

PANEL STUDY OF INCOME DYNAMICS
1999 PUBLIC RELEASE FAMILY FILE

TABLE OF CONTENTS

I. INTRODUCTION

A. WHAT'S NEW FOR 1999

1. THE TWO-YEAR INTERVIEW CYCLE AND NEW QUESTIONS
2. ADDITIONAL NEW IMMIGRANT SAMPLE ADDITION
3. OCCUPATION, INDUSTRY AND OTHER CODES

II. CHARACTERISTICS

A. FILES AND FORMAT

B. VARIABLE NAMES, POSITIONS AND GENERATED VARIABLES FOR THE FAMILY FILE

C. DOCUMENTATION AND CODES

1. DOCUMENTATION
2. 1999 QUESTIONNAIRE
3. CODE CATEGORIES

D. BRINGING FORWARD BACKGROUND INFORMATION FOR HEAD AND WIFE/"WIFE"

E. PROBLEM VARIABLES, MISSING VARIABLES

F. COMPARABILITY WITH 1993 AND EARLIER PUBLIC RELEASE II FILES

G. NOTES ON NON-QUESTIONNAIRE VARIABLES IN 1999

Release 7 notes: January 2016: We have added one new variable Splitoff Indicator (ER13005E). Additionally we have corrected the top code values for A20 and A21 (ER13041 and ER13042) to match the codebook description. We have corrected the bottom code for H14/H38 (ER15544/ER15652) and H17/H41 (ER15547/ER15655) for Head and Wife/"Wife" where 1 means one cigarette or fewer per day.

A. What's New for 1999

Release 4 of the 1999 Public Release I Family File incorporates variables that had been available previously in supplemental files: Head/Wife/"Wife" work hours and work weeks, Head/Wife/"Wife" annual income variables, completed education of Head/Wife/"Wife", and a recalculated 1997 core family/immigrant 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').

In addition, a 1999 core/immigrant family cross-sectional weight has been constructed. For details about this new weight, go to the PSID 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'.

The core content of the 1999 questionnaire omits very few of the items asked in 1997. The 1997 questions about incoming and outgoing child support were dropped and background information was not re-asked for some Heads and Wives/"Wives", but otherwise the questionnaires include the same items.

However, the 1999 questionnaire is much larger than ever before, with two thirds again as many variables as in 1997.

1. The Two-Year Interview Cycle and New Questions

The transition to biennial interviewing was a big change for the PSID. Because nobody wants to miss out on information about the "off" year, a section was added to try to capture at least minimal information on income and work for the family. These questions are located in Section R of the questionnaire.

Many other new sets of questions were added for 1999: a module on utilities costs (Section A); a section on vehicles owned by the family (Section V) and some questions about expenses (Section X), both located at the end of the housing questions in Section A; information about food stamp receipt for the "off" year and additional questions about the effects of poverty on food consumption, mainly located at the end of Section F; detailed questions about pensions belonging to Head and Wife/"Wife" (Section P), the wealth sequence that has traditionally been asked every five years (Section W), and a small group of questions about spending (end of Section W), all located immediately following Section G; a revised and expanded health series for Head and Wife/"Wife", questions about effects of diet, and a sequence on health insurance coverage and medical expenses, all in Section H.

2. Additional New Immigrant Sample Addition

The PSID added 441 immigrant families in 1997, but the sampling effort had yielded some additional cases that weren't eligible at the time of the 1997 interview because they had not yet been resident in the U.S. for two years. In

1999 we located, persuaded and interviewed seventy more new immigrant families. These additional families are included on the data files with all the other PSID respondents.

The 1997 immigrant families and individuals were assigned a unique range of values so that they can easily be distinguished from the other PSID samples. The 1968 Interview (ID) Number for the 1997 sample addition is in the range 3001-3441. For the 1999 addition, we simply continued this series, with values of 3442 through 3511.

All new heads and wives/"wives", both immigrant and core sample, were asked in what year they came to the U.S. to stay. These questions served as screening questions for the immigrant sample and were useful in development of 1999 sample weights for both the core and the immigrants.

3. Occupation, Industry and Other Codes

Since hand coding was completed before the family file was released, all occupation and industry codes are available on the file. This includes the three-digit codes for Head and Wife/"Wife" on main and extra jobs, reasons for leaving jobs, the two-digit industry codes for family-owned businesses, and the field of endeavor for non-academic degrees and certificates.

II. CHARACTERISTICS

A. Files and Format

The 1999 Public Release I family data consists of one data file with 3,535 variables and 6,997 records. The file contains one record for each family interviewed in 1999 and includes all family-level variables collected in 1999.

In addition to the data file and this document, we also include SAS, SPSS and STATA data definition statements.

