HEALTH AND RETIREMENT STUDY
2000 Core
Final, Version 1.0
September 2002
(Last revised April 2003)
Data Description and Usage

TABLE OF CONTENTS

Data Description and Usage

1. Overview

The 2000 HRS Core (Final, Version 1.0) data release contains data obtained as part of the Health and Retirement Study (HRS), a national longitudinal study of the economic, health, marital, and family status, as well as public and private support systems, of older Americans. The HRS is a rich source of longitudinal, cross-sectional data for researchers and policymakers who study aging.

Funding was provided by the National Institute on Aging at NIH (U01 AGO9740), with supplemental support from the Social Security Administration. The survey was conducted by the Institute for Social Research (ISR), the University of Michigan.

By receiving the data, which have been freely provided, you must agree to use it for research and statistical purposes only and make no effort to identify the respondents therein. In addition, you must agree to send us a copy of any publications you produce based on the data. See Obtaining the Data for additional details.

1A.   The Sample Interviewed in 2000

The HRS 1998 data collection combined, for the first time, the AHEAD and the HRS sub-samples and added two additional sub-samples. The AHEAD sub-sample was interviewed previously in 1993 and 1995. The HRS sub-sample was interviewed previously in 1992, 1994 and 1996. In 1998, and every two years thereafter, the HRS, AHEAD and two new sub-samples – War Babies (WB), and Children of the Depression (CODA) - were interviewed. For more details about the sample, see the The Sample Design. The data collection period for the 2000 interview was February 2000 through January 2001

1B.   2000 Questionnaire Sections

The 2000 HRS data collection instrument or questionnaire contained the following sections.

```
Section PR  Preload
Section CS  Cover Screen
Section A   Demographics
Section B   Health Status
Section PC  Proxy Cognition
Section C   Cognition
Section D   Family Structure
Section E   Health Care Utilization
Section F   Housing
Section G   Employment (includes G, GG, GH, GD and GJ)
Section H   Expectations
Section J   Assets and Income
Section N   Widowhood and Asset Changes
Section R   Health Insurance and Event History (includes R and EV)
Section M   Modules and Respondent Cognition (includes RC, M1 to M5 and M7
            to M12
Section T   Other (Time Stamps, Thumbnails)
```

See Modules for the 2000 Data Collection for details about the 2000 modules.
See Cognition Sections for details about the cognition sections.

1C.   Levels of Files

In the 2000 data collection instrument, most questions were asked of all
respondents. Some questions were asked about the household. For two-respondent
households, these questions were asked just of a designated financial
respondent, family respondent, or coversheet respondent (the first respondent
interviewed) on behalf of the entire household.

In addition to the familiar household-level and respondent-level files, the HRS
2000 Final Core Data Release, version 1, contains files at five other levels:
household-member-and-child, sibling, helper, transfer-to-child and transfer-
from-child.

1C1.   Household-Level Files

Household-level files contain questions that were asked about the household of a
designated coversheet, financial or family respondent. A coversheet respondent
answered family questions (section CS) on behalf of the entire household.
Similarly, a family respondent answered family questions (section D) on behalf
of the entire household and a financial respondent answered household-level
financial questions (sections F, J, and N) on behalf of the entire household.
The household-level files contain one record for each household in which at
least one interview was obtained in 2000.

1C2.   Respondent-Level Files

Respondent-level files contain questions that were asked of all respondents
about themselves (or asked of a proxy about the respondent if the respondent was
not able to give an interview). The files contain one record for each
respondent or proxy who gave an interview in 2000.

1C3.   Household-Member-and-Child-Level Files

These files contain information provided by the family respondent (section D and
sections PR and CS) or financial respondent (section J) about each household
member or child of the respondent or of the respondent's spouse. If the
designated family respondent did not provide an interview, information from the
non-family respondent was used for PR and CS sibling files. The files contain
one record for each household member or child.

1C4.   Sibling-Level Files

These files contain information provided by the family respondent about the
respondent's siblings and siblings of the respondent's spouse/partner (sections
PR and D). If the designated family respondent did not provide an interview,
information from the non-family respondent was used for PR sibling file. The
files contain one record for each sibling or spouse's/partner's sibling.

1C5.   Helper-Level File

This file contains information provided by each respondent about helpers other
than the respondent's spouse or partner. A helper may be a person who or
organization that was reported by the respondent as providing help with ADLs or
IADLs. If the helper was a child, the record contains information about the help
provided by the child and the child's spouse or partner, if any. The file
contains one record for each helper or, if the helper is a married or partnered
child, the helping couple, for each respondent reporting help. If a child

helped both mom and dad, the file will contain two records – one of mom's report of the child's helping her and one for dad's report of the child's helping him.

1C6. Transfer-to-Child-Level File

This file contains information provided by the family respondent about transfers of money to a child or grandchild. The file contains one record for each transfer to a child or grandchild.

1C7. Transfer-from-Child-Level-File

This file contains information provided by the family respondent about transfers of money from children or grandchildren. The file contains one record for each transfer from a child or grandchild.

2. File Naming Conventions

Files are named beginning with "H00" for HRS 2000. Next a letter or two designating the questionnaire section. Next a separator, "_" and then one or two letters designating the level,

        H for household-level,
        R for respondent-level,
        MC for household-member-and-child-level,
        SB for sibling-level,
        HP for helper-level,
        TC for transfer-to-child-level, and
        FC for transfer-from-child-level.

For example, putting it all together, files named H00A_R include variables from section A at the respondent-level.

The following extensions are used for the six different types of files that are distributed.

        .DA for data files,
        .SAS for SAS program statements,
        .SPS for SPSS program statements,
        .DO for Stata DO statements,
        .DCT for Stata dictionary statements, and
        .TXT for codebook files.

One of each of these file types is provided for each of 29 parts of the 2000 HRS Core (Final, Version 1.0) data release. For example,

        H00A_R.DA contains respondent data from section A,
        H00A_R.SAS contains corresponding SAS program statements,
        H00A_R.SPS contains corresponding SPSS program statements,
        H00A_R.DO contains corresponding Stata DO statements,
        H00A_R.DCT contains corresponding Stata dictionary statements, and
        H00A_R.TXT contains the ASCII codebook.

3. Data Files

The 2000 HRS Core (Final, Version 1.0) data is distributed in 29 parts. The parts are listed below along with the number of cases (N), number of variables (NV), and the primary identifiers (IDS). The records in the data files are sorted in order by these primary identifiers.

The 2000 HRS Core data are provided in ASCII format, with fixed-length records.
You'll want to use associated SAS, SPSS or Stata program statements to read the
data into the analysis package of your choice.

Household-level-files, IDS=HHID GSUBHH

```
    H00CS_H      N=13214 NV=66       IDS=HHID GSUBHH
    H00D_H       N=13214 NV=545      IDS=HHID GSUBHH
    H00F_H       N=13214 NV=183      IDS=HHID GSUBHH
    H00J_H       N=13214 NV=595      IDS=HHID GSUBHH
    H00N_H       N=13214 NV=113      IDS=HHID GSUBHH
```

Respondent-level-files

```
    H00PR_R      N=19580 NV=253      IDS=HHID PN
    H00CS_R      N=19580 NV=96       IDS=HHID PN
    H00A_R       N=19580 NV=98       IDS=HHID PN
    H00B_R       N=19580 NV=190      IDS=HHID PN
    H00C_R       N=19580 NV=97       IDS=HHID PN
    H00PC_R      N=19580 NV=74       IDS=HHID PN
    H00D_R       N=19580 NV=16       IDS=HHID PN
    H00E_R       N=19580 NV=295      IDS=HHID PN
    H00G_R       N=19580 NV=1515     IDS=HHID PN
    H00H_R       N=19580 NV=54       IDS=HHID PN
    H00J_R       N=19580 NV=84       IDS=HHID PN
    H00M_R       N=19580 NV=358      IDS=HHID PN
    H00N_R       N=19580 NV=86       IDS=HHID PN
    H00R_R       N=19580 NV=219      IDS=HHID PN
    H00T_R       N=19580 NV=58       IDS=HHID PN
```

Household-member-and-child-level-files

```
    H00PR_MC     N=46023 NV=33       IDS=HHID GSUBHH OPN
    H00CS_MC     N=46023 NV=18       IDS=HHID GSUBHH OPN
    H00D_MC      N=46023 NV=48       IDS=HHID GSUBHH OPN
    H00J_MC      N=46023 NV=28       IDS=HHID GSUBHH OPN
```

Sibling-level-files

```
    H00PR_SB     N=17590 NV=22       IDS=HHID GSUBHH OPN
    H00D_SB      N=17590 NV=35       IDS=HHID GSUBHH OPN
```

Transfer-to-child-level-file

```
    H00D_TC      N=5839  NV=27       IDS=HHID GSUBHH OPN GTRANNUM
```

Transfer-from-child-level-file

```
    H00D_FC      N=852   NV=27       IDS=HHID GSUBHH OPN GTRANNUM
```

Helper-level-file

```
    H00E_HP      N=4156  NV=49       IDS=HHID PN OPN
```

4. Identification Variables

Identification variables for HRS 2000 are stored in character format. For further information about merging 2000 data, see Examples of Merging.

4A. Primary Identification Variables

Five variables, HHID, GSUBHH, PN, OPN, and GTRANNUM are used in various combinations to uniquely identify the five different level datasets that comprise this data release.

4A1. HHID – Household Identifier

In the initial wave of data collection (in 1992 for the HRS sub-sample, in 1993 for the AHEAD sub-sample, and in 1998 for the WB and CODA sub-samples) each sample household was assigned a Household Identifier. HHID is stable across waves of data collection and uniquely identifies the original household and any households derived from that household in subsequent waves of data collection. HHID has six-digits.

4A2. GSUBHH – 2000 Sub-household Identifier

In combination with HHID, GSUBHH uniquely identifies a household at the time of the 2000 data collection. Sub-household identifiers can be different at each wave. GSUBHH has one-digit. For more information, see Examples of Sub-Household and Respondent Person Number and Other Person Number Assignments.

4A3. PN – Person Number

In combination with HHID, PN uniquely identifies a respondent or respondent's spouse or partner. PNs are unique within an original household (HHID). The PN assigned to a particular respondent does not change across waves. PN has three-digits.

4A4. OPN – Other Person Number.

In combination with HHID and GSUBHH, OPN uniquely identifies another person -- household member, child, sibling, or helper -- at the time of the 2000 data collection. OPN has three-digits. See Datasets Including GSUBHH and OPN as Primary Identifiers for additional information.

4A5. GTRANNUM – 2000 Transfer Number

In combination with HHID and GSUBHH, GTRANNUM uniquely identifies transfer to or from a child or grandchild. GTRANNUM has one digit.

4B. Primary Identification Variables for Datasets at Each of the Seven Levels

Two identifiers uniquely identify records in the
   o   respondent-level datasets:

     1) HHID   HOUSEHOLD IDENTIFIER
     2) PN   PERSON NUMBER

Three identifiers uniquely identify records in the
   o   helper-level datasets:

     1) HHID   HOUSEHOLD IDENTIFIER
     2) PN   PERSON NUMBER
     3) OPN   OTHER PERSON NUMBER

Two identifiers uniquely identify records in the
    o household-level datasets:

        1) HHID   HOUSEHOLD IDENTIFIER
        2) GSUBHH   2000 SUB-HOUSEHOLD IDENTIFIER

Three identifiers uniquely identify records in the
    o household-member-and-child-level, and
    o sibling-level datasets:

        1) HHID   HOUSEHOLD IDENTIFIER
        2) GSUBHH   2000 SUB-HOUSEHOLD IDENTIFIER
        3) OPN   OTHER PERSON NUMBER

Four identifiers uniquely identify records in the
    o transfer-to-child-level, and
    o transfer-from-child-level datasets:

        1) HHID   HOUSEHOLD IDENTIFIER
        2) GSUBHH   2000 SUB-HOUSEHOLD IDENTIFIER
        3) OPN   OTHER PERSON NUMBER
        4) GTRANNUM 2000 TRANSFER NUMBER

4C.   Secondary Identification Variables

In addition to the primarily identification variables that uniquely identify
records in a dataset, secondary identification variables that allow links to
other datasets are provided. Two sets of secondary identification variables are
provided. They are listed below.

