

HEALTH AND RETIREMENT STUDY
2002 Exit
Final, Version 1.0
December 2006

Data Description and Usage

TABLE OF CONTENTS

| | |
|---|------------|
| TABLE OF CONTENTS..... | III |
| DATA DESCRIPTION AND USAGE | 1 |
| 1. OVERVIEW | 1 |
| 1A. THE SAMPLE INTERVIEWED IN 2002..... | 1 |
| 1B. 2002 QUESTIONNAIRE SECTIONS..... | 2 |
| 1C. LEVELS OF FILES | 2 |
| 1C1. Respondent Level Files | 2 |
| 1C2. Household Member and Child Level Files..... | 3 |
| 1C3. Helper Level File | 3 |
| 1C4. Transfer-to-Child Level File | 3 |
| 1C5. Transfer-from-Child Level File..... | 3 |
| 2. FILE NAMING CONVENTIONS | 3 |
| 3. DATA FILES | 4 |
| 4. IDENTIFICATION VARIABLES..... | 5 |
| 4A. PRIMARY IDENTIFICATION VARIABLES | 5 |
| 4A1. HHID – Household Identifier..... | 5 |
| 4A2. SSUBHH – 2002 Sub-household Identifier | 5 |
| 4A3. PN – Person Number | 5 |
| 4A4. OPN – Other Person Number..... | 5 |
| 4A5. STC_NDX or SFC_NDX – 2002 Transfer Number..... | 5 |
| 4B. PRIMARY IDENTIFICATION VARIABLES FOR DATASETS AT EACH OF THE FIVE LEVELS..... | 5 |
| 4C. SECONDARY IDENTIFICATION VARIABLES | 6 |
| 5. DISTRIBUTION FILES AND DIRECTORY STRUCTURE..... | 6 |
| 5A. DISTRIBUTION FILES | 6 |
| 5B. DIRECTORY STRUCTURE | 7 |
| 6. PROGRAM STATEMENTS..... | 7 |
| 6A. USING THE FILES WITH SAS..... | 7 |
| 6B. USING THE FILES WITH SPSS | 8 |
| 6C. USING THE FILES WITH STATA | 8 |
| 7. DOCUMENTATION | 8 |
| 7A. CODEBOOK..... | 8 |
| 7A1. Variable Names..... | 9 |
| 7B. OTHER TYPES OF DOCUMENTATION..... | 10 |
| 7B1. 2000 HRS Final Release Data Description | 10 |
| 7B2. Box and Arrow Questionnaire | 10 |
| 8. ADDITIONAL NOTES..... | 10 |
| 8A. CHANGES IN DATA RESULTING FROM RESPONDENT COMMENTS | 11 |
| 8B. ESTATE CALLBACK | 11 |
| 8C. UNFOLDING BRACKET VARIABLES AND IMPUTATIONS | 11 |
| 8D. IMPROPER HOUSEHOLD SPLITTING BETWEEN CORE AND EXIT INSTRUMENTS | 14 |
| 8E. DROPPED PRELOAD IN THE 2002 EXIT INSTRUMENT | 15 |
| 8F. EXIT RE-INTERVIEWS | 15 |
| 8G. DROPPED CASE | 15 |
| 8H. SUMMARY OF DATA MODEL (DM) CHANGES DURING 2002 EXIT PRODUCTION | 15 |

