

January 2016

**PANEL STUDY OF INCOME DYNAMICS
2009 PUBLIC RELEASE FAMILY FILE DOCUMENTATION**

TABLE OF CONTENTS

I. INTRODUCTION -- WHAT'S NEW FOR 2009

A. QUESTIONNAIRE CHANGES

B. NEW GENERATED VARIABLES

II. DATA CHARACTERISTICS

A. FILES AND FORMAT

B. VARIABLE NAMES, POSITIONS AND GENERATED VARIABLES

C. HAND CODED AND GENERATED VARIABLES

D. DOCUMENTATION AND CODES

1. 2009 Questionnaire

2. Codebook and Code Categories

RELEASE 2 NOTES

November 2013: An error was corrected for how Completed Education for Head (ER46981) and for Wife/"Wife" (ER46982) were calculated for old Heads and Wives/"Wive's". Since all education questions were re-asked in 2009, completed education summary variables should have been re-calculated based on this new information for old Heads and Wives/"Wives" collected in 2009. The correction included bringing the corrected 2009 data forward to 2011.

We are also correcting the summary variable Births to OFUMS only - 2007 (ER47011).

In addition, in Release 2 we are suppressing two previously publicly released variables Beale Rural-Urban Code (ER46975) and Size of Largest City in County (ER46976).

RELEASE 3 NOTES

February 2014: In Release 3 we have added one new coded variable in the philanthropy section M12code (ER46659A) and Collapsed Beale (ER46975A).

RELEASE 4 NOTES

January 2016: For Release 4 we have added 74 new variables in the T-2 (2007) Income section: Number of Jobs for Head (ER46670A) and Wife/"Wife" (ER46681A), months employed for Head (ER46670B-ER46670M) and Wife/"Wife" (ER46681B-ER46681M), months unemployed for Head (ER46666A-ER46666L) and Wife/"Wife" (ER46677A-ER46677L), and months out of the labor force for Head (ER46668A-ER46668L) and Wife/"Wife" (ER46679A-ER46679L).

I. INTRODUCTION – WHAT'S NEW FOR 2009

A. Questionnaire Changes

Most sets of questions from the 2007 wave continue to be part of the 2009 instrument, but a few sequences were dropped, some were added and others revised.

In Section A, a life satisfaction question was added (A3). Also added to the housing section is a series of questions asking homeowners whether they are behind on their mortgage payments and whether they have been faced with foreclosure (A27FOR1-6). Similar questions about asking renters about previous foreclosures are found at (A37FOR1-7).

In the employment sections, BC for Head and DE for Wife/"Wife", summary confirmations about weeks and hours worked across all jobs were added at BC14a-f for Head and DE14a-f for Wife/"Wife".

The flow off the questionnaire was altered so that only a fraction of the Heads and Wives/"Wives" were asked the job level questions BC/DE42. Only those cases where we could not determine the weeks from the start and stop dates were asked these helper questions that are not being released. They are being used during editing to determine the value of weeks across all jobs at the BC/DE14 series. Deleted questions in BC and DE are BC/DE49 and BC/DE57-60.

In Section F, the frame of reference changed for F14 and F16. In 2007, these food stamp questions asked about 'last year', in 2009 they were revised to say 'last month.' Some questions about car ownership were dropped. Deleted questions include those asking about who usually drives the car as well as some of the purchasing/leasing questions (F51, F52, F56 and F62-63).

Questions in Section W, (W129-W144) about intergenerational help have been dropped.

The health section, Section H, we continue to ask the childhood health calendar (H32-H45). For each condition and type of report about Head or Wife/"Wife" we have yes/no indicators for each condition at each age. There are a total of six sets of Childhood Health Calendar reports. The series was asked about Head and Wife/"Wife" and depending on who the respondent was, the information is stored in the appropriate set.