Secondary Identification Variables for
    o respondent-level and
    o helper-level datasets:

        GSUBHH   2000 SUB-HOUSEHOLD IDENTIFIER
        FSUBHH   1998 SUB-HOUSEHOLD IDENTIFIER
        GPN_SP   2000 SPOUSE/PARTNER PERSON NUMBER
        GCSR   2000 WHETHER COVERSHEET RESPONDENT
        GFAMR   2000 WHETHER FAMILY RESPONDENT
        GFINR   2000 WHETHER FINANCIAL RESPONDENT
        GQDT   2000 Q_NAIRE FILE
        GQNR   2000 QUESTIONNAIRE

Secondary Identification Variables for
    o household-level,
    o household-member-or-child-level,
    o sibling-level,
    o transfer-from-child-level, and
    o transfer-to-child-level datasets:

        FSUBHH   1998 SUB-HOUSEHOLD IDENTIFIER
        GPN_CS   2000 COVERSHEET RESP PERSON NUMBER
        GPN_FAM   2000 FAMILY RESP PERSON NUMBER
        GPN_FIN   2000 FINANCIAL RESP PERSON NUMBER
        GPN_NCS   2000 NON-COVERSHEET RESP PERSON NUMBER
        GPN_NFAM 2000 NON-FAMILY RESP PERSON NUMBER
        GPN_NFIN 2000 NON-FINANCIAL RESP PERSON NUMBER

```
GQDT      2000 Q_NAIRE FILE
GQNR_CS   2000 QUESTIONNAIRE - COVERSHEET RESP
GQNR_FAM  2000 QUESTIONNAIRE - FAMILY RESP
GQNR_FIN  2000 QUESTIONNAIRE - FINANCIAL RESP
```

4D. Datasets Including GSUBHH and OPN as Primary Identifiers

```
When working with datasets including GSUBHH and OPN as primary identifiers --
household-member-and-child-level, sibling-level, transfer-from-child-level, or
transfer-to-child-level datasets – it is essential to use GSUBHH to distinguish
other persons.

Sometimes records with the same HHID and OPN will be separate reports about the
same person.  For example, see Couple with Children and Siblings Divorces. where
the couple with children divorces, the record where HHID is 89012, GSUBHH is 1,
and OPN is 101 is the report about the child from the mother, and the record
where HHID is 89012, GSUBHH is 2, and OPN is 101 is the report about the same
child from the father.

At other times records with the same HHID and OPN will be reports about
different persons.  For example, see Couple Divorces, One Respondent Remarries,
Both Split-off Households Have New Members. where the couple divorces and both
split-off households have new members, the record where HHID is 90123, GSUBHH is
2, and OPN is 151 is for the ex-wife's mother, while the record where HHID is
90123, GSUBHH is 1, and OPN is 151 is for the ex-husband's new stepchild.

See also the section Linking Other Persons across Time.
```

5. Distribution Files and Directory Structure

5A. Distribution Files

```
The files are packaged for download from our Web site in two different ways – as
one big .zip file that contains seven smaller .zip files and one .pdf file or
the eight smaller files available individually for separate download.  The
combined file is h00core.zip.

The individual .zip files for separate download are:

    Data files
        h00da.zip containing data files.

    Program statement files
        h00sas.zip containing SAS data descriptors.
        h00sps.zip containing SPSS data descriptors.
        h00sta.zip containing Stata data descriptors.

    Documentation files
        h00cb.zip containing the codebook.
        h00qn.zip containing the questionnaire.
        h00spc.zip containing Surveycraft programming specifications.
        h00dd.pdf containing this document.

You'll typically want to download at least the data files, one set of program
statement files, and the codebook files.
```

5B.   Directory Structure

While a particular setup is not required for using HRS files, we have
traditionally suggested a directory structure. By using this directory
structure, you will not have to change the path name in your data descriptor
files. If you use a different structure, just change the directory references
in the program statement files.

```
Directory                  Contents

c:\HRS2000                 Files downloaded from Web site
c:\hrs2000\codebook        Unzipped files from h00cb.zip
c:\hrs2000\data            Unzipped files from h00da.zip
c:\hrs2000\qnaire          Unzipped files from h00qn.zip
c:\hrs2000\sas             Unzipped files from h00sas.zip
c:\hrs2000\spss            Unzipped files from h00sps.zip
c:\hrs2000\stata           Unzipped files from h00sta.zip
c:\hrs2000\scraft          Unzipped files from h00prn.zip
```

Decompress the selected .zip files into the appropriate subdirectories. You
will need about 191 MB of free space on your storage device to store the 29 .DA
files

6. Program Statements

Each data file comes with associated SPSS, SAS or Stata program statements to
read the data. Files containing SPSS statements are named with a .SPS extension,
those with SAS statements with a .SAS extension, and those with Stata statements
with .DO and .DCT extensions.

The statement files are named beginning with the same prefix as the
corresponding data file. For example, SAS statements in the file H00A_R.SAS go
with the H00A_R.DA data file.

6A.   Using the Files with SAS

To create a SAS system file for a particular dataset, two file types must be
present for that dataset -- .SAS program statement files and .DA data files.

To create a SAS system file, load the *.SAS file into the SAS Program Editor.

If the *.SAS file is located in "c:\hrs2000\sas" and the data file is located in
"c:\hrs2000\data", you can run the file as is. A SAS system file (*.SD2 or
*.SAS7BDAT) will be saved to directory "c:\hrs2000\sas".

If the files are not located in the specified directories, you will need to edit
the *.SAS file to reflect the proper path names prior to running the file.

6B.   Using the Files with SPSS

To create an SPSS system file for a particular dataset, two file types must be
present for that dataset -- .SPS program statement files and .DA data files.

To create an SPSS system file, open the *.SPS file in SPSS as an SPSS Syntax
File.

If the *.SPS file is located in "c:\hrs2000\spss" and the data file is located
in "c:\hrs2000\DATA", you can run the file as is. An SPSS system file (*.SAV)
will be saved to directory "c:\hrs2000\spss".

If the files are not located in the specified directories, you will need to edit
the *.SPS file to reflect the proper path names prior to running the file.

6C. Using the Files with Stata

To use Stata with a particular dataset, the following three file types must be
present for that dataset -- .DCT files, .DO files, and .DA data files.

Files with the suffix .DA contain the raw data for Stata to read. Files with
the suffix .DCT are Stata dictionaries used by Stata to describe the data. Files
with the suffix .DO are short Stata programs ("do files") which you may use to
read in the data. Load the .DO file into Stata and then submit it.

If the *.DO and *.DCT files are located in "c:\hrs2000\stata" and the data file
is located in "c:\hrs2000\data", you can run the .DO file as is.

If the files are not located in these directories, you must edit the *.DO and
*.DCT files to reflect the proper path names before you run the files.

Note that the variable names provided in the .DCT files are uppercase. If you
prefer lower case variable names, you may wish to convert the .DCT files to
lower case prior to use. You may do this by reading the .DCT file into a text
or word processing program and changing the case. For instance in Microsoft
Word, Edit, Select All, Format, Change Case, lowercase.

7. Documentation

There are several types of documentation available for use with the 2000 HRS
Core (Final, Version 1.0) data release. In addition to this document, they
include a codebook, programming specifications, and a box-and-arrow
questionnaire.

7A. Codebook

The HRS 2000 Codebook is provided as a series of 29 ASCII text files and one
master code text file. There is a codebook file corresponding to each data
file. Each variable has its own codebook entry. Examples are included below.

==

```
G1654      C1.RATE MEMORY
           Section: C              Level: Respondent        CAI Reference: Q1654
           Type: Numeric           Width: 1                 Decimals: 0

           C1.
                   Part of this study is concerned with people's memory, and
                   ability to think about things.  First, how would you rate
                   your memory at the present time?  Would you say it is
                   excellent, very good, good, fair or poor?
           ..................................................................
            1070           1. EXCELLENT
            4222           2. VERY GOOD
            7748           3. GOOD
            3727           4. FAIR
             728           5. POOR
              11           8. DK (don't know); NA (not ascertained)
              11           9. RF (refused)
```

```
          2064    Blank. INAP (Inapplicable); [Q514:CS1A] IS (2 OR 3); partial
                         interview

===============================================================================
          ASSIGNMENT STATEMENTS
          *    If {When Q160 is (Q474), [Q183:PR183] IS (1) AND [Q182:PR182]
               IS (5)} then [G595.CS15Y11.VIRGIN R FAM/FIN] = 1
          *    If {When Q160 is (Q474), [Q183:PR183] IS (5) AND [Q182:PR182]
               IS (1)} then [G595.CS15Y11.VIRGIN R FAM/FIN] = 2
          *    If {When Q160 is (Q474), [Q183:PR183] IS (1) AND [Q182:PR182]
               IS (1)} then [G595.CS15Y11.VIRGIN R FAM/FIN] = 3
          *    If {When Q160 is (Q474), [Q183:PR183] IS (5) AND [Q182:PR182]
               IS (5)} then [G595.CS15Y11.VIRGIN R FAM/FIN] = 4
          - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
G595      CS15Y11.VIRGIN R FAM/FIN
          Section: CS          Level: Respondent      CAI Reference: Q595
          Type: Numeric        Width: 1               Decimals: 0

          User note:  CS15Y11.VIRGIN R FAM/FIN is assigned a value of 0.
          ......................................................................
           121          0. NOT INTERVIEWED
          3812          1. FINANCIAL R
          3890          2. FAMILY R
          8967          3. FINANCIAL & FAMILY R
          2791          4. NON-FINANCIAL & NON-FAMILY R

===============================================================================

G3078     F6.HOME VALUE $
          Section: F           Level: Household       CAI Reference: Q3078
          Type: Numeric        Width: 8               Decimals: 0

          F6.
                   What is its present value?  I mean, about what would it
                   bring if it were sold today?

                   DO NOT PROBE DK/RF

                   IWER: USE '0' FOR NOTHING


                   AMOUNT:
          ......................................................................
             2                    0. Actual value
            17             1-4999. Actual value (verified)
          7743        5000-500000. Actual value
           137   500001-99999996. Actual value (verified)
          1046           99999998. DK (don't know); NA (not ascertained)
           107           99999999. RF (refused)
          4108               Blank. INAP (Inapplicable); [Q760:CS22Y37] IS
                                    (5); [Q558:CS11] IS (1) AND [Q794:CS26]
                                    IS (NE 1); [Q3061:F3] IS (2); [Q3061:F3]
                                    IS (7 OR DK OR RF); [Q3064:F4] IS (3);
                                    [Q3064:F4] IS (7 OR DK OR RF); [Q3069:F5]
                                    IS (4); [Q3069:F5] IS (7 OR DK OR RF);
                                    [Q3071:F5b] IS (NE DK AND NE RF);
                                    [Q3072:F5c] IS (3 OR DK OR RF); [Q3073:F5d]
```

```
                              IS (NE 5); [Q3074:F5e] IS (A); [Q3075:F5f]
                              IS (A); partial interview; no financial
                              Respondent
```

```
 ==============================================================================
```

7A1.   Variable and Question Information

A horizontal line of equal signs sets off codebook entries from each other. Variable and question information is separated from codeframe information by a line of periods.

7A1a.   Assignment Statements and User Notes with Assignments

Assignment statements are included for variables that were assigned a value by the CAI program. No question may have been asked, or a question may have been asked only of respondents not assigned a value. If present, they are followed by a line of dashes and precede any other variable information.

If more than one assignment statement is present, assignments are made by the first condition that was satisfied, i.e., they may be read as a series of IF-THEN-ELSE statements.

User notes of the form "User note: CS15Y11.VIRGIN R FAM/FIN is assigned a value of 0." denote the assignment of a value to a variable if the variable did not receive a value from a previous assignment statements, if any.

7A1b.   Variable Information

The variable information is contained on three lines. The first eight columns of the first line contain the variable name. Variable names that are specific to the 2000 HRS cross-section begin with a "G", i.e., " G1654". Other variable naming conventions are described below. To the right of the variable name is the variable label.

The next two lines contain six fields with additional variable information.

7A1b1.   Section

This is the section letter of the data files. This differs slightly from the letter of questionnaire section. For instance included in data file for section G are questions labeled G, GG, GH, GJ and GD. Questionnaire section EV is found in data file section R. Data file section M includes questions labeled RC, M1 to M5 and M7 to M12.

7A1b2.   Level

This indicates whether the data were collected at the household-level, respondent-level or another level.

        "Household" for household-level files,
        "Respondent" for respondent-level files,
        "HHMemberChild" for household member/child-level files,
        "Sibling" for sibling-level files,
        "ToChild" for transfer-to-child-level files,
        "FromChild" for transfer-from-child-level files, and
        "Helper" for helper-level files.

7A1b3.   CAI Reference

The CAI (Computer Assisted Interviewing) Reference begins with the letter "Q".
For data collected as part of the interview and not masked, the variable name
and CAI Reference differ only in this first letter and in any extension to the
variable name distinguishing multiple mention and looped variables. These
numbers are in the range 1 to 7373. The 10000 and 30000 series are used for
masked variables, and derived variables with alterative formats. The 9000
series is used as the CAI Reference for identification variables. The 20000
series is used for data release version variables.

