

PSID File Structure and Merging PSID Data Files

07/05/2024

This document is prepared to assist users in merging various PSID files to create analytical extract. Please note that the PSID Data Center automatically merges PSID, CDS and TAS data, taking care of many the merges described below. However, we continue to provide this document because it can help users understand the structure of the PSID data.

This information is presented in four separate sections:

1. PSID file structure,
2. Two methods of assembling a cross-year family-individual file,
3. Assembling a cross-year family file
4. Single-year family files and single-year family-individual files.

1 PSID File Structure

The traditional cross-year family-individual file used for the PSID through 1989 has been replaced by separate single-year family files and a cross-year individual file. For instance, through the 2021 data collection year there are 42 single-year family files containing family-level variables collected in each wave of the study from 1968 through 2021 and a single cross-year individual file containing all individual-level variables collected from 1968 to 2021 for both respondents and non-respondents. Thus the "main" PSID data files include two types of data files – a) single-year family files and b) a cross-year individual file.

1.1 The single-year family files

Each single-year family file contains one record for each family interviewed in the specified year. The thirty-nine single-year family files (one for each year of the study from 1968 through 2021) contain all of the family-level variables collected in each wave. The records in each file are identified by the family Interview Number for that year, in sort order by that variable, and contain the family-level variables for that year.

Annual family files – contain family-level data collected in a single wave

```
+-----+
|68fam|
+-----+
Format:      family data 1968
Records:     one record for each family in 1968
Ids:         1968 family Interview Number
Sort order:  1968 family Interview Number
N:           4,802 families
```

```

+-----+
|69fam|
+-----+
Format:      family data 1969
Records:     one record for each family in 1969
Ids:         1969 family Interview Number
Sort order:  1969 family Interview Number
N:           4,460 families
.
.
+-----+
|21fam|
+-----+
Format:      family data 2021
Records:     one record for each family in 2021
Ids:         2021 family Interview Number
Sort order:  2021 family Interview Number
N:           9,207 families

```

1.2 The cross-year individual file

The cross-year individual file contains one record for each person ever in a PSID family from the beginning of the study through the current year. The records in the cross-year individual file are identified by 1968 family Interview Number (ER30001) and Person Number (ER30002) and are in sort order by these variables. The file also contains the Interview Number of the family with which the person was associated in each year after 1968 and all other individual-level variables from 1968 through 2021.

1968-2021 Cross-Year Individual File – Contains All Individual-Level Data Collected From 1968-2021

```

+-----+ +-----+ +-----+
|sortid's| |68ind|69ind|...|21ind|
+-----+ +-----+ +-----+
Format:      individual data for 1968-2021
Records:     one record for each person ever-in through 2021
Ids:         1968 family Interview Number and Person Number
Sort order:  1968 family Interview Number and Person Number
N:           84,121 persons

```

2 Assembling a Cross-Year Family-Individual File

Few analysts will want to analyze the full data file for all persons ever in the study, and so your first step is to decide which variables, individuals and years of data interest you.

The basic principle in merging data from a single-year family file with data from the cross-year individual file involves matching the two files using annual Interview Numbers for the year in which the family variables were collected. Thus it is critical that the annual Interview Number variables be retained as part of any subset of data, either family or individual. The chart below shows the family Interview Number variables for the single-year family files and cross-year individual file.

Family interview numbers in single-year family files and in cross-year individual file

Year File	Family File	Individual
1968	V3	ER30001
1969	V442	ER30020
1970	V1102	ER30043
1971	V1802	ER30067
1972	V2402	ER30091
1973	V3002	ER30117
1974	V3402	ER30138
1975	V3802	ER30160
1976	V4302	ER30188
1977	V5202	ER30217
1978	V5702	ER30246
1979	V6302	ER30283
1980	V6902	ER30313
1981	V7502	ER30343
1982	V8202	ER30373
1983	V8802	ER30399
1984	V10002	ER30429
1985	V11102	ER30463
1986	V12502	ER30498
1987	V13702	ER30535
1988	V14802	ER30570
1989	V16302	ER30606
1990	V17702	ER30642
1991	V19002	ER30689
1992	V20302	ER30733
1993	V21602	ER30806
1994	ER2002	ER33101
1995	ER5002	ER33201
1996	ER7002	ER33301
1997	ER10002	ER33401
1999	ER13002	ER33501
2001	ER17002	ER33601
2003	ER21002	ER33701
2005	ER25002	ER33801
2007	ER36002	ER33901
2009	ER42002	ER34001
2011	ER47302	ER34101
2013	ER53002	ER34201
2015	ER60002	ER34301
2017	ER66002	ER34501
2019	ER72002	ER34701
2021	ER78002	ER34901

Note that not each record in the cross-year individual file will have a matching record in every single-year family file. This happens when an individual who was once part of a responding family moves away or dies and is no longer associated with a family in the study; the person is said to be non-response. The non-response person's Interview Number in the cross-year individual file is filled with 0s (as are the other variables) for years in which no data were collected about him or her.

When merging the cross-year individual file with a single-year family file, both SPSS and SAS will fill in system missing values for the family variables for individuals who were not associated with a responding family. Depending on your particular analysis needs, you may or may not wish to include individuals with missing family-year records. Provide appropriate instructions to the programs you use for merging to include or exclude individuals with missing family-year records.

We can think of several approaches to creating a cross-year family-individual file from the components. Two are described and illustrated below. SAS and SPSS statements provided in the SAS and SPSS sub-directories can be used to help construct the programs.

2.1 Method 1 - Merge Using Family Data Added Sequentially To Cross-Year Individual Data

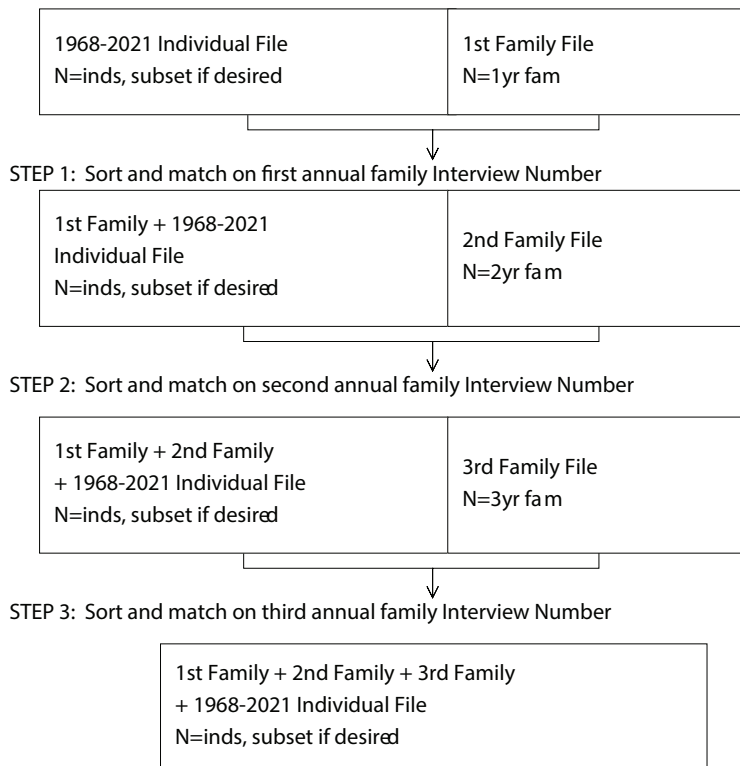
First select individuals and variables from the cross-year individual file (remembering to retain all relevant annual family Interview Number variables) and then match that data with the desired variables from a single-year family file, matching on the appropriate annual family Interview Number variable, using a one-to-many match.

Next, match the resulting file (which now contains one record for each individual with selected variables from the cross-year individual file and the first family file) with a second family file matching on the appropriate annual family Interview Number variable, using a one-to-many match.

Repeat with additional single-year family files until all required family data are obtained and merged with the cross-year individual data, as the diagram below shows.

See [SAS](#), [STATA](#) or [SPSS](#) example for an illustration of this approach using three years of family data.

Merge Using Family Data Added Sequentially To Cross-Year Individual Data



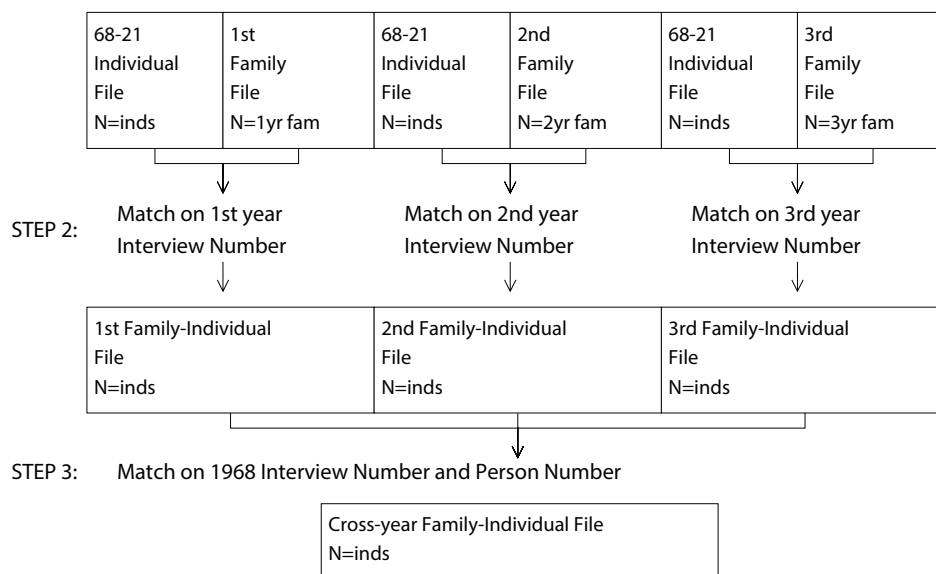
2.2 Method 2 - Merge Using Multiple Family-Individual Files

Alternatively, you could do a series of one-to-many matches of the single-year family files and the cross-year individual file matching on the appropriate annual family Interview Number and then merge the resulting single-year family-individual files in a one-to-one match using the 1968 Interview Number and Person Number. Detailed steps are noted below.

- Step1: Subset annual family Interview Number and other selected variables and select cases from cross-year individual file.
- Step2a: Subset selected variables from the year-n family file.
- Step2b: Sort subsetting year-n family file from Step 2a by year-n family Interview Number.
- Step2c: Sort subsetting cross-year individual file from Step 1 by year-n family Interview Number.
- Step2d: Merge sorted cross-year individual file from Step 2c with sorted year-n subsetting family file from 2b (a one-to-many, family-to-individual, match) matching on the year-n family Interview Number.
- Step2e: Sort resulting year-n family-individual file from Step 2d by the individual identifiers, 68 family Interview Number (ER30001) and Person Number (ER30002).
- ... Repeat Steps 2a-2e for all other years.
- Step3: Merge family-individual files from Step 2e by the individual identifiers, 68 family Interview Number (ER30001) and Person Number (ER30002).

See [SAS](#), [STATA](#) or [SPSS](#) example for an illustration of this approach using 41 waves of family data.

Illustration Of Merge Using Multiple Family-Individual Files



3 Assembling A Cross-Year Family File

To assemble a 1992 cross-year family file from these files, a procedure similar to one of the above would be followed, but only the cross-year individual records of the 1992 head would be selected from the cross-year individual file. Merge data from the single-year family files using the annual family Interview Number variables to match as described in Method 1 or Method 2 above to create a merged 1968-2021 family-level file for currently responding families.

Each member of a family has a family Interview Number for each wave with a value identical to the values of that data item for all the other family members in that family that year. In addition, except in 1968, each individual is annually assigned a unique sequence number, which indicates the person's position and status for any given year's list of family members. Thus, the first person listed, always the Head of the family, is 01, the second person listed is 02, and so on.

To create a 1992 cross-year family-level file, select from the cross-year individual file those cases where ER30734 (1992 Sequence Number) is equal to 01, since each family must have at least one member, although it may or may not have more.¹

To create other years' cross-year family-level files, the Sequence Number variable for the latest desired year of data should be used and merges done with the appropriate single-year family files. Again, this produces a file of families who were response through the latest year and eliminates families who had already become nonresponding.

4 Single-Year Family Files And Single-Year Family-Individual Files

Producing single-year family files for cross-sectional analysis is simplicity itself. Simply use the single-year file.

Single-year family-individual files are also relatively simple. Select all individuals whose Sequence Number for the desired year is non-zero (for 1968, use ER30003, Relationship to Head, instead) and match the family Interview Number for that year from the individual file with the family Interview Number from the corresponding family file. The family Interview Numbers in the family and individual files are listed in a table in Section "2 Assembling a Cross-Year Family-Individual File", above.

5 Additional Help

If you have any remaining questions, please [contact us](#).

¹Variable ER30734, Sequence Number, should be used instead of ER30735, Relationship to Head, because although each family has one and only one current Head (i.e., where ER30734 = 01-20 and ER30735 = 10), it is possible that the prior year's Head has moved out since the previous interview and a new Head is present for the current interview. Relationship to Head for movers-out is coded with reference to the previous year's Head, so for both the current Head and the previous Head, ER30735 = 10.

There is not an 1968 Sequence Number variable; use ER30003, Relationship to Head, instead. There was only one Head per household in 1968.

A Save Example Programs

Please *right click* file name and select “Save Embedded File to Disk...” to save example source codes (or use *Attachments Panel* in your PDF viewer). If you are unable to save individual files using a web browser, please save FileStructure.pdf and use a pdf viewer on your local computer to right click and save these programs.

1. [SAS Example - Method 1.sas](#)
2. [SAS Example - Method 2.sas](#)
3. [STATA Example - Method 1.do](#)
4. [STATA Example - Method 2.do](#)
5. [SPSS Example - Method 1.sps](#)
6. [SPSS Example - Method 2.sps](#)

B EXAMPLES: Merging PSID Data

B.1 SAS

B.1.1 SAS Example - Method 1

```
*-----*
| This example program demonstrates a relatively simple method for
| merging PSID data. It uses data from 3 different years, subset-
| ting criteria, and the compress and tagsort options.
|
| When working with PSID data, the amount of available system disk
| space and memory is often an important consideration. The follow-
| ing options can be used to enhance system performance and control
| the amount of disk space and memory used:
|
| 1) LENGTH statements specify the number of bytes used for storing
| variables and can significantly reduce the size of a data set.
| The precision of a numeric variable is dependent on its length.
| Length specifications for numeric variables are host-specific.
| Do not shorten length specifications for numeric variables con-
| taining fractions.
|
| 2) The COMPRESS= data set or system option can decrease the size
| of a data set and reduce the number of input/output operations.
|
| 3) The SORTSIZE= option specifies the maximum amount of memory
| available to PROC SORT. Sortsize parameters are host-specific.
|
| 4) The TAGSORT= option can be used with PROC SORT to reduce the
| amount of temporary disk space used. When the total length of
| BY variables is small compared with the record length, temp-
| orary disk space is reduced significantly, however, processing
| time can be much higher.
|
| 5) The BUFNO= data set or system option specifies the number of
| page buffers to use when reading from or writing to a SAS data
| set. Increasing the number of available buffers uses more
| memory while reducing the number of input/output operations.
|-----*
```

```

|
| 6) The BUFSIZE= data set or system option specifies the permanent
| page size for a SAS data set. Increasing the page size uses
| more memory while reducing the number of input/output operat-
| ions. Using a buffer size larger than necessary is inefficient.
|
*-----;
* Read in cross-year individual file and select variables
  from 1990-1992 needed for analysis;
LIBNAME PSID "[FOLDER NAME]";
DATA IND90_92(COMPRESS=YES RENAME=(ER30642=ID90 ER30689=ID91 ER30733=ID92));
ATTRIB
  ER30001 LABEL="1968 INTERVIEW NUMBER" format=f4.
  ER30002 LABEL="PERSON NUMBER" 68" format=f3.
  ER30642 LABEL="1990 INTERVIEW NUMBER" format=f5.
  ER30643 LABEL="SEQUENCE NUMBER" 90" format=f2.
  ER30644 LABEL="RELATION TO HEAD" 90" format=f2.
  ER30645 LABEL="AGE OF INDIVIDUAL" 90" format=f3.
  ER30653 LABEL="EMPLOYMENT STAT-IND" 90" format=f1.
  ER30657 LABEL="COMPLETED EDUC-IND" 90" format=f2.
  ER30659 LABEL="TOT TXBL INCOME-IND" 90" format=f6.
  ER30689 LABEL="1991 INTERVIEW NUMBER" format=f4.
  ER30690 LABEL="SEQUENCE NUMBER" 91" format=f2.
  ER30691 LABEL="RELATION TO HEAD" 91" format=f2.
  ER30692 LABEL="AGE OF INDIVIDUAL" 91" format=f3.
  ER30699 LABEL="EMPLOYMENT STAT-IND" 91" format=f1.
  ER30703 LABEL="COMPLETED EDUC-IND" 91" format=f2.
  ER30705 LABEL="TOT LABOR INCOME-IND" 91" format=f6.
  ER30707 LABEL="TOT ASSET INCOME-IND" 91" format=f6.
  ER30733 LABEL="1992 INTERVIEW NUMBER" format=f4.
  ER30734 LABEL="SEQUENCE NUMBER" 92" format=f2.
  ER30735 LABEL="RELATION TO HEAD" 92" format=f2.
  ER30736 LABEL="AGE OF INDIVIDUAL" 92" format=f3.
  ER30744 LABEL="EMPLOYMENT STAT" 92" format=f1.
  ER30748 LABEL="COMPLETED EDUCATION" 92" format=f2.
  ER30750 LABEL="TOT LABOR INCOME" 92" format=f6.
  ER30752 LABEL="TOT ASSET INCOME" 92" format=f6.
  ER30805 LABEL="COMBINED IND WEIGHT" 92" format=f7.3
  ER32000 LABEL="SEX OF INDIVIDUAL" format=f1.
  ER32022 LABEL="# LIVE BIRTHS TO THIS INDIVIDUAL" format=f2.
  ER32049 LABEL="LAST KNOWN MARITAL STATUS" format=f1.
;
INFILE "[PATH]\IND2021ER.txt" LRECL = 6117 ;
INPUT
  ER30001 2 - 5
  ER30002 6 - 8
  ER30642 1528 - 1532
  ER30643 1533 - 1534
  ER30644 1535 - 1536
  ER30645 1537 - 1539
  ER30653 1555 - 1555
  ER30657 1563 - 1564
  ER30659 1566 - 1571
  ER30689 1647 - 1650
  ER30690 1651 - 1652
  ER30691 1653 - 1654
  ER30692 1655 - 1657
  ER30699 1672 - 1672
  ER30703 1680 - 1681
  ER30705 1683 - 1688
  ER30707 1690 - 1695
  ER30733 1764 - 1767
  ER30734 1768 - 1769
  ER30735 1770 - 1771
  ER30736 1772 - 1774
  ER30744 1790 - 1790

