January, 2016

PANEL STUDY OF INCOME DYNAMICS
2003 PUBLIC RELEASE 6 FAMILY FILE

TABLE OF CONTENTS

_____

RELEASE 4 NOTES
May 2008: Variable ER23268 has been reconstructed; variable ER24179 has been recalculated. Variable ER24180, a family cross-sectional weight, has been added.

RELEASE 5 NOTES
November 2013: Correcting Births to OFUMS in 2001 (ER24178) and suppressing previously released variables Beale Rural-Urban Code (ER24144), and Size of Largest City the county (ER24145).

RELEASE 6 NOTES
February 2014: Added four new coded variables in the philanthropy section: M12code, M41code, M43code and M52code (ER23542A, ER23664A, ER23674A, ER23697A). Added fourteen new T-2 (Section R) variables (ER23718A-ER23718N) and Collapsed Beale (ER24144A).

RELEASE 7 NOTES
January 2016: For Release 7 we have added 96 new variables in the T-2 (2001) Income section, 86 of which are about time unemployed for Head (ER23702A1-ER23702B5) and Wife/"Wife" (ER23702F4-ER23702G8), time out of the labor force for Head (ER23702B6-ER23702D1) and Wife/"Wife" (ER23702G9-ER23702J4), and time employed for Head (ER23702D2-ER23702E7) and Wife/"Wife" (ER23702J5-ER23702L1). The other 10 new variables are T-2 hours worked, labor income, and accuracy checks for Head (ER23702E8-ER23702F3) and for Wife/"Wife" (ER23702L2-ER23702L6).

I. INTRODUCTION

A. What's New for 2003

1. 2003 Questionnaire Changes

Most of the sets of questions added in the 1999 wave continue to be part of the 2003 instrument, but a few sequences and sections were dropped, some were added and others revised. Question F14b, regarding whether food stamp recipients use a plastic EBT card or paper coupons, was dropped for 2003. A question series asking about Heads' and Wives'/"Wives'" home use of computers was new for 2003. See questionnaire questions A47a-A47i (ER21099-ER21116).

In the income section (Section G), questions about business profits and losses (G11-G11b) were altered; the PSID now asks for gross receipts, expenses, and net profit or loss. Wife's/"Wife's" asset income questions (G59a-d; 2003 ER22335-ER22401) were reorganized to strictly parallel Head's.

In the health section (Section H), questions were added about recent health comparisons, childhood health levels, and hospitalizations for Head (H1a-d and H8-H8a; ER23011-ER23013 and ER23089-ER23090) and Wife/"Wife" (H25a-d and H32-H32a; ER23138-ER23140 and ER23217-ER23218), and IADLs were resurrected (H11a-m and H35a-m; ER23106-ER23117 and ER23233-ER23244).

A large overhaul was done to questions from 2001 Section T, on philanthropic giving, at the behest of its developers at the Center for Philanthropy at Indiana University-Purdue University. The 2003 version (Section M) asks much more detail about time spent by Heads and Wives/"Wives" doing volunteer work for nonprofit organizations.

Responses to new indicators in Section R, Welfare Reform, about meals for the elderly and school lunches and breakfasts (R78a, R80a and R81a) are located on the cross-year individual file. Reasons for denial of assistance (R72a-l, in the series ER23865-ER24044) were included.

The major transformation of the 2003 questionnaire was the conversion of the employment sections (Sections B and C for Heads; Sections D and E for Wives/"Wives") to an Event History Calendar (EHC) format. Briefly, EHC is a data collection method that asks the respondent for dates of landmark events such as births of children, moves to new housing, etc., as assists in dating employment spells and periods of time off. One significant feature of the change is that the PSID no longer divides all jobs into main versus side or secondary ("extra") jobs.

Unemployment time and time laid off from work had previously been included in the same set of variables but were separated with the move to EHC. Details about beginning and ending wage rates and weekly work hours, promotions, and occupational changes with the same employer were dropped, but information about longevity with the employer, current pay, union membership, etc., continue to be included. See questionnaire sections BC and DE for further information.

