January, 2016

PANEL STUDY OF INCOME DYNAMICS
1995 PUBLIC RELEASE FAMILY FILE

CONTENT OUTLINE

I. INTRODUCTION
  A.  What's new for 1995
    1. Education supplement
    2. Supplemental question on food shortage
    3. New questions about recipients of Supplemental Security
       Income

II.  CHARACTERISTICS
  A.  Files and format
  B.  Documentation
  C.  Variable numbers, positions and generated variables
  D.  Bringing forward background information for Head and
      Wife/"Wife"
  E.  Problem variables, missing variables
  F.  Comparability with 1992 final release file and earlier final
      release files
  G.  Additional notes: Sample supplements in 1993, 1994 and 1995


      ...................................................................
Release 6 notes: January 2016: Added generated variables splitoff indicator (ER5005F), 1968 Family Identifier (ER5005G), Collapsed Beale Code (ER6997F), Number of Individual records (ER5005H), DE1 Checkpoint (ER5560A) and F1 Checkpoint (ER6053A), A45 (ER5066A) and A46 (ER5066B). In addition we corrected the top code amount for A20 (ER5032) and A21 (ER5033) to match the codebook values.

Release 5 notes: December 2013: Added generated variables such as Family Composition Change (ER5004A), DUFU variables (ER6999F-ER6999R), Splitoff variables (ER6999D-ER6999E), Births to HD/WF/OFUM (ER6999S-ER6999V), Generated Marital Status (ER6999A), Change in Marital Status (ER6999B), Couple status of Head (ER6999C), Region where HD grew up (ER6997C), Head Geographic Mobility (ER6997D), and Current Region (ER6997E). Interview length, mode of interview, language of interview, calls for iw (ER5005B-ER5005E). Added additional KL variables that were missing; L2 State Father Grew up (ER6787A), L3 State Mother Grew up (ER6787B), and L8 State Head Grew up (ER6789A). Brought forward the background section for old Heads and Wives in section KL. Added missing variable E66A UNEMPLOYED FEB (ER5931A).Corrected Job Search variables that had been erroneously filled with zeros (ER5094-ER5096 ER5317-ER5319 ER5588-ER5590 ER5811-ER5813).


I.  INTRODUCTION

A.  What's new for 1995

Release 4 of the 1995 Public Release I Family File incorporates variables that had been available previously in supplemental files:  Head and Wife/"Wife" occupation and industry codes, reason a job ended, industry codes for businesses owned, Head and Wife/"Wife" field of certificates earnings, Head's father's occupation, Head's first occupation, Head and Wife/"Wife" work hours and work weeks, Head, Wife/"Wife" annual income variables, needs standards, state code, completed education of Head and Wife/"Wife", and a recalculated 1995 core family longitudinal weight (for details about the construction of this weight, go to the PSID website, psidonline.isr.umich.edu:  select 'Data & Documentation', then select 'Sample Weights', then 'PSID Longitudinal Weights;, and 'PSID Revised Longitudinal Weights 1993-2005').

Questions A45-A47 were reinstated: whether might move, definiteness of move, and why might move.

The following sections were dropped in 1995: the Wealth Section, the Section on Expectations, and Male Head's Military Service.

1. Education supplement

A supplemental module on education, conducted in 1995, covered a wide range of issues regarding the schooling experiences of family members between the ages of 5 and 49. The module gathered retrospective reports about attendance in public vs. private schools (non-religious, Catholic or other religion), highest grade completed, last date of attendance, high school graduation or GED, special education (including gifted), whether an individual has repeated a grade and extracurricular activities. Questions about pre-school experiences including involvement in Head Start, nursery school and day care programs were asked. In addition, there were questions about behavior problems in school (including suspension and expulsion), contact with police (not including minor traffic offenses) and time spent in correctional institutions such as reform school or prison. Data for this supplement are included in the 1968-1995 public release I individual file (ER33220-ER33274).