```



```

ER30748 1798 - 1799
ER30750 1801 - 1806
ER30752 1808 - 1813
ER30805 1907 - 1913
ER32000 2057 - 2057
ER32022 2111 - 2112
ER32049 2187 - 2187
;
IF ER30645=999 THEN ER30645=.;
IF ER30657=99 THEN ER30657=.;
IF ER30692=999 THEN ER30692=.;
IF ER30703=99 THEN ER30703=.;
IF ER30736=999 THEN ER30736=.;
IF ER30748=99 THEN ER30748=.;
IF ER32022=98 THEN ER32022=.;
IF ER32049=8 THEN ER32049=.;
* Select those who were ever heads or wives/"wives" between 1990 and 1992;
IF (ER30643 EQ 01 AND ER30644 EQ 10) OR
(ER30643 EQ 02 AND (ER30644 EQ 20 OR ER30644 EQ 22)) OR
(ER30690 EQ 01 AND ER30691 EQ 10) OR
(ER30690 EQ 02 AND (ER30691 EQ 20 OR ER30691 EQ 22)) OR
(ER30734 EQ 01 AND ER30735 EQ 10) OR
(ER30734 EQ 02 AND (ER30735 EQ 20 OR ER30735 EQ 22));
RUN ;
* Read in 1990 family file and select variables needed for analysis;
DATA FAM90 (COMPRESS=YES RENAME=(V17702=ID90));
INFILE "[PATH]\FAM1990.txt" LRECL=2332;
INPUT
V17702 2 - 6
V17836 281 - 286
V18262 1175 - 1177
V18564 1656 - 1658
V18814 2046 - 2046
V18878 2188 - 2193
V18887 2230 - 2234 .2
V18888 2235 - 2239 .2
;
LABEL
V17702="1990 INTERVIEW NUMBER"
V17836="WIFE 89 LABOR/WAGE"
V18262="C9-10 OCC-LAST JOB (H-U)"
V18564="E9-10 OCC-LAST JOB (W-U)"
V18814="M32 RACE OF HEAD (1 MEN)"
V18878="TOTAL HEAD LABOR Y 89"
V18887="HEAD 89 AVG HRLY EARNING"
V18888="WIFE 89 AVG HRLY EARNING";
IF V18262=999 THEN V18262=.;
IF V18564=999 THEN V18564=.;
IF V18814=9 THEN V18814=.;
RUN ;
* Sort fam90 and ind90_92 by id90;
PROC SORT DATA=FAM90 TAGSORT;
BY ID90;
RUN ;
PROC SORT DATA=IND90_92 TAGSORT;
BY ID90;
RUN ;
* Merge fam90 and ind90_92 by id90;
DATA PSID.FAM_IND (COMPRESS=YES);
MERGE FAM90 IND90_92 (IN=IND90_92);
BY ID90;
IF IND90_92;
RUN ;
* Read in 1991 family file and select variables needed for analysis;
DATA FAM91 (COMPRESS=YES RENAME=(V19002=ID91));

```

```

INFILE "[PATH]\FAM1991.txt" LRECL = 2336 ;
INPUT
  V19002      2 - 5
  V19136     281 - 286
  V19562    1175 - 1177
  V19864    1656 - 1658
  V20114    2046
  V20178    2188 - 2193
  V20187    2230 - 2234 .2
  V20188    2235 - 2239 .2
;
LABEL
  V19002="1991 INTERVIEW NUMBER"
  V19136="WIFE 90 LABOR/WAGE"
  V19562="C9-10 OCC-LAST JOB (H-U)"
  V19864="E9-10 OCC-LAST JOB (W-U)"
  V20114="L32 RACE OF HEAD (1 MEN)"
  V20178="TOTAL HEAD LABOR Y 90"
  V20187="HEAD 90 AVG HRLY EARNING"
  V20188="WIFE 90 AVG HRLY EARNING";
  IF V19562=999 THEN V19562=.;
  IF V19864=999 THEN V19864=.;
  IF V20114=9 THEN V20114=.;
RUN ;
* Sort fam91 and psid.fam_ind by id91;
PROC SORT DATA=FAM91 TAGSORT;
BY ID91;
RUN ;
PROC SORT DATA=PSID.FAM_IND TAGSORT;
BY ID91;
RUN ;
* Merge fam91 and psid.fam_ind by id91;
DATA PSID.FAM_IND(COMPRESS=YES);
MERGE FAM91 PSID.FAM_IND(IN=FAM_IND);
BY ID91;
IF FAM_IND;
RUN ;
* Read in 1992 family file and select variables needed for analysis;
DATA FAM92(COMPRESS=YES RENAME=(V20302=ID92));
INFILE "[PATH]\FAM1992.txt" LRECL = 2325 ;
INPUT
  V20302      2 - 5
  V20436     282 - 287
  V20862    1189 - 1191
  V21164    1670 - 1672
  V21420    2066
  V21484    2172 - 2177
  V21493    2218 - 2222 .2
  V21494    2223 - 2227 .2
;
LABEL
  V20302="1992 INTERVIEW NUMBER"
  V20436="WIFE 91 LABOR/WAGE"
  V20862="C9-10 OCC-LAST JOB (H-U)"
  V21164="E9-10 OCC-LAST JOB (W-U)"
  V21420="M32 RACE OF HEAD (1 MEN)"
  V21484="TOTAL HEAD LABOR Y 91"
  V21493="HEAD 91 AVG HRLY EARNING"
  V21494="WIFE 91 AVG HRLY EARNING";
  IF V20862=999 THEN V20862=.;
  IF V21164=999 THEN V21164=.;
  IF V21420=9 THEN V21420=.;
RUN ;
* Sort fam92 and psid.fam_ind by id92;
PROC SORT DATA=FAM92 TAGSORT;

```

```
BY ID92;
RUN ;
PROC SORT DATA=PSID.FAM_IND TAGSORT;
BY ID92;
RUN ;
* Merge fam92 and psid.fam_ind by id92;
DATA PSID.FAM_IND(COMPRESS=YES);
MERGE FAM92 PSID.FAM_IND(IN=FAM_IND);
BY ID92;
IF FAM_IND;
RUN;
```

B.1.2 SAS Example - Method 2

```
*-----*
|           MERGE USING MULTIPLE FAMILY-INDIVIDUAL FILES           |
| Step 1:   Subset family-level id's and selected variables       |
|           and select cases from xyr-individual file           |
| Step 2a:  Subset year-n family file                             |
| Step 2b:  Sort year-n family file from 2a by year-n family id  |
| Step 2c:  Sort xyr-individual file from step 1 by year-n family id |
| Step 2d:  Merge sorted xyr-individual file from 2c             |
|           with sorted year-n subsetted family file from 2b    |
|           (a one-to-many [family-to-individual] match)        |
| Step 2e:  Sort resulting year-n family-individual file from 2d |
|           by individual ids                                     |
| ...       Repeat Steps 2 for all other years                   |
| Step 3:   Merge family-individual files from step 2e          |
*-----*
LIBNAME mrg '[Add Your Path Here]';
/* ===== */
/* step 1: subset family-level id's and individual variables */
/* and select cases from cross-year individual record */
/* ===== */
DATA mrg.indvars (COMPRESS=YES);
ATTRIB
ER30001 LABEL="1968 INTERVIEW NUMBER" format=f4.
ER30002 LABEL="PERSON NUMBER 68" format=f3.
ER30020 LABEL="1969 INTERVIEW NUMBER" format=f4.
ER30043 LABEL="1970 INTERVIEW NUMBER" format=f4.
ER30067 LABEL="1971 INTERVIEW NUMBER" format=f4.
ER30091 LABEL="1972 INTERVIEW NUMBER" format=f4.
ER30117 LABEL="1973 INTERVIEW NUMBER" format=f4.
ER30138 LABEL="1974 INTERVIEW NUMBER" format=f4.
ER30160 LABEL="1975 INTERVIEW NUMBER" format=f4.
ER30188 LABEL="1976 INTERVIEW NUMBER" format=f4.
ER30217 LABEL="1977 INTERVIEW NUMBER" format=f4.
ER30246 LABEL="1978 INTERVIEW NUMBER" format=f4.
ER30283 LABEL="1979 INTERVIEW NUMBER" format=f4.
ER30313 LABEL="1980 INTERVIEW NUMBER" format=f4.
ER30343 LABEL="1981 INTERVIEW NUMBER" format=f4.
ER30373 LABEL="1982 INTERVIEW NUMBER" format=f4.
ER30399 LABEL="1983 INTERVIEW NUMBER" format=f4.
ER30429 LABEL="1984 INTERVIEW NUMBER" format=f4.
ER30463 LABEL="1985 INTERVIEW NUMBER" format=f4.
ER30498 LABEL="1986 INTERVIEW NUMBER" format=f4.
ER30535 LABEL="1987 INTERVIEW NUMBER" format=f4.
ER30570 LABEL="1988 INTERVIEW NUMBER" format=f4.
ER30606 LABEL="1989 INTERVIEW NUMBER" format=f4.
ER30642 LABEL="1990 INTERVIEW NUMBER" format=f5.
ER30689 LABEL="1991 INTERVIEW NUMBER" format=f4.
ER30733 LABEL="1992 INTERVIEW NUMBER" format=f4.
ER30806 LABEL="1993 INTERVIEW NUMBER" format=f5.
ER33101 LABEL="1994 INTERVIEW NUMBER" format=f5.
ER33201 LABEL="1995 INTERVIEW NUMBER" format=f5.
ER33301 LABEL="1996 INTERVIEW NUMBER" format=f4.
ER33401 LABEL="1997 INTERVIEW NUMBER" format=f5.
ER33501 LABEL="1999 INTERVIEW NUMBER" format=f5.
ER33601 LABEL="2001 INTERVIEW NUMBER" format=f4.
ER33701 LABEL="2003 INTERVIEW NUMBER" format=f5.
ER33801 LABEL="2005 INTERVIEW NUMBER" format=f5.
ER33901 LABEL="2007 INTERVIEW NUMBER" format=f5.
ER34001 LABEL="2009 INTERVIEW NUMBER" format=f5.
```

```

ER34101 LABEL="2011 INTERVIEW NUMBER" format=f5.
ER34201 LABEL="2013 INTERVIEW NUMBER" format=f5.
ER34301 LABEL="2015 INTERVIEW NUMBER" format=f5.
ER34501 LABEL="2017 INTERVIEW NUMBER" format=f5.
ER34701 LABEL="2019 INTERVIEW NUMBER" format=f5.
ER34901 LABEL="2021 INTERVIEW NUMBER" format=f5.
/* your individual variable labels here */
;
INFILE "[PATH]\IND2021ER.txt" LRECL = 6117 ;
INPUT
/* 1968 ID */ ER30001 2 - 5
/* 1968 PN */ ER30002 6 - 8
/* 1969 */ ER30020 44 - 47
/* 1970 */ ER30043 97 - 100
/* 1971 */ ER30067 152 - 155
/* 1972 */ ER30091 207 - 210
/* 1973 */ ER30117 265 - 268
/* 1974 */ ER30138 317 - 320
/* 1975 */ ER30160 370 - 373
/* 1976 */ ER30188 436 - 439
/* 1977 */ ER30217 503 - 506
/* 1978 */ ER30246 571 - 574
/* 1979 */ ER30283 648 - 651
/* 1980 */ ER30313 718 - 721
/* 1981 */ ER30343 788 - 791
/* 1982 */ ER30373 858 - 861
/* 1983 */ ER30399 919 - 922
/* 1984 */ ER30429 992 - 995
/* 1985 */ ER30463 1077 - 1080
/* 1986 */ ER30498 1167 - 1170
/* 1987 */ ER30535 1259 - 1262
/* 1988 */ ER30570 1348 - 1351
/* 1989 */ ER30606 1438 - 1441
/* 1990 */ ER30642 1528 - 1532
/* 1991 */ ER30689 1647 - 1650
/* 1992 */ ER30733 1764 - 1767
/* 1993 */ ER30806 1914 - 1918
/* 1994 */ ER33101 2201 - 2205
/* 1995 */ ER33201 2300 - 2304
/* 1996 */ ER33301 2493 - 2496
/* 1997 */ ER33401 2577 - 2581
/* 1999 */ ER33501 2662 - 2666
/* 2001 */ ER33601 2856 - 2859
/* 2003 */ ER33701 3038 - 3042
/* 2005 */ ER33801 3193 - 3197
/* 2007 */ ER33901 3492 - 3496
/* 2009 */ ER34001 3770 - 3774
/* 2011 */ ER34101 3992 - 3996
/* 2013 */ ER34201 4229 - 4233
/* 2015 */ ER34301 4495 - 4499
/* 2017 */ ER34501 4869 - 4873
/* 2019 */ ER34701 5249 - 5253
/* 2021 */ ER34901 5683 - 5687

/* your individual variable locations here */
;
/* your individual missing data here */
;
/* your filter for case selection, if any, here
this selects the first 10 cases for testing setup */
/* if _n_ gt 10 then stop; */
RUN;
/* ===== */
/* step 2 for 1968 family file (n=4802) */
/* ===== */

```

```

DATA fam68; INFILE "[PATH]\FAM1968.txt" LRECL=771;
INPUT
  V3 7-10
  /*      your 1968 variable locations here */
;
LABEL
  V3 ="INTERVIEW NUMBER 68 1:6-9"
  /*      your 1968 variable labels here      */
;
  /*      your 1968 missing data here      */
RUN ;
PROC SORT DATA=fam68 TAGSORT; BY V3; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER30001; RUN;
DATA mrg.famind68(COMPRESS=YES);
  MERGE fam68(rename=V3=ER30001) mrg.indvars(IN=INDOK);
  BY ER30001; IF INDOK;
RUN;
PROC SORT DATA=mrg.famind68 TAGSORT; BY ER30001 ER30002;
RUN ;
/* ===== */
/*      step 2 for 1969 family file (n=4460)      */
/* ===== */
DATA fam69; INFILE "[PATH]\FAM1969.txt" LRECL=1079;
INPUT
  V442 2-5
  /*      your 1969 variable locations here */
;
LABEL
  V442 ="1969 INT NUMBER 11:6-9"
  /*      your 1969 variable labels here      */
;
  /*      your 1969 missing data here      */
RUN ;
PROC SORT DATA=fam69 TAGSORT; BY V442; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER30020; RUN;
DATA mrg.famind69(COMPRESS=YES);
  MERGE fam69(rename=V442=ER30020) mrg.indvars(IN=INDOK);
  BY ER30020; IF INDOK;
RUN;
PROC SORT DATA=mrg.famind69 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/*      step 2 for 1970 family file (n=4645)      */
/* ===== */
DATA fam70; INFILE "[PATH]\FAM1970.txt" LRECL=1199;
INPUT
  V1102 2-5
  /*      your 1970 variable locations here */
;
LABEL
  V1102 ="1970 INT # 21:6-9"
  /*      your 1970 variable labels here      */
;
  /*      your 1970 missing data here      */
RUN ;
PROC SORT DATA=fam70 TAGSORT; BY V1102; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER30043; RUN;
DATA mrg.famind70(COMPRESS=YES);
  MERGE fam70(rename=V1102=ER30043) mrg.indvars(IN=INDOK);
  BY ER30043; IF INDOK;
RUN;
PROC SORT DATA=mrg.famind70 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/*      step 2 for 1971 family file (n=4840)      */
/* ===== */

```

```

/* ===== */
DATA fam71; INFILE "[PATH]\FAM1971.txt" LRECL=974;
INPUT
  V1802 2-5
  /*      your 1971 variable locations here */
;
LABEL
  V1802 ="71 ID NO."
  /*      your 1971 variable labels here */
;
  /*      your 1971 missing data here */
RUN ;
PROC SORT DATA=fam71 TAGSORT; BY V1802; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER30067; RUN;
DATA mrg.famind71(COMPRESS=YES);
  MERGE fam71(rename=V1802=ER30067) mrg.indvars(IN=INDOK);
  BY ER30067; IF INDOK;
RUN;
PROC SORT DATA=mrg.famind71 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/*      step 2 for 1972 family file (n=5060) */
/* ===== */
DATA fam72; INFILE "[PATH]\FAM1972.txt" LRECL=1020;
INPUT
  V2402 2-5
  /*      your 1972 variable locations here */
;
LABEL
  V2402 ="1972 INT #          46:6-9"
  /*      your 1972 variable labels here */
;
  /*      your 1972 missing data here */
RUN ;
PROC SORT DATA=fam72 TAGSORT; BY V2402; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER30091; RUN;
DATA mrg.famind72(COMPRESS=YES);
  MERGE fam72(rename=V2402=ER30091) mrg.indvars(IN=INDOK);
  BY ER30091; IF INDOK;
RUN;
PROC SORT DATA=mrg.famind72 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/*      step 2 for 1973 family file (n=5285) */
/* ===== */
DATA fam73; INFILE "[PATH]\FAM1973.txt" LRECL=582;
INPUT
  V3002 2-5
  /*      your 1973 variable locations here */
;
LABEL
  V3002 ="1973 INT #          59:6-9"
  /*      your 1973 variable labels here */
;
  /*      your 1973 missing data here */
RUN ;
PROC SORT DATA=fam73 TAGSORT; BY V3002; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER30117; RUN;
DATA mrg.famind73(COMPRESS=YES);
  MERGE fam73(rename=V3002=ER30117) mrg.indvars(IN=INDOK);
  BY ER30117; IF INDOK;
RUN;
PROC SORT DATA=mrg.famind73 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */

```

```

/*          step 2 for 1974 family file (n=5517)          */
/* ===== */
DATA fam74; INFILE "[PATH]\FAM1974.txt"    LRECL=633;
INPUT
  V3402 2-5
  /*          your 1974 variable locations here          */
;
LABEL
  V3402 ="1974 ID NUMBER"
  /*          your 1974 variable labels here          */
;
  /*          your 1974 missing data here          */
RUN ;
PROC SORT DATA=fam74 TAGSORT; BY V3402; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER30138; RUN;
DATA mrg.famind74(COMPRESS=YES);
  MERGE fam74(rename=V3402=ER30138) mrg.indvars(IN=INDOK);
  BY ER30138; IF INDOK;
RUN;
PROC SORT DATA=mrg.famind74 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/*          step 2 for 1975 family file (n=5725)          */
/* ===== */
DATA fam75; INFILE "[PATH]\FAM1975.txt"    LRECL=777;
INPUT
  V3802 2-5
  /*          your 1975 variable locations here          */
;
LABEL
  V3802 ="1975 INT #          80:6-9"
  /*          your 1975 variable labels here          */
;
  /*          your 1975 missing data here          */
RUN ;
PROC SORT DATA=fam75 TAGSORT; BY V3802; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER30160; RUN;
DATA mrg.famind75(COMPRESS=YES);
  MERGE fam75(rename=V3802=ER30160) mrg.indvars(IN=INDOK);
  BY ER30160; IF INDOK;
RUN;
PROC SORT DATA=mrg.famind75 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/*          step 2 for 1976 family file (n=5862)          */
/* ===== */
DATA fam76; INFILE "[PATH]\FAM1976.txt"    LRECL=1499;
INPUT
  V4302 2-5
  /*          your 1976 variable locations here          */
;
LABEL
  V4302 ="1976 ID NUMBER          6V2"
  /*          your 1976 variable labels here          */
;
  /*          your 1976 missing data here          */
RUN ;
PROC SORT DATA=fam76 TAGSORT; BY V4302; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER30188; RUN;
DATA mrg.famind76(COMPRESS=YES);
  MERGE fam76(rename=V4302=ER30188) mrg.indvars(IN=INDOK);
  BY ER30188; IF INDOK;
RUN;
PROC SORT DATA=mrg.famind76 TAGSORT; BY ER30001 ER30002;
RUN;