2. Generated Variables

a. Family composition variables
        Family Composition Change (ER21007) is available again this year, as is
        another important variable for identifying changes in family
        composition, Splitoff Indicator (ER21005). Note that the Splitoff

Indicator indicates a splitoff family only in the first year such a
family forms; thereafter, these families receive code values that
designate reinterviews.

b. Family level income/employment variables

Family income, work hours, and wage rate variables that had been
included in the 1997 and 1999 family files but were removed for 2001 are
once more part of the family file. These variables were not computed
for Release 1 but are available beginning with Release 3. Wealth
composite variables remain in a separate data file within the Data
Center.

c. Weights

Weights are available as part of the 2003 family file. The
core/immigrant family longitudinal weight (ER24179) has been
recalculated for Release 4 and subsequent releases. For details, see
our internet site, PSIDONLINE.ISR.UMICH.EDU: select 'Data &
Documentation', then select 'Sample Weights', then 'PSID Longitudinal
Weights', and 'PSID Revised Longitudinal Weights 1993-2005'.

Also beginning with Release 4, a new family cross-sectional weight,
ER24180, has been constructed. Information on this new weight is also
available on our internet site psidonline.isr.umich.edu): select 'Data
& Documentation', then select 'Sample Weights', then 'PSID Cross-
sectional Weights', and 'Cross-sectional Analysis of PSID Data and
Cross-sectional Weights 1997-2005'.

d. "New" variables

Some generated variables that are included in the 2003 list of
variables contained zeroes for Releases 1 and 2 but are now filled with
actual values.

REGION HD GREW UP	ER24146
HD GEOGRAPHIC MOBILITY	ER24147
COMPLETED ED-HD	ER24148
COMPLETED ED-WF	ER24149
YEAR NEW HEAD IN FU	ER24153
YEAR NEW WIFE IN FU	ER24154
BIRTHS TO HEAD ONLY LAST YEAR	ER24171
BIRTHS TO WIFE ONLY LAST YEAR	ER24172
BIRTHS TO HEAD AND WIFE LAST YEAR	ER24173
BIRTHS TO OFUMS ONLY LAST YEAR	ER24174
BIRTHS TO HEAD ONLY TWO YEARS AGO	ER24175
BIRTHS TO WIFE ONLY TWO YEARS AGO	ER24176
BIRTHS TO HEAD AND WIFE TWO YEARS AGO	ER24177
BIRTHS TO OFUMS ONLY TWO YEARS AGO	ER24178

In addition, background information (Sections K and L, for
Wives/"Wives" and Heads, respectively), which is only asked when a
person newly acquires a relationship of Head or Wife/"Wife", was copied
forward from prior waves for those Heads and Wives/"Wives" who
maintained their relationship classification from the prior wave.

3. Occupation, Industry and Other Codes

As in prior waves, all occupation and industry codes are available on the file. This includes the three-digit codes for Head and Wife/"Wife" on up to four jobs, one-digit reason for job termination, the two-digit industry codes for family-owned businesses, and the field of endeavor for non-academic degrees and certificates.

For 2003, the PSID made the switch to the 2000 Census Occupation and Industry Code. Note that since 1997, background information for Heads and Wives/"Wives" was coded using the 1970 code. Since background information, which includes occupations and industries of each parent's and own first job, is asked only when a Head or Wife/"Wife" changes status, codes for 2003 Heads, Wives and "Wives" who were most recently asked these questions in 1997 through 2001 would contain values incompatible with 2003. These older values have been converted for 2003 data.

II. DATA CHARACTERISTICS

A. Files and Format

The 2003 family data file consists of one data file with 3,295 variables and 7,822 records. The family file contains one record for each family interviewed in 2003 and includes all family-level variables collected in 2003.

Using the web-based Data Center is the most efficient way to obtain the data. At this time, we still 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 data file in ASCII format and SAS, SPSS and STATA data definition statements.

These 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 these data definition files.

B. Variable Names, Positions and Generated Variables for the Family File

The 2003 family variable names are in the range ER21001-ER24180.

