August 2019

PANEL STUDY OF INCOME DYNAMICS
2017 PUBLIC RELEASE FAMILY FILE DOCUMENTATION

_____

TABLE OF CONTENTS

_____

Release 2 Note: Users should be aware that the 2017 CORE/IMMIGRANT FAM WEIGHT NUMBER 1 (ER71570) has zero values for the new Immigrant 2017 sample (ER66009=4001-4462). For Release 2 we have added a new variable 2017 CROSS-SECTIONAL FAMILY WEIGHT (ER71571) which encompasses the Core 68/Immigrant 97/99 and Immigrant 17 Samples.

I. INTRODUCTION – WHAT'S NEW FOR 2017

A. Questionnaire Changes

The most significant change for the 2017 Family File is the addition of a new Immigrant cohort for 2017. These individuals were screened in a separate data collection effort in 2016 and those who met the screening criteria were included in the 2017 main data collection effort. In summary, at the time of the 2016 screening, individuals were selected if the Reference Person and/or the Spouse/Partner were born prior to 1997 (the year of our last cohort addition) and were not living in the United States at that time or were born after 1997 to parents who were not living in the United States in 1997. The purpose of the new immigrant refresher sample is to ensure that the PSID remains representative of the current U.S. population. This new 2017 Immigrant cohort has been assigned 1968 Interview Numbers (ER66009) in the range of 4001-4462. We have added 452 new 2017 Immigrant families whose first wave's interview is 2017.

Many sets of questions from the 2015 Wave continue to be part of the 2017 instrument. At the same time, there are notable additions, deletions, and revisions which are detailed below, by section.

Section A: Housing, Utilities, and Computer Use

A new question was added asking how much rental value might be if home were rented for those who are current homeowners (A20A1/A20A1PER). In addition, given low reporting, in the past several waves, for multiple foreclosures (A37A-A37G), the foreclosure series was shortened, from collecting up to two reports, to collecting a single report. Lastly, in the A20 series about home ownership, the unfolding brackets, for present value of apartment/mobile home/home, have different amounts this wave.

Section BCDE: Reference Person/Spouse/Partner Employment
Starting in 2013 and continuing for the 2017 Wave if Reference Person/Spouse/Partner had the same "Current Main Job" (CMJ), certain information about employer was preloaded in this section and not re-asked. This wave, we added an additional question specifically about CMJ hours if Reference Person/Spouse/Partner had one or more current jobs (BC/DE39ACKPT and BC/DE39A). Furthermore, for this wave, we utilized the four-digit 2010 US Census Occupation codes and the four-digit 2012 US Census Industry codes instead of the 3-digit 2000 US Census code frames we had been using. Also, a new series, Time Use at Work, was added at BC/DE60ACKPT-BC/DE60E. Finally, in the series asking about methods used for job seeking, a new category was added to include using social media (BCDE65). Follow-up questions about the use of the internet for job/work searches (BCDE65ACKPT-BC65B) were also added.

Section F: Expenditures: Childcare, Food, Vehicles, and Transportation
In conjunction with the Time Use at Work series added to BCDE above, additional time use questions were also added to Section F (F1A-F1K). However, users should note that F1A variables for Reference Person and Spouse/Partner were asked in prior waves as F2 and F3, respectively, so F1A is not technically a new question. Additionally, a slight alteration to the Food Security series was made so that this module is skipped if the Respondent is a Proxy reporter (FOOD1CKPT). Also, a new question and accompanying unfolding brackets were added to the consumption series about computing expenses (F86A-FA86A3) as was its corresponding imputed variable (COMPUTING EXPENDITURE 2016: ER71522; see section B. Changes to Generated Variables below).

Section G: Income
Only one variable was added to the release file: G43, Number of Non-VA Pensions, for Spouse/Partner.

Section R: T-2 Income
Several questions from this section were dropped: R20-R55. We now only ask R2, Labor Income for Reference Person/Spouse/Partner.

Section W: Wealth
An additional question about whether the family has a checking or savings account was added to Section W (W27A).

Section P: Pensions
One item of note in this section is that the unfolding bracketed amounts in the P20b-e series were changed.

Section H: Health
A new code value option of "995: Less than one year old" was added to questions in the Chronic Conditions series (e.g., H6A1). Additionally, there is a new series regarding changes in thinking or memory asked of Reference Persons/Spouses/Partners age 65 or older (H5NCKPT-H5N9). This screening series referred to as *The AD8: The Washington University Dementia Screening Test,* also referred to as *The Eight-item Interview to Differentiate Aging and Dementia,* is a copyrighted instrument of Washington University. All rights reserved. Copyright 2008 by Washington University in St. Louis Missouri, USA. These values, for Reference Person/Spouse/Partner, can also be found on the 1968-2017 Cross-Year Individual File (ER34581-ER34589) because they were also asked of OFUMs. Also, two new questions about mobility were added at H8A3 and H8A4. Lastly, some of the unfolding bracket questions regarding medical expenses were deleted this wave (H68, H73, and H75).