2. Supplemental question on food shortage

The PSID staff and the PSID Board of Overseers encourage researchers to propose supplemental questions to the PSID that are suitable for inclusion in a panel study and add value to existing and planned PSID questions. Proposed supplementary questions are selected by the PSID Board of Overseers after review of proposals explaining the scientific merit of the proposed supplement. Short supplementary questions are paid for by the PSID core budget. In 1995 the supplemental question was:

> F23. In the last twelve months, did you ever run out of the foods that you needed to make a meal and didn't have the money to get more?

This variable is included in the 1995 public release I family file as ER6091.

3. New questions about recipients of Supplemental Security Income

Permanent additions to the questions on Supplemental Security Income were added. These questions were included to exclude amounts received on behalf of someone else. If the receipt of SSI income was reported for the Head, Wife/"Wife" or any OFUM (other family unit member), a follow-up question about whether it was received for that person or for someone else was asked. These variables are included in the 1995 public release I family file as ER6277 and ER6584 for the Head and Wife/"Wife", respectively.

II. CHARACTERISTICS

A. Files and format

The zip package consists of the 1995 Public Release I Family File and a number of files associated with it. In addition to this file, the 1995 public release I family file package includes:

| File name | Contents |
|---|---|
| FAM1995ER.TXT | 1995 family data |
| FAM1995ER.SAS | SAS statements for family variables |
| FAM1995ER.SPS | SPSS statements for family variables |
| FAM1995ER.DO | STATA statements for family variables |
| FAM1995ER_CODEBOOK_public.PDF | Codebook for family variables |

The contents of these files are described in detail below.

Data File

The data are in raw ASCII form.  The 1995 public release I family data file contains one record for each family interviewed in 1995.  The file includes all family-level variables collected in 1995.  Refer to one of the corresponding files -- .SAS, .SPS, or .DO -- for record format layout information, variable names and variable labels.

Data Definition Statement Files

These files contain SAS, SPSS, or STATA data definition statements that provide variable names, locations, and variable labels.  Missing data statements have not been provided.  You should check the questionnaire and frequencies or means for each variable you intend to include in your analysis to determine which code values should be defined as missing.

The SAS, SPSS, and STATA statements are NOT intended to represent completed 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.

B.  Documentation

The 1995 Public Release I Family File Codebook is in PDF format.  Unweighted frequencies appear to the left of the codeframes.

Codes

In general, codes follow our traditional structure, although "don't know" responses are now largely distinguished from other missing data responses.  If the questionnaire does not indicate otherwise, 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.  A few fields contained non-numeric characters, and these have also been converted to zeros for the early release file.  If a variable contains a code value that is neither included in the questionnaire nor one of the zero, eight or nine codes just mentioned, you should assume missing data for that value.

The inevitable exception: codes 21 through 24 for month variables in event dating questions were not printed in the questionnaire but were used throughout the CATI application to indicate mentions of season only.  These codes follow:

  21.  DK month, but season was winter
  22.  DK month, but season was spring
  23.  DK month, but season was summer
  24.  DK month, but season was autumn

C.  Variable numbers, positions and generated variables

All variable numbers for the 1995 Public Release I family file are prefaced with "ER", rather than "V", to assist both analysts and study staff to determine whether reference is to the public release I file or to the final release version.  The 1995 public release I family variables are in the range ER5001-ER7000B.

Background information has not been asked about Heads and Wive/"Wives" each and every year.  (Note, however, that in 1995 we asked about the education of all 1995 Heads and Wives/"Wives" in Section M of the questionnaire; these

data are included in the cross-year public release I individual file.)  We ask the questions about new Heads and new Wives/"Wives" only.  During processing, we have traditionally "brought forward" the 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.

D.  Bringing forward background information for Head and Wife/"Wife"

In Release 5, we have brought the background section forward for all Heads and Wives/"Wives".

E.  Problem variables, missing variables

The 1995 public release I family file includes many series of variables concerning monthly dating of events during the prior calendar year.  For example, ER5118-ER5129 indicate the months during which the Head worked on his or her present main job in 1994 (questionnaire question B39).  The "strings" consist of a set of twelve dummy variables, one for each month.  A code value of 0 indicates that the activity did not occur during this month; a code value of 1 indicates that it did.  However, due to a programming error, the month of January in each monthly "string" is suspect because it can contain a value of 1 when the value should be 0.  If you are using any of these strings you may want to inspect the string for unusual patterns and recode the January value if it seems appropriate.  For instance, cases with values of 1 for January but not for February may warrant special handling.  Otherwise, annual incomes could be miscalculated if the monthly string were used for computation.

