdi_benefit_stdate_k

Label

PHUS: SSDI benefit start date 1

Concept

Type

numeric

Files

F1 F2

Full Description

This start date variable tells when the first SSDI benefit payment was recorded in the Payment History Update System, the administrative database maintained by SSA to track actual payments made to beneficiaries. This start date can differ from the MBR start date 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 2014.

Groups

MBR/PHUS Variables
Disability Benefit

Variable Name

phus_ssdi_benefit_totamt_k

Label

PHUS: SSDI total monthly benefit k (where k=1 to 4)

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 month k of own SSDI benefits receipt.

Groups

MBR/PHUS Variables
Disability Benefit

Variable Name

phus_widowsp_benefit_stdate

Label

PHUS: startdate of widowed spouse benefits

Concept

Type

numeric

Files

F1 F2

Full Description

This start date 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 start date can differ from the MBR start date 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 2014.

Groups

MBR/PHUS Variables

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 eligiblity 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.

Groups

MBR/PHUS Variables

Variable Name

pos_phus_agedsp_benefit_totamt

Label

PHUS: indicator of positive benefit amount

Concept

Type

numeric

Files

F1 F2

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.

Values (2 total)

0	Does not have a positive aged spouse benefit amount
1	Has a positive aged spouse benefit amount

Groups

MBR/PHUS Variables
Aged Spouse Benefit

Variable Name

pos_phus_retire_benefit_totamt

Label

PHUS: indicator of positive benefit amount

Concept

Type

numeric

Files

F1 F2

Full Description

Indicator that the PHUS recorded a positive retirement benefit amount (and consequently had a non-missing PHUS start date) at some point after the MBR eligibility start date.

Values (2 total)

0	Did not receive a positive retirement benefit amount at some point after MBR eligibility date
1	Received a positive retirement benefit amount at some point after MBR eligibility date

Groups

MBR/PHUS Variables

Variable Name

pos_phus_ssdi_benefit_totamt_k

Label

PHUS: indicator of positive SSDI benefit amount for application k

Concept

Type

numeric

Files

F1 F2

Full Description

Indicator that the PHUS recorded a positive SSDI benefit amount (and consequently had a non-missing PHUS start date) at some point after the MBR eligibility start date. Individuals in the SSB may have information on up to 4 SSDI applications (k = 1 to 4).

Values (2 total)

0	Did not receive positive SSDI amount
1	Received positive SSDI amount

Groups

MBR/PHUS Variables
Disability Benefit

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.

Values (2 total)

0	Did not receive positive widowed spouse benefit
1	Received positive widowed spouse benefit

Groups

MBR/PHUS Variables

Variable Name

race

Label

Race

Concept

Type

numeric

Files

F1 F2

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)

1	White
2	Black
3	Other

Value Ranges

Value Range

Range: [1 , 3]

Groups

Demographic Variables

Variable Name

rot

Label

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 researhers 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]

Groups

Identifiers

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.

Groups

Identifiers

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.

Groups

Identifiers

Variable Name

spouse_personid

Label

Personid of spouse

Concept

Type

numeric

Files

F1 F2

Full Description

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 recorded 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. In the SSB, the marital link between SIPP respondents denoted by personid and spouse_personid has been synthesized so that family characteristics cannot be used as a blocking variable in a re-identification attack.

Groups

Identifiers

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.

Groups

Supplemental Security Record Variables

Variable Name

ssr_ssi_applied

Label

SSR: Applied for SSI benefits

Concept

Type

numeric

Files

F1 F2

Full Description

An indicator for whether the individual has a recorded application for Supplemental Security Income (SSI).

Values (2 total)

0	Does not have a recorded SSI application
1	Has a recorded SSI application

Groups

Supplemental Security Record Variables

Variable Name

ssr_ssi_benefit

Label

SSR: Received SSI benefits

Concept

Type

numeric

Files

F1 F2

Full Description

An indicator for whether the applicant ever received SSI benefits.

Values (2 total)

0	Has never received SSI benefits
1	Has received SSI benefits

Groups

Supplemental Security Record Variables

Variable Name

ssr_ssi_benefit_fed_totamt

Label

SSR: Total federal benefit amount

Concept

Type

numeric

Files

F1 F2

Full Description

The total SSI benefit amount from federal sources.