Section J: Marriages and Children
There were no changes to the questions about marriage this wave. In the section about children, however, new skips were added (OS62CKPT and OS63CKPT) for targeting only the mother or only the father regarding wanting another child if the Respondent was a mother or a father.

Section KL: Background of Reference Persons/Spouses/Partners
Question KL78C was deleted this wave; otherwise, this section remains unchanged. As in 2015, users should note that several of the education variables are also repeated on the Cross-Year Individual File where parallel questions are asked of Big OFUMs (OFUMs age 16+). Refer to the 1968-2017 Cross-Year Individual File documentation and codebook for further information

Section IMM: New Immigrant Section
This Section is newly added for 2017 to gather information about English language proficiency and languages used at home and outside the home (ER70981-ER71038). This section was only asked for those in the Immigrant 2017cohort.

Section M: Philanthropy Section
Several questions have been deleted in this section, including the unfolding bracket amounts for the M2-M12 series. The module about Reference Person/Spouse/Partner religious attendance (M56A-M57B) that had been asked in prior waves was restored to the 2017 instrument.

Section IO: Interviewer Observations Section
One series in the Interviewer Observations Section, provided on the Family File, includes indicators about which other person/helpers assisted with each section of the interview (IO28). Based on changes to the 2017 instrument, this series was modified accordingly. For example, questions about whether Respondent received help and who assisted Respondent in answering questions in the New Immigrant Section were added.

B. Changes to Generated Variables

We have added two new generated variables to assist with the identification of the 2017 Immigrant cohort added this wave, Whether Immigrant 2017 Sample Family (ER70979) and Immigrant 2016 Screening Status for this Family Unit (ER70980).

In 2017, we continue to provide generated variables for the Food Security series. Using these assessment questions, constructed variables include: a raw score, a scaled score, and a food security category based on the raw score. Missing values were imputed for all items and both a household-level scaled score and category were generated, as well as a child scaled score and category. All food security data processing was done following the methodology outlined by the USDA-ERS, in the Guide to Measuring Household Food Security – 2000; Measuring Children's Food Security in U.S. Households, 1995-99; and the U.S. Household Food Security Survey Module: Three-Stage Design, with Screeners, September 2012.

Background information (Section K, for Spouses/Partners, and Section L, for Reference Person) is only asked when a person is a "new" Reference Person/Spouse/Partner. Values are copied forward, from prior data collections, for those Reference Persons/Spouses/Partners who maintain their relationship classification from the preceding wave. For 2017, same as for the 2015 Wave, we asked for an education update from our returning Reference Persons and Spouses/Partners. If they reported additional education attainment, we updated their education variables in the background section and reported the year of the education update on the Family File (see variables "Year Highest Education Updated", ER70753, for Spouses/Partners, and ER70891, for Reference Persons).

In 2017, we continue to provide a set of variables generated from the Employment EHC start and stop dates for the "off" year, 2015. These variables include whether Reference Person/Spouse/Partner was employed, unemployed, or out of the labor force, and the number of weeks (ER71068-ER71160). This range of variables also includes whether Reference Person/Spouse/Partner was employed, unemployed, or out of the labor force, by month, in 2015.

The Wealth Section contains imputed versions of items from Section W, which were released, in waves prior to 2009, as supplements. These imputed variables continue to be part of the Family File and are located at ER71427–ER71486.

We continue to provide set of imputed variables based on consumption questions mainly derived from Section A and Section F. They are located in the Family File at ER71487–ER71527. These imputed versions of consumption categories had been released in prior waves as a separate file. Waves 1999-2013 can still be found on the supplemental data page on our website http://simba.isr.umich.edu/Zips/ZipMain.aspx (see Consumption Expenditure Data: 1999-2013). We have a new variable, ER71522, which is the corresponding imputed version of the new question (F86A) added in Section F on computing expenditures.

For the 2017 wave we continue to calculate the Beale Rural-Urban Code and continue to use the 2013 classification scheme as published by the USDA. We no longer provide a Collapsed Beale code for public release but have replaced it with a Metro/Non-Metro Indicator (ER71531) instead.

II. DATA CHARACTERISTICS

A. Files and Format

The 2017 Family Data File consists of one data file with 9,607 records and 5,571 variables. The Family File contains one record for each family interviewed in 2017 and includes all family-level variables collected in 2017, including extensive information about Reference Person (and Spouse/Partner).

Using the web-based Data Center is the most efficient way to obtain the data (https://simba.isr.umich.edu/default.aspx). At this time, we also provide the data in a .zip package (https://simba.isr.umich.edu/Zips/ZipMain.aspx). 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. Additionally, missing data statements are not provided as part of the files.

B. Variable Names and Positions

The 2017 Family File variable names are in the range ER66001–ER71571.

The 2017 Family 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 wave-specific codebooks to obtain information about the parameters for their creation. Some examples of variables that may be constructed using component variables provided in this data file include annualized mortgage, rent payments, and food costs.