The data in the .zip package are in raw ASCII form, and the file is named FAM1999ER.TXT. Refer to the corresponding .SAS, .SPS or .DO file for record format layout information, variable names and variable labels. If you download the data in ASCII form from the PSID Data Center, you will receive SAS, SPSS or STATA statements, as you request.

Whether downloaded from the .zip package or the PSID Data Center, the .SAS, .SPS, and .DO files include SAS, SPSS and STATA data definition statements that provide variable names, locations, and variable labels.

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.

B. Variable Names, Positions and Generated Variables

All variable names for the 1999 Public Release I family file are prefaced with "ER" rather than "V" to assist both analysts and study staff in determining whether reference is to the Release I or the eventual Release II version. The 1999 family variables are in the range ER13001-ER16519.

Variables not included in this file for which component items are available include:

- * imputed and annualized mortgage and rent payments, and
- * imputed and annualized food costs.

Since component items exist on the file, you can generate these items. Needless to say, imputations have NOT been done for missing data. To create variables from the 1999 Public Release I data that resemble those on final files from 1993 and earlier waves, we suggest you consult the 1993 codebooks, where we provide information about their creation.

Most background items have not been asked about all Heads and Wives/"Wives" each and every year. We ask the questions 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. For the first time, the 1999 Release I file HAS undergone this "bringing forward" from the 1997 Release I data. The good news is that all items asked afresh in 1997 have been copied to equivalent variables for 1999 if the Head or Wife/"Wife" was the same person. So, for example, race is available for all Heads and Wives/"Wives" with no further ado. But we also have bad news, or at least not completely good news. Education was not re-asked in 1997, and so this information is still not available for some Heads and Wives/"Wives" in the 1999 file. See Section D below for a detailed description of how you can "bring forward" the missing background items yourself.

Other variables are not generatable because income components of individuals other than Head and Wife/"Wife" are not included in the 1999 Release I data. Variables not included in this file which cannot be generated from available information include:

- * detailed incomes for other family members (although collective transfer and taxable totals ARE included),
- * Head's geographic mobility, and
- * county unemployment rate.

C. Documentation and Codes

1. Documentation

In addition to this file, ER99FAMDOC.TXT, the .zip file contains a traditional codebook for the 1999 family file, FAM1999ER_CODEBOOK.PDF, with unweighted frequencies to the left of the code frames.

2. 1999 Questionnaire

Our website includes an HTML depiction of the computer-assisted interviewing application for 1999. This replaces the PDF-format questionnaires we have provided for the past several years. Use the labels from the SAS and SPSS data definition statements to match variables with questions in the CAI application. The HTM application contains codes for most data items.

3. Code Categories

The HTM version of the 1999 interviewing application provides most codes for the family data. 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 neither included in the application nor one of the "zero", "eight" or "nine" codes just mentioned, assume missing data for that value. We will correct such cases for Release II, but time constraints do not permit this sort of cleaning for Release I.

This year, quite a number of generated variables are included on the Release I family files but are not part of the application. See Part G for further help.

D. Bringing Forward Background Information for Head and Wife/"Wife"

In Release 5, the background section for Head and Wife/"Wife" has been brought forward.

E. Problem Variables, Missing Variables

In general, we are very pleased with our Release I data quality for 1999. We have conquered past problems with data extraction that had caused unfortunate gaps in the data items prior to 1997.

Since for the most part these data have not yet been edited, outrageous answers given by respondents or entered by interviewers remain. However, the user may be assured that the early-release data are a faithful representation of collected information.

We're aware of specifics on one variable with wildly improbable or incorrect values (6 cases):

ER13037 A16 ACTUAL # ROOMS has values of 33, 55, 70, 77, 88 and 97.

One case (ER13002=2552) is missing background information for the wife because of an error in the 1997 data.

Values for variables referring to third and fourth extra jobs for working wives/"wives" (ER13928-ER13975) contain only zeroes. Similarly, nonworking wives/"wives" reported no second through fourth extra jobs (ER14156-ER14227).

F. Comparability with 1993 Public Release II File and Earlier Public Release II Files

Beginning with the 1993 wave, the data were collected using CATI (Computer Assisted Telephone Interviewing). This meant 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 1994-1999 Public Release I and the 1993 Public Release II family 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. Typical responses are \$500 per month, \$100 per week, or don't know how much per month. In addition, the 1993 Public Release II file includes an annualized rent amount, free of missing data, and an accuracy variable indicating whether an imputation was necessary to arrive at the annualization. This is more in line with 1992 and earlier procedures. But the Public Release I files have not yet been processed through the imputation algorithm. Therefore, many of the 1994 through 1999 Public Release I variables are not directly equivalent to variables from

1992 and earlier waves.