```



```

/* ===== */
/*      step 2 for 1977 family file (n=6007)      */
/* ===== */
DATA fam77; INFILE "[PATH]\FAM1977.txt"      LRECL=963;
INPUT
  V5202 2-5
;
LABEL
  V5202 ="1977 ID"
  /*      your 1977 variable labels here      */
;
  /*      your 1977 missing data here      */
RUN ;
PROC SORT DATA=fam77 TAGSORT; BY V5202; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER30217; RUN;
DATA mrg.famind77(COMPRESS=YES);
  MERGE fam77(rename=V5202=ER30217) mrg.indvars(IN=INDOK);
  BY ER30217; IF INDOK;
RUN;
PROC SORT DATA=mrg.famind77 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/*      step 2 for 1978 family file (n=6154)      */
/* ===== */
DATA fam78; INFILE "[PATH]\FAM1978.txt"      LRECL=988;
INPUT
  V5702 2-5
  /*      your 1978 variable locations here      */
;
LABEL
  V5702 ="1978 ID"
  /*      your 1978 variable labels here      */
;
  /*      your 1978 missing data here      */
RUN ;
PROC SORT DATA=fam78 TAGSORT; BY V5702; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER30246; RUN;
DATA mrg.famind78(COMPRESS=YES);
  MERGE fam78(rename=V5702=ER30246) mrg.indvars(IN=INDOK);
  BY ER30246; IF INDOK;
RUN;
PROC SORT DATA=mrg.famind78 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/*      step 2 for 1979 family file (n=6373)      */
/* ===== */
DATA fam79; INFILE "[PATH]\FAM1979.txt"      LRECL=1062;
INPUT
  V6302 2-5
  /*      your 1979 variable locations here      */
;
LABEL
  V6302 ="1979 ID"
  /*      your 1979 variable labels here      */
;
  /*      your 1979 missing data here      */
RUN ;
PROC SORT DATA=fam79 TAGSORT; BY V6302; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER30283; RUN;
DATA mrg.famind79(COMPRESS=YES);
  MERGE fam79(rename=V6302=ER30283) mrg.indvars(IN=INDOK);
  BY ER30283; IF INDOK;
RUN;
PROC SORT DATA=mrg.famind79 TAGSORT; BY ER30001 ER30002;
RUN;

```

```

/* ===== */
/*      step 2 for 1980 family file (n=6533)      */
/* ===== */
DATA fam80; INFILE "[PATH]\FAM1980.txt"    LRECL=1123;
INPUT
  V6902 2-5
  /*      your 1980 variable locations here      */
;
LABEL
  V6902 ="1980 INTERVIEW NUMBER"
  /*      your 1980 variable labels here      */
;
  /*      your 1980 missing data here      */
RUN ;
PROC SORT DATA=fam80 TAGSORT; BY V6902; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER30313; RUN;
DATA mrg.famind80(COMPRESS=YES);
  MERGE fam80(rename=V6902=ER30313) mrg.indvars(IN=INDOK);
  BY ER30313; IF INDOK;
RUN;
PROC SORT DATA=mrg.famind80 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/*      step 2 for 1981 family file (n=6620)      */
/* ===== */
DATA fam81; INFILE "[PATH]\FAM1981.txt"    LRECL=1250;
INPUT
  V7502 2-5
  /*      your 1981 variable locations here      */
;
LABEL
  V7502 ="1981 INTERVIEW NUMBER"
  /*      your 1981 variable labels here      */
;
  /*      your 1981 missing data here      */
RUN ;
PROC SORT DATA=fam81 TAGSORT; BY V7502; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER30343; RUN;
DATA mrg.famind81(COMPRESS=YES);
  MERGE fam81(rename=V7502=ER30343) mrg.indvars(IN=INDOK);
  BY ER30343; IF INDOK;
RUN;
PROC SORT DATA=mrg.famind81 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/*      step 2 for 1982 family file (n=6742)      */
/* ===== */
DATA fam82; INFILE "[PATH]\FAM1982.txt"    LRECL=1117;
INPUT
  V8202 2-5
  /*      your 1982 variable locations here      */
;
LABEL
  V8202 ="1982 INTERVIEW NUMBER"
  /*      your 1982 variable labels here      */
;
  /*      your 1982 missing data here      */
RUN ;
PROC SORT DATA=fam82 TAGSORT; BY V8202; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER30373; RUN;
DATA mrg.famind82(COMPRESS=YES);
  MERGE fam82(rename=V8202=ER30373) mrg.indvars(IN=INDOK);
  BY ER30373; IF INDOK;
RUN;
PROC SORT DATA=mrg.famind82 TAGSORT; BY ER30001 ER30002;

```

```

RUN;
/* ===== */
/*      step 2 for 1983 family file (n=6852)      */
/* ===== */
DATA fam83; INFILE "[PATH]\FAM1983.txt"    LRECL=1315;
INPUT
  V8802 2-5
  /*      your 1983 variable locations here      */
;
LABEL
  V8802 ="1983 INTERVIEW NUMBER"
  /*      your 1983 variable labels here      */
;
  /*      your 1983 missing data here      */
/* ===== */
RUN;
PROC SORT DATA=fam83 TAGSORT; BY V8802; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER30399; RUN;
DATA mrg.famind83(COMPRESS=YES);
MERGE fam83(rename=V8802=ER30399) mrg.indvars(IN=INDOK);
BY ER30399; IF INDOK;
RUN;
PROC SORT DATA=mrg.famind83 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/*      step 2 for 1984 family file (n=6918)      */
/* ===== */
DATA fam84; INFILE "[PATH]\FAM1984.txt"    LRECL=2009;
INPUT
  V10002 2-5
  /*      your 1984 variable locations here      */
;
LABEL
  V10002 ="1984 INTERVIEW NUMBER"
  /*      your 1984 variable labels here      */
;
  /*      your 1984 missing data here      */
/* ===== */
RUN;
PROC SORT DATA=fam84 TAGSORT; BY V10002; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER30429; RUN;
DATA mrg.famind84(COMPRESS=YES);
MERGE fam84(rename=V10002=ER30429) mrg.indvars(IN=INDOK);
BY ER30429; IF INDOK;
RUN;
PROC SORT DATA=mrg.famind84 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/*      step 2 for 1985 family file (n=7032)      */
/* ===== */
DATA fam85; INFILE "[PATH]\FAM1985.txt"    LRECL=2472;
INPUT
  V11102 2-5
  /*      your 1985 variable locations here      */
;
LABEL
  V11102 ="1985 INTERVIEW NUMBER"
  /*      your 1985 variable labels here      */
;
  /*      your 1985 missing data here      */
/* ===== */
RUN;
PROC SORT DATA=fam85 TAGSORT; BY V11102; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER30463; RUN;
DATA mrg.famind85(COMPRESS=YES);
MERGE fam85(rename=V11102=ER30463) mrg.indvars(IN=INDOK);
BY ER30463; IF INDOK;
RUN;

```

```

PROC SORT DATA=mrg.famind85 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/*      step 2 for 1986 family file (n=7018)      */
/* ===== */
DATA fam86; INFILE "[PATH]\FAM1986.txt"      LRECL=2234;
  INPUT
    V12502 2-5
    /*      your 1986 variable locations here      */
  ;
  LABEL
    V12502 ="1986 INTERVIEW NUMBER"
    /*      your 1986 variable labels here      */
  ;
    /*      your 1986 missing data here      */
RUN ;
PROC SORT DATA=fam86 TAGSORT; BY V12502; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER30498; RUN;
DATA mrg.famind86(COMPRESS=YES);
  MERGE fam86(rename=V12502=ER30498) mrg.indvars(IN=INDOK);
  BY ER30498; IF INDOK;
RUN;
PROC SORT DATA=mrg.famind86 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/*      step 2 for 1987 family file (n=7061)      */
/* ===== */
DATA fam87; INFILE "[PATH]\FAM1987.txt"      LRECL=2053;
  INPUT
    V13702 2-5
    /*      your 1987 variable locations here      */
  ;
  LABEL
    V13702 ="1987 INTERVIEW NUMBER"
    /*      your 1987 variable labels here      */
  ;
    /*      your 1987 missing data here      */
RUN ;
PROC SORT DATA=fam87 TAGSORT; BY V13702; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER30535; RUN;
DATA mrg.famind87(COMPRESS=YES);
  MERGE fam87(rename=V13702=ER30535) mrg.indvars(IN=INDOK);
  BY ER30535; IF INDOK;
RUN;
PROC SORT DATA=mrg.famind87 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/*      step 2 for 1988 family file (n=7114)      */
/* ===== */
DATA fam88; INFILE "[PATH]\FAM1988.txt"      LRECL=2730;
  INPUT
    V14802 2-5
    /*      your 1988 variable locations here      */
  ;
  LABEL
    V14802 ="1988 INTERVIEW NUMBER"
    /*      your 1988 variable labels here      */
  ;
    /*      your 1988 missing data here      */
RUN ;
PROC SORT DATA=fam88 TAGSORT; BY V14802; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER30570; RUN;
DATA mrg.famind88(COMPRESS=YES);
  MERGE fam88(rename=V14802=ER30570) mrg.indvars(IN=INDOK);
  BY ER30570; IF INDOK;

```

```

RUN;
PROC SORT DATA=mrg.famind88 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/*      step 2 for 1989 family file (n=7114)      */
/* ===== */
DATA fam89; INFILE "[PATH]\FAM1989.txt"    LRECL=2517;
INPUT
  V16302 2-5
  /*      your 1989 variable locations here      */
;
LABEL
  V16302 ="1989 INTERVIEW NUMBER"
  /*      your 1989 variable labels here      */
;
  /*      your 1989 missing data here      */
;
RUN;
PROC SORT DATA=fam89 TAGSORT; BY V16302; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER30606; RUN;
DATA mrg.famind89(COMPRESS=YES);
  MERGE fam89(rename=V16302=ER30606) mrg.indvars(IN=INDOK);
  BY ER30606; IF INDOK;
RUN;
PROC SORT DATA=mrg.famind89 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/*      step 2 for 1990 family file (n=9371)      */
/* ===== */
DATA fam90; INFILE "[PATH]\FAM1990.txt"    LRECL=2332;
INPUT
  V17702 2-6
  /*      your 1990 variable locations here      */
;
LABEL
  V17702 ="1990 INTERVIEW NUMBER 90"
  /*      your 1990 variable labels here      */
;
  /*      your 1990 missing data here      */
;
RUN;
PROC SORT DATA=fam90 TAGSORT; BY V17702; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER30642; RUN;
DATA mrg.famind90(COMPRESS=YES);
  MERGE fam90(rename=V17702=ER30642) mrg.indvars(IN=INDOK);
  BY ER30642; IF INDOK;
RUN;
PROC SORT DATA=mrg.famind90 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/*      step 2 for 1991 family file (n=9363)      */
/* ===== */
DATA fam91; INFILE "[PATH]\FAM1991.txt"    LRECL=2336;
INPUT
  V19002 2-5
  /*      your 1991 variable locations here      */
;
LABEL
  V19002 = "1991 INTERVIEW NUMBER"
  /*      your 1991 variable labels here      */
;
  /*      your 1991 missing data here      */
;
RUN;
PROC SORT DATA=fam91 TAGSORT; BY V19002; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER30689; RUN;
DATA mrg.famind91(COMPRESS=YES);
  MERGE fam91(rename=V19002=ER30689) mrg.indvars(IN=INDOK);

```

```

        BY ER30689; IF INDOK;
RUN;
PROC SORT DATA=mrg.famind91 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/*      step 2 for 1992 family file (n=9829)      */
/* ===== */
DATA fam92; INFILE "[PATH]\FAM1992.txt" LRECL=2325;
INPUT
    V20302 2-5
    /*      your 1992 variable locations here      */
;
LABEL
    V20302 = "1992 INTERVIEW NUMBER"
    /*      your 1992 variable labels here      */
;
    /*      your 1992 missing data here      */
RUN;
PROC SORT DATA=fam92 TAGSORT; BY V20302; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER30733; RUN;
DATA mrg.famind92(COMPRESS=YES);
MERGE fam92(rename=V20302=ER30733) mrg.indvars(IN=INDOK);
BY ER30733; IF INDOK;
RUN;
PROC SORT DATA=mrg.famind92 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/*      step 2 for 1993 family file (n=9977)      */
/* ===== */
DATA fam93; INFILE "[PATH]\FAM1993.txt" LRECL=3329;
INPUT
    V21602 2-6
    /*      your 1993 variable locations here      */
;
LABEL
    V21602 = "1993 INTERVIEW NUMBER"
    /*      your 1993 variable labels here      */
;
    /*      your 1993 missing data here      */
RUN;
PROC SORT DATA=fam93 TAGSORT; BY V21602; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER30806; RUN;
DATA mrg.famind93(COMPRESS=YES);
MERGE fam93(rename=V21602=ER30806) mrg.indvars(IN=INDOK);
BY ER30806; IF INDOK;
RUN;
PROC SORT DATA=mrg.famind93 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/*      step 2 for 1994 family file (n=10764)      */
/* ===== */
DATA fam94; INFILE "[PATH]\FAM1994ER.txt" LRECL=4379;
INPUT
    ER2002 2-6
    /*      your 1994 variable locations here      */
;
LABEL
    ER2002 = "1994 INTERVIEW #"
    /*      your 1994 variable labels here      */
;
    /*      your 1994 missing data here      */
RUN;
PROC SORT DATA=fam94 TAGSORT; BY ER2002; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER33101; RUN;
DATA mrg.famind94(COMPRESS=YES);

```

```

        MERGE fam94(rename=ER2002=ER33101) mrg.indvars(IN=INDOK);
        BY ER33101; IF INDOK;
RUN;
PROC SORT DATA=mrg.famind94 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/*      step 2 for 1995 family file (n=10401)      */
/* ===== */
DATA fam95; INFILE "[PATH]\FAM1995ER.txt"    LRECL=4131;
INPUT
    ER5002 2-6
    /*      your 1995 variable locations here      */
;
LABEL
    ER5002 = "1995 INTERVIEW #"
    /*      your 1995 variable labels here          */
;
    /*      your 1995 missing data here            */
RUN ;
PROC SORT DATA=fam95 TAGSORT; BY ER5002; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER33201; RUN;
DATA mrg.famind95(COMPRESS=YES);
    MERGE fam95(rename=ER5002=ER33201) mrg.indvars(IN=INDOK);
    BY ER33201; IF INDOK;
RUN;
PROC SORT DATA=mrg.famind95 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/*      step 2 for 1996 family file (n=8511)      */
/* ===== */
DATA fam96; INFILE "[PATH]\FAM1996ER.txt"    LRECL=4614;
INPUT
    ER7002 2-5
    /*      your 1996 variable locations here      */
;
LABEL
    ER7002 = "1996 INTERVIEW #"
    /*      your 1996 variable labels here          */
;
    /*      your 1996 missing data here            */
RUN ;
PROC SORT DATA=fam96 TAGSORT; BY ER7002; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER33301; RUN;
DATA mrg.famind96(COMPRESS=YES);
    MERGE fam96(rename=ER7002=ER33301) mrg.indvars(IN=INDOK);
    BY ER33301; IF INDOK;
RUN;
PROC SORT DATA=mrg.famind96 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/*      step 2 for 1997 family file (n=6747)      */
/* ===== */
DATA fam97; INFILE "[PATH]\FAM1997ER.txt"    LRECL=4473;
INPUT
    ER10002 2-6
    /*      your 1997 variable locations here      */
;
LABEL
    ER10002 = "1997 INTERVIEW #"
    /*      your 1997 variable labels here          */
;
    /*      your 1997 missing data here            */
RUN ;
PROC SORT DATA=fam97 TAGSORT; BY ER10002; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER33401; RUN;

```

```

DATA mrg.famind97 (COMPRESS=YES);
    MERGE fam97 (rename=ER10002=ER33401) mrg.indvars (IN=INDOK);
    BY ER33401; IF INDOK;
RUN;
PROC SORT DATA=mrg.famind97 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/*      step 2 for 1999 family file (n=6997)      */
/* ===== */
DATA fam99; INFILE "[PATH]\FAM1999ER.txt"    LRECL=6938;
    INPUT
        ER13002 2-6
        /*      your 1999 variable locations here      */
    ;
    LABEL
        ER13002 = "1999 FAMILY INTERVIEW (ID) NUMBER"
        /*      your 1999 variable labels here      */
    ;
        /*      your 1999 missing data here      */
RUN ;
PROC SORT DATA=fam99 TAGSORT; BY ER13002; RUN;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER33501; RUN;
DATA mrg.famind99 (COMPRESS=YES);
    MERGE fam99 (rename=ER13002=ER33501) mrg.indvars (IN=INDOK);
    BY ER33501; IF INDOK;
RUN:
PROC SORT DATA=mrg.famind99 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/*      step 2 for 2001 family file (n=7406)      */
/* ===== */
DATA fam01; INFILE "[PATH]\FAM2001ER.txt"    LRECL=6751;
    INPUT
        ER17002 2-5
        /*      your 2001 variable locations here      */
    ;
    LABEL
        ER17002 = "2001 FAMILY INTERVIEW (ID) NUMBER"
        /*      your 2001 variable labels here      */
    ;
        /*      your 2001 missing data here      */
RUN ;
PROC SORT DATA=fam01 TAGSORT; BY ER17002; RUN ;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER33601; RUN;
DATA mrg.famind01 (COMPRESS=YES);
    MERGE fam01 (rename=ER17002=ER33601) mrg.indvars (IN=INDOK);
    BY ER33601; IF INDOK;
RUN;
PROC SORT DATA=mrg.famind01 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/*      step 2 for 2003 family file (n=7822)      */
/* ===== */
DATA fam03; INFILE "[PATH]\FAM2003ER.txt"    LRECL=6196;
    INPUT
        ER21002 2-6
        /*      your 2003 variable locations here      */
    ;
    LABEL
        ER21002 = "2003 FAMILY INTERVIEW (ID) NUMBER"
        /*      your 2003 variable labels here      */
    ;
        /*      your 2003 missing data here      */
RUN ;
PROC SORT DATA=fam03 TAGSORT; BY ER21002; RUN ;