7A1b4.   Type

Variable type -- numeric or character. Most variables are stored as numeric
with the exception of the identification variables, which are stored as
character variables.

7A1b5.   Width

Variable width in the ASCII data file.

7A1b6.   Decimals

The number of decimals. Usually zero.

7A1c.   Question Text and User Note

This is followed, for most variables, by the question text and, for a few
variables, by a user note containing supplemental information provided by HRS
staff.

Question text in upper and lower case was that read to the respondent.
Interviewer instructions that appeared on the CAI screen but were not read to
the respondent and rules that determined what text was read to the respondent
are in upper case. Text fills, e.g., answers to previous questions that were
part of the question text are designated with a Q-number and question reference
or brief description, e.g., "since Q218-PREV WAVE IW MONTH".

7A2.   Codeframe Information

After a line of periods, codeframe information for each variable is provided.

7A2a.   Code Categories or Ranges

For categorical values single code values are provided. For continuous
variables, ranges are designated in the codebook, e.g., "1993-2000"; however,
actual, continuous values, e.g., 1993, 1994, 1995, etc., are provided in the
data. Blanks are used throughout for INAP.

7A2b.   Frequencies

To the left of each code is its frequency. The frequencies are unweighted.
They can be used to gain an initial overview of the data and to verify that data
read into your statistical package accurately.

7A2c.   Code Text

In general, code text picked up from the data collection instrument
specifications is in all capital letters. Code text not from the data
collection instrument is in upper and lower case letters. These include codes
added as the result of the comment review process. See Changes in Data because
of Respondent Comments for further description of comment review process.

For values that were determined to be unlikely, the data collection instrument required verification of the value. These are designated in the codebook as "Actual value (verified)".

7A2c1.   INAP Code Text

Extensive documentation of the INAP category derived from the data collection instrument program specifications is provided. The rules specify the CAI reference and do not include loop or multiple mention information.

In a series of variables derived from multiple-mention question, the INAP category of mentions other than the first will include respondents who provided no second, third, etc. response even though the INAP text does not explicitly state this.

When used in the codebook, typically in the INAP text, but also in question text that includes rules for alternative wording and in Assignment Statements, the (A) symbol means "answered". For example, "[Q1981:D01a] IS (A)" can be read as "if Q1981 was answered". The text (NE "|") means "not blank". For example, "[Q2951:E158a] IS (NE "|" OR DK OR RF)" can be read as "if Q2951 is not blank or don't know or refused".

In the INAP text, 0 means a value of zero or blank or not answered. For instance, the phrase "[Q5002:H7] IS (00)" would include both a value of 0 and a value of blank for G5002, the variable with the CAI reference Q5002.

The rules in the INAP text do not reference codes added to the codeframe as a result of the comment review process. Nor to they reflect occasional, minor changes made to variable labels in processing.

Assignment statements may also return blanks; these blanks will not be documented in the INAP code text.

7A3.   *Special Considerations*

Several special codebook conventions and types of variables are described below.

7A3a.   Multiple-response and Looped Variables

Variables from questions that allowed multiple-mentions have a "M" in their name and a suffix that designates the mention, e.g., G1109M1, G1109M1, G1109M3.

Variables from questions that were in a loop have a "_" in their name and a suffix that designates the loop, e.g., G2301_1, G2301_2.

Null multiple mention variables and variables from null loops beyond the first mention or first loop are not included in the data. If null mentions or null loops are deleted, user notes indicating the number of mentions or loops possible and the number with non-null data are provided, e.g., "User note: Up to twelve responses were allowed for this question, the actual maximum number of responses was five."

Variables from questions in loops that allowed multiple mentions end with a three-digit number indicating the loop and the mention. For example, eighty variables would result for a question that was in a loop of twenty and allowed up to four mentions. The resulting variables would be, for example, G1988001 to G1988004 for the four mentions from the first loop, G1988005 to G1988008 for the

four mentions from the second loop and so forth, with G1988077 to G1988080 for the four mentions from the twentieth loop.

A user note is provided indicating the maximum and the actual number of loops and mentions, e.g., "User note: Up to twenty loops with four mentions were allowed for this question. The maximum number of loops used was five and the maximum number of responses was four."

7A3b.   Masked Variables

Because we are dedicated to protecting the anonymity of our respondents and to protecting the confidentiality of the information they provide, a number of variables have been masked or are simply not included in the final release public dataset. Some of these variables may be made available to analysts as restricted data. See our Web site for details.

Names, addresses, days of birth, information on geographical relocation and similar variables are not included in publicly released files. See also Variables Not Included in This Release.

Geographical locations are recoded to a level no more detailed than U.S. Census Region and Division. Data on the highest educational degree earned have been further grouped together to increase cell sizes. Denominations have been grouped. Industry and occupation codes have been recoded into a limited number of categories from the original three-digit U.S. Census code.

The names of variables that were masked for confidentiality end in the letter "M"; for example, variable " G1061M" is the masked version of variable "G1061". The following variables have been masked for confidentiality.

```
    G1061M    A2A.REGION - US BORN                          H00A_R
    G1063M    A2B.REGION - NON US BORN                      H00A_R
    G1077M    A3C.R HIGH DEGREE - MASKED                    H00A_R
    G1084M    A4F.FATHER USUAL OCCUPATION - MASKED          H00A_R
    G1090M1M  A7A.TYPE HISPANIC - MASKED                    H00A_R
    G1092M    A8.R RACE - MASKED                            H00A_R
    G1122M    A27.WHERE LIVE WHEN IN SCH - REGION           H00A_R
    G1140M    A36A.DENOMINATION - MASKED                    H00A_R

    G1377M1M  B19a-1. OTHER HEALTH SPEC-MENT 1 - MASKED     H00B_R
    G1377M2M  B19a-2. OTHER HEALTH SPEC-MENT 2 - MASKED     H00B_R

    G805M     CS33.CURRENT RESIDENCE - REGION               H00CS_H
    G812M     CS35.OTH RESIDENCE - REGION                   H00CS_H

    G792M     CS25B.FACILITY WHERE LIVE - REGION            H00CS_R

    G2156M    D106.REGION THEY LIVE                         H00D_H
    G2169M    D115.REGION MOTHER LIVE                       H00D_H
    G2182M    D125.REGION FATHER LIVE                       H00D_H
    G2343M    D106-2.REGION THEY-IN-LAW LIVE                H00D_H
    G2356M    D115-2.REGION MOTHER-IN-LAW LIVE              H00D_H
    G2369M    D125-2.REGION FATHER-IN-LAW LIVE              H00D_H

    G2236M    D143c.SIB/S-IN-LAW HIGHEST DEGREE-MASKED      H00D_SB

    G3234M    F40B.REGION OF 2ND HOME                       H00F_H
```

```
G3434M    G23.PREVIOUS WAVE WORK DONE - MASKED          H00G_R
G3436M    G25.CURRENT OCC  - MASKED                     H00G_R
G3503M    G41A.IND - MASKED                             H00G_R
G3505M    G42.OCC - MASKED                              H00G_R
G3960M    GG4.IND - MASKED                              H00G_R
G3962M    GG6.OCC - MASKED                              H00G_R
G4101M    GH5.IND - MASKED                              H00G_R
G4103M    GH6.OCC - MASKED                              H00G_R
G4308M1M GD1A-1.CAUSE PROB - MASKED                     H00G_R
G4308M2M GD1A-2.CAUSE PROB - MASKED                     H00G_R
G4339M    GD12D.OCCUPATION - MASKED                     H00G_R
G4340M    GD12E.INDUSTRY - MASKED                       H00G_R
G4554M1M GJ1A-1.CAUSE PROB - MASKED                     H00G_R
G4554M2M GJ1A-2.CAUSE PROB - MASKED                     H00G_R
G4559M1M GJ1F-1.CAUSE PROB - MASKED                     H00G_R
G4559M2M GJ1F-2.CAUSE PROB - MASKED                     H00G_R
G4630M1M GJ29A-1.CAUSE PROB - MASKED                    H00G_R
G4630M2M GJ29A-2.CAUSE PROB - MASKED                    H00G_R
G4690M    GJ40.OCCUPATION - MASKED                      H00G_R
G4691M    GJ41.INDUSTRY - MASKED                        H00G_R
G4731M    GJ61.OCCUPATION - MASKED                      H00G_R
G4732M    GJ62.INDUSTRY - MASKED                        H00G_R
G4767M    GJ80.OCCUPATION - MASKED                      H00G_R
G4768M    GJ81.INDUSTRY - MASKED                        H00G_R
G4781M    GJ90.PREVIOUS WAVE WORK DONE - MASKED         H00G_R
G4782M    GJ91.INDUSTRY - MASKED                        H00G_R

G5012M    H10B.MOVE TO ANOTHER REGION                   H00H_R

G6207M    N63c.REGION DIED                              H00N_R
```

7A3c.   Variables with Alternative Format

A few variables have been derived. They have an "A" extension and are listed
below. They appear in the codebook following the associated variable.

Much of Section A was asked only of new respondents. For your convenience, past
wave data were merged for a few variables for previously interviewed
respondents. For example, variable G1053A contains data from preload variables
G29_1-G29_3 and G82_1-G82_3 and coversheet variable G474 combined with data from
G1053. Previous years' data for G1089A and G1092A was obtained from the tracker
file.

```
G475A    HRS2000 VERSION                               H00CS_R

G1051A    A1.BIRTHDATE MONTH - COMBINED                H00A_R
G1053A    A1B.BIRTHDATE YEAR - COMBINED                H00A_R
G1074A    A3.R EDUCATION - COMBINED                    H00A_R
G1089A    A7.HISPANIC - COMBINED                       H00A_R
G1092A    A8.R RACE - COMBINED MASKED                  H00A_R

G2947A    HELPER RELATIONSHIP COMBINED SOURCE          H00E_HP

G7373A    VQ FINAL VERSION                             H00T_R
```

7A3d.   Banked-screen Questions

For some questions one question resulted in two variables. For instance, in
Section G there are many questions that asked about amounts where the answers
were given in the form of "amount" and "per". For example, the question G144
asking, "On average, how much were you earning before deductions during these
months?" results in two variables G3880 - G144.EARNINGS, and G3881 - G144A.PER
where G3880 is a dollar amount and G3881 is a time period. In the codebook
you'll find the question text with the first of these two variables, G3880, and
a line "AMOUNT: PER:". The second of these two variables, G3881, includes no
question text in the codebook. See the previous variable, G3880, for the
question text.

Similarly, through out the codebook you'll find questions that asked about the
date of an event that generated two variables, month and year. For example, the
question A40a asking "In what month and year were you married?" followed in the
codebook with "MONTH: YEAR:" generated two variables, G1148 - A40A.POST W1
MARRIAGE MO and G1149 - A40C.YEAR. In the codebook the first variable of this
pair includes the question text; the second does not.

Also sometimes one CAI screen contained many questions. For instance, a series
of questions, labeled B19b to B19k, asked about preventative health behaviors.
The question text for all of the questions in the series appears with the first
question. Subsequent questions contain no question text. The variable labels
for the subsequent questions will provide information to know which question was
asked.

7A4.   Master Code

A master code file contains detailed codeframes used in several sections of the
codebook. The master codes include health conditions, health conditions -
alphabetical list, occupation codes, industry codes, and state and country
codeframes.

7B.   Other Types of Documentation

In addition to this document and the codebook, two other types of documentation
are available.

7B1.   Programming Specifications

HRS 2000 data were collected using the Surveycraft CAI (Computer Assisted
Interviewing) program. Surveycraft uses program specification files to
determine what the program is supposed to do. The file "H00SPC.TXT" is a basic
translation of the HRS 2000 Surveycraft program specifications.

This file retains much of the Surveycraft program language syntax and is, in
places, not easy to decipher. However, it is the ultimate source for how the
data were collected and the flow of the data collection instrument.

7B2.   Box and Arrow Questionnaire

The research community has referred to the type of documentation that describes
the questions asked in the interview as a "questionnaire". Since the 2000 HRS
data were collected using a CAI program, a traditional hard-copy questionnaire
was not produced as part of the data collection phase. However, we have
provided a version of the traditional box-and-arrow questionnaire to help
document the asking sequence of the questions.

8. Additional Notes

Found here are miscellaneous additional notes regarding HRS 1998 Final Data
Release, Version 2. If we become aware of additional issues, they will be
posted on our Web site in the Data Alerts section.

8A.   Basic Demographic Variables

We have included many variables in the preload and coversheet sections because
they determine the flow of the subsequent interview. Some of these variables
were updated by information obtained later in the interview. We have noted this
by including "UPDATED" in the label. However, all this can become confusing.
In addition to the "A" variables listed above, Variables with Alternative
Format, we recommend you use these variables for your analysis.

```
        G599        CS15Y14.NEW SP COUPLENESS                  H00CS_R
        G757        CS22Y34.R SEX                              H00CS_R
        G764        CS22Y41.SP GENDER                          H00CS_R


        G1158       A43A.MARITAL STATUS_ASSIGNED               H00A_R
```

8B.   Selecting the Correct Variable for the Respondent from Preload Information

The preload section contains multiple instances of a number of variables. The
information below will allow you to select the variable with the information the
respondent.