| | |
|--|-----------|
| 9. OBTAINING THE DATA | 16 |
| 9A. REGISTRATION AND DOWNLOADING THE DATA..... | 16 |
| 9B. CONDITIONS OF USE..... | 16 |
| 9C. PUBLICATIONS BASED ON DATA | 17 |
| 10. IF YOU NEED TO KNOW MORE | 17 |
| 10A. HRS INTERNET SITE..... | 17 |
| 10B. CONTACT INFORMATION | 17 |
| 11. APPENDIX | 18 |
| 11A. EXAMPLES OF SUB-HOUSEHOLD AND RESPONDENT PERSON NUMBER AND OTHER PERSON NUMBER ASSIGNMENTS..... | 18 |
| 11A1. <i>Married Couple Stays Married</i> | 18 |
| 11A2. <i>Couple Divorces</i> | 19 |
| 11A3. <i>One or Both Respondents Die</i> | 19 |
| 11A4. <i>Single Respondent Marries</i> | 20 |
| 11A5. <i>Couple Divorces, One Respondent Remarries and Divorces</i> | 20 |
| 11A6. <i>Couple Divorces and Marries Again</i> | 21 |
| 11B. CROSS WAVE MERGE USING 2002 HRS EXIT DATA AND 2000 CORE DATA..... | 22 |
| 11B1. <i>SAS Example of R to R Merge - 2000 Exit to 1998 Core (for only those in the 2000 Exit data file)</i> | 22 |
| 11B2. <i>Stata Example of R to R Merge - 2002 Exit to 2000 Core (for only those in the 2002 Exit data)</i> | 22 |
| 11B3. <i>SPSS Example of R to R Merge - 2002 Exit to 2000 Core (for only those in the 2002 Exit data file)</i> | 22 |
| 11C. MERGE OF 2002 EXIT INTERVIEW TO DATA OF SURVIVING SPOUSE OR PARTNER IN 2002 CORE..... | 23 |
| 11C1. <i>SAS Example of 2002 Exit to 2002 Core surviving spouse Merge</i> | 23 |
| 11C2. <i>Stata Example of 2002 Exit to 2002 Core surviving spouse Merge</i> | 23 |
| 11C3. <i>SPSS Example of 2002 Exit to 2002 Core surviving spouse Merge</i> | 23 |

Data Description and Usage

1. Overview

The 2002 HRS Exit (Final, Version 1.0) data release consists of 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 National Institute of Aging provided funding (NIH U01 AG09740), with supplemental support from the Social Security Administration. The Institute for Social Research (ISR) Survey Research Center (SRC) at the University of Michigan conducted the survey.

The current release of the 2002 Exit (Final, Version 1.0) has 1,501 respondents. An Exit Interview is attempted with a "proxy informant" for panel members who have died. To the extent possible, proxy informants are knowledgeable about the health, family, and financial situation of the deceased (often the proxy is a widow, widower, or some other family member). The content of the 2002 Exit interview is similar to the 2002 core (or living) interview. As such, the 2002 Exit data are amenable to answering researchers' questions concerning medical care, expenditures, how assets are distributed following death, and family decision-making during the latter part of life.

By receiving the data, which have been freely provided, you agree to use them for research and statistical purposes only and make no effort to identify the respondents. In addition, you 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 2002*

The data collection period for the 2002 Exit interview was August 2002 through March 2003. As of 2002, 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 of 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 of 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 Baby (WB) sub-sample consists of people who were born in 1942 through 1947, were household residents of the conterminous U.S. in the spring of 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 (CODA) sub-sample consists of people who were born in 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). Cases from each of the cohorts (HRS, AHEAD, CODA, and WB) are present in the HRS 2002 Exit Data. For more details about the sample, see our Web site.

1B. 2002 Questionnaire Sections

The content of the 2002 Exit data collection instrument is roughly equivalent to the content of the HRS 2000 Exit instrument. However, in 2002 the data collection instrument required new programming due to a change in the software used to collect the data. Consequently, users who are familiar with the HRS will notice that the section headings and variable names have changed markedly between HRS 2000 Exit and HRS 2002 Exit. The following table presents the sections of the questionnaire as they are in the HRS 2002 Exit instrument versus their appearance in the HRS 2000 Exit.

| 2002 Section | Content | 2000 Section |
|--------------|------------------------------------|--------------|
| PR | Preload | PR |
| A | CoverScreen | CS |
| B | Demographics | A |
| C | Physical Health | B |
| D | Cognition | PC/C/RC |
| E | Family Structure and Transfers | D1 |
| G | Functional Limitations and Helpers | E2 |
| J | Employment | G |
| N | Health Services and Insurance | E1/R1 |
| T | Wills and Life Insurance | J2/R2 |
| W | Event History | EV |
| Y | Time Calculations | TS |

1C. Levels of Files

In the 2002 Exit Questionnaire, most questions were asked of all Respondents.

In addition to the familiar respondent-level files, the 2002 HRS Exit (Final, Version 1.0), contains files at four other levels: household-member-and-child, helper, transfer-to-child, and transfer-from-child.

1C1. Respondent Level Files

Respondent-level files contain questions that were asked of all proxy respondents about the deceased respondent. The files contain one record for each exit proxy who gave an interview in 2002 Exit interview.

1C2. Household Member and Child Level Files

These files contain characteristics about household members, children and siblings.¹ The information can come from Section A (coverscreen), Section E or the preload section.