Groups

Supplemental Security Record Variables

Variable Name

ssr_ssi_ceased

Label

SSR: Benefits ceased

Concept

Type

numeric

Files

F1 F2

Full Description

SSI benefits recorded as ceased on the SSR; see SSR: last payment date.

Values (2 total)

0	SSI benefit has not ceased
1	SSI benefit has ceased

Groups

Supplemental Security Record Variables

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.

Values (16 total)

1	Diseases of the blood
2	Circulatory system
3	Congenital anomalies
4	Digestive system
5	Endocrine, nutritional, and metabolic diseases
6	Genitourinary system
7	Infectious and parasitic diseases
8	Injuries
9	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

Groups

Supplemental Security Record Variables

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.

Groups

Supplemental Security Record Variables

Variable Name

ssr_ssi_last_pmt_dt

Label

SSR: Last payment date

Concept

Type

numeric

Files

F1 F2

Full Description

The date of last recorded payment of SSI benefits. This variable is saved as a SAS date.

Groups

Supplemental Security Record Variables

Variable Name

ssr_ssi_type

Label

SSR: Type of benefit

Concept

Type

numeric

Files

F1 F2

Full Description

Type of SSI benefit applied for or received by the individual.

Values (3 total)

1	Aged individual
2	Disabled or blind individual
3	Disabled or blind child

Groups

Supplemental Security Record Variables

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
4	Arizona
5	Arkansas
6	California
8	Colorado
9	Connecticut
10	Delaware
11	DC
12	Florida
13	Georgia
15	Hawaii

16 Idaho *see description
17 Illinois
18 Indiana
19 Iowa *see description
20 Kansas
21 Kentucky
22 Louisiana
23 Maine *see description
24 Maryland
25 Massachusetts
26 Michigan
27 Minnesota
28 Mississippi *see description
29 Missouri
30 Montana *see description
31 Nebraska
32 Nevada
33 New Hampshire
34 New Jersey
35 New Mexico *see description
36 New York
37 North Carolina
38 North Dakota *see description
39 Ohio
40 Oklahoma
41 Oregon
42 Pennsylvania

44	Rhode Island
45	South Carolina
46	South Dakota *see description
47	Tennessee
48	Texas
49	Utah
50	Vermont *see description
51	Virginia
53	Washington
54	West Virginia *see description
55	Wisconsin
56	Wyoming *see description
61	(1990, 1991, 1992, 1993, 1996, 2001 panels) Maine, Vermont *see description
62	(1990, 1991, 1992, 1993 panels) Iowa, North Dakota, South Dakota (1996, 2001 panels) North Dakota, South Dakota, Wyoming *see description
63	(1990, 1991, 1992, 1993 panels) Alaska, Idaho, Montana, Wyoming *see description
90	(1984 panel only) Idaho, New Mexico, South Dakota, Wyoming *see description
91	(1984 panel only) Mississippi, West Virginia *see description

Groups

Geographic Variables

Variable Name	time_arrive_usa
Label	Time Period of Arrival to US (Foreign Born)
Concept	
Type	numeric
Files	F1 F2

Full Description

This variable gives the year block when a respondent immigrated in the United States and is in-scope when foreign_born=1. The year of arrival to the U.S. is from the migration history wave 2 topical module (TM8736 for 1984, 1990-1993; emoveus for 1996, 2001, 2004, and 2008). Year blocks are generally five years but are as wide as 9 years and as short as 2 years. Categories are: 1=1959 or earlier; 2=1960-1964; 3=1965-1969; 4=1970-1974; 5=1974-1979; 6=1980-1981; 7=1982-1984; 8=1985-1993; 9=1994-1999; 10=2000-2004; 11=2005-2009; 12=2010-2014.

Values (13 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
11	2005-2009
12	2010-2014

sysmiss	Structurally missing, out of scope for question (foreign_born=0)

Value Ranges

Value Range

Range: [1 , 10]

Groups

Demographic Variables
Geographic Variables

Variable Name

total_der_fica_YYYY

Label

DER: Total earnings from FICA-covered jobs

Concept

Type

numeric

Files

F1 F2

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 1978-2014.

Groups

Detailed Earnings Record Variables

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 1978-2014.