```



```

PROC SORT DATA=mrg.indvars TAGSORT; BY ER33701; RUN;
DATA mrg.famind03(COMPRESS=YES);
MERGE fam03(rename=ER21002=ER33701) mrg.indvars(IN=INDOK);
BY ER33701; IF INDOK;
RUN ;
PROC SORT DATA=mrg.famind03 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/* step 2 for 2005 family file (n=8002) */
/* ===== */
DATA fam05; INFILE "[PATH]\FAM2005ER.txt" LRECL=6333;
INPUT
ER25002 2-6
/* your 2005 variable locations here */
;
LABEL
ER25002 = "2005 FAMILY INTERVIEW (ID) NUMBER"
/* your 2005 variable labels here */
;
/* your 2005 missing data here */
RUN ;
PROC SORT DATA=fam05 TAGSORT; BY ER25002; RUN ;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER33801; RUN ;
DATA mrg.famind05(COMPRESS=YES);
MERGE fam05(rename=ER25002=ER33801) mrg.indvars(IN=INDOK);
BY ER33801; IF INDOK;
RUN ;
PROC SORT DATA=mrg.famind05 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/* step 2 for 2007 family file (n=8289) */
/* ===== */
DATA fam07; INFILE "[PATH]\FAM2007ER.txt" LRECL=8314;
INPUT
ER36002 2-6
/* your 2007 variable locations here */
;
LABEL
ER36002 = "2007 FAMILY INTERVIEW (ID) NUMBER"
/* your 2007 variable labels here */
;
/* your 2007 missing data here */
RUN ;
PROC SORT DATA=fam07 TAGSORT; BY ER36002;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER33901;
DATA mrg.famind07(COMPRESS=YES);
MERGE fam07(rename=ER36002=ER33901) mrg.indvars(IN=INDOK);
BY ER33901; IF INDOK;
RUN ;
PROC SORT DATA=mrg.famind07 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/* step 2 for 2009 family file (n=8690) */
/* ===== */
DATA fam09; INFILE "[PATH]\FAM2009ER.txt" LRECL=8223;
INPUT
ER42002 2-6
/* your 2009 variable locations here */
;
LABEL
ER42002 = "2009 FAMILY INTERVIEW (ID) NUMBER"
/* your 2009 variable labels here */
;
/* your 2009 missing data here */
RUN ;

```

```

PROC SORT DATA=fam09 TAGSORT; BY ER42002;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER34001;
DATA mrg.famind09(COMPRESS=YES);
MERGE fam09(rename=ER42002=ER34001) mrg.indvars(IN=INDOK);
BY ER34001; IF INDOK;

RUN ;
PROC SORT DATA=mrg.famind09 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/* step 2 for 2011 family file (n=8907) */
/* ===== */
DATA fam11; INFILE "[PATH]\FAM2011ER.txt" LRECL=8339;
INPUT
ER47302 2-6
/* your 2011 variable locations here */
;
LABEL
ER47302 = "2011 FAMILY INTERVIEW (ID) NUMBER"
/* your 2011 variable labels here */
;
/* your 2011 missing data here */
RUN ;
PROC SORT DATA=fam11 TAGSORT; BY ER47302;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER34101;
DATA mrg.famind11(COMPRESS=YES);
MERGE fam11(rename=ER47302=ER34101) mrg.indvars(IN=INDOK);
BY ER34101; IF INDOK;

RUN ;
PROC SORT DATA=mrg.famind11 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/* step 2 for 2013 family file (n=9063) */
/* ===== */
DATA fam13; INFILE "[PATH]\FAM2013ER.txt" LRECL=8485;
INPUT
ER53002 2-6
/* your 2013 variable locations here */
;
LABEL
ER53002 = "2013 FAMILY INTERVIEW (ID) NUMBER"
/* your 2013 variable labels here */
;
/* your 2013 missing data here */
RUN ;
PROC SORT DATA=fam13 TAGSORT; BY ER53002;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER34201;
DATA mrg.famind13(COMPRESS=YES);
MERGE fam13(rename=ER53002=ER34201) mrg.indvars(IN=INDOK);
BY ER34201; IF INDOK;

RUN ;
PROC SORT DATA=mrg.famind13 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/* step 2 for 2015 family file (n=9048) */
/* ===== */
DATA fam15; INFILE "[PATH]\FAM2015ER.txt" LRECL=9250;
INPUT
ER60002 2-6
/* your 2015 variable locations here */
;
LABEL
ER60002 = "2015 FAMILY INTERVIEW (ID) NUMBER"
/* your 2015 variable labels here */
;
/* your 2015 missing data here */

```

```

RUN ;
PROC SORT DATA=fam15 TAGSORT; BY ER60002;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER34301;
DATA mrg.famind15(COMPRESS=YES);
MERGE fam15(rename=ER60002=ER34301) mrg.indvars(IN=INDOK);
BY ER34301; IF INDOK;
RUN ;
PROC SORT DATA=mrg.famind15 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/* step 2 for 2017 family file (n=9607) */
/* ===== */
DATA fam17; INFILE "[PATH]\FAM2017ER.txt" LRECL=9462;
INPUT
ER66002 2-6
/* your 2017 variable locations here */
;
LABEL
ER66002 = "2017 FAMILY INTERVIEW (ID) NUMBER"
/* your 2017 variable labels here */
;
/* your 2017 missing data here */
RUN ;
PROC SORT DATA=fam17 TAGSORT; BY ER66002;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER34501;
DATA mrg.famind17(COMPRESS=YES);
MERGE fam17(rename=ER66002=ER34501) mrg.indvars(IN=INDOK);
BY ER34501; IF INDOK;
RUN ;
PROC SORT DATA=mrg.famind17 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/* step 2 for 2019 family file (n=9569) */
/* ===== */
DATA fam19; INFILE "[PATH]\FAM2019ER.txt" LRECL=9652;
INPUT
ER72002 2-6
/* your 2019 variable locations here */
;
LABEL
ER72002 = "2019 FAMILY INTERVIEW (ID) NUMBER"
/* your 2019 variable labels here */
;
/* your 2019 missing data here */
RUN ;
PROC SORT DATA=fam19 TAGSORT; BY ER72002;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER34701;
DATA mrg.famind19(COMPRESS=YES);
MERGE fam19(rename=ER72002=ER34701) mrg.indvars(IN=INDOK);
BY ER34701; IF INDOK;
RUN ;
PROC SORT DATA=mrg.famind19 TAGSORT; BY ER30001 ER30002;
RUN;
/* ===== */
/* step 2 for 2021 family file (n=9207) */
/* ===== */
DATA fam21; INFILE "[PATH]\FAM2021ER.txt" LRECL=7890;
INPUT
ER78002 2-6
/* your 2021 variable locations here */ ;
LABEL
ER78002 = "2021 FAMILY INTERVIEW (ID) NUMBER"
/* your 2021 variable labels here */ ;
/* your 2021 missing data here */
RUN ;

```

```

PROC SORT DATA=fam21 TAGSORT; BY ER78002;
PROC SORT DATA=mrg.indvars TAGSORT; BY ER34901;
DATA mrg.famind21 (COMPRESS=YES);
  MERGE fam21 (rename=ER78002=ER34901) mrg.indvars (IN=INDOK);
  BY ER34901; IF INDOK;
RUN ; PROC SORT DATA=mrg.famind21 TAGSORT; BY ER30001 ER30002;
RUN;

/* ===== */
/* step 3: merge familiy-individual files on individual identifiers */
/* ===== */
DATA mrg.xyrfind (COMPRESS=YES);
  MERGE
    mrg.famind68 mrg.famind69
    mrg.famind70 mrg.famind71 mrg.famind72 mrg.famind73 mrg.famind74
    mrg.famind75 mrg.famind76 mrg.famind77 mrg.famind78 mrg.famind79
    mrg.famind80 mrg.famind81 mrg.famind82 mrg.famind83 mrg.famind84
    mrg.famind85 mrg.famind86 mrg.famind87 mrg.famind88 mrg.famind89
    mrg.famind90 mrg.famind91 mrg.famind92 mrg.famind93 mrg.famind94
    mrg.famind95 mrg.famind96 mrg.famind97 mrg.famind99
    mrg.famind01 mrg.famind03 mrg.famind05 mrg.famind07 mrg.famind09
    mrg.famind11 mrg.famind13 mrg.famind15 mrg.famind17 mrg.famind19
    mrg.famind21
  ;
  BY ER30001 ER30002;
RUN;

```

B.2 STATA

B.2.1 STATA Example - Method 1

```
/*-----*
| This example program demonstrates a relatively simple method for |
| merging PSID data. It uses data from 3 different years, subset- |
| ting criteria, and the compress and tagsort options.           |
|                                                                 |
| When working with PSID data, the amount of available system disk |
| space and memory is often an important consideration.         |
|                                                                 |
| merge command uses syntax for STATA 12, for earlier versions of STATA |
| it should be modified accordingly                               |
*-----*/
#delimit ;
cd "[FOLDER NAME]" ;
tempfile IND90_92 FAM90 FAM91 FAM92 ;
* Read in cross-year individual file and select variables
*   from 1990-1992 needed for analysis;
infix
  ER30001      2 - 5
  ER30002      6 - 8
  ER30642     1528 - 1532
  ER30643     1533 - 1534
  ER30644     1535 - 1536
  ER30645     1537 - 1539
  ER30653     1555 - 1555
  ER30657     1563 - 1564
  ER30659     1566 - 1571
  ER30689     1647 - 1650
  ER30690     1651 - 1652
  ER30691     1653 - 1654
  ER30692     1655 - 1657
  ER30699     1672 - 1672
  ER30703     1680 - 1681
  ER30705     1683 - 1688
  ER30707     1690 - 1695
  ER30733     1764 - 1767
  ER30734     1768 - 1769
  ER30735     1770 - 1771
  ER30736     1772 - 1774
  ER30744     1790 - 1790
  ER30748     1798 - 1799
  ER30750     1801 - 1806
  ER30752     1808 - 1813
  ER30805     1907 - 1913
  ER32000     2057 - 2057
  ER32022     2111 - 2112
  ER32049     2187 - 2187
using "[PATH]\IND2021ER.txt", clear
;
label variable ER30001      "1968 INTERVIEW NUMBER" ;
label variable ER30002      "PERSON NUMBER" ;
label variable ER30642      "1990 INTERVIEW NUMBER" ;
label variable ER30643      "SEQUENCE NUMBER90" ;
label variable ER30644      "RELATION TO HEAD" ;
label variable ER30645      "AGE OF INDIVIDUAL" ;
label variable ER30653      "EMPLOYMENT STAT-IND" ;
label variable ER30657      "COMPLETED EDUC-IND" ;
label variable ER30659      "TOT TXBL INCOME-IND" ;
label variable ER30689      "1991 INTERVIEW NUMBER" ;
label variable ER30690      "SEQUENCE NUMBER91" ;
label variable ER30691      "RELATION TO HEAD" ;
label variable ER30692      "AGE OF INDIVIDUAL" ;
```

```

label variable ER30699      "EMPLOYMENT STAT-IND      91" ;
label variable ER30703      "COMPLETED EDUC-IND      91" ;
label variable ER30705      "TOT LABOR INCOME-IND     91" ;
label variable ER30707      "TOT ASSET INCOME-IND     91" ;
label variable ER30733      "1992 INTERVIEW NUMBER" ;
label variable ER30734      "SEQUENCE NUMBER92" ;
label variable ER30735      "RELATION TO HEAD        92" ;
label variable ER30736      "AGE OF INDIVIDUAL        92" ;
label variable ER30744      "EMPLOYMENT STAT          92" ;
label variable ER30748      "COMPLETED EDUCATION     92" ;
label variable ER30750      "TOT LABOR INCOME         92" ;
label variable ER30752      "TOT ASSET INCOME         92" ;
label variable ER30805      "COMBINED IND WEIGHT      92" ;
label variable ER32000      "SEX OF INDIVIDUAL" ;
label variable ER32022      "# LIVE BIRTHS TO THIS INDIVIDUAL" ;
label variable ER32049      "LAST KNOWN MARITAL STATUS" ;
rename ER30642 ID90 ;
rename ER30689 ID91 ;
rename ER30733 ID92 ;
replace ER30645=. if ER30645==999 ;
replace ER30657=. if ER30657==99 ;
replace ER30692=. if ER30692==999 ;
replace ER30703=. if ER30703==99 ;
replace ER30736=. if ER30736==999 ;
replace ER30748=. if ER30748==99 ;
replace ER32022=. if ER32022==98 ;
replace ER32049=. if ER32049==8 ;
* Select those who were ever heads or wives/"wives" between 1990 and 1992 ;
keep if (
  (ER30643 == 01 & ER30644 == 10) |
  (ER30643 == 02 & (ER30644 == 20 | ER30644 == 22)) |
  (ER30690 == 01 & ER30691 == 10) |
  (ER30690 == 02 & (ER30691 == 20 | ER30691 == 22)) |
  (ER30734 == 01 & ER30735 == 10) |
  (ER30734 == 02 & (ER30735 == 20 | ER30735 == 22))
) ;
sort ER30001 ER30002 ;
save "'IND90_92'" ;
* Read in 1990 family file and select variables needed for analysis ;
infix
  V17702      2 - 6
  V17836      281 - 286
  V18262      1175 - 1177
  V18564      1656 - 1658
  V18814      2046 - 2046
  V18878      2188 - 2193
  V18887      2230 - 2234
  V18888      2235 - 2239
using "[PATH]\FAM1990.txt", clear
;
label variable V17702 "1990 INTERVIEW NUMBER" ;
label variable V17836 "WIFE 89 LABOR/WAGE" ;
label variable V18262 "C9-10 OCC-LAST JOB (H-U)" ;
label variable V18564 "E9-10 OCC-LAST JOB (W-U)" ;
label variable V18814 "M32 RACE OF HEAD (1 MEN)" ;
label variable V18878 "TOTAL HEAD LABOR Y 89" ;
label variable V18887 "HEAD 89 AVG HRLY EARNING" ;
label variable V18888 "WIFE 89 AVG HRLY EARNING" ;
rename V17702 ID90 ;
replace V18262=. if V18262==999 ;
replace V18564=. if V18564==999 ;
replace V18814=. if V18814==9 ;
sort ID90 ;
save "'FAM90'" ;
* Merge fam90 and ind90_92 by id90 ;

```

```

merge 1:m ID90 using "IND90_92" , keep(using matched) ;
drop _merge ;
save "FAM_IND" , replace ;
* Read in 1991 family file and select variables needed for analysis;
infix
    V19002      2 - 5
    V19136     281 - 286
    V19562     1175 - 1177
    V19864     1656 - 1658
    V20114     2046
    V20178     2188 - 2193
    V20187     2230 - 2234
    V20188     2235 - 2239
using "[PATH]\FAM1991.txt", clear
;
label variable V19002 "1991 INTERVIEW NUMBER" ;
label variable V19136 "WIFE 90 LABOR/WAGE" ;
label variable V19562 "C9-10 OCC-LAST JOB (H-U)" ;
label variable V19864 "E9-10 OCC-LAST JOB (W-U)" ;
label variable V20114 "L32 RACE OF HEAD (1 MEN)" ;
label variable V20178 "TOTAL HEAD LABOR Y 90" ;
label variable V20187 "HEAD 90 AVG HRLY EARNING" ;
label variable V20188 "WIFE 90 AVG HRLY EARNING" ;
rename V19002 ID91 ;
replace V19562=. if V19562==999 ;
replace V19864=. if V19864==999 ;
replace V20114=. if V20114==9 ;
sort ID91 ;
save "FAM91" ;
* Merge fam91 and fam_ind by id91 ;
merge 1:m ID91 using FAM_IND , keep(using matched);
drop _merge ;
save "FAM_IND" , replace ;
* Read in 1992 family file and select variables needed for analysis;
infix
    V20302      2 - 5
    V20436     282 - 287
    V20862     1189 - 1191
    V21164     1670 - 1672
    V21420     2066
    V21484     2172 - 2177
    V21493     2218 - 2222
    V21494     2223 - 2227
using "[PATH]\FAM1992.txt", clear
;
label variable V20302 "1992 INTERVIEW NUMBER" ;
label variable V20436 "WIFE 91 LABOR/WAGE" ;
label variable V20862 "C9-10 OCC-LAST JOB (H-U)" ;
label variable V21164 "E9-10 OCC-LAST JOB (W-U)" ;
label variable V21420 "M32 RACE OF HEAD (1 MEN)" ;
label variable V21484 "TOTAL HEAD LABOR Y 91" ;
label variable V21493 "HEAD 91 AVG HRLY EARNING" ;
label variable V21494 "WIFE 91 AVG HRLY EARNING" ;
rename V20302 ID92 ;
replace V20862=. if V20862==999 ;
replace V21164=. if V21164==999 ;
replace V21420=. if V21420==9 ;
sort ID92 ;
save "FAM92" ;
* Merge fam92 and fam_ind by id92 ;
merge 1:m ID92 using FAM_IND , keep(using matched);
drop _merge ;
sort ID92 ID91 ID90 ER30001 ER30002 ;
save "FAM_IND" , replace ;