For the variables in the list below, one of six variables contains information
about the respondent. For instance, there are six variables that may contain
information about the respondent's birth year, G29_1 to G29_3 for respondents in
household 1 and G82_1 to G82_3 for respondents in household 2. In order to know
which variable to use, you should use two additional variables from the
coversheet section the first, G480, to determine whether to use the variables
for household 1 or household 2, and the second, G474, to determine which of the
three variables to use.

```
G480      CS0Y25.WHICH HH
      1. Household 1
      2. Household 2

G474      CS0Y5.R INDEX
      1. OLD R (1ST SLOT)
      2. OLD R (2ND SLOT)
      3. NEW R (3RD SLOT)

      G19_1-G19_3         PR19.HH1 R INDEX
      G25_1-G25_3         PR25.WHICH ENTRY COHORT
      G26_1-G26_3         PR26.HH1 R PREV WAVE IW
      G29_1-G29_3         PR29.HH1 R YR BIRTH - UPDATED
      G30_1-G30_3         PR30.HH1 R MO BIRTH - UPDATED
      G32_1-G32_3         PR32.HH1: R ORIG SAMP MEMB
      G33_1-G33_3         PR33.HH1: R AGE ELIG
      G35_1-G35_3         PR35.HH1 R SEX - UPDATED
      G36_1-G36_3         PR36.HH1 R MARITAL STAT
      G37_1-G37_3         PR37.R COUPLENESS - UPDATED
      G38_1-G38_3         PR38.S/P DIE DURING STUDY
      G39_1-G39_3         PR39.HH INTERVIEWED IN PRIOR WAVE
      G40_1-G40_3         PR40.HH1 R SAME SPOUSE/PARTNER - UPDATED
```

```
G41_1-G41_3          PR41.HH1 PREV WAVE FAMILY/NONFAM-UPDATED
G42_1-G42_3          PR42.HH1 PREV WAVE FINAN/NONFIN-UPDATED
G43_1-G43_3          PR43.HH1 R WHICH HOUSEHOLD - UPDATED
G44_1-G44_3          PR44.HH1 1ST/2ND IW IN HH
G45_1-G45_3          PR45.R RELATION TO R - UPDATED
G46_1-G46_3          PR46.R LIVING - UPDATED
G47_1-G47_3          PR47.HH1 R IN NURSING HOME - UPDATED

G72_1-G72_3          PR72.HH2 R INDEX
G78_1-G78_3          PR78.HH2 WHICH ENTRY COHORT
G79_1-G79_3          PR79.HH2 R PREV WAVE IW - UPDATED
G82_1-G82_3          PR82.HH2 BIRTH YR - UPDATED
G83_1-G83_3          PR83.HH2 R MO BIRTH - UPDATED
G85_1-G85_3          PR85.HH2: R ORIG SAMP MEMB
G86_1-G86_3          PR86.HH2: R AGE ELIG
G88_1-G88_3          PR88.HH2 R SEX - UPDATED
G89_1-G89_3          PR89.HH2 R MARITAL - UPDATED
G90_1-G90_3          PR90.HH2 R COUPLENESS - UPDATED
G91_1-G91_3          PR91.S/P DIED DURING STUDY
G92_1-G92_3          PR92.R`S HH IW`ED IN PRIOR WAVE
G93_1-G93_3          PR93.HH2 R SAME SPOUSE/PARTNER - UPDATED
G94_1-G94_3          PR94.HH2 PREV WAVE FAM/NONFAM - UPDATED
G95_1-G95_3          PR95.HH2 PREV WAVE FIN/NONFIN - UPDATED
G97_1-G97_3          PR97.HH2 R 1ST/2ND IW IN HH - UPDATED
G98_1-G98_3          PR98.R RELATION TO R - UPDATED
G99_1-G99_3          PR99.R LIVING - UPDATED
G100_1-G100_3        PR100.HH2 R IN NURSING HOME - UPDATED
```

For the variables listed below, one of three variables contains information about the respondent. For instance, there are three variables, G178_1 to G178_3, that may contain information about the respondent's previous wave coupleness. You should use G474 to determine which of the three variables to use.

```
G160_1-G160_3        PR160.VIRGIN R INDEX
G166_1-G166_3        PR166.VIRGIN WHICH ENTRY COHORT
G170_1-G170_3        PR170.VIRGIN R PREV WAVE YR BIRTH
G171_1-G171_3        PR171.VIRGIN R MO BIRTH
G178_1-G178_3        PR178.VIRGIN R COUPLENESS
G179_1-G179_3        PR179.VIRG S/P DIED/NO NEW S/P YET
```

For the variables listed below, one of two variables contains information about the respondent. For instance, there are two variables, G50 and G103, that may contain information about the respondent's previous wave second residence. You should use G480 to determine which of the two variables to use.

```
G50                  PR50.HH1 PREV WAVE HAS 2ND RESIDENCE
G53                  PR53.HH1 2ND OWN RESIDENCE - UPDATED
G56                  PR56.HH1 R MOVED – UPDATED

G103                 PR103.HH2 PREV WAVE HAS 2ND RESIDENCE
G106                 PR106.HH2 2ND OWN RESIDENCE - UPDATED
G109                 PR109.HH2 1ST R MOVED - UPDATED
```

8C. Variables Combined in Household Member Child Data

The variables in the first column in the table below may be referenced in ASSIGNEMENT STATEMENTS, question text, or INAP text but do not appear in the Household-Member-Child file. The equivalent variables that do appear in the file are listed in the second column.

```
Q57 PR57.HH2 FAM MEM INDEX              Q4 PR4.HH1 FAM MEM INDEX
Q58 PR58.HH2 MEM RES/NR STAT            Q5 PR5.HH1 MEMBER RES/NR STAT
Q59 PR59.HH2 FAM LINE NO_LINEP          Q6 PR6.HH1 FAM LINE NO_LINEP
Q60 PR60.HH2 MEM STATUS                 Q7 PR7.HH1 STATUS
Q63 PR63.HH2 MEM SEX                    Q10 PR10.HH1 MEM SEX
Q64 PR64.HH2 REL TO IDFM                Q11 PR11.HH1 REL TO IDFM
Q65 PR65.HH2 REL TO IDNFM               Q12 PR12.HH1 REL TO IDNFM
Q66 PR66.HH2 MARITAL-CHILD              Q13 PR13.HH1 MARITAL-CHILD
Q67 PR67.HH2 SP STATUS                  Q14 PR14.HH1 SP STATUS
Q69 PR69.HH2 MEM YR BORN                Q16 PR16.HH1 MEM YR BORN
Q70 PR70.HH2 MEM SP YR BORN             Q17 PR17.HH1 SP YR BORN
Q71 PR71.HH2 PREV WAVE SP RES/NR STAT   Q18 PR18.HH1 SP RES/NR STAT
```

8D. Households with No or Two Coversheet or Family or Financial Respondents

As noted earlier in this document, the data collection design was to have asked most questions of all respondents, some questions of just a designated coversheet or family or financial respondent on behalf of the household. However, occasionally that is not what happened. For some households we did not obtain an interview from a family or financial respondent. And for a few other households we obtained interviews with two coversheet or family or financial respondents.

8D1. Households with No Coversheet or Family or Financial Respondent

Household records are provided for all households. Twenty-eight households had no coversheet respondent; 107 households had no family respondent; 97 households had no financial respondent. The household records for these households contain null values for the missing information.

Households missing a coversheet or family or financial respondent can identified, respectively, by values of "Blank. No coversheet/family/financial respondent" in these variables:

```
GPN_CS - 2000 COVERSHEET RESP PERSON NUMBER
GPN_FAM - 2000 FAMILY RESP PERSON NUMBER
GPN_FIN - 2000 FINANCIAL RESP PERSON NUMBER
```

in the household record.

8D2. Households with Two Family or Financial Respondents

Ninety-five households had two coversheet respondents; eleven households had two family respondents; 10 households had two financial respondents. The responses of one of these two respondents were selected for inclusion in the household-level files.

The respondents whose responses were used in the household records as well as those whose responses were not used are indicated, respectively, by the values of

```
GCSR - 2000 WHETHER COVERSHEET RESPONDENT
```

```
      GFINR - 2000 WHETHER FINANCIAL RESPONDENT
      GFAMR - 2000 WHETHER FAMILY RESPONDENT
```

in the respondent record.

8E.   Unfolding Bracket Variables and Imputations

For this data release, imputations for most unfolding bracket variables have
been and provided separately (instead of creating these summary bracket
variables as was done in some earlier waves). See the section Imputation File
for more details about the contents of this file.

Typically, a series of unfolding bracket questions followed a lead-in question
asking for an amount. If an actual amount was not given, a series questions
asked -- if the amount was less, more, or about a specified amount. Most often
the series consisted of two to four unfolding bracket questions.

The codeframes for the unfolding bracket questions were such as this:

```
      1. LESS THAN $5,000
      3. ABOUT $5,000
      5. MORE THAN $5,000
      8. DK (don't know); NA (not ascertained)
      9. RF (refused)
```

For some unfolding bracket series, a preload variable assigned an entry point,
low, medium or high, which was used to determine the asking order of the
unfolding bracket questions. For instance, consider the following set of
questions.

```
      0)    G5082 J8.R HOW MUCH WAGES AND SALARY
            J8. About how much wage and salary income did you receive in [Q779-
            CS24Y1.LCY], before taxes and other deductions?

      1)    G5083 J8A.R DK-5K        (low)
            J8a. Did it amount to less than $5,000, more than $5,000, or what?

      2)    G5084 J8Y1A.R DK-25000  (medium)
            J8Y1a. Did it amount to less than $25,000, more than $25,000, or
            what?

      3)    G5085 J8B.R DK-50000    (high)
            J8b. (Did it amount to) less than $50,000, more than $50,000, or
            what?

      4)    G5086 J8C.R DK-100000   (very high)
            J8c. (Did it amount to) less than $100,000, more than $100,000, or
            what?

      5)    G5087 J8Y1B.R DK-25000  (medium)
            J8Y1b. (Did it amount to) less than $25,000, more than $25,000, or
            what?

      6)    G5088 J8D.R DK-5000     (low)
            J8d. (Did it amount to) less than $5,000, more than $5,000, or what?
```

The order in which these follow-up questions were asked was determined by the value of a pre-load variable.

o For respondents with a low entry point, questions 1) low 2) medium 3) high 4) very high could be asked.
o For respondents with a medium entry point, questions 2) medium 3) high 4) very high 6) low could be asked.
o For respondents with a high entry point, questions 3) high 4) very high 5) medium 6) low could be asked.

This resulted in the following patterns of response for these six follow-up questions.