1C3. Helper Level File

This file contains information provided by each exit proxy about the deceased respondent's helpers. A helper may be a person or organization that was reported by the proxy as providing help with ADLs or IADLs in Section G of the 2002 Exit questionnaire. The helper file contains one record for each helper.

1C4. Transfer-to-Child Level File

This file contains information provided by the exit proxy about transfers of money to a child or grandchild. The file contains one record for each transfer to a child or grandchild. Transfers to children are addressed in Section E.

1C5. Transfer-from-Child Level File

This file contains information provided by the exit proxy about transfers of money from children or grandchildren. The file contains one record for each transfer from a child or grandchild. Transfers from children are addressed in Section E.

2. File Naming Conventions

Files are named beginning with "x02" for HRS 2002 Exit, followed by a letter (or two) designating the questionnaire section. A separator, "_" and then one or two letters designating the level, follows the section letter designator.

- R for Respondent Level
- MC for Household Member and Child Level
- HP for Helper Level
- TC for Transfer-to-Child Level
- FC for Transfer-from-Child Level

For example, the file x02A_R includes variables from section A (coverscreen) at the respondent level. Alternatively, file x02E_TC includes variables from section E and the data are at the monetary Transfer-to-Child 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 State DO statements,
- .DCT for Stata dictionary statements, and
- .TXT for codebook files

¹ Siblings were added to the Household Member and Child Level Files in the 2002 Exit Early release.

One of each of these file types is provided for each of the 17 data files for the 2002 HRS Exit (Final, Version 1.0) data release. For example,

X02A_R.DA contains respondent data from section A,
X02A_R.SAS contains corresponding SAS program statements,
X02A_R.SPS contains corresponding SPSS program statements,
X02A_R.DO contains corresponding Stata DO statements,
X02A_R.DCT contains corresponding Stata dictionary statements, and
X02A_R.TXT contains the ASCII codebook.

3. Data Files

The 2002 HRS Exit (Final, Version 1.0) data are distributed in 17 data files. The files are listed below along with the number of cases (N), number of variables (NV), and the primary identifiers (IDS).

The 2002 HRS Exit data are provided in ASCII format, with fixed-length records. Use associated SAS, SPSS or Stata program statements to read the data into the analysis package of your choice.

Respondent Level Files: PRIMARY IDS = HHID, PN

| | | |
|---------|--------|--------|
| X02PR_R | N=1501 | NV=78 |
| X02A_R | N=1501 | NV=116 |
| X02B_R | N=1501 | NV=77 |
| X02C_R | N=1501 | NV=70 |
| X02D_R | N=1501 | NV=70 |
| X02E_R | N=1501 | NV=36 |
| X02G_R | N=1501 | NV=96 |
| X02J_R | N=1501 | NV=143 |
| X02N_R | N=1501 | NV=254 |
| X02T_R | N=1501 | NV=478 |
| X02W_R | N=1501 | NV=14 |
| X02IO_R | N=1501 | NV=106 |
| X02Y_R | N=1501 | NV=19 |

Household Member and Child Level File: PRIMARY IDS = HHID, SSUBHH, OPN

| | | |
|----------|--------|-------|
| X02PR_MC | N=8292 | NV=22 |
|----------|--------|-------|

Transfer-to-Child Level File: PRIMARY IDS = HHID, SSUBHH, OPN, STC_NDX

| | | |
|---------|-------|-------|
| X02E_TC | N=306 | NV=16 |
|---------|-------|-------|

Transfer-from-Child Level File: PRIMARY IDS = HHID, SSUBHH, OPN, SFC_NDX

| | | |
|---------|-------|-------|
| X02E_FC | N=141 | NV=17 |
|---------|-------|-------|

Helper Level File: PRIMARY IDS = HHID, PN, OPN

| | | |
|---------|--------|-------|
| X02G_HP | N=2736 | NV=25 |
|---------|--------|-------|

4. Identification Variables

Identification variables for HRS 2002 Final Exit (Version 1.0) are stored in character format.

4A. Primary Identification Variables

Several variables, HHID, SSUBHH, PN, OPN, STC_NDX, and SFC_NDX are used in various combinations to uniquely identify records in 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. SSUBHH – 2002 Sub-household Identifier

In combination with HHID, SSUBHH uniquely identifies a household at the time of the 2002 data collection. Sub-household identifiers can be different at each wave due to dissolution or reconstitution of a household (e.g. divorce, separation, marriage, or death). SSUBHH has one-digit. In the 2002 Exit, SSUBHH is either 3 or 4. 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. 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, SSUBHH, and PN, OPN uniquely identify another person (a household member, child, sibling, or helper), at the time of the 2002 data collection. OPN has three-digits.