```

B.2.2 STATA Example - Method 2

```
/*-----*
|           MERGE USING MULTIPLE FAMILY-INDIVIDUAL FILES           |
| Step 1:   Subset family-level id's and selected variables       |
|           and select cases from xyr-individual file             |
| Step 2a:  Subset year-n family file                             |
| Step 2b:  Sort year-n family file from 2a by year-n family id  |
| Step 2c:  Sort xyr-individual file from step 1 by year-n family id |
| Step 2d:  Merge sorted xyr-individual file from 2c             |
|           with sorted year-n subsetted family file from 2b    |
|           (a one-to-many [family-to-individual] match)         |
| Step 2e:  Sort resulting year-n family-individual file from 2d |
|           by individual ids                                     |
| ...       Repeat Steps 2 for all other years                   |
| Step 3:   Merge family-individual files from step 2e           |
| merge command uses syntax for STATA 12, for earlier versions of STATA |
| it should be modified accordingly                               |
/*-----*/
#delimit ;
cd "[FOLDER NAME]" ;
tempfile
  FAM68 FAM69 FAM70 FAM71 FAM72 FAM73 FAM74 FAM75 FAM76 FAM77 FAM78 FAM79
  FAM80 FAM81 FAM82 FAM83 FAM84 FAM85 FAM86 FAM87 FAM88 FAM89
  FAM90 FAM91 FAM92 FAM93 FAM94 FAM95 FAM96 FAM97 FAM99
  FAM01 FAM03 FAM05 FAM07 FAM09 FAM11 FAM13 FAM15 FAM17 FAM19 FAM21
;
/* ===== */
/*      step 1: subset family-level id's and individual variables      */
/*      and select cases from cross-year individual record              */
/* ===== */
infix
  /* 1968 ID */ ER30001    2 - 5
  /* 1968 PN */ ER30002    6 - 8
  /* 1969 */ ER30020    44 - 47
  /* 1970 */ ER30043    97 - 100
  /* 1971 */ ER30067   152 - 155
  /* 1972 */ ER30091   207 - 210
  /* 1973 */ ER30117   265 - 268
  /* 1974 */ ER30138   317 - 320
  /* 1975 */ ER30160   370 - 373
  /* 1976 */ ER30188   436 - 439
  /* 1977 */ ER30217   503 - 506
  /* 1978 */ ER30246   571 - 574
  /* 1979 */ ER30283   648 - 651
  /* 1980 */ ER30313   718 - 721
  /* 1981 */ ER30343   788 - 791
  /* 1982 */ ER30373   858 - 861
  /* 1983 */ ER30399   919 - 922
  /* 1984 */ ER30429   992 - 995
  /* 1985 */ ER30463  1077 - 1080
  /* 1986 */ ER30498  1167 - 1170
  /* 1987 */ ER30535  1259 - 1262
  /* 1988 */ ER30570  1348 - 1351
  /* 1989 */ ER30606  1438 - 1441
  /* 1990 */ ER30642  1528 - 1532
  /* 1991 */ ER30689  1647 - 1650
  /* 1992 */ ER30733  1764 - 1767
  /* 1993 */ ER30806  1914 - 1918
  /* 1994 */ ER33101  2201 - 2205
```



```

/* 1995 */ ER33201 2300 - 2304
/* 1996 */ ER33301 2493 - 2496
/* 1997 */ ER33401 2577 - 2581
/* 1999 */ ER33501 2662 - 2666
/* 2001 */ ER33601 2856 - 2859
/* 2003 */ ER33701 3038 - 3042
/* 2005 */ ER33801 3193 - 3197
/* 2007 */ ER33901 3492 - 3496
/* 2009 */ ER34001 3770 - 3774
/* 2011 */ ER34101 3992 - 3996
/* 2013 */ ER34201 4229 - 4233
/* 2015 */ ER34301 4495 - 4499
/* 2017 */ ER34501 4869 - 4873
/* 2019 */ ER34701 5249 - 5253
/* 2021 */ ER34901 5683 - 5687

/* your individual variable locations here */
using "[PATH]\IND2021ER.txt" , clear
;
label variable ER30001 "1968 INTERVIEW NUMBER 68" ;
label variable ER30002 "PERSON NUMBER 68" ;
label variable ER30020 "1969 INTERVIEW NUMBER 69" ;
label variable ER30043 "1970 INTERVIEW NUMBER 70" ;
label variable ER30067 "1971 INTERVIEW NUMBER 71" ;
label variable ER30091 "1972 INTERVIEW NUMBER 72" ;
label variable ER30117 "1973 INTERVIEW NUMBER 73" ;
label variable ER30138 "1974 INTERVIEW NUMBER 74" ;
label variable ER30160 "1975 INTERVIEW NUMBER 75" ;
label variable ER30188 "1976 INTERVIEW NUMBER 76" ;
label variable ER30217 "1977 INTERVIEW NUMBER 77" ;
label variable ER30246 "1978 INTERVIEW NUMBER 78" ;
label variable ER30283 "1979 INTERVIEW NUMBER 79" ;
label variable ER30313 "1980 INTERVIEW NUMBER 80" ;
label variable ER30343 "1981 INTERVIEW NUMBER 81" ;
label variable ER30373 "1982 INTERVIEW NUMBER 82" ;
label variable ER30399 "1983 INTERVIEW NUMBER 83" ;
label variable ER30429 "1984 INTERVIEW NUMBER 84" ;
label variable ER30463 "1985 INTERVIEW NUMBER 85" ;
label variable ER30498 "1986 INTERVIEW NUMBER 86" ;
label variable ER30535 "1987 INTERVIEW NUMBER 87" ;
label variable ER30570 "1988 INTERVIEW NUMBER 88" ;
label variable ER30606 "1989 INTERVIEW NUMBER 89" ;
label variable ER30642 "1990 INTERVIEW NUMBER 90" ;
label variable ER30689 "1991 INTERVIEW NUMBER 91" ;
label variable ER30733 "1992 INTERVIEW NUMBER 92" ;
label variable ER30806 "1993 INTERVIEW NUMBER 93" ;
label variable ER33101 "1994 INTERVIEW NUMBER 94" ;
label variable ER33201 "1995 INTERVIEW NUMBER 95" ;
label variable ER33301 "1996 INTERVIEW NUMBER 96" ;
label variable ER33401 "1997 INTERVIEW NUMBER 97" ;
label variable ER33501 "1999 INTERVIEW NUMBER 99" ;
label variable ER33601 "2001 INTERVIEW NUMBER 01" ;
label variable ER33701 "2003 INTERVIEW NUMBER 03" ;
label variable ER33801 "2005 INTERVIEW NUMBER 05" ;
label variable ER33901 "2007 INTERVIEW NUMBER 07" ;
label variable ER34001 "2009 INTERVIEW NUMBER 09" ;
label variable ER34101 "2011 INTERVIEW NUMBER 11" ;
label variable ER34201 "2013 INTERVIEW NUMBER 13" ;
label variable ER34301 "2015 INTERVIEW NUMBER 15" ;
label variable ER34501 "2017 INTERVIEW NUMBER 17" ;
label variable ER34701 "2019 INTERVIEW NUMBER 19" ;
label variable ER34901 "2021 INTERVIEW NUMBER 21" ;

/* your individual variable labels here */
/* your individual missing data here */

```

```

/*      your filter for case selection, if any, here
      this selects the first 10 cases for testing setup */
sort ER30001 ;
save INDVARS , replace;
/* ===== */
/*      Step 2 for 1968 family file (n=4802)      */
/* ===== */
infix
  V3 7-10
  /* your 1968 variable locations here */
using "[PATH]\FAM1968.txt" , clear
;
  label variable V3 "INTERVIEW NUMBER 68 1:6-9" ;
  /* your 1968 variable labels here */
  /* your 1968 missing data here */
  rename V3 ER30001 ;
  sort ER30001 ;
save "'FAM68'" ;
merge 1:m ER30001 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND68 , replace;
/* ===== */
/*      step 2 for 1969 family file (n=4460)      */
/* ===== */
infix
  V442 2-5
  /* your 1969 variable locations here */
using "[PATH]\FAM1969.txt" , clear
;
  label variable V442 "1969 INT NUMBER 11:6-9" ;
  /* your 1969 variable labels here */
  /* your 1969 missing data here */
  rename V442 ER30020 ;
  sort ER30020 ;
save "'FAM69'" ;
merge 1:m ER30020 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND69 , replace;
/* ===== */
/*      step 2 for 1970 family file (n=4645)      */
/* ===== */
infix
  V1102 2-5
  /* your 1970 variable locations here */
using "[PATH]\FAM1970.txt" , clear
;
  label variable V1102 "1970 INT # 21:6-9" ;
  /* your 1970 variable labels here */
  /* your 1970 missing data here */
  rename V1102 ER30043 ;
  sort ER30043 ;
save "'FAM70'" ;
merge 1:m ER30043 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND70 , replace;
/* ===== */
/*      step 2 for 1971 family file (n=4840)      */
/* ===== */
infix
  V1802 2-5
  /* your 1971 variable locations here */
using "[PATH]\FAM1971.txt" , clear

```

```

;
  label variable V1802 "71 ID NO." ;
  /* your 1971 variable labels here */
  /* your 1971 missing data here */
  rename V1802 ER30067 ;
  sort ER30067 ;
save "FAM71" ;
merge 1:m ER30067 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND71 , replace;
/* ===== */
/*      step 2 for 1972 family file (n=5060)      */
/* ===== */
infix
  V2402 2-5
  /* your 1972 variable locations here */
using "[PATH]\FAM1972.txt" , clear
;
  label variable V2402 "1972 INT # 46:6-9" ;
  /* your 1972 variable labels here */
  /* your 1972 missing data here */
  rename V2402 ER30091 ;
  sort ER30091 ;
save "FAM72" ;
merge 1:m ER30091 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND72 , replace;
/* ===== */
/*      step 2 for 1973 family file (n=5285)      */
/* ===== */
infix
  V3002 2-5
  /* your 1973 variable locations here */
using "[PATH]\FAM1973.txt" , clear
;
  label variable V3002 "1973 INT # 59:6-9" ;
  /* your 1973 variable labels here */
  /* your 1973 missing data here */
  rename V3002 ER30117 ;
  sort ER30117 ;
save "FAM73" ;
merge 1:m ER30117 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND73 , replace;
/* ===== */
/*      step 2 for 1974 family file (n=5517)      */
/* ===== */
infix
  V3402 2-5
  /* your 1974 variable locations here */
using "[PATH]\FAM1974.txt" , clear
;
  label variable V3402 "1974 ID NUMBER" ;
  /* your 1974 variable labels here */
  /* your 1974 missing data here */
  rename V3402 ER30138 ;
  sort ER30138 ;
save "FAM74" ;
merge 1:m ER30138 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND74 , replace;

```

```

/* ===== */
/*      step 2 for 1975 family file (n=5725)      */
/* ===== */
infix
  V3802 2-5
  /* your 1975 variable locations here */
using "[PATH]\FAM1975.txt" , clear
;
  label variable V3802 "1975 INT # 80:6-9" ;
  /* your 1975 variable labels here */
  /* your 1975 missing data here */
  rename V3802 ER30160 ;
  sort ER30160 ;
save "FAM75" ;
merge 1:m ER30160 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND75 , replace;
/* ===== */
/*      step 2 for 1976 family file (n=5862)      */
/* ===== */
infix
  V4302 2-5
  /* your 1976 variable locations here */
using "[PATH]\FAM1976.txt" , clear
;
  label variable V4302 "1976 ID NUMBER 6V2" ;
  /* your 1976 variable labels here */
  /* your 1976 missing data here */
  rename V4302 ER30188 ;
  sort ER30188 ;
save "FAM76" ;
merge 1:m ER30188 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND76 , replace;
/* ===== */
/*      step 2 for 1977 family file (n=6007)      */
/* ===== */
infix
  V5202 2-5
using "[PATH]\FAM1977.txt" , clear
;
  label variable V5202 "1977 ID" ;
  /* your 1977 variable labels here */
  /* your 1977 missing data here */
  rename V5202 ER30217 ;
  sort ER30217 ;
save "FAM77" ;
merge 1:m ER30217 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND77 , replace;
/* ===== */
/*      step 2 for 1978 family file (n=6154)      */
/* ===== */
infix
  V5702 2-5
  /* your 1978 variable locations here */
using "[PATH]\FAM1978.txt" , clear
;
  label variable V5702 "1978 ID" ;
  /* your 1978 variable labels here */
  /* your 1978 missing data here */
  rename V5702 ER30246 ;

```

```

        sort ER30246 ;
save "FAM78" ;
merge 1:m ER30246 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND78 , replace;
/* ===== */
/*      step 2 for 1979 family file (n=6373)      */
/* ===== */
infix
    V6302 2-5
    /* your 1979 variable locations here */
using "[PATH]\FAM1979.txt" , clear
;
    label variable V6302 "1979 ID" ;
    /* your 1979 variable labels here */
    /* your 1979 missing data here */
    rename V6302 ER30283 ;
    sort ER30283 ;
save "FAM79" ;
merge 1:m ER30283 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND79 , replace;
/* ===== */
/*      step 2 for 1980 family file (n=6533)      */
/* ===== */
infix
    V6902 2-5
    /* your 1980 variable locations here */
using "[PATH]\FAM1980.txt" , clear
;
    label variable V6902 "1980 INTERVIEW NUMBER" ;
    /* your 1980 variable labels here */
    /* your 1980 missing data here */
    rename V6902 ER30313 ;
    sort ER30313 ;
save "FAM80" ;
merge 1:m ER30313 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND80 , replace;
/* ===== */
/*      step 2 for 1981 family file (n=6620)      */
/* ===== */
infix
    V7502 2-5
    /* your 1981 variable locations here */
using "[PATH]\FAM1981.txt" , clear
;
    label variable V7502 "1981 INTERVIEW NUMBER" ;
    /* your 1981 variable labels here */
    /* your 1981 missing data here */
    rename V7502 ER30343 ;
    sort ER30343 ;
save "FAM81" ;
merge 1:m ER30343 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND81 , replace;
/* ===== */
/*      step 2 for 1982 family file (n=6742)      */
/* ===== */
infix
    V8202 2-5

```

```

/* your 1982 variable locations here */
using "[PATH]\FAM1982.txt" , clear
;
label variable V8202 "1982 INTERVIEW NUMBER" ;
/* your 1982 variable labels here */
/* your 1982 missing data here */
rename V8202 ER30373 ;
sort ER30373 ;
save "FAM82" ;
merge 1:m ER30373 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND82 , replace;
/* ===== */
/* step 2 for 1983 family file (n=6852) */
/* ===== */
infix
V8802 2-5
/* your 1983 variable locations here */
using "[PATH]\FAM1983.txt" , clear
;
label variable V8802 "1983 INTERVIEW NUMBER" ;
/* your 1983 variable labels here */
/* your 1983 missing data here */
rename V8802 ER30399 ;
sort ER30399 ;
save "FAM83" ;
merge 1:m ER30399 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND83 , replace;
/* ===== */
/* step 2 for 1984 family file (n=6918) */
/* ===== */
infix
V10002 2-5
/* your 1984 variable locations here */
using "[PATH]\FAM1984.txt" , clear
;
label variable V10002 "1984 INTERVIEW NUMBER" ;
/* your 1984 variable labels here */
/* your 1984 missing data here */
rename V10002 ER30429 ;
sort ER30429 ;
save "FAM84" ;
merge 1:m ER30429 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND84 , replace;
/* ===== */
/* step 2 for 1985 family file (n=7032) */
/* ===== */
infix
V11102 2-5
/* your 1985 variable locations here */
using "[PATH]\FAM1985.txt" , clear
;
label variable V11102 "1985 INTERVIEW NUMBER" ;
/* your 1985 variable labels here */
/* your 1985 missing data here */
rename V11102 ER30463 ;
sort ER30463 ;
save "FAM85" ;
merge 1:m ER30463 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;

```

```

drop _merge ;
save FAMIND85 , replace;
/* ===== */
/*      step 2 for 1986 family file (n=7018)      */
/* ===== */
infix
    V12502 2-5
    /* your 1986 variable locations here */
using "[PATH]\FAM1986.txt" , clear
;
    label variable V12502 "1986 INTERVIEW NUMBER" ;
    /* your 1986 variable labels here */
    /* your 1986 missing data here */
    rename V12502 ER30498 ;
    sort ER30498 ;
save "FAM86" ;
merge 1:m ER30498 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND86 , replace;
/* ===== */
/*      step 2 for 1987 family file (n=7061)      */
/* ===== */
infix
    V13702 2-5
    /* your 1987 variable locations here */
using "[PATH]\FAM1987.txt" , clear
;
    label variable V13702 "1987 INTERVIEW NUMBER" ;
    /* your 1987 variable labels here */
    /* your 1987 missing data here */
    rename V13702 ER30535 ;
    sort ER30535 ;
save "FAM87" ;
merge 1:m ER30535 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND87 , replace;
/* ===== */
/*      step 2 for 1988 family file (n=7114)      */
/* ===== */
infix
    V14802 2-5
    /* your 1988 variable locations here */
using "[PATH]\FAM1988.txt" , clear
;
    label variable V14802 "1988 INTERVIEW NUMBER" ;
    /* your 1988 variable labels here */
    /* your 1988 missing data here */
    rename V14802 ER30570 ;
    sort ER30570 ;
save "FAM88" ;
merge 1:m ER30570 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND88 , replace;
/* ===== */
/*      step 2 for 1989 family file (n=7114)      */
/* ===== */
infix
    V16302 2-5
    /* your 1989 variable locations here */
using "[PATH]\FAM1989.txt" , clear
;
    label variable V16302 "1989 INTERVIEW NUMBER" ;

```

```

/* your 1989 variable labels here */
/* your 1989 missing data here */
rename V16302 ER30606 ;
sort ER30606 ;
save "FAM89" ;
merge 1:m ER30606 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND89 , replace;
/* ===== */
/*      step 2 for 1990 family file (n=9371)      */
/* ===== */
infix
    V17702 2-6
    /* your 1990 variable locations here */
using "[PATH]\FAM1990.txt" , clear
;
    label variable V17702 "1990 INTERVIEW NUMBER 90" ;
    /* your 1990 variable labels here */
    /* your 1990 missing data here */
    rename V17702 ER30642 ;
    sort ER30642 ;
save "FAM90" ;
merge 1:m ER30642 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND90 , replace;
/* ===== */
/*      step 2 for 1991 family file (n=9363)      */
/* ===== */
infix
    V19002 2-5
    /* your 1991 variable locations here */
using "[PATH]\FAM1991.txt" , clear
;
    label variable V19002 "1991 INTERVIEW NUMBER" ;
    /* your 1991 variable labels here */
    /* your 1991 missing data here */
    rename V19002 ER30689 ;
    sort ER30689 ;
save "FAM91" ;
merge 1:m ER30689 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND91 , replace;
/* ===== */
/*      step 2 for 1992 family file (n=9829)      */
/* ===== */
infix
    V20302 2-5
    /* your 1992 variable locations here */
using "[PATH]\FAM1992.txt" , clear
;
    label variable V20302 "1992 INTERVIEW NUMBER" ;
    /* your 1992 variable labels here */
    /* your 1992 missing data here */
    rename V20302 ER30733 ;
    sort ER30733 ;
save "FAM92" ;
merge 1:m ER30733 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND92 , replace;
/* ===== */
/*      step 2 for 1993 family file (n=9977)      */
/* ===== */