LOW ENTRY POINT

G5083 low 5K	G5084 medium 25K	G5085 high 50K	G5086 v high 100K	G5087 medium 50K	G5088 low 5K	Frequency
1	.	.	.	.	.	31
3	.	.	.	.	.	4
5	1	.	.	.	.	40
5	3	.	.	.	.	3
5	5	1	.	.	.	18
5	5	3	.	.	.	2
5	5	5	1	.	.	1
5	5	5	5	.	.	2
5	5	5	9	.	.	2
5	5	9	.	.	.	3
5	8	.	.	.	.	4
5	9	.	.	.	.	8
8	.	.	.	.	.	14
9	.	.	.	.	.	61

MEDIUM ENTRY POINT

G5083 low 5K	G5084 medium 25K	G5085 high 50K	G5086 v high 100K	G5087 medium 50K	G5088 low 5K	Frequency
.	1	.	.	.	1	23
.	1	.	.	.	3	5
.	1	.	.	.	5	54
.	1	.	.	.	8	3
.	1	.	.	.	9	1
.	3	.	.	.	.	7
.	5	1	.	.	.	21
.	5	3	.	.	.	1
.	5	5	1	.	.	5
.	5	5	5	.	.	2
.	5	8	.	.	.	1
.	5	9	.	.	.	4
.	8	.	.	.	.	11
.	9	.	.	.	.	43

HIGH ENTRY POINT

```
G5083    G5084    G5085    G5086    G5087    G5088    Frequency
 low      medium   high     v high   medium   low
 5K       25K      50K      100K     50K      5K
-------------------------------------------------------------
   .        .        1        .        1        1        18
   .        .        1        .        1        3         8
   .        .        1        .        1        5        69
   .        .        1        .        3        .        10
   .        .        1        .        5        .        23
   .        .        1        .        8        .         5
   .        .        1        .        9        .         4
   .        .        3        .        .        .         3
   .        .        5        1        .        .         5
   .        .        5        3        .        .         1
   .        .        5        5        .        .         2
   .        .        5        9        .        .         3
   .        .        8        .        .        .         9
   .        .        9        .        .        .        47
```

8F. Changes in Data because of Respondent Comments

```
The interviewer recorded comments the respondent made during the course of the
interview.  HRS staff reviewed these comments for selected questions.  The coded
answer was changed if it was determined that the comment changed the substance
of the recorded answer.  Occasionally additional codes were added to an existing
codeframe.  The text of added codeframes appears in the codebook in upper and
lower case.  Double reporting of income or assets in Section J was corrected
wherever possible.

When a change was made as a result of the comment review process,
inconsistencies with subsequent variables may result.  A limited number of
changes for consistency were made to immediately subsequent variables within the
section. The INAP text in the codebook does not include codeframes added in the
comment review process.  No consistency changes were made to variables in other
sections. If any comment change affected a subsequent branch-point, the branch-
point was not changed.

If a change was made to an actual amount variable followed by a series of
unfolding bracket questions, the answers to the unfolding bracket questions were
retained.
```

8G. Modules for the 2000 Data Collection

```
There were new modules for 2000 HRS.  They included modules about Medicare
knowledge, alternative medicine, planning and expectations for retirement,
social and economic altruism, benevolence and obligation, health plan booklet,
health utilities index, risk tolerance, alcohol consumption and IADL measures,
proxy validation, valuing health.

Module 1: Medicare knowledge: Module 1 asks questions and presents hypothetical
situations to ascertain respondents' knowledge about HMO and non-HMO Medicare
and about sources of their information about Medicare.

Module 2: Alternative medicine: This module covers respondents' recent and past
use of herbal or other dietary supplements and medications, treatments by
chiropractors, massage therapists, or acupuncturists, and spiritual practices
that may be related to health.
```

Module 3: Planning and expectations for retirement: Module 3 asks about activities R may have undertaken to plan for retirement; the questions are slightly different for those who are already fully retired than for those still anticipating full retirement. A short set of questions at the end is designed to get at R's propensity towards planning ahead.

Module 4: Economic altruism: This module is designed to ascertain a measure of the R's willingness to give regular financial assistance to help out relatives and friends at varying levels of need, as well as to charities.

Module 5: Benevolence and obligation: This module asks about R's self-perception as a giver to others, and his or her reasons for giving, especially as it relates to giving to family members.

Module 6: Request for health plan booklet: Module 6 is the request for R's health plan booklet to help assess the possibility of using this method to obtain details of R's health insurance coverage. This module appears in Section R, questions R84A. HAVE INSURANCE BOOKLET through R84G. REASSURANCE.

Module 7: Health utilities index: This module implements the Mark III version of the Health Utilities Index. Problems with vision, hearing, mobility, hand and arm use, mental functioning, general discomfort, and outlook on life are assessed. Domain scores and overall utility preference scores will be computed.

Module 8: Tolerance for large and small risks: Module 8 gets at R's willingness to take financial risks by posing a set of hypothetical situations and asking R to report what he or she would do in those situations. Part of the module is asked in section H, the "expectations" section, questions H18. RISK AVERSION-1 through H19. PLANNING/SAVING/SPENDING.

Module 9: Alcohol consumption and HRS Wave 1 IADL measures: A very short module with a few questions about difficulties of everyday activities and alcohol consumption, and one question in section B, B21C1. DEFINE A DRINK.

Module 10: Proxy validation: Module 10 repeats selected questions asked in sections B, C, and E about R's health and daily activities, but asks the R to answer these questions about his or her spouse/partner; the intention here is to be able to look at the agreement between R and Spouse reports of health problems of one of them.

Module 11: Social altruism: This module asks about R's connectedness to other people and the emotional support available to R from their spouse/partner, adult children, parents and parents-in-law, and friends.

Module 12: Valuing health: Module 12 asks respondents to compare their present state of health to perfect health using a standard time- tradeoff question valuing health in terms of years of life, and a similar willingness-to-pay question valuing health in terms of money.

Included in section M, although not a module, is a section with questions labeled RC. It is a section of the interview that is administered to sample respondents for whom proxy interviews were done, but who were willing and able to do this section themselves.

8H.   Cognition Sections

The questionnaire contained three series of cognition questions found in sections C – Cognition, PC - Proxy Cognition and RC - Respondent Cognition. Section C was asked of respondents who answered the interview for themselves. If a proxy answered the interview, section PC was asked of the proxy about the sample respondent and, if the interview was not a "cognitive interview" (G517=4) and the sample respondent agreed, section RC was asked of the sample respondent directly.

8I.   Two Versions of Masked Occupation Code in 1998 Only

For AHEAD data collections in 1993 and 1995, occupation was coded using the SRC 2-Digit Occupation Code (adapted from the 1980 U. S. Census Occupation Code) and masked for public release. For HRS data collections in 1992, 1994 and 1996, occupation was coded using the 3-digit 1980 U. S. Census Occupation Code and masked for public release. The AHEAD and HRS masked occupation codeframes are not the same.

For 1998, the first year that the AHEAD and HRS sub-samples were merged for data collection, we released occupation masked in two ways for the AHEAD sub-sample – using both the AHEAD and HRS masking schemes. Variables masked in the AHEAD tradition are named with a "AM" at the end of the variable name and variables masked in the HRS tradition are named with a "HM" at the end of the variable name, e.g., "F3185AM" and "F3185HM". In 2000 and subsequent years we will use just the HRS masking scheme for all sub-samples. See our Web site for more detailed information about how occupation and industry were coded in each wave of the study.

8J.   Variables Not Included in This Release

A few variables have not been included in the 2000 HRS Core (Final, Version 1.0) data release. Some variables have not been included for reasons of confidentiality. They may be made available to analysts as restricted data. See our Web site for details.

```
31     PR31.HH1 R DY BIRTH
84     PR84.HH2 R DY BIRTH
172    PR172.VIRGIN R DY BIRTH
521    CS2a2.DAY BORN
534    CS2Y3.R DAY BORN
537    CS2Y6.R DAY BORN
1118   A26c.YEARS LIVED MAIN HU AREA
1119   A26Y1c.(A26c) SINCE YEAR LIVE MAIN HU
1271   B5g.CANCER START
6208   N63d.CAUSE OF DEATH
```

Some open-ended questions were not coded during the initial round of coding.

```
6372   R74.WHEN ELIGIBLE
6698   M2-3hs. WHAT OTHERS?
6750   M2-5s1. OTHERS HERBS USED
7142   M12-6. MOST IMPORTANT MED PROB
```

9. Longitudinal Issues

9A.   The Sample Design

As of 2000, the HRS sample was comprised of four sub-samples.

> The first sub-sample, the HRS sub-sample, consists of people who were born 1931 through 1941 and were household residents of the conterminous U.S. in the spring 1992, and their spouses or partners at the time of the initial interview in 1992 or at the time of any subsequent interview. The HRS sub-sample was interviewed in 1992 and every two years thereafter.

> The AHEAD sub-sample consists of people who were born in 1923 or earlier, were household residents of the conterminous U.S. in the spring 1992, and were still household residents at the time of their first interview in 1993 or 1994, and their spouses or partners at the time of the initial interview or at the time of any subsequent interview. The AHEAD sub-sample was interviewed in 1993-94, 1995-96, 1998 and every two years thereafter.

> The War Babies sub-sample consists of people who were born 1942 through 1947, were household residents of the conterminous U.S. in the spring 1992, who, at that time, did not have a spouse or partner born before 1924 or between 1931 and 1941, and were still household residents at the time of the first interview in 1998, and their spouses or partners at the time of the initial interview or at the time of any subsequent interview. The War Baby sub-sample was interviewed in 1998 and every two years thereafter.

> The Children of the Depression sub-sample consists of people who were born 1924 through 1930, were household residents of the conterminous U.S. when first interviewed in 1998, and who, at that time, did not have a spouse or partner who was born before 1924 or between 1931 and 1947, and their spouses or partners at the time of the initial interview or at the time of any subsequent interview. The Children of the Depression sub-sample was interviewed in 1998 and every two years thereafter.

Original sample members are those selected as described above and their spouses or partners at the time of the initial interview in 1992 (HRS), 1993 (AHEAD) or 1998 (CODA or WB). For more details about the sample, see our Web site.

9B.   Linking Respondents across Time

Respondent records from all prior waves may be linked to respondent records from 2000 by HHID and PN. The sub-household identifiers, ASUBHH for 1992, BSUBHH for 1993, W2SUBHH for 1994, DSUBHH for 1995, ESUBHH for 1996, FSUBHH for 1998 and GSUBHH for 2000 can be used to link cross-sectional household data with the cross-sectional respondent data.

9C.   Linking Other Persons across Time

Linking other person records across time is more complex. You can use HHID, the previous wave's SUBHH and OPN to link one wave's other person record with the corresponding other person record from the previous wave. For example, to link 2000 other person records with 1998 other person records use HHID, FSUBHH, and OPN.

For some 2000 other person records this will result in two 2000 records matching one 1998 record. In an example like <u>Couple with Children and Siblings Divorces.</u> where a couple with children divorces, the 2000 record where HHID is 89012, GSUBHH is 1, and OPN is 101 is the report about the child from the mother and the record where HHID is 89012, GSUBHH is 2, and OPN is 101 is the report about the same child from the father. The value of FSUBHH for each of these 2000 records would 0. Each of these 2000 records would match the 1998 record about the child before the couple divorced as the 1998 record would have a HHID of 089012, a FSUBHH of 0 and a OPN of 101.

For other 2000 other person records this will result in no 2000 record matching a 1998 record. In an example like <u>Couple Divorces, One Respondent Remarries, Both Split-off Households Have New Members.</u> where the couple divorces and both split-off households have new members, the record where HHID is 90123, GSUBHH is 2, and OPN is 151 is for the ex-wife's cohabiting mother, while the record where HHID is 90123, GSUBHH is 1, and OPN is 151 is for the ex-husband's new stepchild. The value of FSUBHH for each of these 2000 records would 0. Neither of these 2000 records would match a 1998 record before the couple divorced because no 1998 record would have a HHID of 089012, a FSUBHH of 0 and a OPN of 151.

However, be aware that the technique of matching OPN records by HHID, previous wave SUBHH and OPN does not work well for OPNs who drop out for one or more waves and later reappear, for waves of data when the value of 999 was used for many OPNs, for persons who assumed the OPN number of their deceased spouse, for persons with more than one OPN or for OPNs used by more than one person.

9D. Summary of Changes from Prior Waves

Over the course of the study there have been a number of changes in key variables. There are no new changes of this type for 2000. However, since the analyst may be working with some or all of the earlier waves of data, changes from previous years are documented below.

9D1. Identification Variable HHID Has One Extra Digit

A "0" was added to the front of the identification variable HHID for HRS sub-sample households, in order to make it possible to combine HRS and AHEAD sub-samples for the 1998 data collection. The change makes the old HRS five-digit HHID six digits long. For example, old HHID "12345" becomes "012345" as of HRS 1996 and subsequent years.

Similarly, a "20" was added to the front of the identification variable HHID for AHEAD sub-sample households, in order to make it possible to combine HRS and AHEAD sub-samples for the 1998 data collection. The change makes the old AHEAD four-digit HHID six digits long. For example, old HHID "1234" becomes "201234".

You will want to be aware of these changes when preparing merges with previous years' data that you have downloaded earlier with five-digit or four-digit HHIDs. Early data sets currently on our Web site have been retrofitted with six-digit HHIDs, albeit numeric, see below.

9D2. Character Type Identification Variables

In some prior releases of HRS (or AHEAD), all variables were stored in NUMERIC format. Releasing identification variables in numeric format caused some problems.

o All digits of an identifier matter. The character string "000031010" is
   not the same thing as the numeric string "31010".

o When we began to deal with complex respondent identifiers such as those in
   HRS and attempted to create one identification variable by combing
   several, we ran the risk of corrupting the identifier due to loss of
   precision when treating it as a numeric value.

Thus starting with the 1995 AHEAD, we release all primary and secondary
identification variables in character format and no longer create combined
identification variables.

Also released as character variables are variables that include an other person
number that may be used for merging with a primary or secondary identification
variable. Typically these are variables from questions asking who or which
person did something. Examples are:

```
G1982     D01B.LIVES NEAREST
G1987_1   D03BS.WHICH CHILD MARRIAGE CHANGE
G2260001  D146B.WHICH SIB HELPED
G2880     E99.IADLS-WHO HELP-1
```

Most software packages will require that the identification variables in each
dataset be of the same type. We suggest that analysts of older HRS or AHEAD
datasets convert the numeric identifiers to character type.