Other variables in the 1995 public release I family file with suspect code distributions include:

ER6795 and ER6797 (questions L15 and L17), literacy of  Head's father and mother, respectively, have 21 cases with values of 9 that should probably have had values of 0  instead.

The following variables have all missing data:

   ER5054 (question A36), reason why the family neither owns  nor rents the HU; and

   ER6720 (question G111), a checkpoint for number of  dependents.

Missing variables include:

   employment status for individuals other than the Head or  Wife/"Wife";

   question G113, the number of persons dependent on this family for more than half of their support; and

   questions G9a-G9d, whether Head and Wife/ "Wife" spent time working at a business and, if so, whether they reported those work hours.

F.  Comparability with 1992 final release file and earlier final release files

   Beginning with the 1993 wave, the data were collected using CATI (Computer Assisted Telephone Interviewing).

This means that information about each question was collected electronically by the interviewer and, in effect, was coded at the time of data collection.  Conversion to standardized units of measurement, formerly performed as part of our coding operation has not yet been done.  As a result, the data in the public release I files much more directly resemble the answers to questionnaire questions than the 1992 and early years' data did.  For example, instead of one variable indicating monthly rental expense, rent costs now exist as two variables: one for the dollar amount and one for the time unit, e.g., $500 per month and $100 per week are typical of responses to the question about rent payments

Beginning with the 1990 data, we faced problems with the size of our merged cross-year record formats -- for 1990 the logical record length of a merged cross-year record would have exceed 32,767. The 1989 files were the last released the theretofore traditional cross-year family-individual and cross-year family format.

G.  Additional notes:  Sample supplements in 1993, 1994 and 1995

We had added a Latino sample of 2,043 families to the PSID in 1990. This sample is described in detail in the 1990 documentation.  It was derived from a sample selected and interviewed by Temple University Institute for Survey Research for their Latino National Political Survey (LNPS). The Latino addition was made congruent with our usual ID scheme and unique identifier formats. Latino sample cases are easily identified in the family and individual files by the code values for 1968 ID Number (V20302 in the 1992 family file and V30001 in the 1968-1992 individual file) -- the Latino sample has code values in the range 7001-9043.

In 1992 several different kinds of recontacts were attempted. These are described in detail in the 1992 family documentation, but briefly, three groups were selected:

    1) all 1991 nonresponse;
    2) a random subset of SRC and Census sample members who had become
        nonresponse in 1990 or earlier; and
    3) all of Temple University's Latino sample persons who were
        not successfully interviewed by us in 1990.

The successfully recontacted Latino families have 1968 ID Numbers in the range 9044-9308.

Our recontact effort for 1993 included the resurrection of many nonresponse sample persons who shared a 1968 ID number with families still responding in 1992, similar to the second group selected for 1992 as described above. But in contrast to this 1992 group, priority was given to families with connected individuals under age 18. All sample individuals within such a family were selected for recontact, even if they themselves were older.

The main focus of the 1994 recontact effort was to follow nonsample ex-spouses of sample members; these ex-spouses had one or more children with the sample members, and at least one of those children was expected to be under age 18 by 1994. In addition, recontacts were attempted with persons who had become nonresponse in 1992 or 1993, with nonresponse core sample persons who had no other family members still responding by 1993 (some of whom had become nonresponse as early as 1969), and with some children formerly designated nonsample but born to sample members since the study began.

The 1993, 1994 and 1995 waves included a change in PSID following rules. We now follow all sample persons who leave home, regardless of age. So, for example, when a sample male Head leaves his nonsample wife and their sample children, we attempt an interview not only with him but also with her because her household contains their sample children. Beginning with the 1994 data collection, we also now consider as sample those children who are born to a sample parent in a year when the sample parent was not in an interviewed family.