Groups

Detailed Earnings Record Variables

Variable Name

totearn_Y_M

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 month M of year Y. M goes from 1 to 12 and represents calendar month. Y goes from 1 to 6 and represents the number calendar year for the given SIPP panel. Months that are outside the time frame covered by an individual's SIPP panel will always be missing and out of universe. The corresponding SIPP variables are PP-EARN1, PP-EARN2, PP-EARN3, PP-EARN4 in the 1984 panel; EARN for the 1990, 1991, 1992, and 1993 panels; and TPEARN for the 1996, 2001, 2004, and 2008 panels. This variable includes both jobs and businesses (self-employment) earnings and is top-coded.

Groups

Income Variables

Variable Name

totearn_ser_YYYY

Label

SER: Capped Earnings from all FICA-covered jobs

Concept

Type

numeric

Files

F1 F2

Full Description

Person-level annual earnings that were taxed by FICA; these variables include earnings only up to the FICA taxable maximum and cover the years 1951-2014. These earnings are the inputs for calculating the OASDI benefit a person and his or her spouse will receive upon retirement or disability.

Groups

Summary Earnings Record Variables

Variable Name

totalhrs_Y_M

Label

Total Usual Weekly Hours Worked at All Jobs

Concept

Type

numeric

Files

F1 F2

Full Description

This variable is calculated based on usual hours worked per week and weeks worked with pay in month M of year Y. M goes from 1 to 12 and represents the calendar month. Y goes from 1 to 6 and represents the number calendar year for the given SIPP panel. Months that are outside the time frame covered by an individual's SIPP panel will always be missing and out of universe. The SIPP asks for usual hours worked per week for up to two jobs. The usual hours for job 1 and job 2 are combined and then multiplied by the weeks worked with pay that month (see wkswp_Y_M). The usual hours worked SIPP variables are WS12024 and WS22024 in the 1984 panel; WS12025 and WS22125 in the 1990, 1991, 1992, and 1993 panels; and EJBHRS1 and EJBHRS2 in the 1996, 2001, 2004, and 2008 panels. Total usual hours worked per week is capped at 99 in the GSF and SSB; therefore, the cap for totalhrs_Y_M is 99 multiplied by wkswp_Y_M.

Groups

Labor Force Variables

Variable Name

totinc_Y_M

Label

Total Personal Income

Concept

Type

numeric

Files

F1 F2

Full Description

Total personal income summed from all sources in month M of year Y. M goes from 1 to 12 and represents the calendar month. Y goes from 1 to 6 and represents the number calendar year for a given SIPP panel. Months that are outside the time frame covered by an individual's SIPP panel will always be missing and out of universe. The relevant SIPP variable is PPTOTIN1, PPTOTIN2, PPTOTIN3, and PPTOTIN4 in the 1984 panel; TOTINC in the 1990, 1991, 1992, and 1993 panels; and TPTOTINC in the 1996, 2001, 2004, and 2008 panels.

Groups

Income Variables

Variable Name

totnetworth

Label

Total Net Worth

Concept

Type

numeric

Files

F1 F2

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).

Groups

Economic Variables

Variable Name

vcamt_Y_M

Label

Amount of veterans compensation or pension benefit

Concept

Type

numeric

Files

Full Description

This variable is the amount of veterans compensation or pension benefits that a respondent received in month M of year Y. M goes from 1 to 12 and represents the calendar month. Y goes from 1 to 6 and represents the number calendar year for a given SIPP panel. The variable comes from the core file. The relevant SIPP variables are I08AMT1-I08AMT4 in 1984; S08AMT in the 1990, 1991, 1992, and 1993 panels; T08AMT in the 1996, 2001, 2004, and 2008 panels.

Groups

Variable Name

vetrecip_Y_M

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 month M of year Y. M goes from 1 to 12 and represents the calendar month. Y goes from 1 to 6 and represents the number calendar year for a given SIPP panel. The variable comes from the core file. The relevant SIPP variables are 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)

0	Did not receive veterans compensation or veterans benefits
1	Received veterans compensation or veterans benefits

Groups

Variable Name

wcamt_Y_M

Label

Amount of Workers Compensation Received

Concept

Type

numeric

Files

Full Description

This variable is the amount of workers compensation benefits that a respondent received in month M of year Y. M goes from 1 to 12 and represents the calendar month. Y goes from 1 to 6 and represents the number calendar year for a given SIPP panel. The variable comes from the core file. The relevant SIPP variables are I10AMT1-I10AMT4 in 1984; S10AMT in the 1990, 1991, 1992, and 1993 panels; T10AMT in the 1996, 2001, 2004, and 2008 panels.