```



```

/* ===== */
infix
  V21602 2-6
  /* your 1993 variable locations here */
using "[PATH]\FAM1993.txt" , clear
;
  label variable V21602 "1993 INTERVIEW NUMBER" ;
  /* your 1993 variable labels here */
  /* your 1993 missing data here */
  rename V21602 ER30806 ;
  sort ER30806 ;
save "FAM93" ;
merge 1:m ER30806 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND93 , replace;
/* ===== */
/*      step 2 for 1994 family file (n=10764)      */
/* ===== */
infix
  ER2002 2-6
  /* your 1994 variable locations here */
using "[PATH]\FAM1994ER.txt" , clear
;
  label variable ER2002 "1994 INTERVIEW #" ;
  /* your 1994 variable labels here */
  /* your 1994 missing data here */
  rename ER2002 ER33101 ;
  sort ER33101 ;
save "FAM94" ;
merge 1:m ER33101 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND94 , replace;
/* ===== */
/*      step 2 for 1995 family file (n=10401)      */
/* ===== */
infix
  ER5002 2-6
  /* your 1995 variable locations here */
using "[PATH]\FAM1995ER.txt" , clear
;
  label variable ER5002 "1995 INTERVIEW #" ;
  /* your 1995 variable labels here */
  /* your 1995 missing data here */
  rename ER5002 ER33201 ;
  sort ER33201 ;
save "FAM95" ;
merge 1:m ER33201 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND95 , replace;
/* ===== */
/*      step 2 for 1996 family file (n=8511)      */
/* ===== */
infix
  ER7002 2-5
  /* your 1996 variable locations here */
using "[PATH]\FAM1996ER.txt" , clear
;
  label variable ER7002 "1996 INTERVIEW #" ;
  /* your 1996 variable labels here */
  /* your 1996 missing data here */
  rename ER7002 ER33301 ;
  sort ER33301 ;

```

```

save "FAM96" ;
merge 1:m ER33301 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND96 , replace;
/* ===== */
/*      step 2 for 1997 family file (n=6747)      */
/* ===== */
infix
    ER10002 2-6
    /* your 1997 variable locations here */
using "[PATH]\FAM1997ER.txt" , clear
;
    label variable ER10002 "1997 INTERVIEW #" ;
    /* your 1997 variable labels here */
    /* your 1997 missing data here */
    rename ER10002 ER33401 ;
    sort ER33401 ;
save "FAM97" ;
merge 1:m ER33401 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND97 , replace;
/* ===== */
/*      step 2 for 1999 family file (n=6997)      */
/* ===== */
infix
    ER13002 2-6
    /* your 1999 variable locations here */
using "[PATH]\FAM1999ER.txt" , clear
;
    label variable ER13002 "1999 FAMILY INTERVIEW (ID) NUMBER" ;
    /* your 1999 variable labels here */
    /* your 1999 missing data here */
    rename ER13002 ER33501 ;
    sort ER33501 ;
save "FAM99" ;
merge 1:m ER33501 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND99 , replace;
/* ===== */
/*      step 2 for 2001 family file (n=7406)      */
/* ===== */
infix
    ER17002 2-5
    /* your 2001 variable locations here */
using "[PATH]\FAM2001ER.txt" , clear
;
    label variable ER17002 "2001 FAMILY INTERVIEW (ID) NUMBER" ;
    /* your 2001 variable labels here */
    /* your 2001 missing data here */
    rename ER17002 ER33601 ;
    sort ER33601 ;
save "FAM01" ;
merge 1:m ER33601 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND01 , replace;
/* ===== */
/*      step 2 for 2003 family file (n=7822)      */
/* ===== */
infix
    ER21002 2-6
    /* your 2003 variable locations here */

```

```

using "[PATH]\FAM2003ER.txt" , clear
;
    label variable ER21002 "2003 FAMILY INTERVIEW (ID) NUMBER" ;
    /* your 2003 variable labels here */
    /* your 2003 missing data here */
    rename ER21002 ER33701 ;
    sort ER33701 ;
save "FAM03" ;
merge 1:m ER33701 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND03 , replace;
/* ===== */
/*          step 2 for 2005 family file (n=8002)          */
/* ===== */
infix
    ER25002 2-6
    /* your 2005 variable locations here */
using "[PATH]\FAM2005ER.txt" , clear
;
    label variable ER25002 "2005 FAMILY INTERVIEW (ID) NUMBER" ;
    /* your 2005 variable labels here */
    /* your 2005 missing data here */
    rename ER25002 ER33801 ;
    sort ER33801 ;
save "FAM05" ;
merge 1:m ER33801 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND05 , replace;
/* ===== */
/*          step 2 for 2007 family file (n=8289)          */
/* ===== */
infix
    ER36002 2-6
    /* your 2007 variable locations here */
using "[PATH]\FAM2007ER.txt" , clear
;
    label variable ER36002 "2007 FAMILY INTERVIEW (ID) NUMBER" ;
    /* your 2007 variable labels here */
    /* your 2007 missing data here */
    rename ER36002 ER33901 ;
    sort ER33901 ;
save "FAM07" ;
merge 1:m ER33901 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND07 , replace;
/* ===== */
/*          step 2 for 2009 family file (n=8690)          */
/* ===== */
infix
    ER42002 2-6
    /* your 2009 variable locations here */
using "[PATH]\FAM2009ER.txt" , clear
;
    label variable ER42002 "2009 FAMILY INTERVIEW (ID) NUMBER" ;
    /* your 2009 variable labels here */
    /* your 2009 missing data here */
    rename ER42002 ER34001 ;
    sort ER34001 ;
save "FAM09" ;
merge 1:m ER34001 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;

```

```

save FAMIND09 , replace;
/* ===== */
/*      step 2 for 2011 family file (n=8907)      */
/* ===== */
infix
    ER47302 2-6
    /* your 2011 variable locations here */
using "[PATH]\FAM2011ER.txt" , clear
;
    label variable ER47302 "2011 FAMILY INTERVIEW (ID) NUMBER" ;
    /* your 2011 variable labels here */
    /* your 2011 missing data here */
    rename ER47302 ER34101 ;
    sort ER34101 ;
save "'FAM11'" ;
merge 1:m ER34101 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND11 , replace;
/* ===== */
/*      step 2 for 2013 family file (n=9063)      */
/* ===== */
infix
    ER53002 2-6
    /* your 2013 variable locations here */
using "[PATH]\FAM2013ER.txt" , clear
;
    label variable ER53002 "2013 FAMILY INTERVIEW (ID) NUMBER" ;
    /* your 2013 variable labels here */
    /* your 2013 missing data here */
    rename ER53002 ER34201 ;
    sort ER34201 ;
save "'FAM13'" ;
merge 1:m ER34201 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND13 , replace;
/* ===== */
/*      step 2 for 2015 family file (n=9048)      */
/* ===== */
infix
    ER60002 2-6
    /* your 2015 variable locations here */
using "[PATH]\FAM2015ER.txt" , clear
;
    label variable ER60002 "2015 FAMILY INTERVIEW (ID) NUMBER" ;
    /* your 2015 variable labels here */
    /* your 2015 missing data here */
    rename ER60002 ER34301 ;
    sort ER34301 ;
save "'FAM15'" ;
merge 1:m ER34301 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND15 , replace;
/* ===== */
/*      step 2 for 2017 family file (n=9607)      */
/* ===== */
infix
    ER66002 2-6
    /* your 2017 variable locations here */
using "[PATH]\FAM2017ER.txt" , clear
;
    label variable ER66002 "2017 FAMILY INTERVIEW (ID) NUMBER" ;
    /* your 2017 variable labels here */

```

```

/* your 2017 missing data here */
rename ER66002 ER34501 ;
sort ER34501 ;
save "FAM17" ;
merge 1:m ER34501 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND17 , replace;
/* ===== */
/*      step 2 for 2019 family file (n=9569)      */
/* ===== */
infix
ER72002 2-6
/* your 2019 variable locations here */
using "[PATH]\FAM2019ER.txt" , clear
;
label variable ER72002 "2019 FAMILY INTERVIEW (ID) NUMBER" ;
/* your 2019 variable labels here */
/* your 2019 missing data here */
rename ER72002 ER34701 ;
sort ER34701 ;
save "FAM19" ;
merge 1:m ER34701 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND19 , replace;

/* ===== */
/*      step 2 for 2021 family file (n=9207)      */
/* ===== */
infix
ER78002 2-6
/* your 2021 variable locations here */
using "[PATH]\FAM2021ER.txt" , clear
;
label variable ER78002 "2021 FAMILY INTERVIEW (ID) NUMBER" ;
/* your 2021 variable labels here */
/* your 2021 missing data here */
rename ER78002 ER34901 ;
sort ER34901 ;
save "FAM21" ;
merge 1:m ER34901 using INDVARS , keep(using matched) ;
sort ER30001 ER30002 ;
drop _merge ;
save FAMIND21 , replace;

/* ===== */
/*      step 3: merge familiy-individual files on individual identifiers      */
/* ===== */
use FAMIND68 , clear ;
merge ER30001 ER30002
using
FAMIND69 FAMIND70 FAMIND71 FAMIND72 FAMIND73 FAMIND74 FAMIND75
FAMIND76 FAMIND77 FAMIND78 FAMIND79 FAMIND80 FAMIND81 FAMIND82
FAMIND83 FAMIND84 FAMIND85 FAMIND86 FAMIND87 FAMIND88 FAMIND89
FAMIND90 FAMIND91 FAMIND92 FAMIND93 FAMIND94 FAMIND95 FAMIND96
FAMIND97 FAMIND99 FAMIND01 FAMIND03 FAMIND05 FAMIND07 FAMIND09
FAMIND11 FAMIND13 FAMIND15 FAMIND17 FAMIND19 FAMIND21
;
drop _merge* ;
save XYRFIND , replace;

```

B.3 SPSS

B.3.1 SPSS Example - Method 1

```
*-----*
| This example program demonstrates a relatively simple method for |
| merging PSID data. It uses data from 3 different years, subsetting |
| criteria, and the compress option (to conserve disk space).      |
*-----*
* Read in 1990-1992 cross-year individual file and select variables
needed for analysis.
FILE HANDLE IND90_92 / NAME="[PATH]\IND2021ER.txt" LRECL = 6117.
DATA LIST FILE=IND90_92 /
  ER30001      2 - 5
  ER30002      6 - 8
  ER30642     1528 - 1532
  ER30643     1533 - 1534
  ER30644     1535 - 1536
  ER30645     1537 - 1539
  ER30653     1555 - 1555
  ER30657     1563 - 1564
  ER30659     1566 - 1571
  ER30689     1647 - 1650
  ER30690     1651 - 1652
  ER30691     1653 - 1654
  ER30692     1655 - 1657
  ER30699     1672 - 1672
  ER30703     1680 - 1681
  ER30705     1683 - 1688
  ER30707     1690 - 1695
  ER30733     1764 - 1767
  ER30734     1768 - 1769
  ER30735     1770 - 1771
  ER30736     1772 - 1774
  ER30744     1790 - 1790
  ER30748     1798 - 1799
  ER30750     1801 - 1806
  ER30752     1808 - 1813
  ER30805     1907 - 1913
  ER32000     2057 - 2057
  ER32022     2111 - 2112
  ER32049     2187 - 2187
.
VARIABLE LABELS
  ER30001 "1968 INTERVIEW NUMBER"
  ER30002 "PERSON NUMBER 68"
  ER30642 "1990 INTERVIEW NUMBER"
  ER30643 "SEQUENCE NUMBER 90"
  ER30644 "RELATION TO HEAD 90"
  ER30645 "AGE OF INDIVIDUAL 90"
  ER30653 "EMPLOYMENT STAT-IND 90"
  ER30657 "COMPLETED EDUC-IND 90"
  ER30659 "TOT TXBL INCOME-IND 90"
  ER30689 "1991 INTERVIEW NUMBER"
  ER30690 "SEQUENCE NUMBER 91"
  ER30691 "RELATION TO HEAD 91"
  ER30692 "AGE OF INDIVIDUAL 91"
  ER30699 "EMPLOYMENT STAT-IND 91"
  ER30703 "COMPLETED EDUC-IND 91"
  ER30705 "TOT LABOR INCOME-IND 91"
  ER30707 "TOT ASSET INCOME-IND 91"
  ER30733 "1992 INTERVIEW NUMBER"
  ER30734 "SEQUENCE NUMBER 92"
  ER30735 "RELATION TO HEAD 92"
  ER30736 "AGE OF INDIVIDUAL 92"
```

```
ER30744 "EMPLOYMENT STAT 92"  
ER30748 "COMPLETED EDUCATION 92"  
ER30750 "TOT LABOR INCOME 92"  
ER30752 "TOT ASSET INCOME 92"  
ER30805 "COMBINED IND WEIGHT 92"  
ER32000 "SEX OF INDIVIDUAL"  
ER32022 "# LIVE BIRTHS TO THIS INDIVIDUAL"  
ER32049 "LAST KNOWN MARITAL STATUS"
```

MISSING VALUES

```
ER30645 (999)  
ER30657 (99)  
ER30692 (999)  
ER30703 (99)  
ER30736 (999)  
ER30748 (99)  
ER32022 (98)  
ER32049 (8)
```

RENAME VARIABLES

```
(ER30642=ID90)  
(ER30689=ID91)  
(ER30733=ID92)
```

* This example selects those who were ever heads or wives/"wives"
between 1990 and 1992.

```
SELECT IF (ER30643 EQ 01 AND ER30644 EQ 10) OR  
  (ER30643 EQ 02 AND (ER30644 EQ 20 OR ER30644 EQ 22)) OR  
  (ER30690 EQ 01 AND ER30691 EQ 10) OR  
  (ER30690 EQ 02 AND (ER30691 EQ 20 OR ER30691 EQ 22)) OR  
  (ER30734 EQ 01 AND ER30735 EQ 10) OR  
  (ER30734 EQ 02 AND (ER30735 EQ 20 OR ER30735 EQ 22))
```

* Sort active file by id90.

SORT CASES BY ID90.

SAVE OUTFILE="IND90_92.SYS" / COMPRESS.

EXECUTE.

* Read in 1990 family file and select variables needed for analysis.

FILE HANDLE FAM90 / NAME="[PATH]\FAM1990.txt" LRECL=2332 .

DATA LIST FILE=FAM90 /

```
V17702 2 - 6  
V17836 281 - 286  
V18262 1175 - 1177  
V18564 1656 - 1658  
V18814 2046 - 2046  
V18878 2188 - 2193  
V18887 2230 - 2234 (2)  
V18888 2235 - 2239 (2)
```

VARIABLE LABELS

```
V17702 "1990 INTERVIEW NUMBER"  
V17836 "WIFE 89 LABOR/WAGE"  
V18262 "C9-10 OCC-LAST JOB (H-U)"  
V18564 "E9-10 OCC-LAST JOB (W-U)"  
V18814 "M32 RACE OF HEAD (1 MEN)"  
V18878 "TOTAL HEAD LABOR Y 89"  
V18887 "HEAD 89 AVG HRLY EARNING"  
V18888 "WIFE 89 AVG HRLY EARNING"
```

MISSING VALUES

```
V18262 (999)  
V18564 (999)  
V18814 (9)
```

RENAME VARIABLES

```

(V17702=ID90)
.
* Sort active file by id90.
SORT CASES BY ID90.
* Merge ind90_92.sys and active file by id90.
MATCH FILES FILE="IND90_92.SYS" / TABLE=* / BY ID90.
* Sort active file by id91.
SORT CASES BY ID91.
SAVE OUTFILE="FAM_IND.SYS" / COMPRESS.
* Read in 1991 family file and select variables needed for analysis.
FILE HANDLE FAM91 / NAME="[PATH]\FAM1991.txt" LRECL = 2336.
DATA LIST FILE=FAM91 /
  V19002 2 - 5
  V19136 281 - 286
  V19562 1175 - 1177
  V19864 1656 - 1658
  V20114 2046
  V20178 2188 - 2193
  V20187 2230 - 2234 (2)
  V20188 2235 - 2239 (2)
.
VARIABLE LABELS
  V19002 "1991 INTERVIEW NUMBER"
  V19136 "WIFE 90 LABOR/WAGE"
  V19562 "C9-10 OCC-LAST JOB (H-U)"
  V19864 "E9-10 OCC-LAST JOB (W-U)"
  V20114 "L32 RACE OF HEAD (1 MEN)"
  V20178 "TOTAL HEAD LABOR Y 90"
  V20187 "HEAD 90 AVG HRLY EARNING"
  V20188 "WIFE 90 AVG HRLY EARNING".
MISSING VALUES
  V19562 (999)
  V19864 (999)
  V20114 (9).
RENAME VARIABLES
  (V19002=ID91).
* Sort active file by id91.
SORT CASES BY ID91.
* Merge fam_ind.sys and active file by id91.
MATCH FILES FILE="FAM_IND.SYS" / TABLE=* / BY ID91.
* Sort active file by id92.
SORT CASES BY ID92.
SAVE OUTFILE="FAM_IND.SYS" / COMPRESS.
* Read in 1992 family file and select variables needed for analysis.
FILE HANDLE FAM92 / NAME="[PATH]\FAM1992.txt" LRECL = 2294.
DATA LIST FILE=FAM92 /
  V20302 2 - 5
  V20436 282 - 287
  V20862 1189 - 1191
  V21164 1670 - 1672
  V21420 2066
  V21484 2172 - 2177
  V21493 2218 - 2222 (2)
  V21494 2223 - 2227 (2)
.
VARIABLE LABELS
  V20302 "1992 INTERVIEW NUMBER"
  V20436 "WIFE 91 LABOR/WAGE"
  V20862 "C9-10 OCC-LAST JOB (H-U)"
  V21164 "E9-10 OCC-LAST JOB (W-U)"
  V21420 "M32 RACE OF HEAD (1 MEN)"
  V21484 "TOTAL HEAD LABOR Y 91"
  V21493 "HEAD 91 AVG HRLY EARNING"
  V21494 "WIFE 91 AVG HRLY EARNING".
MISSING VALUES

```



```
V20862 (999)
V21164 (999)
V21420 (9).
RENAME VARIABLES
(V20302=ID92).
* Sort active file by id92.
SORT CASES BY ID92.
* Merge fam_ind.sys and active file by id92.
MATCH FILES FILE="FAM_IND.SYS" / TABLE=* / BY ID92.
SAVE OUTFILE="FAM_IND.SYS" / COMPRESS.
```

B.3.2 SPSS Example - Method 2

```
*-----*
|
| MERGE USING MULTIPLE FAMILY-INDIVIDUAL FILES
|
| Step 1: Subset family-level id's and selected variables
| and subset cases from xyr-individual file
|
| Step 2a: Subset year-n family file
| Step 2b: Sort year-n family file from 2a by year-n family id
| Step 2c: Sort xyr-individual-id file from step 1 by year-n family id
| Step 2d: Merge sorted xyr-individual-id file from 2c
| with sorted year-n subsetted family file from 2b
| (a one-to-many [family-to-individual] match)
| Step 2e: Sort resulting year-n family-individual file from 2d
| by individual ids
|
| ... Repeat Steps 2 for all other years
|
| Step 3: Merge family-individual files from step 2e
|
*-----*
FILE HANDLE indvars / NAME="[PATH]\IND2021ER.txt" LRECL = 6117.
DATA LIST FILE=indvars /
ER30001 2 - 5
ER30002 6 - 8
ER30020 44 - 47
ER30043 97 - 100
ER30067 152 - 155
ER30091 207 - 210
ER30117 265 - 268
ER30138 317 - 320
ER30160 370 - 373
ER30188 436 - 439
ER30217 503 - 506
ER30246 571 - 574
ER30283 648 - 651
ER30313 718 - 721
ER30343 788 - 791
ER30373 858 - 861
ER30399 919 - 922
ER30429 992 - 995
ER30463 1077 - 1080
ER30498 1167 - 1170
ER30535 1259 - 1262
ER30570 1348 - 1351
ER30606 1438 - 1441
ER30642 1528 - 1532
ER30689 1647 - 1650
ER30733 1764 - 1767
ER30806 1914 - 1918
ER33101 2201 - 2205
ER33201 2300 - 2304
ER33301 2493 - 2496
ER33401 2577 - 2581
ER33501 2662 - 2666
ER33601 2856 - 2859
ER33701 3038 - 3042
ER33801 3193 - 3197
ER33901 3492 - 3496
ER34001 3770 - 3774
ER34101 3992 - 3996
ER34201 4229 - 4233
ER34301 4495 - 4499
```