For the most part, we leave the method of this transformation to you. However,
we do have a couple of suggestions.

o Before reading in datasets, you can change the variable type in the
   program statements themselves.

```
Package  Type       Statement for variable V1
-------  ---------  ------------------------------
SAS      numeric    V1 1-4
         character  V1 $ 1-4
-------  ---------  ------------------------------
Stata    numeric    _column(1) byte V1 %4.0f "LABEL"
         character  _column(1) str4 V1 %4s   "LABEL"
-------  ---------  ------------------------------
SPSS     numeric    V1 1-4
         character  V1 1-4 (A)
```

o Each software package has different ways of converting variable types in
   an existing dataset.

   In SAS, for example, to convert numeric variable HHID to character
   variable HHID, you can use this code.

```
     *      numeric to character;
     CHHID=put(HHID,z6.0);
     attrib
     CHHID label='HOUSEHOLD IDENTIFIER' length=$6 format=$char6.;
     drop HHID;
     rename CHHID=HHID;
```

In STATA to convert numeric variable HHID to character variable HHID, you
can use this code.

```
*        numeric to string
gen str6 chhid=substr(string(hhid+1000000),2,6)
drop hhid
rename chhid hhid
```

9D3. INAP Codes Stored as Blanks

Beginning with the 1995 AHEAD data, INAP codes were stored in the ASCII data
files as blanks. In 1992 and 1994, INAPs were assigned to a number, usually
zero or a value ending in "96". In 1993, INAPs were assigned to a system
missing value. If you are merging the 1998 data with data from 1992, 1993 or
1994, you will want to be aware of these differences.

10.   Other Data Files

10A. Tracker File

We attempt to aid you in tracking the HRS sample longitudinally through use of
our "tracker file." The tracker file contains a record for every individual who
was interviewed one or more times during the period 1992 through the most recent
year as well as a record for the non-responding spouse or partner of an original
sample member. The primary identification variables for the tracker file are
HHID and PN.

The tracker file contains basic demographic information as well as some basic
information obtained from the National Death Index (NDI). In addition it
contains basic information about the interview year such as whether a person
gave an interview in a particular wave so that it can be used to understand the
interview history of respondents. The tracker file also contains cross-
sectional weights for each interview year 1992 through the most recent year for
both respondent and household level analysis.

The current version of the tracker file, the Cross-Wave: Tracker File (v.2.0),
is based on 1998 HRS Core (Early) (v.1.0) and previous years' interviews.
Version 3.0 of the tracker file, based on information from the 2000 HRS Core
(Final, Version 1.0) and previous years' interviews will be available soon. The
tracker file is available at our Web site.

10B. Imputation File

Imputed variables based on the 2000 HRS Core (Early, v.1.0) data release may be
found at our Web site in the 2000 HRS Core Imputations (Early, v.1.0,
10/26/2001). Imputations are provided for the variables satisfying our ad hoc
selection rules listed below.

1) The Bracket Rule: In general, all the variables with brackets will
   imputed, and included in the imputation file.
2) The Income/Asset Component Rule: A variable without brackets will be
   imputed if it is determined to be an integral component of household
   income, housing equity, or net wealth.
3) The Imputability Rule: A variable eligible for imputation based on Rule
   1 or 2 will not be imputed if there are not enough valid observations.

Also included are four aggregate variables -- main home equity, and second home equity, total assets, and household income. The data are preliminary in the sense that they were generated based on the early core data release.

The N's in the early imputation files and final core files differ slightly. The values for GSUBHH are different for about fifteen households. The early imputation section D file is released at the respondent-level, but the variables contained therein map to corresponding variables included final core section D household, transfer-to-child, transfer-from-child-level files.

2000 HRS Core Imputations (Early, v.1.0) Files

        Respondent-level files (N=19,581)
                Section D   Family Structure
                Section E   Health Care Utilization
                Section R   Health Insurance

        Household-level files (N=13,118)
                Section F   Housing
                Section J   Assets and Income
                Section N   Asset Changes

Imputation files based on the 2000 HRS Core (Final, Version 1.0) data release will be available at a later date.

11.   Obtaining the Data

11A. Registration and Downloading the Data

HRS data are available for free to researchers and analysts at the HRS Web site. In order to obtain public release data, you must first register at our Web site. Once you have completed the registration process, your username and password will be sent to you via e-mail. Your username and password are required to download any data files.

By registering all users, we are able to document for our sponsors the size and diversity of our user community allowing us to continue to collect these important data. Registered users receive user support, information related to errors in the data, future releases, workshops, and publication lists. The information you provide will not be used for any commercial use, and will not be redistributed to third parties.

11B. Conditions of Use

By registering, you agree to the Conditions of Use governing access to Health and Retirement public release data. You must agree to

    o   not attempt to identify respondents
    o   not transfer data to third parties except as specified
    o   not share your username and password
    o   include specified citations in work based on HRS data
    o   provide information to us about publications based on HRS data
    o   report apparent errors in the HRS data or documentation files
    o   notify us of changes in your contact information

For more information concerning privacy issues and conditions of use, please read "Conditions of Use for Public Data Files" and "Privacy and Security Notice" at the Public File Download Area of the HRS Web site.

11C. Publications Based on Data

As part of the data registration process, you agree to include specified citations and to inform HRS of any papers, publications, or presentations based on HRS data. Please send a copy of any publications you produce based on HRS data, with a bibliographical reference, if appropriate, to the address below.

> Health and Retirement Study
> Attn: Papers and Publications
> The Institute for Social Research, Room 3050
> P.O. Box 1248
> Ann Arbor, MI (USA) 48106-1248

Alternately, you may contact us by e-mail at hrsquest@isr.umich.edu with "Attn: Papers and Publications" in the subject line.

12.   If You Need to Know More

This document is intended to serve as a brief overview and to provide guidelines to using the 2000 HRS Core (Final, Version 1.0) data. If you have questions or concerns that are not adequately covered here or on our Web site, or if you have any comments, please contact us. We will do our best to provide answers.

12A. HRS Internet Site

Health and Retirement Study public release data and additional information about the study are available on the Internet. To access the data and other relevant information, point your Web browser to the HRS Web site. Our new, as of 9/1/2002, URL is:

> http://hrsonline.isr.umich.edu/

12B. Contact Information

If you need to contact us, you may do so by one of the methods listed below.

> Internet: Help Desk at our Web site
>
> E-mail: hrsquest@isr.umich.edu
>
> Postal service:
>   Health and Retirement Study
>   The Institute for Social Research, Room 3050
>   The University of Michigan
>   P.O. Box 1248
>   Ann Arbor, MI 48106-1248
>
> FAX: (734) 647-1186

Appendices

A. Examples of Merging

Many analyses require variables that appear in separate files. Sometimes you will need to obtain variables from files at different levels that contain different numbers of records. Before you can do your analysis work, the files will need to be merged in an appropriate manner. Prior to doing any data management you should ask yourself several questions.

- o What are the variables of interest? Identifying the variables needed for an analysis allows you to subset files to include only the necessary variables, weights, and identification variables. Smaller files are more manageable.
- o What should be the level of the analysis file – will it have one record per respondent, one record per household or what?
- o What identification variables will be required to merge the various files that contain variables needed for your analysis?
- o Will the merge be a one-to-one matching of records or a one-to-many matching?

We've provided sample setups for several types of merging. We hope that these will be of help to you as you think of your particular analytic needs.

A1.   Respondent-level Files

To create a respondent-level file with data from two or more respondent-level files, merge the respondent-level files using HHID and PN. This is a one-to-one match. Each input file contains 19,580 records. A respondent-level output file with 19,580 respondent records results.

A2.   Household-level Files

To create a household-level file with data from two or more household-level files, merge the household-level files using HHID and GSUBHH. This is a one-to-one match. Each input file contains 13,214 records. A household-level output file with 13,214 household records results

A3.   Respondent-level Files with Household-level Data

To create a respondent-level file including data from a household-level file, merge the respondent-level file(s) and the household-level file(s) using HHID and GSUBHH. This is a one-to-many match (one household-to-many respondents). Household-level input files contain 13,214 records; respondent-level input files contain 19,580 records. A respondent-level output file with 19,580 respondent records results.

Since you are creating a respondent-level output file, be sure to keep PN, as well as HHID and GSUBHH, in the output file even though PN is not used for the merging.

A4.   Household-level Files with Respondent-level Data

To create a household-level file including variables from both respondents requires several steps.

First, separate the respondent-level records into two groups, one for the family respondent, the first person, and one for the non-family respondent, the second person.

We suggest you use GFAMR – 2000 WHETHER FAMILY RESPONDENT to create the two groups of respondents. The first group will contain 13,107 records (107 households did not have a family respondent); the second group will contain 6,473 records.

If you choose to use another variable to create two respondent groups, be sure each respondent group does not have more than one person from any one household (otherwise data will be lost).

Second, rename all variables for the second respondent group (except HHID and GSUBHH) to avoid overlap in the combined file.

Third, merge the two respondent files and the household-level file using HHID and GSUBHH.

Although the respondent groups contain fewer than 13,214 records, this is a one-to-one match. A household-level file with 13,214 household records results. The household-level output file includes household variables, a set of variables for the first respondent and a set of variables for the second respondent.

We recommend that you keep PN for the first and second respondent even though PN is not used for the merging and the output file is a household-level file.

A5. Creating a Respondent-level File with Parent Data from a Household-level File

Questions about parents of both the respondent and the respondent's spouse or partner were asked of the family respondent. This information is distributed at a household-level. You may wish to create a respondent-level file with information about the respondent's parents, whether the respondent was the family respondent or not. In order to do this, you will have to obtain the information about the parents from the household record and merge the variables to the proper respondent record.

In the example below, the file "famr" will have 13,107 observations one for each family respondent. The file "nfamr" will have 6,473 observations one for each non-family respondent. The combined file "resp" will have 19,580 observations one for each respondent.

A6. SAS Examples

A6a.   SAS Example of Respondent to Respondent Merge

```
*--------------------------------------------------------------------------;
*      sort each input dataset by HHID PN
       subset variables;

proc sort
       data=in.h00a_r
       out=h00a_r(keep=hhid pn g1051 g1053);
       by hhid pn;
run;

proc sort
```

```
             data=in.h00b_r
             out=h00b_r(keep=hhid pn g1229 g1233);
             by hhid pn;
      run;

      proc sort
             data=in.h00c_r
             out=h00c_r(keep=hhid pn g1654 g1655);
             by hhid pn;
      run;

      *-----------------------------------------------------------------------;
      *      merge three datasets by HHID PN to create a respondent file
             this is a one-to-one merge;
      data resp;
             merge h00a_r h00b_r h00c_r;
             by hhid pn;
      run;
```

A6b.   SAS Example of Household to Household Merge

```
      *-----------------------------------------------------------------------;
      *      sort each input dataset by HHID GSUBHH
             subset variables;
      proc sort
             data=in.h00f_h
             out=h00f_h(keep=hhid gsubhh g3060 g3061);
             by hhid gsubhh;
      run;

      proc sort
      data=in.h00j_h
             out=h00j_h(keep=hhid gsubhh g5068);
             by hhid gsubhh;
      run;

      *-----------------------------------------------------------------------;
      *      merge two datasets by HHID GSUBHH to create a household file
             this is a one-to-one merge;
      data hhold;
             merge h00f_h h00j_h;
             by hhid gsubhh;
      run;
```

A6c.   SAS Example of Household to Respondent Merge

```
      *-----------------------------------------------------------------------;
      *      sort each input dataset by HHID GSUBHH
             subset variables;

      proc sort
             data=in.h00a_r
             out=h00a_r(keep=hhid gsubhh pn g1051 g1053);
             by hhid gsubhh;
      run;
      proc sort
             data=in.h00j_h
             out=h00j_h(keep=hhid gsubhh g5068);
```

```
        by hhid gsubhh;
run;
*----------------------------------------------------------------------;
*       merge two datasets by HHID GSUBHH to create a respondent file
        this is a one-to-many merge;
data resp;
        merge h00a_r h00j_h;
        by hhid gsubhh;
run;
```

A6d.   SAS Example of Respondent to Household Merge

```
*----------------------------------------------------------------------;
*       file of family respondents;
proc sort
        data=in.h00a_r(keep=hhid gsubhh gfamr pn g1051 g1053
        rename=(gfamr=gfamra pn=pna g1051=g1051a g1053=g1053a)
        where=(gfamra in(1 2 3 4)))
        out=h00a1;
        by hhid gsubhh;
run;


*----------------------------------------------------------------------;
*       file of non-family respondents;
proc sort
        data=in.h00a_r(keep=hhid gsubhh gfamr pn g1051 g1053
        rename=(gfamr=gfamrb pn=pnb g1051=g1051b g1053=g1053b)
        where=(gfamrb not in(1 2 3 4)))
        out=h00a2;
        by hhid gsubhh;
run;


*----------------------------------------------------------------------;
*       variables from a household-level file;
proc sort
        data=in.h00j_h(keep=hhid gsubhh g5068)
        out=h00j;
        by hhid gsubhh;
run;


*----------------------------------------------------------------------;
*       household-level file with information from two respondents;
data hhold;
        merge
        h00a1 h00a2 h00j;
        by hhid gsubhh;
run;
```