C. Hand-Coded and Generated Variables

The PSID hand codes several data items for the Family File. For example, Reference Person's and Spouse's/Partner's occupations and industries, to a maximum of four jobs apiece, are coded, for 2017, using the four-digit codes from the 2010 version of the US Census Occupation Codes and the 2012 version for Industry Codes. 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. Additionally, the field of endeavor for Reference Person's and Spouse's/Partner's non-academic degrees and certificates is coded for as many as three each. Any education received outside the U.S. is coded for Reference Person, Spouse/Partner, and their parents, as well.

Background items, such as education, are collected only for "new" Reference Persons and Spouses/Partners in a given wave. During processing, we have traditionally "brought forward" background information from previous waves for Reference Persons or Spouses/Partners who are the same persons as in the prior wave. Beginning in the 2013 Wave and continuing for the 2017 Wave, we asked for an education update for our returning Reference Persons/Spouses/Partners. If they reported additional education attainment, their education variables were updated in the background section and year of the education update reported on the Family File (see variables "Year Highest Education Updated", ER70753, for Spouses/Partners, and ER70891, for Reference Persons). 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 ER71543, for Reference Persons, and ER71544, for Spouses/Partners. Completed Education of Reference Person and Spouse/Partner variables (ER71538 and ER71539, respectively) are generated from the background information.

Family composition and change variables include Family Composition Change (ER66007), Splitoff Indicator (ER66005), Reference Person-Spouse/Partner Sample Status (ER71545), and variables about births to Reference Person, Spouse/Partner, and other family members during the prior calendar year, 2016 (ER71561–ER71564), and during the "off" year, 2015 (ER71566–ER71568). Note that the Splitoff Indicator is only assigned to a family in the year that family first moves out of the main family and forms its own separate household; after that one wave of being a splitoff, these families receive code values that designate reinterviews. For 2017, a new code value of 5 was added to the Splitoff Indicator to designate a new 2017 Immigrant Family (wave one interview for that group). Two other variables concerning splitoffs are the number of splitoffs arising from a main family (ER71546) and the Family Interview Number of the main family from which a splitoff family originated (ER71547).

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 (ER71548–ER71559), the Household ID number (ER71560), and the number of persons not included in any PSID Family Unit who are sharing the Household Unit (ER66023). 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 at https://psidonline.isr.umich.edu/data/Documentation/Fam/1993/93guide.pdf).

The marital status variables consist of Reference Person's current marital status (ER66024), the generated form of marital status comparable with years prior to 1977 (ER71540), change in marital status of Reference Person between waves (ER71541), and couple status of Reference Person (ER71542).

Location variables include PSID/GSA and FIPS state codes (ER66003 and ER66004); Current Region, Beale Rural-Urban code, and Size of the Largest City in the County (ER71530–ER71533). The Beale Rural-Urban Code (ER71532) and Size of Largest City in the County (ER71533) are suppressed to protect the anonymity of our respondents and may be obtained under a restricted data contract (see our website under "Restricted Use", for details: http://simba.isr.umich.edu/data/data.aspx). A Metro/Non-Metro Indicator based on the Beale Rural-Urban Code (ER71531) is a public release variable we provide for users' convenience. Finally, we continue to provide two derived variables, from background information: Region where Reference Person (ER71534) or Spouse/Partner Grew Up (ER71536) and Reference Person's (ER71535) and Spouse's/Partner's (ER71537) 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, 2016 (ER71528). Additionally, since the PSID has switched to biennial interviewing, comparable needs standard has also been generated for the "off" year, 2015 (ER71529).

The variable indicating whether a PSID family lives in institutional housing is ER66008. The variable indicating the total number of data records from the Cross-Year Individual File that are associated with a panel family is ER71226.

Weights are available as part of the 2017 Family File at ER71570 and ER71571. Users should note that the new immigrant 2017 sample (ER66009=4001-4462) has values of zeros for this variable as no historical data is available for these cases and, therefore, they currently should not be included in any longitudinal analysis. For Release 2 we have added a new variable 2017 CROSS-SECTIONAL FAMILY WEIGHT (ER71571) which encompasses the Core 68/Immigrant 97/99 and Immigrant 17 Samples.

Finally, the imputed work hours and income variables can be found at ER71227–ER71426.

D. Documentation and Codes

1. 2017 Questionnaire

Our website includes an HTML depiction of the computer-assisted interviewing (CAI) applications for 1993-2001 while Box-and-Arrow questionnaires are available, for all waves, in PDF format (https://psidonline.isr.umich.edu/Guide/documents.aspx). 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 (https://simba.isr.umich.edu/default.aspx). Variable descriptions are complete for all of the 2017 Family data items. The downloadable .zip package of the Family data, on our website, contains a .pdf version of the entire Family File codebook for 2017(https://simba.isr.umich.edu/zips/zipmain.aspx).

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 (e.g., those that do not apply) 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.