Groups

Variable Name

wkcomp_Y_M

Label

Indicator for receipt of workers compensation

Concept

Type

Files

Full Description

This variable indicates that a respondent received workers compensation benefits in month M of year Y. M goes from 1 to 12 and represents the calendar month. Y goes from 1 to 6 and represents the number calendar year for a given SIPP panel. The variable comes from the core file. The relevant SIPP variables are I10REC1-I10REC4 in 1984; R10 in the 1990, 1991, 1992, and 1993 panels; ER10 in the 1996, 2001, 2004, and 2008 panels.

Values (2 total)

0	Did not receive workers compensation benefits
1	Received workers compensation benefits

Groups

Variable Name

wksjob_Y_M

Label

Weeks at a Job

Concept

Type

numeric

Files

F1 F2

Full Description

Total number of weeks that the respondent held a job in month M of year Y. M goes from 1 to 12 and represents the calendar month. Y goes from 1 to 6 and represents the number calendar year for a given SIPP panel. Months that are outside the time frame covered by an individual's SIPP panel will always be missing and out of universe. The relevant SIPP variable is WKSJB (1984 panel), WKSJOB (1990, 1991, 1992, 1993), or RMWKWJB (1996, 2001, 2004, 2008). Values range from 0 to 5.

Groups

Labor Force Variables

Variable Name

wkswp_Y_M

Label

Weeks With Pay

Concept

Type

numeric

Files

F1 F2

Full Description

Total number of weeks worked with pay in Month M of year Y. M goes from 1 to 12 and represents the calendar month. Y goes from 1 to 6 and represents the number calendar year for a given SIPP panel. Months that are outside the time frame covered by an individual's SIPP panel will always be missing and out of universe. The wkswp_Y_M variable is calculated by subtracting the weeks without pay in month M of year Y from wksjob_Y_M. The relevant SIPP variable for weeks without pay is WKSWOP (1984, 1990, 1991, 1992, and 1993 panels), or RMWKSAB (1996, 2001, 2004, 2008). Values range from 0 to 5.

Groups

Labor Force Variables

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)

0
1
2
3
4

Value Ranges

Value Range

Range: [0 , 4]

Groups

Summary Earnings Record Variables

Variable Name

year_bach

Label

Year of Bachelors Degree

Concept

Type

numeric

Files

F1 F2

Full Description

This variable gives the year that post-high school education ended and comes from the Education and Training History topical module in wave 3 of the 1984 panel and wave 2 of the other SIPP panels. Individuals must have educ_5cat=4 or higher to be in-scope for this variable. The relevant variable is TM8036 for the 1984 panel; TM8434 for 1990, 1991, 1992, 1993; and TBACHYR for 1996, 2001, 2004, and 2008.

Groups

Education Variables

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 ended and comes from the Education and Training History topical module in wave 3 of the 1984 panel and wave 2 of the 1990, 1991, 1992, 1993, 1996, 2001, 2004, and 2008 panels. The relevant SIPP variables are TM8420 for the 1984 and 1990-1993 panels and TCOLLSTR for the 1996, 2001, 2004, and 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)

Groups

Education Variables

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 post-high school education ended and comes from the Education and Training History topical module in wave 3 of the 1984 panel and wave 2 of the other SIPP panels. The relevant SIPP variables are TM8404 and TM8412 for 1984, 1990-1993 panels; variable TLSTSCHL and THSYR for 1996-2008 panels. This variable is in scope for all levels of educ_5cat.

Groups

Education Variables

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 Education and Training History topical module in wave 3 of the 1984 panel and wave 2 of the other SIPP panels. The relevant SIPP variables are TM8030 for 1984; TM8426 and TM8440 for the 1990-1993 panels; variables TLASTCOL, TVOCYR, TASSOCYR, TBACHYR, TADVNCYR for the 1996, 2001, 2004, and 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)

Groups

Education Variables

Variable Name

date_enter_sipp

Label

Wave Entered SIPP

Concept

numeric

Type

Files

Full Description

Date of the wave when the respondent entered the SIPP.