The 2003 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 mortgage amounts, rent payments and food costs.

Background items such as education are collected for new Heads and new Wives/"Wives" only. 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. Beginning with Release 3 of the 2003 data file, we have completed this "bringing forward" from prior waves.

C. Documentation and Codes

1. Documentation

In addition to this file, DOC.TXT, the .zip file contains a traditional codebook
for the 2003 family file, FAM2003ER_CODEBOOK.PDF, with unweighted frequencies to
the left of the code frames. A variable means report is contained in
FAM2003ER_VDM.TXT.

2. 2003 Questionnaire

The Data Center has been modified to allow users to create and download
customized codebooks that provide variable-specific documentation. As of this
writing, all variable descriptions in codebooks through the 2003 wave are
complete.

The PSID website includes PDF-format box-and-arrow questionnaires and HTML
versions of the computer-assisted interviewing (CAI) applications at:

        http://psidonline.isr.umich.edu/data/zipCore.aspx

Use the labels from the SAS, SPSS, and STATA data definition statements to match
variables with questions in the questionnaire or CAI application. The
questionnaire and the CAI application contain codes for quick-and-dirty use with
most data items, although these should be cross-checked against the codebook.

3. Code Categories

Please refer to the 2003 family file codebook for descriptions of code
categories. In general, codes follow our traditional structure, 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
not included in the code category, assume missing data for that value.

See Part E below for additional information on generated variables.

D. Problem Variables and Missing Variables

All variables are complete beginning with Release 4. ??

E. Generated Variables

A number of generated variables are included in the 2003 family file.

One such group is location data. We include PSID/GSA and FIPS state codes
(ER21003 and ER21004). We also include the PSID's traditional variables for
region, Beale rural-urban code, and size of the largest city in the county
(ER24143-ER24145, respectively). The codes for FIPS and PSID/GSA codes are
included in Appendix 1 of the 1985 documentation, available in .pdf format on
our website. Codes and generation details about the other three variables are
located in the 2003 codebook.

Two additional location variables were added for 2003. These are Region where Head Grew Up (ER24146) and Head's Geographic Mobility (ER24147), both derived from background information.

We also offer a group of marital status variables: current marital status of Head (ER21023), the generated form of marital status comparable with years prior to 1977 (ER24150), change in marital status of Head between waves (ER24151), and couple status of Head (ER24152).

The variable indicating whether a PSID family lives in institutional housing is present (ER21008), as is the variable indicating the total number of data records from the cross-year individual file that are associated with a panel family (ER24076).

Both USDA and Census needs standards have been generated for the prior calendar year, 2002 (ER24139 and ER24140). Additionally, since the PSID has switched to biennial interviewing, comparable needs standards have also been generated for the "off" year, 2001 (ER24141 and ER24142).

Completed education of Head and Wife/"Wife" (ER24148 and ER24149) and the year in which the background information was most recently asked (ER24153 and ER24154) are present as of Release 3. These four variables are generated from background information.

Two traditional PSID family-level generated variables concerning splitoffs are included: the number of splitoffs arising from a main family (ER24156); the family interview number of the main family from which a splitoff family originated (ER24157).

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 (ER24158-ER24169), the household ID number (ER24170), and the number of persons not included in any PSID family who are sharing the household unit (ER21022). The PSID documentation for 1993 and earlier waves has additional information about multiple PSID families sharing the same household (see Section I, Part 5 of the front matter).

Family Composition Change (ER21007) and Splitoff Indicator (ER21005) are included. Head-Spouse Sample Status (ER24155) had not been generated since 1993. These variables were not computed in Release 1, but have been available beginning with Release 2.

Variables about births to Head, spouse and other family members during the prior calendar year, 2002 (ER24171-ER24174), and during the "off" year, 2001 (ER24175-ER24178), have been added for the first time since 1993. These variables were not computed in Release 1, but have been available beginning with Release 2.

Imputed family income measures and work hours and wage rates for Heads and Wives/"Wives" (ER24077-ER24138) are now located with the main 2003 family data. These variables were not computed in Release 1, but have been available beginning with Release 2.