ER34501 4869 - 4873
ER34701 5249 - 5253
ER34901 5683 - 5687

/* your individual variable locations here */

VARIABLE LABELS

ER30001 "1968 INTERVIEW NUMBER"
ER30002 "PERSON NUMBER 68"
ER30020 "1969 INTERVIEW NUMBER"
ER30043 "1970 INTERVIEW NUMBER"
ER30067 "1971 INTERVIEW NUMBER"
ER30091 "1972 INTERVIEW NUMBER"
ER30117 "1973 INTERVIEW NUMBER"
ER30138 "1974 INTERVIEW NUMBER"
ER30160 "1975 INTERVIEW NUMBER"
ER30188 "1976 INTERVIEW NUMBER"
ER30217 "1977 INTERVIEW NUMBER"
ER30246 "1978 INTERVIEW NUMBER"
ER30283 "1979 INTERVIEW NUMBER"
ER30313 "1980 INTERVIEW NUMBER"
ER30343 "1981 INTERVIEW NUMBER"
ER30373 "1982 INTERVIEW NUMBER"
ER30399 "1983 INTERVIEW NUMBER"
ER30429 "1984 INTERVIEW NUMBER"
ER30463 "1985 INTERVIEW NUMBER"
ER30498 "1986 INTERVIEW NUMBER"
ER30535 "1987 INTERVIEW NUMBER"
ER30570 "1988 INTERVIEW NUMBER"
ER30606 "1989 INTERVIEW NUMBER"
ER30642 "1990 INTERVIEW NUMBER"
ER30689 "1991 INTERVIEW NUMBER"
ER30733 "1992 INTERVIEW NUMBER"
ER30806 "1993 INTERVIEW NUMBER"
ER33101 "1994 INTERVIEW NUMBER"
ER33201 "1995 INTERVIEW NUMBER"
ER33301 "1996 INTERVIEW NUMBER"
ER33401 "1997 INTERVIEW NUMBER"
ER33501 "1999 INTERVIEW NUMBER"
ER33601 "2001 INTERVIEW NUMBER"
ER33701 "2003 INTERVIEW NUMBER"
ER33801 "2005 INTERVIEW NUMBER"
ER33901 "2007 INTERVIEW NUMBER"
ER34001 "2009 INTERVIEW NUMBER"
ER34101 "2011 INTERVIEW NUMBER"
ER34201 "2013 INTERVIEW NUMBER"
ER34301 "2015 INTERVIEW NUMBER"
ER34501 "2017 INTERVIEW NUMBER"
ER34701 "2019 INTERVIEW NUMBER"
ER34901 "2021 INTERVIEW NUMBER"

/* your individual variable labels here */

MISSING VALUES

/* your individual missing values here */

/* your filter for case selection, if any, here */

SAVE OUTFILE="indvars.sys".

* =====

Step 2: subset data from 1968 family file (n=4802)

=====.

FILE HANDLE fam68 / NAME='[PATH]\FAM1968.txt' LRECL=771.

DATA LIST FILE=fam68 /

V3 7-10

```

        /* your 1968 variable locations here */
        .
VARIABLE LABELS
  V3 ="INTERVIEW NUMBER 68 1:6-9"
  /* your 1968 variable labels here */
        .
MISSING VALUES
  /* your 1968 missing data values here */
        .
SORT CASES BY V3.
SAVE OUTFILE="fam68.sys".
GET FILE="indvars.sys".
SORT CASES BY ER30001.
MATCH FILES FILE=* / TABLE="fam68.sys"/ RENAME=(V3=ER30001)
  / BY=ER30001.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind68.sys".
* =====
  Step 2: subset data from 1969 family file (n=4460)
  =====.
FILE HANDLE fam69 / NAME='[PATH]\FAM1969.txt' LRECL=1079.
DATA LIST FILE=fam69 /
  V442 2-5
  /* your 1969 variable locations here */
        .
VARIABLE LABELS
  V442 ="INTERVIEW NUMBER 69 11:6-9"
  /* your 1969 variable labels here */
        .
MISSING VALUES
  /* your 1969 missing data values here */
        .
SORT CASES BY V442.
SAVE OUTFILE="fam69.sys".
GET FILE="indvars.sys".
SORT CASES BY ER30020.
MATCH FILES FILE=* / TABLE="fam69.sys"/ RENAME=(V442=ER30020)
  / BY=ER30020.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind69.sys".
* =====
  Step 2: subset data from 1970 family file (n=4645)
  =====.
FILE HANDLE fam70 / NAME='[PATH]\FAM1970.txt' LRECL=1199.
DATA LIST FILE=fam70 /
  V1102 2-5
  /* your 1970 variable locations here */
        .
VARIABLE LABELS
  V1102 ="1970 INT # 21:6-9"
  /* your 1970 variable labels here */
        .
MISSING VALUES
  /* your 1970 missing data values here */
        .
SORT CASES BY V1102.
SAVE OUTFILE="fam70.sys".
GET FILE="indvars.sys".
SORT CASES BY ER30043.
MATCH FILES FILE=* / TABLE="fam70.sys"/ RENAME=(V1102=ER30043)
  / BY=ER30043.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind70.sys".
* =====
  Step 2: subset data from 1971 family file (n=4840)
  =====.

```

```

=====
FILE HANDLE fam71 / NAME='[PATH]\FAM1971.txt' LRECL=974.
DATA LIST FILE=fam71 /
  V1802 2-5
  /* your 1971 variable locations here */
.
VARIABLE LABELS
  V1802 ="1971 ID NO."
  /* your 1971 variable labels here */
.
MISSING VALUES
  /* your 1971 missing data values here */
.
SORT CASES BY V1802.
SAVE OUTFILE="fam71.sys".
GET FILE="indvars.sys".
SORT CASES BY ER30067.
MATCH FILES FILE=* / TABLE="fam71.sys"/ RENAME=(V1802=ER30067)
  / BY=ER30067.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind71.sys".
* =====
  Step 2: subset data from 1972 family file (n=5060)
  =====
FILE HANDLE fam72 / NAME='[PATH]\FAM1972.txt' LRECL=1020.
DATA LIST FILE=fam72 /
  V2402 2-5
  /* your 1972 variable locations here */
.
VARIABLE LABELS
  V2402 ="1972 INT # 46:6-9"
  /* your 1972 variable labels here */
.
MISSING VALUES
  /* your 1972 missing data values here */
.
SORT CASES BY V2402.
SAVE OUTFILE="fam72.sys".
GET FILE="indvars.sys".
SORT CASES BY ER30091.
MATCH FILES FILE=* / TABLE="fam72.sys"/ RENAME=(V2402=ER30091)
  / BY=ER30091.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind72.sys".
* =====
  Step 2: subset data from 1973 family file (n=5285)
  =====
FILE HANDLE fam73 / NAME='[PATH]\FAM1973.txt' LRECL=582.
DATA LIST FILE=fam73 /
  V3002 2-5
  /* your 1973 variable locations here */
.
VARIABLE LABELS
  V3002 ="1973 INT # 59:6-9"
  /* your 1973 variable labels here */
.
MISSING VALUES
  /* your 1973 missing data values here */
.
SORT CASES BY V3002.
SAVE OUTFILE="fam73.sys".
GET FILE="indvars.sys".
SORT CASES BY ER30117.
MATCH FILES FILE=* / TABLE="fam73.sys"/ RENAME=(V3002=ER30117)
  / BY=ER30117.

```

```

SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind73.sys".
* =====
      Step 2: subset data from 1974 family file (n=5517)
      =====.
FILE HANDLE fam74 / NAME='[PATH]\FAM1974.txt' LRECL=633.
DATA LIST FILE=fam74 /
      V3402 2-5
      /* your 1974 variable locations here */
.
VARIABLE LABELS
      V3402 ="1974 ID NUMBER"
      /* your 1974 variable labels here */
.
MISSING VALUES
      /* your 1974 missing data values here */
.
SORT CASES BY V3402.
SAVE OUTFILE="fam74.sys".
GET FILE="indvars.sys".
SORT CASES BY ER30138.
MATCH FILES FILE=* / TABLE="fam74.sys" / RENAME=(V3402=ER30138)
      / BY=ER30138.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind74.sys".
* =====
      Step 2: subset data from 1975 family file (n=5725)
      =====.
FILE HANDLE fam75 / NAME='[PATH]\FAM1975.txt' LRECL=777.
DATA LIST FILE=fam75 /
      V3802 2-5
      /* your 1975 variable locations here */
.
VARIABLE LABELS
      V3802 ="1975 INT # 80:6-9"
      /* your 1975 variable labels here */
.
MISSING VALUES
      /* your 1975 missing data values here */
.
SORT CASES BY V3802.
SAVE OUTFILE="fam75.sys".
GET FILE="indvars.sys".
SORT CASES BY ER30160.
MATCH FILES FILE=* / TABLE="fam75.sys" / RENAME=(V3802=ER30160)
      / BY=ER30160.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind75.sys".
* =====
      Step 2: subset data from 1976 family file (n=5862)
      =====.
FILE HANDLE fam76 / NAME='[PATH]\FAM1976.txt' LRECL=1499.
DATA LIST FILE=fam76 /
      V4302 2-5
      /* your 1976 variable locations here */
.
VARIABLE LABELS
      V4302 ="1976 ID NUMBER 6V2"
      /* your 1976 variable labels here */
.
MISSING VALUES
      /* your 1976 missing data values here */
.
SORT CASES BY V4302.
SAVE OUTFILE="fam76.sys".

```

```

GET FILE="indvars.sys".
SORT CASES BY ER30188.
MATCH FILES FILE=* / TABLE="fam76.sys"/ RENAME=(V4302=ER30188)
/ BY=ER30188.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind76.sys".
* =====
Step 2: subset data from 1977 family file (n=6007)
=====
FILE HANDLE fam77 / NAME='[PATH]\FAM1977.txt' LRECL=963.
DATA LIST FILE=fam77 /
V5202 2-5
/* your 1977 variable locations here */
.
VARIABLE LABELS
V5202 ="1977 ID"
/* your 1977 variable labels here */
.
MISSING VALUES
/* your 1977 missing data values here */
.
SORT CASES BY V5202.
SAVE OUTFILE="fam77.sys".
GET FILE="indvars.sys".
SORT CASES BY ER30217.
MATCH FILES FILE=* / TABLE="fam77.sys"/ RENAME=(V5202=ER30217)
/ BY=ER30217.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind77.sys".
* =====
Step 2: subset data from 1978 family file (n=6154)
=====
FILE HANDLE fam78 / NAME='[PATH]\FAM1978.txt' LRECL=988.
DATA LIST FILE=fam78 /
V5702 2-5
/* your 1978 variable locations here */
.
VARIABLE LABELS
V5702 ="1978 ID"
/* your 1978 variable labels here */
.
MISSING VALUES
/* your 1978 missing data values here */
.
SORT CASES BY V5702.
SAVE OUTFILE="fam78.sys".
GET FILE="indvars.sys".
SORT CASES BY ER30246.
MATCH FILES FILE=* / TABLE="fam78.sys"/ RENAME=(V5702=ER30246)
/ BY=ER30246.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind78.sys".
* =====
Step 2: subset data from 1979 family file (n=6373)
=====
FILE HANDLE fam79 / NAME='[PATH]\FAM1979.txt' LRECL=1062.
DATA LIST FILE=fam79 /
V6302 2-5
/* your 1979 variable locations here */
.
VARIABLE LABELS
V6302 ="1979 ID"
/* your 1979 variable labels here */
.
MISSING VALUES

```

```

/* your 1979 missing data values here */
.
SORT CASES BY V6302.
SAVE OUTFILE="fam79.sys".
GET FILE="indvars.sys".
SORT CASES BY ER30283.
MATCH FILES FILE=* / TABLE="fam79.sys" / RENAME=(V6302=ER30283)
/ BY=ER30283.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind79.sys".
* =====
Step 2: subset data from 1980 family file (n=6533)
=====
FILE HANDLE fam80 / NAME='[PATH]\FAM1980.txt' LRECL=1123.
DATA LIST FILE=fam80 /
V6902 2-5
/* your 1980 variable locations here */
.
VARIABLE LABELS
V6902 ="1980 INTERVIEW NUMBER"
/* your 1980 variable labels here */
.
MISSING VALUES
/* your 1980 missing data values here */
.
SORT CASES BY V6902.
SAVE OUTFILE="fam80.sys".
GET FILE="indvars.sys".
SORT CASES BY ER30313.
MATCH FILES FILE=* / TABLE="fam80.sys" / RENAME=(V6902=ER30313)
/ BY=ER30313.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind80.sys".
* =====
Step 2: subset data from 1981 family file (n=6620)
=====
FILE HANDLE fam81 / NAME='[PATH]\FAM1981.txt' LRECL=1250.
DATA LIST FILE=fam81 /
V7502 2-5
/* your 1981 variable locations here */
.
VARIABLE LABELS
V7502 ="1981 INTERVIEW NUMBER"
/* your 1981 variable labels here */
.
MISSING VALUES
/* your 1981 missing data values here */
.
SORT CASES BY V7502.
SAVE OUTFILE="fam81.sys".
GET FILE="indvars.sys".
SORT CASES BY ER30343.
MATCH FILES FILE=* / TABLE="fam81.sys" / RENAME=(V7502=ER30343)
/ BY=ER30343.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind81.sys".
* =====
Step 2: subset data from 1982 family file (n=6742)
=====
FILE HANDLE fam82 / NAME='[PATH]\FAM1982.txt' LRECL=1117.
DATA LIST FILE=fam82 /
V8202 2-5
/* your 1982 variable locations here */
.
VARIABLE LABELS

```



```

V8202 ="1982 INTERVIEW NUMBER"
/* your 1982 variable labels here */
.
MISSING VALUES
/* your 1982 missing data values here */
.
SORT CASES BY V8202.
SAVE OUTFILE="fam82.sys".
GET FILE="indvars.sys".
SORT CASES BY ER30373.
MATCH FILES FILE=* / TABLE="fam82.sys"/ RENAME=(V8202=ER30373)
/ BY=ER30373.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind82.sys".
* =====
Step 2: subset data from 1983 family file (n=6852)
=====
FILE HANDLE fam83 / NAME='[PATH]\FAM1983.txt' LRECL=1315.
DATA LIST FILE=fam83 /
V8802 2-5
/* your 1983 variable locations here */
.
VARIABLE LABELS
V8802 ="1983 INTERVIEW NUMBER"
/* your 1983 variable labels here */
.
MISSING VALUES
/* your 1983 missing data values here */
.
SORT CASES BY V8802.
SAVE OUTFILE="fam83.sys".
GET FILE="indvars.sys".
SORT CASES BY ER30399.
MATCH FILES FILE=* / TABLE="fam83.sys"/ RENAME=(V8802=ER30399)
/ BY=ER30399.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind83.sys".
* =====
Step 2: subset data from 1984 family file (n=6918)
=====
FILE HANDLE fam84 / NAME='[PATH]\FAM1984.txt' LRECL=2009.
DATA LIST FILE=fam84 /
V10002 2-5
/* your 1984 variable locations here */
.
VARIABLE LABELS
V10002 ="1984 INTERVIEW NUMBER"
/* your 1984 variable labels here */
.
MISSING VALUES
/* your 1984 missing data values here */
.
SORT CASES BY V10002.
SAVE OUTFILE="fam84.sys".
GET FILE="indvars.sys".
SORT CASES BY ER30429.
MATCH FILES FILE=* / TABLE="fam84.sys"/ RENAME=(V10002=ER30429)
/ BY=ER30429.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind84.sys".
* =====
Step 2: subset data from 1985 family file (n=7032)
=====
FILE HANDLE fam85 / NAME='[PATH]\FAM1985.txt' LRECL=2472.
DATA LIST FILE=fam85 /

```

```

V11102 2-5
/* your 1985 variable locations here */
.
VARIABLE LABELS
V11102 ="1985 INTERVIEW NUMBER"
/* your 1985 variable labels here */
.
MISSING VALUES
/* your 1985 missing data values here */
.
SORT CASES BY V11102.
SAVE OUTFILE="fam85.sys".
GET FILE="indvars.sys".
SORT CASES BY ER30463.
MATCH FILES FILE=* / TABLE="fam85.sys"/ RENAME=(V11102=ER30463)
/ BY=ER30463.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind85.sys".
* =====
Step 2: subset data from 1986 family file (n=7018)
=====
FILE HANDLE fam86 / NAME='[PATH]\FAM1986.txt' LRECL=2234.
DATA LIST FILE=fam86 /
V12502 2-5
/* your 1986 variable locations here */
.
VARIABLE LABELS
V12502 ="1986 INTERVIEW NUMBER"
/* your 1986 variable labels here */
.
MISSING VALUES
/* your 1986 missing data values here */
.
SORT CASES BY V12502.
SAVE OUTFILE="fam86.sys".
GET FILE="indvars.sys".
SORT CASES BY ER30498.
MATCH FILES FILE=* / TABLE="fam86.sys"/ RENAME=(V12502=ER30498)
/ BY=ER30498.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind86.sys".
* =====
Step 2: subset data from 1987 family file (n=7061)
=====
FILE HANDLE fam87 / NAME='[PATH]\FAM1987.txt' LRECL=2053.
DATA LIST FILE=fam87 /
V13702 2-5
/* your 1987 variable locations here */
.
VARIABLE LABELS
V13702 ="1987 INTERVIEW NUMBER"
/* your 1987 variable labels here */
.
MISSING VALUES
/* your 1987 missing data values here */
.
SORT CASES BY V13702.
SAVE OUTFILE="fam87.sys".
GET FILE="indvars.sys".
SORT CASES BY ER30535.
MATCH FILES FILE=* / TABLE="fam87.sys"/ RENAME=(V13702=ER30535)
/ BY=ER30535.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind87.sys".
* =====