Unlike data collected through 1993, the Public Release I family data have NOT been cleaned with our manual economic edit process, so you must convert these kinds of amounts into some sort of consistent unit for inter-case comparison and make decisions about handling missing data. In addition, we expect that values for quite a few cases will change when we do perform economic edit operations. For example, all time unit questions include an "other" code, as well as options for missing data; amounts associated with these "other" codes will be recoded from interviewers' notes when the data are cleaned.

You can create cross-year files with variables needed for your particular analysis by merging the necessary information from the appropriate family files. The necessary family identification numbers appear both in the Public Release II and Public Release I cross-year individual files. Since they do not appear on any family data files from 1993 onward, you can obtain them from the Head's record in the 1968-1999 cross-year individual file (see Section D, step 1, above). Detailed instructions for the process of creating the cross-year files are included in the 1993 family documentation volume and are also available at our Internet site; thus, they are not repeated here.

G. Notes on Non-Questionnaire Variables in 1999

The 1999 family file includes quite a number of variables that are not lifted directly from the questionnaire. Some are income summaries and are located at the end of the data file. These income amounts do not quite match the 1999 'Income Plus' files, however, as we recalculated for the early release data and included new immigrant cases (which were excluded from the 'Income Plus' file).

Each of these income components has an associated accuracy code that indicates whether imputations were made in order to calculate totals. Accuracy variable codes are 0 for no imputation and 1 if an imputation was needed.

We are also pleased to include our traditional hand coded items--all occupations and industries, reasons for leaving employers, and fields of endeavor for non-academic certificates received

Values for the ethnicity questions, asked of all Heads and Wives/"Wives" in 1997 and of new Heads and Wives/"Wives" in 1999, have been coded and are included. (The 1997 values were "brought forward" for some Heads and Wives/"Wives" in 1999, along with other background information, so all 1999 Heads and Wives/"Wives" have information here.)

One-digit ethnicity (ER15840 for Wife/"Wife"; ER15932 for Head):

1. "American"
2. Hyphenated American (e.g., African-American, Mexican-American)
3. National origin (e.g., French, German [n.b. PA Dutch=German], Dutch, Iranian, Scots-Irish)

4. Nonspecific Hispanic identity (e.g., Chicano, Latino)
5. Racial (e.g., white or Caucasian, black)
6. Religious (e.g., Jewish, Roman Catholic, Baptist)
7. Other (e.g., "a mutt", "Heinz 57", "good old boy", "hillbilly")
9. NA; DK

Two-digit ethnicity (ER15841-ER15842 for Wife/"Wife"; ER15933-ER15934 for Head):

If a national origin was given for the one-digit variable (code 2 or 3), then up to two mentions are coded here.

1. American (meaning U.S.)
2. American Indian, Eskimo, Aleut
3. British: English, Scottish, Irish, Welsh
4. Western European: French, German, Italian, Portuguese, Spanish, Swiss, Netherlander/Dutch, Belgian, Luxembourger, Liechtensteiner, Sicilian, Corsican, Sardinian, Maltese
5. Eastern European: Romanian, Polish, Russian and former Soviet republics, Greek, Bulgarian, Yugoslavian, Bosnian, Herzegovinian, Croatian, Slovenian, Slovakian, Czech, Albanian
6. Northern European/Scandinavian: Norwegian, Swedish, Finnish, Dane
7. Middle Eastern: Iranian, Turkish, Iraqi, Syrian, Lebanese, Egyptian, Afghani, Pakistani, Omani, Yemenite, Jordanian, Georgian, Armenian, Azerbaijani, Uzbekistani, Tajikistani, Turkmenistani, Saudi Arabian, Israeli
8. East Asian: Chinese, Japanese, Korean
9. South or Southeast Asian: Indian, Vietnamese, Cambodian, Laotian, Thai
10. Pacific Islander: Filipino, Indonesian
11. Canadian
12. Central American: Nicaraguan, Mexican, etc.
13. Caribbean: Cuban, Haitian, etc.
14. South American: Peruvian, Chilean, etc.
15. African
16. Oceania: Australian, New Zealander, New Guinean
97. Other
99. NA, DK
0. INAP.: no second mention; one-digit ethnicity=1,4-7,9

Many new question sequences were added to the interview for 1999. Almost all of the codes for these new variables are present in the CAI documentation on our internet site, but vehicle manufacturer, make and type codes are not. Details were asked for as many as three vehicles.