These reports were carried forward from the 2007 wave if Head or Wife/"Wife" remained the same. In 2009, we may have gathered new types of reports (self report and spousal report, for example) about the same individual if the respondent changes. To help understand which reports are from which wave, we added variables at the start of each report to indicate which year this report originated. For set one for example, (1) Self Report about Head the year of report is stored at (ER44317).

- (1) Self Report about Head (ER44317-ER44634)
- (2) Spousal Report about Head (ER44635-ER44952)
- (3) Other Report about Head (ER44953-ER45270)
- (4) Self Report about Wife/"Wife" (ER45414-ER45731)
- (5) Spousal Report about Wife/"Wife" (ER45732-ER46049)
- (6) Other Report about Wife/"Wife" (ER46050-ER46367)

All Heads and Wives/"Wives" were asked the education series KL43-67 in 2009.

Section R, which asks about income for the "off" year, or t-2, was scaled back significantly. Remaining questions ask whether the family has received certain types of income but the amounts and months collection questions were dropped. Please see the 2009 Box and Arrow documentation at <ftp://ftp.isr.umich.edu/pub/src/psid/questionnaires/q2009.pdf> for details.

Section S, which asked about the effects of Hurricane Katrina is gone.

Some of the questions from the interviewer observation section are included on the public release file. These questions are asked of the interviewer after the interview is completed. Observation questions included on the family file include those asked about the respondents' reaction to requesting their Medicare number (ER46700-ER46705). Also included in this series are the sections of the interview that were completed with a helper (ER46706-ER46759).

B. Changes to Generated Variables

Background information (Section K for Wives/"Wives" and Section L for Heads) is only asked when a person newly acquires a relationship of Head or Wife/"Wife", values are copied forward from prior data collections for those Heads and Wives/"Wives" who maintain their relationship classification from the preceding wave.

A new variable to indicate whether this family has a record in the DUST supplement was added (ER46696).

Now that all wives have been asked questions about where they grew up in the background section, we have now generated Region where Wife Grew Up (ER46979) and Wife's Geographic Mobility (ER46980).

Also new for 2009, is a set of variables generated from the Employment EHC start and stop dates for the "off" year, 2007. These variables include whether Head or Wife were Employed, Unemployed, or Out of the Labor Force, and the number of weeks (ER4665-ER46682).

The USDA needs standards variables have been dropped.

The wealth file contains imputed versions of items from Section W, which has been released in prior waves as a supplement is now part of the 2009 Family File (ER46936- ER46971).

II. DATA CHARACTERISTICS

A. Files and Format

The 2009 family data file consists of one data file with 8,690 records and 5,088 variables. The family file contains one record for each family interviewed in 2009 and includes all family-level variables collected in 2009, including extensive information about Head and Wife/"Wife".

Using the web-based Data Center is the most efficient way to obtain the data. At this time, we also provide the data in a .zip package. If you download the data in ASCII form from the PSID website, you will receive SAS, SPSS or STATA statements, as you request. In the .zip package, we include the entire data file in ASCII format and SAS, SPSS and STATA data definition statements. These statements provide variable names, locations, and variable labels.

The definition statement files are not intended to represent complete and full programs for the respective statistical program packages to run extracts, analysis, etc. You must provide all other statements needed to complete a program. Missing data statements are not provided as part of the files.

B. Variable Names and Positions

The 2009 family variable names are in the range ER42001- ER47012.

The 2009 data file provides many component variables that analysts may use to construct summary variables. We leave it to the individual analyst to make decisions about imputation methods to use, if any, and how to treat missing data. Analysts who wish to construct summary variables that are parallel to those provided in historical waves of the PSID may consult codebooks for those years to obtain information about rules for their creation. Some examples of variables that may be constructed using component variables provided in this data file include annualized mortgage and rent payments and food costs.

C. Hand Coded and Generated Variables

The PSID hand codes several data items for the family file. Head's and Wife's/"Wife's" occupations and industries, to a maximum of four jobs apiece, are coded for 2009 using the three-digit codes from the 2000 Census Occupation and Industry Code. The information for each job also includes a one-digit reason for job termination where applicable. Family-owned businesses are coded using a two-digit industry code, and the field of endeavor for Head's and Wife's/"Wife's" non-academic degrees and certificates is coded for as many as three each. Any education received outside the U.S. is coded for Head, Wife/"Wife" and their parents.