A6e.   SAS Example of Respondent File with Parent Information from a Household-level File

```
*----------------------------------------------------------------------;
*       select parent variables from household file for family r;
data famr;
        set in.h00d_h;
        *       keep id variables;
        keep hhid pn;
        *       keep parent variables of interest;
        keep g2122 g2123;
```

```
            *       assign person number for family r;
            pn=gpn_fam;
            attrib pn label='PERSON NUMBER' format=$char3.;
            *       output records for households with family R;
            if pn ne '';
      run;

      *--------------------------------------------------------------------------;
      *       select parent variables from household file for non-family r;
      data nfamr;
            set in.h00d_h;
            *       keep id variables;
            keep hhid pn;
            *       keep parent variables;
            keep g2122 g2123;
            *       copy in-law variables in to output variable location;
            g2122=g2309;
            g2123=g2310;
            *       assign person number for family r;
            pn=gpn_nfam;
            attrib pn label='PERSON NUMBER' format=$char3.;
            *       output records for households with non-family R;
            if pn ne '';
      run;

      *--------------------------------------------------------------------------;
      *       concatenate files;
      data resp;
            set famr nfamr;
      run;
```

A7. Stata Examples

A7a. *Stata Example of Respondent to Respondent Merge*

```
      set prefix "hrs2000"
      * This is an optional statement
      * Make sure all the data files are in directory "hrs2000"

      use HHID PN G1229 G1233 using h00b_r
      sort HHID PN
      save b_r, replace

      use HHID PN G1654 G1655 using h00c_r
      sort HHID PN
      save c_r, replace

      *       merge three datasets by HHID PN to create a respondent file
      *       this is a one-to-one merge

      use HHID PN G1051 G1053 using h00a_r
      sort HHID PN
      merge HHID PN using b_r
      drop _m
      sort HHID PN
      merge HHID PN using c_r
      drop _m
```

```
                sort HHID PN
                save resp, replace


A7b.    Stata Example of Household to Household Merge
                *----------------------------------------------------------------------------
                *       sort each input dataset by HHID GSUBHH
                *       subset variables

                use HHID GSUBHH G3060 G3061 using h00f_h
                sort HHID GSUBHH
                save f_h, replace

                use HHID GSUBHH G5068 using h00j_h
                sort HHID GSUBHH
                merge HHID GSUBHH using f_h
                drop _m
                sort HHID GSUBHH
                save hhold, replace


A7c.    Stata Example of Household to Respondent Merge
                *----------------------------------------------------------------------------
                *       sort each input dataset by HHID GSUBHH
                *       subset variables

                use HHID GSUBHH G5068 using h00j_h
                sort HHID GSUBHH
                save j_h, replace


                *----------------------------------------------------------------------------
                *       merge two datasets by HHID GSUBHH to create a respondent file
                *       this is a one-to-many merge

                use HHID PN GSUBHH G1051 G1053 using h00a_r
                sort HHID GSUBHH
                merge HHID GSUBHH using j_h
                tab _m
                drop _m
                save resp, replace


A7d.    Stata Example of Respondent to Household Merge
                *----------------------------------------------------------------------------
                *       variables from a household-level file

                use HHID GSUBHH G5068 using h00j_h
                sort HHID GSUBHH
                save j_h, replace


                *----------------------------------------------------------------------------
                *       file of family respondents

                use HHID GSUBHH GFAMR PN G1051 G1053 using h00a_r
                keep if GFAMR==1|GFAMR==2|GFAMR==3|GFAMR==4
                rename GFAMR GFAMRa
                rename PN PNa
                rename G1051 G1051a
                rename G1053 G1053a
```

```
      sort HHID GSUBHH
      save a_r1, replace

      *-----------------------------------------------------------------------
      *       start with file of non-family respondents and merge to create

      *-----------------------------------------------------------------------
      *       household-level file with information from two respondents;

      use HHID GSUBHH GFAMR PN G1051 G1053 using h00a_r
      keep if GFAMR~=1 & GFAMR~=2 & GFAMR~=3 & GFAMR~=4
      rename GFAMR GFAMRb
      rename PN PNb
      rename G1051 G1051b
      rename G1053 G1053b
      sort HHID GSUBHH
      merge HHID GSUBHH using a_r1
      tab _m
      drop _m
      sort HHID GSUBHH
      merge HHID GSUBHH using j_h
      tab _m
      drop _m
      sort HHID GSUBHH
      save hhold, replace
```

A7e.   Stata Example of Respondent File with Parent Information from a Household-level File

```
      *-----------------------------------------------------------------------
      *       select parent variables from household file for family r
      use h00d_h, clear

      * assign person number for family r
      gen str3 PN=GPN_FAM
      label var PN "PERSON NUMBER"

      * keep ID variables and parent variables of interest
      keep HHID PN G2122 G2123

      *       output records for households with family R
      keep if PN~=""
      save fam1, replace

      * select parent variables from household file for non-family r
      use h00d_h, clear
      drop G2122 G2123

      *-----------------------------------------------------------------------
      *       select parent variables from household file for non-family r
      *       assign person number for family r
      gen str3 PN=GPN_NFAM

      * copy in-law variables in to output variable location
      gen G2122=G2309
      gen G2123=G2310

      * keep ID variables and parent variables of interest
```

```
        keep HHID PN G2122 G2123

        *     output records for households with non-family R
        keep if PN~=""
        save fam2, replace

        *-------------------------------------------------------------------------
        * concatenate  files
        use fam1, clear
        append using fam2
        save resp, replace
```

A8.   SPSS Examples

A8a.   SPSS Example of Respondent to Respondent Merge

```
        /*-------------------------------------------------------------------------
        /*    sort each input dataset by HHID PN
        /*    subset variables

        GET FILE 'c:\hrs2000\spss\h00a_r.sav'
             /keep=hhid pn g1051 g1053.
        SORT CASES BY hhid pn.
        save outfile='c:\temp\h00a_r.sav'.

        GET FILE 'c:\hrs2000\spss\h00b_r.sav'
             /keep=hhid pn g1229 g1233.
        SORT CASES BY hhid pn.
        save outfile='c:\temp\h00b_r.sav'.

        GET FILE 'c:\hrs2000\spss\h00c_r.sav'
             /keep=hhid pn g1654 g1655.
        SORT CASES BY hhid pn.
        save outfile='c:\temp\h00c_r.sav'.

        /*-------------------------------------------------------------------------
        /*    merge three datasets by HHID PN to create a respondent file
        /*    this is a one-to-one merge

        MATCH FILES
             /FILE='c:\temp\h00a_r.sav'
             /FILE='c:\temp\h00b_r.sav'
             /FILE='c:\temp\h00c_r.sav'
             /BY hhid pn.
        EXECUTE.

        save outfile='c:\hrs2000\spss\resp.sav'.
```

A8b.   SPSS Example of Household to Household Merge

```
        /*-------------------------------------------------------------------------
        /*    sort each input dataset by HHID GSUBHH
        /*    subset variables

        GET FILE 'c:\hrs2000\spss\h00f_h.sav'
             /keep=hhid gsubhh g3060 g3061.
        SORT CASES BY hhid gsubhh.
        save outfile='c:\temp\h00f_h.sav'.
```

```
GET FILE 'c:\hrs2000\spss\h00j_h.sav'
      /keep=hhid gsubhh g5068.
SORT CASES BY hhid gsubhh.
save outfile='c:\temp\h00j_h.sav'.


/*------------------------------------------------------------------------
/*     merge two datasets by HHID GSUBHH to create a household file
/*     this is a one-to-one merge

MATCH FILES
      /FILE='c:\temp\h00f_h.sav'
      /FILE='c:\temp\h00j_h.sav'
      /BY hhid gsubhh.
EXECUTE.

save outfile='c:\hrs2000\spss\hhold.sav'.
```

A8c.   *SPSS Example of Household to Respondent Merge*

```
/*------------------------------------------------------------------------
/*     sort each input dataset by HHID GSUBHH
/*     subset variables

GET FILE 'c:\hrs2000\spss\h00a_r.sav'
      /keep=hhid gsubhh pn g1051 g1053.
SORT CASES BY hhid gsubhh.
save outfile='c:\temp\h00a_r.sav'.


GET FILE 'c:\hrs2000\spss\h00j_h.sav'
      /keep=hhid gsubhh g5068.
SORT CASES BY hhid gsubhh.
save outfile='c:\temp\h00j_h.sav'.


/*------------------------------------------------------------------------
/*     merge two datasets by HHID GSUBHH to create a respondent file
/*     this is a one-to-many merge

MATCH FILES
      /FILE='c:\temp\h00a_r.sav'
      /TABLE='c:\temp\h00j_h.sav'
      /BY hhid gsubhh.
EXECUTE.

save outfile='c:\hrs2000\spss\resp.sav'.
```

A8d.   *SPSS Example of Respondent to Household Merge*

```
/*------------------------------------------------------------------------
/*     file of family respondents

GET FILE 'c:\hrs2000\spss\h00a_r.sav'
      /keep=hhid gsubhh gfamr pn g1051 g1053.
USE ALL.
SELECT IF(gfamr  = 1 or gfamr = 2 or gfamr = 3 or gfamr=4).
sort cases by hhid gsubhh.
rename variables  gfamr=gfamra.
rename variables  pn=pna.
```

```
        rename variables  g1051=g1051a.
        rename variables  g1053=g1053a.
        EXECUTE .


        save outfile='c:\temp\h00a1.sav'.

        /*-------------------------------------------------------------------------
        /*     file of non-family respondents

        GET FILE 'c:\hrs2000\spss\h00a_r.sav'
             /keep=hhid gsubhh gfamr pn g1051 g1053.
        sort cases by hhid gsubhh.
        USE ALL.
        SELECT IF(gfamr  <> 1 and gfamr <> 2 and gfamr <> 3 and gfamr<>4).
        rename variables  gfamr=gfamrb.
        rename variables  pn=pnb.
        rename variables  g1051=g1051b.
        rename variables  g1053=g1053b.
        EXECUTE .
        save outfile='c:\temp\h00a2.sav'.

        /*-------------------------------------------------------------------------
        /*     variables from a household-level file

        GET FILE 'c:\hrs2000\spss\h00j_h.sav'
             /keep=hhid gsubhh g5068.
        SORT CASES BY hhid gsubhh.
        save outfile='c:\temp\h00j.sav'.

        /*-------------------------------------------------------------------------
        /*     household-level file with information from two respondents

        MATCH FILES
             /FILE='c:\temp\h00a1.sav'
             /FILE='c:\temp\h00a2.sav'
             /FILE="c:\temp\h00j.sav"
             /BY hhid gsubhh.
        EXECUTE.

        save outfile='c:\hrs2000\spss\hhold.sav'.
```

A8e.   SPSS Example of Respondent File with Parent Information from a Household-level File

```
        /*-------------------------------------------------------------------------
        /*     select parent variables from household file for family r

        /*     keep id variables and parent variables of interest
        GET FILE 'c:\hrs2000\spss\h00d_h.sav'
             /keep=hhid gpn_fam g2122 g2123.

        /*     assign person number for family r
        STRING pn (A3).
        COMPUTE pn = gpn_fam .
        VARIABLE LABELS pn 'Person Number' .

        USE ALL.
        /*     output records for households with family R
```

```
SELECT IF(pn <>'').
EXECUTE.
save outfile='c:\temp\famr.sav' /drop= gpn_fam.

/*--------------------------------------------------------------------------
/*    select parent variables from household file for non-family r

/* keep id variables and parent variables of interest
GET FILE 'c:\hrs2000\spss\h00d_h.sav'
     /keep=hhid gpn_nfam g2309 g2310.

/*    assign person number for family r
STRING pn (A3).
COMPUTE pn = gpn_nfam .
VARIABLE LABELS pn 'Person Number' .

/*    copy in-law variables in to output variable location
rename variables g2309=g2122.
rename variables g2310=g2123.
USE ALL.

/*    output records for households with non-family R
SELECT IF(pn <>'').
EXECUTE .

save outfile='c:\temp\nfamr.sav' /drop= gpn_nfam.

/*--------------------------------------------------------------------------
/*    concatenate files
ADD FILES
     /FILE='c:\temp\nfamr.sav'
     /FILE='c:\temp\famr.sav'.
EXECUTE.

save outfile='c:\hrs2000\spss\resp.sav'.
```

B. Examples of Sub-Household and Respondent Person Number and Other Person Number Assignments

In the first year of data collection, all households, consisting of either a single respondent or of two married or partnered respondents, were assigned a SUBHH of 0.