Groups

Variable Name

mdy_lft_sipp

Label

Wave Left SIPP

Concept

numeric

Type

Files

Full Description

Month, day, and year that an individual left their SIPP household. Based on the SIPP variables U-DAYLFT and U-MONLFT (1984 panel); DAYLFT, MONLFT, YRLFT (1990, 1991, 1992, 1993); ULFTDAY and ULFTMON (1996). Exit date information not available for 2001, 2004, or 2008 panels.

Groups

Variable Name

panel_1styear

Label

First calendar year of SIPP panel

Concept

numeric

Type

Files

Full Description

The first calendar year of the panel. Note that several panels begin before their panel value (e.g. the 1996 panel starts at the end of 1995).

Groups

Variable Name

flag_in_tmhist_ms

Label

Flag: In Marital History TM

Concept

numeric

Type

Files

Full Description

Indicates whether or not an individual participated in the marital history topical module. Based on having a non-missing value for the SIPP variable MS_1, MS_2, MS_3, MS_4 (1984 Panel); MS (1990,1991,1992,1993); or EMS (1996,2001,2004,2008).

Values (2 total)

0	Not in the marital history topical module
1	Present in the marital history topical module

Groups

Variable Name

first_birth_year

Label

Year of Birth of First Child

Concept

numeric

Type

Files

Full Description

Birth year of first biological child. Derived from first_admin_birthdate and/or the birth year of first biological child indicated in the SIPP.

Groups

Variable Name

last_birth_year

Label

Year of Birth of Last Child

Concept

numeric

Type

Files

Full Description

Birth year of last-born biological child. Derived from the administrative birthdate of the last biological child and/or SIPP-reported birth year of the last-born biological child.

Groups

Variable Name

halfsamp

Label

Half Sample Code

Concept

numeric

Type

Files

Full Description

Half sample code for variance estimation. In the SIPP, the relevant variables is H1-HSC in the 1984 panel; HHSC in the 1990, 1991, 1992, and 1993 panels; and GHIFSAM for the 1996, 2001, 2004, and 2008 panels. The values are 1 or 2.

Groups

Variable Name

varstrat

Label

Variance Stratum Code

Concept

Type

Files

Full Description

Stratum code for variance estimation. In the SIPP, the relevant variable is H1-HSC for the 1984 panel; HHSC for the 1990, 1991, 1992, and 1993 panels; and GHLFSAM for the 1996, 2001, 2004, and 2008 panels. The values range from 0 to 99.

Groups

Variable Name

mom_personid

Label

SIPP Gold Standard Person ID of Mom

Concept

numeric

Type

Files

Full Description

This variable identifies the personid of an individual's mother within each SSB implicate. In the SSB, the parent-child link between SIPP respondents denoted by personid and mom_personid has been synthesized so that family characteristics cannot be used as a blocking variable in a re-identification attack.

Groups

Variable Name

ser_posearn_YYYY

Label

SER: Positive Earnings

Concept

numeric

Type

Files

Full Description

A binary indicator for having positive earnings (based on the SER) in year YYYY.

Values (2 total)

0	Did not have positive earnings in year YYYY
1	Positive earnings in year YYYY

Groups

Variable Name

mbr_ssdi_ben_dof

Label

MBR: date of filing for SSDI (benefit eventually received)

Concept

numeric

Type

Files

Full Description

Application date of most recent accepted OASDI application. The MBR has information on the OASDI application date of filing for the most recent accepted application and the most recent rejected application. To match with the SSDI application numbers (k = 1 to 4) in the SSB, the date of filing is compared to an individual's SSDI history and other dates associated with an application to assess if it matches one of these four applications.

Groups

Variable Name

mbr_ssdi_deny_dof

Label

MBR: date of filing for SSDI (claim denied)

Concept

numeric

Type

Files

Full Description

Application date of most recent rejected OASDI application. This variable pertains to the most recent rejected OASDI application. The MBR has information on the application date of filing for the most recent accepted application and most recent rejected application. To match with the SSDI application numbers (k = 1 to 4) in the SSB, the date of filing is compared to other dates associated with an application to assess if it matches one of these four applications.