```

```

Step 2: subset data from 1988 family file (n=7114)
=====
FILE HANDLE fam88 / NAME='[PATH]\FAM1988.txt' LRECL=2730.
DATA LIST FILE=fam88 /
V14802 2-5
/* your 1988 variable locations here */
.
VARIABLE LABELS
V14802 ="1988 INTERVIEW NUMBER"
/* your 1988 variable labels here */
.
MISSING VALUES
/* your 1988 missing data values here */
.
SORT CASES BY V14802.
SAVE OUTFILE="fam88.sys".
GET FILE="indvars.sys".
SORT CASES BY ER30570.
MATCH FILES FILE=* / TABLE="fam88.sys"/ RENAME=(V14802=ER30570)
/ BY=ER30570.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind88.sys".
* =====
Step 2: subset data from 1989 family file (n=7114)
=====
FILE HANDLE fam89 / NAME='[PATH]\FAM1989.txt' LRECL=2517.
DATA LIST FILE=fam89 /
V16302 2-5
/* your 1989 variable locations here */
.
VARIABLE LABELS
V16302 ="1989 INTERVIEW NUMBER"
/* your 1989 variable labels here */
.
MISSING VALUES
/* your 1989 missing data values here */
.
SORT CASES BY V16302.
SAVE OUTFILE="fam89.sys".
GET FILE="indvars.sys".
SORT CASES BY ER30606.
MATCH FILES FILE=* / TABLE="fam89.sys"/ RENAME=(V16302=ER30606)
/ BY=ER30606.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind89.sys".
* =====
Step 2: subset data from 1990 family file (n=9371)
=====
FILE HANDLE fam90 / NAME='[PATH]\FAM1990.txt' LRECL=2332.
DATA LIST FILE=fam90 /
V17702 2-6
/* your 1990 variable locations here */
.
VARIABLE LABELS
V17702 ="1990 INTERVIEW NUMBER 90"
/* your 1990 variable labels here */
.
MISSING VALUES
/* your 1990 missing data values here */
.
SORT CASES BY V17702.
SAVE OUTFILE="fam90.sys".
GET FILE="indvars.sys".
SORT CASES BY ER30642.
MATCH FILES FILE=* / TABLE="fam90.sys"/ RENAME=(V17702=ER30642)

```

```

    / BY=ER30642.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind90.sys".
* =====
    Step 2: subset data from 1991 family file (n=9363)
=====
FILE HANDLE fam91 / NAME='[PATH]\FAM1991.txt' LRECL=2336.
DATA LIST FILE=fam91 /
    V19002 2-5
    /* your 1991 variable locations here */
.
VARIABLE LABELS
    V19002 ="1991 INTERVIEW NUMBER"
    /* your 1991 variable labels here */
.
MISSING VALUES
    /* your 1991 missing data values here */
.
SORT CASES BY V19002.
SAVE OUTFILE="fam91.sys".
GET FILE="indvars.sys".
SORT CASES BY ER30689.
MATCH FILES FILE=* / TABLE="fam91.sys" / RENAME=(V19002=ER30689)
    / BY=ER30689.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind91.sys".
* =====
    Step 2: subset data from 1992 family file (n=9829)
=====
FILE HANDLE fam92 / NAME='[PATH]\FAM1992.txt' LRECL=2325.
DATA LIST FILE=fam92 /
    V20302 2-5
    /* your 1992 variable locations here */
.
VARIABLE LABELS
    V20302 ="1992 INTERVIEW NUMBER"
    /* your 1992 variable labels here */
.
MISSING VALUES
    /* your 1992 missing data values here */
.
SORT CASES BY V20302.
SAVE OUTFILE="fam92.sys".
GET FILE="indvars.sys".
SORT CASES BY ER30733.
MATCH FILES FILE=* / TABLE="fam92.sys" / RENAME=(V20302=ER30733)
    / BY=ER30733.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind92.sys".
* =====
    Step 2: subset data from 1993 family file (n=9977)
=====
FILE HANDLE fam93 / NAME='[PATH]\FAM1993.txt' LRECL=3329.
DATA LIST FILE=fam93 /
    V21602 2-6
    /* your 1993 variable locations here */
.
VARIABLE LABELS
    V21602 ="1993 INTERVIEW NUMBER"
    /* your 1993 variable labels here */
.
MISSING VALUES
    /* your 1993 missing data values here */
.
SORT CASES BY V21602.

```

```

SAVE OUTFILE="fam93.sys".
GET FILE="indvars.sys".
SORT CASES BY ER30806.
MATCH FILES FILE=* / TABLE="fam93.sys"/ RENAME=(V21602=ER30806)
  / BY=ER30806.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind93.sys".
* =====
  Step 2: subset data from 1994 family file (n=10764)
  =====.
FILE HANDLE fam94 / NAME='[PATH]\FAM1994ER.txt' LRECL=4379.
DATA LIST FILE=fam94 /
  ER2002 2-6
  /* your 1994 variable locations here */
.
VARIABLE LABELS
  ER2002 ="1994 INTERVIEW #"
  /* your 1994 variable labels here */
.
MISSING VALUES
  /* your 1994 missing data values here */
.
SORT CASES BY ER2002.
SAVE OUTFILE="fam94.sys".
GET FILE="indvars.sys".
SORT CASES BY ER33101.
MATCH FILES FILE=* / TABLE="fam94.sys"/ RENAME=(ER2002=ER33101)
  / BY=ER33101.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind94.sys".
* =====
  Step 2: subset data from 1995 family file (n=10401)
  =====.
FILE HANDLE fam95 / NAME='[PATH]\FAM1995ER.txt' LRECL=4131.
DATA LIST FILE=fam95 /
  ER5002 2-6
  /* your 1995 variable locations here */
.
VARIABLE LABELS
  ER5002 ="1995 INTERVIEW #"
  /* your 1995 variable labels here */
.
MISSING VALUES
  /* your 1995 missing data values here */
.
SORT CASES BY ER5002.
SAVE OUTFILE="fam95.sys".
GET FILE="indvars.sys".
SORT CASES BY ER33201.
MATCH FILES FILE=* / TABLE="fam95.sys"/ RENAME=(ER5002=ER33201)
  / BY=ER33201.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind95.sys".
* =====
  Step 2: subset data from 1996 family file (n=8511)
  =====.
FILE HANDLE fam96 / NAME='[PATH]\FAM1996ER.txt' LRECL=4614.
DATA LIST FILE=fam96 /
  ER7002 2-5
  /* your 1996 variable locations here */
.
VARIABLE LABELS
  ER7002 ="1996 INTERVIEW #"
  /* your 1996 variable labels here */
.

```

```

MISSING VALUES
  /* your 1996 missing data values here */
.
SORT CASES BY ER7002.
SAVE OUTFILE="fam96.sys".
GET FILE="indvars.sys".
SORT CASES BY ER33301.
MATCH FILES FILE=* / TABLE="fam96.sys"/ RENAME=(ER7002=ER33301)
  / BY=ER33301.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind96.sys".
* =====
  Step 2: subset data from 1997 family file (n=6747)
  =====.
FILE HANDLE fam97 / NAME='[PATH]\FAM1997ER.txt' LRECL=4473.
DATA LIST FILE=fam97 /
  ER10002 2-6
  /* your 1997 variable locations here */
.
VARIABLE LABELS
  ER10002 = "1997 INTERVIEW #"
  /* your 1997 variable labels here */
.
MISSING VALUES
  /* your 1997 missing data values here */
.
SORT CASES BY ER10002.
SAVE OUTFILE="fam97.sys".
GET FILE="indvars.sys".
SORT CASES BY ER33401.
MATCH FILES FILE=* / TABLE="fam97.sys"/ RENAME=(ER10002=ER33401)
  / BY=ER33401.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind97.sys".
* =====
  Step 2: subset data from 1999 family file (n=6997)
  =====.
FILE HANDLE fam99 / NAME='[PATH]\FAM1999ER.txt' LRECL=6938.
DATA LIST FILE=fam99 /
  ER13002 2-6
  /* your 1999 variable locations here */
.
VARIABLE LABELS
  ER13002 = "1999 FAMILY INTERVIEW (ID) NUMBER"
  /* your 1999 variable labels here */
.
MISSING VALUES
  /* your 1999 missing data values here */
.
SORT CASES BY ER13002.
SAVE OUTFILE="fam99.sys".
GET FILE="indvars.sys".
SORT CASES BY ER33501.
MATCH FILES FILE=* / TABLE="fam99.sys"/ RENAME=(ER13002=ER33501)
  / BY=ER33501.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind99.sys".
* =====
  Step 2: subset data from 2001 family file (n=7406)
  =====.
FILE HANDLE fam01 / NAME='[PATH]\FAM2001ER.txt' LRECL=6751.
DATA LIST FILE=fam01 /
  ER17002 2-5
  /* your 2001 variable locations here */
.

```

```

VARIABLE LABELS
  ER17002 = "2001 FAMILY INTERVIEW (ID) NUMBER"
  /* your 2001 variable labels here */
.
MISSING VALUES
  /* your 2001 missing data values here */
.
SORT CASES BY ER17002.
SAVE OUTFILE="fam01.sys".
GET FILE="indvars.sys".
SORT CASES BY ER33601.
MATCH FILES FILE=* / TABLE="fam01.sys" / RENAME=(ER17002=ER33601)
  / BY=ER33601.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind01.sys".
* =====
  Step 2: subset data from 2003 family file (n=7822)
  =====.
FILE HANDLE fam03 / NAME='[PATH]\FAM2003ER.txt' LRECL=6196.
DATA LIST FILE=fam03 /
  ER21002 2-6
  /* your 2003 variable locations here */
.
VARIABLE LABELS
  ER21002 = "2003 FAMILY INTERVIEW (ID) NUMBER"
  /* your 2003 variable labels here */
.
MISSING VALUES
  /* your 2003 missing data values here */
.
SORT CASES BY ER21002.
SAVE OUTFILE="fam03.sys".
GET FILE="indvars.sys".
SORT CASES BY ER33701.
MATCH FILES FILE=* / TABLE="fam03.sys" / RENAME=(ER21002=ER33701)
  / BY=ER33701.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind03.sys".
* =====
  Step 2: subset data from 2005 family file (n=8002)
  =====.
FILE HANDLE fam05 / NAME='[PATH]\FAM2005ER.txt' LRECL=6333.
DATA LIST FILE=fam05 /
  ER25002 2-6
  /* your 2005 variable locations here */
.
VARIABLE LABELS
  ER25002 = "2005 FAMILY INTERVIEW (ID) NUMBER"
  /* your 2005 variable labels here */
.
MISSING VALUES
  /* your 2005 missing data values here */
.
SORT CASES BY ER25002.
SAVE OUTFILE="fam05.sys".
GET FILE="indvars.sys".
SORT CASES BY ER33801.
MATCH FILES FILE=* / TABLE="fam05.sys" / RENAME=(ER25002=ER33801)
  / BY=ER33801.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind05.sys".
* =====
  Step 2: subset data from 2007 family file (n=8289)
  =====.
FILE HANDLE fam07 / NAME='[PATH]\FAM2007ER.txt' LRECL=8314.

```

```

DATA LIST FILE=fam07 /
  ER36002 2-6
  /* your 2007 variable locations here */
.
VARIABLE LABELS
  ER36002 = "2007 FAMILY INTERVIEW (ID) NUMBER"
  /* your 2007 variable labels here */
.
MISSING VALUES
  /* your 2007 missing data values here */
.
SORT CASES BY ER36002.
SAVE OUTFILE="fam07.sys".
GET FILE="indvars.sys".
SORT CASES BY ER33901.
MATCH FILES FILE=* / TABLE="fam07.sys"/ RENAME=(ER36002=ER33901)
  / BY=ER33901.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind07.sys".
* =====
  Step 2: subset data from 2009 family file (n=8690)
  =====.
FILE HANDLE fam09 / NAME='[PATH]\FAM2009ER.txt' LRECL=8223.
DATA LIST FILE=fam09 /
  ER42002 2-6
  /* your 2009 variable locations here */
.
VARIABLE LABELS
  ER42002 = "2009 FAMILY INTERVIEW (ID) NUMBER"
  /* your 2009 variable labels here */
.
MISSING VALUES
  /* your 2009 missing data values here */
.
SORT CASES BY ER42002.
SAVE OUTFILE="fam09.sys".
GET FILE="indvars.sys".
SORT CASES BY ER34001.
MATCH FILES FILE=* / TABLE="fam09.sys"/ RENAME=(ER42002=ER34001)
  / BY=ER34001.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind09.sys".
* =====
  Step 2: subset data from 2011 family file (n=8907)
  =====.
FILE HANDLE fam11 / NAME='[PATH]\FAM2011ER.txt' LRECL=8339.
DATA LIST FILE=fam11 /
  ER47302 2-6
  /* your 2011 variable locations here */
.
VARIABLE LABELS
  ER47302 = "2011 FAMILY INTERVIEW (ID) NUMBER"
  /* your 2011 variable labels here */
.
MISSING VALUES
  /* your 2011 missing data values here */
.
SORT CASES BY ER47302.
SAVE OUTFILE="fam11.sys".
GET FILE="indvars.sys".
SORT CASES BY ER34101.
MATCH FILES FILE=* / TABLE="fam11.sys"/ RENAME=(ER47302=ER34101)
  / BY=ER34101.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind11.sys".

```



```

* =====
  Step 2: subset data from 2013 family file (n=9063)
=====
FILE HANDLE fam13 / NAME='[PATH]\FAM2013ER.txt' LRECL=8485.
DATA LIST FILE=fam13 /
  ER53002 2-6
  /* your 2013 variable locations here */
.
VARIABLE LABELS
  ER53002 = "2013 FAMILY INTERVIEW (ID) NUMBER"
  /* your 2013 variable labels here */
.
MISSING VALUES
  /* your 2013 missing data values here */
.
SORT CASES BY ER53002.
SAVE OUTFILE="fam13.sys".
GET FILE="indvars.sys".
SORT CASES BY ER34201.
MATCH FILES FILE=* / TABLE="fam13.sys" / RENAME=(ER53002=ER34201)
  / BY=ER34201.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind13.sys".
* =====
  Step 2: subset data from 2015 family file (n=9048)
=====
FILE HANDLE fam15 / NAME='[PATH]\FAM2015ER.txt' LRECL=9250.
DATA LIST FILE=fam15 /
  ER60002 2-6
  /* your 2015 variable locations here */
.
VARIABLE LABELS
  ER60002 = "2015 FAMILY INTERVIEW (ID) NUMBER"
  /* your 2015 variable labels here */
.
MISSING VALUES
  /* your 2015 missing data values here */
.
SORT CASES BY ER60002.
SAVE OUTFILE="fam15.sys".
GET FILE="indvars.sys".
SORT CASES BY ER34301.
MATCH FILES FILE=* / TABLE="fam15.sys" / RENAME=(ER60002=ER34301)
  / BY=ER34301.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind15.sys".
* =====
  Step 2: subset data from 2017 family file (n=9607)
=====
FILE HANDLE fam17 / NAME='[PATH]\FAM2017ER.txt' LRECL=9462.
DATA LIST FILE=fam17 /
  ER66002 2-6
  /* your 2017 variable locations here */
.
VARIABLE LABELS
  ER66002 = "2017 FAMILY INTERVIEW (ID) NUMBER"
  /* your 2017 variable labels here */
.
MISSING VALUES
  /* your 2017 missing data values here */
.
SORT CASES BY ER66002.
SAVE OUTFILE="fam17.sys".
GET FILE="indvars.sys".
SORT CASES BY ER34501.

```

```

MATCH FILES FILE=* / TABLE="fam17.sys"/ RENAME=(ER66002=ER34501)
  / BY=ER34501.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind17.sys".
* =====
  Step 2: subset data from 2019 family file (n=9569)
  =====.
FILE HANDLE fam19 / NAME='[PATH]\FAM2019ER.txt' LRECL=9652.
DATA LIST FILE=fam19 /
  ER72002 2-6
  /* your 2019 variable locations here */
.
VARIABLE LABELS
  ER72002 = "2019 FAMILY INTERVIEW (ID) NUMBER"
  /* your 2019 variable labels here */
.
MISSING VALUES
  /* your 2019 missing data values here */
.
SORT CASES BY ER72002.
SAVE OUTFILE="fam19.sys".
GET FILE="indvars.sys".
SORT CASES BY ER34701.
MATCH FILES FILE=* / TABLE="fam19.sys"/ RENAME=(ER72002=ER34701)
  / BY=ER34701.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind19.sys".
* =====
  Step 2: subset data from 2021 family file (n=9207)
  =====.
FILE HANDLE fam21 / NAME='[PATH]\FAM2021ER.txt' LRECL=7890.
DATA LIST FILE=fam21 /
  ER78002 2-6
  /* your 2021 variable locations here */
.
VARIABLE LABELS
  ER78002 = "2021 FAMILY INTERVIEW (ID) NUMBER"
  /* your 2021 variable labels here */
.
MISSING VALUES
  /* your 2021 missing data values here */
.
SORT CASES BY ER78002.
SAVE OUTFILE="fam21.sys".
GET FILE="indvars.sys".
SORT CASES BY ER34901.
MATCH FILES FILE=* / TABLE="fam21.sys"/ RENAME=(ER78002=ER34901)
  / BY=ER34901.
SORT CASES BY ER30001 ER30002.
SAVE OUTFILE="famind21.sys".
* =====
  Step 3: match familiy-individual files by individual identifiers
  =====.
MATCH FILES
  FILE=famind68.sys /
  FILE=famind69.sys /
  FILE=famind70.sys /
  FILE=famind71.sys /
  FILE=famind72.sys /
  FILE=famind73.sys /
  FILE=famind74.sys /
  FILE=famind75.sys /
  FILE=famind76.sys /
  FILE=famind77.sys /
  FILE=famind78.sys /

```

```
FILE=famind79.sys /
FILE=famind80.sys /
FILE=famind81.sys /
FILE=famind82.sys /
FILE=famind83.sys /
FILE=famind84.sys /
FILE=famind85.sys /
FILE=famind86.sys /
FILE=famind87.sys /
FILE=famind88.sys /
FILE=famind89.sys /
FILE=famind90.sys /
FILE=famind91.sys /
FILE=famind92.sys /
FILE=famind93.sys /
FILE=famind94.sys /
FILE=famind95.sys /
FILE=famind96.sys /
FILE=famind97.sys /
FILE=famind99.sys /
FILE=famind01.sys /
FILE=famind03.sys /
FILE=famind05.sys /
FILE=famind07.sys /
FILE=famind09.sys /
FILE=famind11.sys /
FILE=famind13.sys /
FILE=famind15.sys /
FILE=famind17.sys /
FILE=famind19.sys /
FILE=famind21.sys /
BY=ER30001 ER30002.
SAVE OUTFILE='xyrfind.sys'.
* =====.
```