Background items such as education are collected only for new Heads and new Wives/"Wives" in a given wave. During processing, we have traditionally "brought forward" background information from previous waves for Heads or Wives/"Wives" who are the same persons as in the prior year. In every wave, each set of background variables is preceded by a variable indicating whether data needed to be brought forward. The wave in which the section was most recently asked is indicated by ER46986 for Heads and ER46987 for Wives/"Wives".

Completed education of Head and Wife/"Wife" (ER46981 and ER46982) are generated from the background information.

Family composition and change variables include Family Composition Change (ER42007), Splitoff Indicator (ER42005), Head-Spouse Sample Status (ER46988), and variables about births to Head, spouse and other family members during the prior calendar year, 2008 (ER47004- ER47007), and during the "off" year, 2007 (ER47008- ER47010). Note that the Splitoff Indicator is only assigned to a family in the year they first move out of the main family and form their own separate household; after that one wave of being a splitoff, these families receive code values that designate reinterviews.

Two other variables concerning splitoffs are the number of splitoffs arising from a main family (ER46989) and the family interview number of the main family from which a splitoff family originated (ER46990).

The PSID produces sets of variables about families sharing the same household: family ID numbers, relationships, and sizes of up to four other PSID families sharing the HU (ER46991- ER47002), the household ID number (ER47003), and the number of persons not included in any PSID family who are sharing the household unit (ER42020). The PSID documentation for 1993 and earlier waves has additional information about multiple PSID families sharing the same household (see "Linking Data" in Section I, Part 5 of the 1993 Guide).

The marital status variables group consists of Head's current marital status (ER42023), the generated form of marital status comparable with years prior to 1977 (ER46983), change in marital status of Head between waves (ER46984), and couple status of Head (ER46985).

Location variables include PSID/GSA and FIPS state codes (ER42003 and ER42004); Current Region, Beale Rural-Urban code, and Size of the Largest City in the County (ER46974- ER46976); and two derived from background information, Region where Head (ER46977) or Wife Grew Up (ER46979) and Head's (ER46978) and Wife's (ER46980) Geographic Mobility. The codes for the FIPS and PSID/GSA codes are included in Appendix 1 of the 1985 documentation, available in .pdf format on our website.

The Census needs standard was generated for the prior calendar year, 2008 (ER46972). Additionally, since the PSID has switched to biennial interviewing, comparable needs standard was also been generated for the "off" year, 2007 (ER46973).

The variable indicating whether a PSID family lives in institutional housing is present (ER42008), as is the variable indicating the total number of data records from the cross-year individual file that are associated with a panel family (ER46760).

Weights are available as part of the 2009 Family File in ER47012.

The imputed work hours and income variables are found at (ER46761- ER46935).

D. Documentation and Codes

1. 2009 Questionnaire

Our website includes an HTML depiction of the computer-assisted interviewing (CAI) applications for 1993-2001. Box-and-arrow questionnaires are available for all waves in PDF format. Use the SAS, SPSS and STATA data definition statements to match variables with questions in the CAI application.

2. Codebook and Code Categories

The Data Center allows users to create and download customized codebooks that provide variable-specific documentation; variable descriptions are complete for all of the 2009 family data items. The downloadable .zip package of the family data on our website contains a .pdf version of the entire family codebook for 2009.

In general, codes follow our traditional schema, although "don't know" responses are now largely distinguished from other missing data responses. Generally, code 8 (or 98 or 998, etc.) represents "don't know" and code 9 (or 99 or 999, etc.) represents other missing data or a refusal. Inappropriate questions are padded with zeroes.

If a variable contains a code value that is neither included in the codebook nor one of the "zero", "eight" or "nine" codes just mentioned, assume missing data for that value.