In subsequent waves, a SUBHH of 0 indicates that the original household has not split due to divorce or separation of spouses or partners, although one member of a couple may have died or a single respondent may have become married or partnered.

A value of 1 or 2 indicates a household in which the original couple split, divorced or separated. One of the original couple is assigned a SUBHH of 1; the other is assigned a SUBHH of 2.

A value of 5 or 6 indicates a previously split household split a second time. One of the couple from a SUBHH 1 or 2 retains a SUBHH of 1 or 2; the other is assigned a SUBHH 5 or 6.

A value of 7 indicates respondents from split household reunited[1].

It is important to understand these assignments when you merge records from different waves of the study.

B1.   Married Couple Stays Married.

Two respondents in a sample household are married at the time of the first cross-section. Each respondent is assigned a HHID of 012345 and a SUBHH of 0. One respondent has a PN of 010, the other a PN of 020.

At the time of the second cross-section the two respondents are still married, and each retains their HHID of 012345 and their SUBHH of 0 and his and her PN of 010 and 020, respectively.

```
Time 1
Household records
     HHID=012345 ASUBHH=0
Respondent records
     HHID=012345 PN=010 ASUBHH=0
     HHID=012345 PN=020 ASUBHH=0

Time 2
Household records
     HHID=012345 CSUBHH=0
Respondent records
     HHID=012345 PN=010 CSUBHH=0
     HHID=012345 PN=020 CSUBHH=0
```

---

[1] In addition, a SUBHH of 3 or 4 indicates the "household" a deceased respondent who is considered to be in a household of his or her own. These values do not occur in these files because all records in these files are from living respondents.

B2.   Couple Divorces.

Two respondents in a sample household are married at the time of the first cross-section. Each respondent is assigned a HHID of 023456 and a SUBHH of 0. One respondent has a PN of 010, the other a PN of 020.

By the time of the second cross-section, the couple has divorced. Both respondents retain the HHID of 023456, but one is assigned a SUBHH of 1 and the other is assigned a SUBHH of 2. Each original respondent retains his and her PN of 010 and 020, respectively.

```
Time 1
Household records
     HHID=023456 ASUBHH=0
Respondent records
     HHID=023456 PN=010 ASUBHH=0
     HHID=023456 PN=020 ASUBHH=0

Time 2
Household records
     HHID=023456 CSUBHH=1
     HHID=023456 CSUBHH=2
Respondent records
     HHID=023456 PN=010 CSUBHH=1
     HHID=023456 PN=020 CSUBHH=2
```

B3.   One or Both Respondents Die.

Two respondents in a sample household are married at the time of the first cross-section. Each respondent is assigned a HHID of 034567 and a SUBHH of 0. One respondent has a PN of 010, the other a PN of 020.

One respondent dies before the next wave. At the next wave, both respondents retain their HHID of 034567. The living respondent retains her SUBHH of 0; the deceased respondent is assigned a SUBHH of 3. (If both respondents die, one would be assigned a SUBHH of 3 and the other would be assigned a SUBHH of 4.) Each original respondent retains his and her PN of 010 and 020, respectively.

```
Time 1
Household records
     HHID=034567 ASUBHH=0
Respondent records
     HHID=034567 PN=010 ASUBHH=0
     HHID=034567 PN=020 ASUBHH=0

Time 2
Household records
     HHID=034567 CSUBHH=0
     HHID=034567 CSUBHH=3 (in exit interview)
Respondent records
     HHID=034567 PN=010 CSUBHH=3 (in exit interview)
     HHID=034567 PN=020 CSUBHH=0
```

B4.   Single Respondent Marries.

A respondent who has never been married is in the first cross-section. The respondent is assigned a HHID of 045678 and a SUBHH of 0 and a PN of 010.

At the time of the second cross-section, the respondent has married. Both the
respondent and her new spouse are assigned a HHID of 045678 and a SUBHH of 0
because the household was not divided. The original respondent retains her PN of
010. Her new spouse is assigned PN of 011.

```
Time 1
Household records
     HHID=045678 ASUBHH=0
Respondent records
     HHID=045678 PN=010 ASUBHH=0


Time 2
Household records
     HHID=045678 CSUBHH=0
Respondent records
     HHID=045678 PN=010 CSUBHH=0
     HHID=045678 PN=011 CSUBHH=0
```

B5.   Couple Divorces, One Respondent Remarries and Divorces.

Two respondents in a sample household are married at the time of the first
cross-section. Each respondent is assigned a HHID of 056789 and a SUBHH of 0.
One respondent has a PN of 010, the other a PN of 020.

By the time of the second cross-section, the couple has divorced and he has
remarried. Both original respondents retain the HHID of 056789, but she is
assigned a SUBHH of 1 and he is assigned a SUBHH of 2. His new spouse is also
assigned the HHID of 056789 and the SUBHH of 2. Each original respondent retains
his and her PN of 010 and 020, respectively. His new spouse is assigned PN of
011.

By the time of the third cross-section, that new couple has gotten divorced. All
respondents retain the HHID of 056789. The original sample member ex-wife has
the SUBHH of 1. The original sample member ex-husband has a SUBHH of 2, and the
non-original sample member, his second ex-wife, is assigned the SUBHH of 5. Each
original respondent retains his and her PN of 010 and 020, respectively. His
second ex-wife retains her PN of 011.

```
Time 1
Household records
     HHID=056789 ASUBHH=0
Respondent records
     HHID=056789 PN=010 ASUBHH=0
     HHID=056789 PN=020 ASUBHH=0


Time 2
Household records
     HHID=056789 CSUBHH=1
     HHID=056789 CSUBHH=2
Respondent records
     HHID=056789 PN=010 CSUBHH=2
     HHID=056789 PN=011 CSUBHH=2
     HHID=056789 PN=020 CSUBHH=1


Time 3
Household records
```

```
     HHID=056789 FSUBHH=1
     HHID=056789 FSUBHH=2
     HHID=056789 FSUBHH=5
Respondent records
     HHID=056789 PN=010 FSUBHH=2
     HHID=056789 PN=011 FSUBHH=5
     HHID=056789 PN=020 FSUBHH=1
```

B6. Couple Divorces and Marries Again.

Two respondents in a sample household are married at the time of the first cross-section. Each respondent is assigned a HHID of 067890 and a SUBHH of 0. One respondent has a PN of 010, the other a PN of 020.

By the time of the second cross-section, the couple has divorced. Both respondents retain the HHID of 067890, but one is assigned a SUBHH of 1 and the other is assigned a SUBHH of 2. Each original respondent retains his and her PN of 010 and 020, respectively.

By the time of the third cross-section, the respondents have remarried each other. Both are assigned the HHID of 067890 and the SUBHH of 7. Each original respondent retains his and her PN of 010 and 020, respectively.

```
Time 1
Household records
     HHID=067890 ASUBHH=0
Respondent records
     HHID=067890 PN=010 ASUBHH=0
     HHID=067890 PN=020 ASUBHH=0

Time 2
Household records
     HHID=067890 CSUBHH=1
     HHID=067890 CSUBHH=2
Respondent records
     HHID=067890 PN=010 CSUBHH=1
     HHID=067890 PN=020 CSUBHH=2

Time 3
Household records
     HHID=067890 FSUBHH=7
Respondent records
     HHID=067890 PN=010 FSUBHH=7
     HHID=067890 PN=020 FSUBHH=7
```

B7. Married Couple with Children and Siblings.

At the time of the first cross-section, sample household with a HHID of 078901 contains two respondents assigned PNs of 010 and 020, respectively. Associated with the household are three children with OPNs of 101, 102, and 103, and two siblings with OPNs of 051 and 052. All seven persons will keep those same PNs and OPNs across time. A friend who lives with the respondents is assigned an OPN of 080. The friend will keep her OPN of 080 across time only if she is a household member at each wave.

```
Time 1
Household records
     HHID=078901 ASUBHH=0
```

```
Respondent records
     HHID=078901 PN=010 ASUBHH=0
     HHID=078901 PN=020 ASUBHH=0
Household member/child records
     HHID=078901 ASUBHH=0 OPN=101 (child)
     HHID=078901 ASUBHH=0 OPN=102 (child)
     HHID=078901 ASUBHH=0 OPN=103 (child)
     HHID=078901 ASUBHH=0 OPN=051 (sibling)
     HHID=078901 ASUBHH=0 OPN=052 (sibling)
     HHID=078901 ASUBHH=0 OPN=080 (friend)
```

B8. Couple with Children and Siblings Divorces.

At the time of the first cross-section, sample household with a HHID of 089012
contains two respondents assigned PNs of 010 and 020, respectively. Associated
with the household are two children with OPNs of 101, and 102, and three
siblings, her two brothers with OPNs of 051 and 052 and his sister with an OPN
of 061. A friend who lives with the respondents is assigned an OPN of 080.

By the time of the second cross-section, the couple has divorced. Both
respondents retain the HHID of 089012, but he is assigned a SUBHH of 2 and she
is assigned a SUBHH of 1. Each original respondent retains his and her PN of 010
and 020, respectively.

The two children appear both in their father's SUBHH 2 and also in their
mother's SUBHH 1 with their respective OPNs, 101, and 102. The three siblings
appear in the SUBHH of their respective siblings and maintain their respective
OPNs. The ex-wife's brothers appear as part of her SUBHH 1 with their OPNs of
051 and 052, respectively. The ex-husband's sister appears as part of his SUBHH
2 with her OPN of 061. The friend has moved out and does not appear in either
household.

```
Time 1
Household records
     HHID=089012 ASUBHH=0
Respondent records
     HHID=089012 PN=010 ASUBHH=0
     HHID=089012 PN=020 ASUBHH=0
Household member/child records
     HHID=089012 ASUBHH=0 OPN=101 (child)
     HHID=089012 ASUBHH=0 OPN=102 (child)
     HHID=089012 ASUBHH=0 OPN=051 (her brother)
     HHID=089012 ASUBHH=0 OPN=052 (her brother)
     HHID=089012 ASUBHH=0 OPN=061 (his sister)
     HHID=089012 ASUBHH=0 OPN=080 (friend)

Time 2
Household records
     HHID=089012 CSUBHH=1
     HHID=089012 CSUBHH=2
Respondent records
     HHID=089012 PN=010 CSUBHH=2
     HHID=089012 PN=020 CSUBHH=1
Household member/child records
     HHID=089012 CSUBHH=1 OPN=101 (child)
     HHID=089012 CSUBHH=1 OPN=102 (child)
     HHID=089012 CSUBHH=1 OPN=051 (her brother)
```

```
       HHID=089012 CSUBHH=1 OPN=052 (her brother)
       HHID=089012 CSUBHH=2 OPN=101 (child)
       HHID=089012 CSUBHH=2 OPN=102 (child)
       HHID=089012 CSUBHH=2 OPN=061 (his sister)
```

B9. Couple Divorces, One Respondent Remarries, Both Split-off Households Have New Members.

```
Two respondents in a sample household are married at the time of the first
cross-section.  Each respondent is assigned a HHID of 090123 and a SUBHH of 0.
One respondent has a PN of 010, the other a PN of 020.

By the time of the second cross-section, the couple has divorced.  She has moved
in with her mother.  He has married a woman with two children.  At the second
cross-section, both original respondents retain the HHID of 090123, but he is
assigned a SUBHH of 1 and she is assigned a SUBHH of 2. Each original respondent
retains his and her PN of 010 and 020, respectively.  His new spouse and new
stepchildren are assigned the HHID of 090123 and the SUBHH of 1. His new spouse
is assigned an PN of 011.  His new stepchildren are assigned OPNs of 151 and
152.  Her mother is assigned the HHID of 090123 and the SUBHH of 2 and an OPN of
151.

Time 1
Household records
       HHID=090123 ASUBHH=0
Respondent records
       HHID=090123 PN=010 ASUBHH=0
       HHID=090123 PN=020 ASUBHH=0

Time 2
Household records
       HHID=090123 CSUBHH=1
       HHID=090123 CSUBHH=2
Respondent records
       HHID=090123 PN=010 CSUBHH=1
       HHID=090123 PN=011 CSUBHH=1 (new wife)
       HHID=090123 PN=020 CSUBHH=2
Household member/child records
       HHID=090123 CSUBHH=1 OPN=151 (his stepchild)
       HHID=090123 CSUBHH=1 OPN=152 (his stepchild)
       HHID=090123 CSUBHH=2 OPN=151 (her mom)
```