Groups

Variable Name

mbr_ssdi_bdof_app

Label

Application associated with date of filing

Concept

numeric

Type

Files

Full Description

This variable provides the application number (k = 1 to 4) that matches with the date of filing for benefits. This variable pertains to mbr_ssdi_ben_dof. The MBR has information, when applicable, on the date of filing of the most recent accepted OASDI application and the date of filing for most recent rejected OASDI application. If an individual's SSDI history and one of their SSDI applications in the SSB aligns with the timing of the date of filing for the most recent accepted application, mbr_ssdi_bdof_app will indicate which application number k.

Values (4 total)

1	Date of filing fits with SSDI application 1
2	Date of filing fits with SSDI application 2
3	Date of filing fits with SSDI application 3
4	Date of filing fits with SSDI application 4

Groups

Variable Name

mbr_ssdi_ddof_app

Label

Application associated with date of filing (denied app)

Concept

numeric

Type

Files

Full Description

This variable provides the application number (k = 1 to 4) that matches with the date of filing for benefits. This variable pertains to mbr_ssdi_deny_dof. The MBR has information on dates of filing for most recent accepted OASDI application and most recent rejected OASDI application. If one of the four SSDI applications included in the SSB aligns with the individual's SSDI history and the timing of an SSDI application, mbr_ssdi_ddof_app will indicate which application.

Values (4 total)

1	Date of filing fits with application 1
2	Date of filing fits with application 2
3	Date of filing fits with application 3
4	Date of filing fits with application 4

Groups

Variable Name

flag_enter_late

Label

Flag: Entered SIPP Late

Concept

numeric

Type

Files

Full Description

A flag indicating if the respondent entered the SIPP after wave 1.

Values (2 total)

0	Did not enter SIPP late
1	Entered SIPP late

Groups

Variable Name

flag_lft_sipp

Label

Flag: Left Before End of SIPP

Concept

numeric

Type

Files

Full Description

Indicator for individual in the SIPP who leave the survey. Variable is based on a positive value for U-REALFT (1984 Panel); REALFT (1990, 1991, 1992, 1993); ULFTMAIN (1996, 2001, 2004, 2008).

Values (2 total)

0	Did not leave SIPP early
1	Left SIPP early

Groups

Variable Name

life_ins_in_scope

Label

In-Scope for Life Insurance Topical Module

Concept

numeric

Type

Files

Full Description

Indicator for being in-scope for the main life insurance topical module. This occurs in Wave 3 for the 1996, 2001, and 2004 SIPP panels and Wave 4 for the 1984, 1990, 1991, 1992, 1993, and 2008 panels.

Values (2 total)

0	Out of scope
1	In-Scope

Groups

Variable Name

pension_in_scope_age

Label

In-Scope for Pension (Level I)

Concept

numeric

Type

Files

Full Description

This variable indicates that an individual was in scope for the pension questions because he or she at least 15 years old 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 and 1992 panels (TM8324); wave 7 in the 1991 panel (TM8324); wave 9 in the 1993 panel (TM8324); wave 7 in the 1996, 2001, and 2004 panels (EARPUNV, RMNJBBS); wave 3 in the 2008 panel (EARPUNV, RMNJBBS).

Values (2 total)

0	Out of scope
1	In-Scope

Groups

Variable Name

fert_in_scope

Label

In-Scope for Fertility Topical Module

Concept

numeric

Type

Files

Full Description

Indicator for being in-scope for the fertility topical module of the SIPP. The fertility topical module occurs in wave 8 of the 1984 panel; wave 2 of the 1990, 1991, 1992, 1993, 1996, 2001, 2004, and 2008 panels.

Values (2 total)

0	Not In-Scope
1	In-Scope

Groups

Variable Name

brthmn_sipp

Label

Month of Birth - SIPP

Concept

numeric

Type

Files

Full Description

Respondent's birth month, as given in SIPP. This date is used for universe determination for SIPP variables.

Groups

Variable Name

brthyr_sipp

Label

Year of Birth - SIPP

Concept

numeric

Type

Files

Full Description

Respondent's birth year, as given in SIPP. This date is used for universe determination for SIPP variables.

Groups