Codes for manufacturer and make (ER13101-ER13102, ER13131-ER13132, and ER13161-ER13162)--the first two digits are the values for the manufacturer variable, and the second pair identify the make within manufacturer:

- 10 01. AMC
- 10 02. AMC Eagle (1980-1988)
- 10 03. AMC Jeep (1970-1987)
- 10 05. AMC Renault
- 10 06. AMC Studebaker
- 10 98. Other AMC; AMC, DK make

- 11 01. Chrysler
- 11 02. Chrysler Dodge
- 11 03. Chrysler Eagle (1987-1998)
- 11 04. Chrysler Jeep (after 1992)
- 11 05. Chrysler Plymouth
- 11 98. Chrysler, DK make

- 12 01. Ford
- 12 02. Ford Lincoln
- 12 03. Ford Mercury
- 12 04. Ford Jaguar
- 12 98. Ford, DK make

- 13 01. GM Buick
- 13 02. GM Cadillac
- 13 03. GM Chevrolet
- 13 04. GM GMC
- 13 05. GM Oldsmobile
- 13 06. GM Pontiac
- 13 07. GM Saturn
- 13 08. GM Saab
- 13 98. GM, DK make

- 19 98. Other domestic manufacturer (Delorean, Hudson, Nash, etc.);
domestic, DK manufacturer

- 21 98. Motor home (Winnebago, Coachman, etc.)

- 30 01. Honda
- 30 02. Honda Acura
- 30 98. Honda, DK make

- 31 01. Nissan
- 31 02. Nissan Infiniti
- 31 03. Nissan Datsun
- 31 98. Nissan, DK make

- 32 01. Toyota
- 32 02. Toyota Lexus
- 32 98. Toyota, DK make

- 33 01. Hyundai
- 34 01. Isuzu
- 35 01. Kia
- 36 01. Mazda
- 37 01. Mitsubishi
- 38 01. Suburu
- 39 01. Suzuki

- 49 05. Daihatsu
- 49 06. Daewoo
- 49 98. Other Asian manufacturer; Asian, DK manufacturer

- 80 01. BMW

- 81 01. Mercedes-Benz

- 82 01. Porsche

- 83 01. Volkswagen
- 83 02. Volkswagen Audi
- 83 98. Volkswagen, DK make

- 84 01. Volvo

- 97 35. Fiat
- 97 46. Land Rover
- 97 60. Peugeot
- 97 98. Other European manufacturer; European, DK manufacturer

- 98 98. DK manufacturer and make
- 99 99. NA manufacturer and make

- 00 00. Inap.: no first vehicle; no second vehicle; no third vehicle

Codes for type of vehicle (ER13104, ER13134, and ER13164):

1. Car
2. Pickup or truck
3. Van
4. Utility vehicle
5. Motor home

8. DK
9. NA
0. Inap.: no first vehicle; no second vehicle; no third vehicle

For the first time this year, a number of generated variables that will be part of the 1999 Release II file and that have traditionally been part of 1993 and earlier Release II files are included.

One such group is location data. We include ER13004 and ER13005, PSID/GSA and FIPS state codes. We also include the PSID's traditional variables for region, Beale rural-urban code, and size of the largest city in the county (ER16430-ER16432, 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 1993 codebook for V23327, V23330 and V21605.

We also offer a marital status group: marital status of Head (ER13021), the generated form of marital status comparable with years prior to 1977 (ER16423), change in marital status of Head between waves (ER16424), and couple status of Head (ER16425). For documentation, see the 1993 codebook for variables V22412, V23336, V23337, and V23338.

The variable indicating whether a PSID family is living in institutional housing is present (ER13017, equivalent to 1993 V21609), as is the variable indicating the total number of data records from the cross-year individual file that are associated with a panel family (ER16422, equivalent to 1993 V23321).

Both USDA and Census needs standards have been generated for the prior calendar year, 1998 (ER16426 and ER14627; see 1993 V23325 and V23326). Additionally, since the PSID has switched to biennial interviewing, comparable needs standards have also been generated for the "off" year, 1997 (ER14628 and ER14629).

Two traditional PSID family-level generated variables concerning splitoffs are, for a main family, the number of splitoffs arising from it and, for a splitoff, the family interview number of the main family whence it split (ER16433 and ER16434, the same as 1993 variables V23342 and V23343).

The PSID documentation for 1993 and earlier waves has information about multiple PSID families sharing the same household. See Section I, Part 5, of the front matter in any recent PSID Public Release II documentation for specifics about this phenomenon. The PSID produces sets of variables about these sharing families: family ID numbers, relationships, and sizes of up to four other PSID families sharing the HU (ER16435-ER16446, comparable to 1993 V23344-V23355), the household ID number (ER16447, the 1999 version of 1993 V23356), and the number of persons not included in any PSID family who are sharing the household (ER13015, comparable to 1993 V22411).