4A5. STC_NDX or SFC_NDX – 2002 Transfer Number

In combination with HHID and SSUBHH, STC_NDX or SFC_NDX uniquely identifies transfer to or from a child or grandchild. STC_NDX (index for transfers to children) and SFC_NDX (index for transfers from children) have one digit.

4B. Primary Identification Variables for Datasets at Each of the Five 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 Dataset:

- 1) HHID HOUSEHOLD IDENTIFIER
- 2) PN PERSON NUMBER

- 3) OPN OTHER PERSON NUMBER

Three identifiers uniquely identify records in the

- o Household Member and Child Level Dataset:

- 1) HHID HOUSEHOLD IDENTIFIER
- 2) SSUBHH 2002 SUB-HOUSEHOLD IDENTIFIER
- 3) OPN OTHER PERSON NUMBER

Four identifiers uniquely identify records in the

- o Transfer-to-Child Level Dataset
- o Transfer-from-Child Level Dataset:

- 1) HHID HOUSEHOLD IDENTIFIER
- 2) SSUBHH 2002 SUB-HOUSEHOLD IDENTIFIER
- 3) OPN OTHER PERSON NUMBER
- 4) STC_NDX/SFC_NDX 2002 TRANSFER NUMBER

4C. Secondary Identification Variables

Secondary identification variables include GSUBHH, SSUBHH, and SPN_SP. The secondary identification variables can be used to link the 2002 Exit data with core data from previous waves, or to link a deceased respondent with data from their surviving spouse or partner.

Secondary Identification Variables for

- o Respondent Level Datasets
- o Helper Level Datasets:

- 1) GSUBHH 2000 SUB-HOUSEHOLD IDENTIFIER
- 2) SPN_SP 2002 SPOUSE/PARTNER PERSON NUMBER
- 3) SSUBHH 2002 SUB-HOUSEHOLD IDENTIFIER

Secondary Identification Variables for

- o Household Member and Child Level Datasets
- o Transfer-from-Child Level Datasets
- o Transfer-to-Child Level Datasets:

- 1) GSUBHH 2000 SUB-HOUSEHOLD IDENTIFIER

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 large .zip file that contains seven smaller .zip files, one .pdf file, and one .csv file, or the nine smaller files available individually for separate download. The combined file is x02exit.zip.

The individual .zip files for separate download are:

Data file
x02da.zip containing data files.

Program statement files

x02sas.zip containing SAS data descriptors.
x02sps.zip containing SPSS data descriptors.
x02sta.zip containing Stata data descriptors.

Documentation files

x02cb.zip containing the codebook.
h02qn.zip containing the questionnaire.
x02dd.pdf containing this document.

5B. Directory Structure

While a particular setup is not required for using the 2002 HRS Final Exit files, if the following directory structure is used, then no changes to the path name in the data descriptor files are necessary. If you use a different structure, change the directory references in the distribution files.

| Directory | Contents |
|--------------------------------|---|
| c:\x2002\ c:\x2002\codebook | Files downloaded from Web site Unzipped files from x02cb.zip |
| c:\x2002\data | Unzipped files from x02da.zip |
| c:\x2002\qnaire | Unzipped files from x02qn.zip |
| c:\x2002\sas | Unzipped files from x02sas.zip |
| c:\x2002\spss | Unzipped files from x02sps.zip |
| c:\x2002\stata | Unzipped files from x02sta.zip |

Decompress the selected .zip files into the appropriate subdirectories. You will need about 5.0 MB of free space on your storage device to store the 17 .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 x02A_R.SAS go with the x02A_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:\x2002\sas" and the data file is located in "c:\x2002\data", you can run the file as is. A SAS system file (*.SD2 or *.SAS7BDAT) will be saved to directory "c:\x2002\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:\x2002\spss" and the data file is located in "c:\x2002\data", you can run the file as is. An SPSS system file (*.SAV) will be saved to directory "c:\x2002\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:\x2002\stata" and the data file is located in "c:\x2002\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, go to Edit, Select All; then go to Format, Change Case, and select lowercase.

7. Documentation

There are several types of documentation available for use with the 2002 HRS Exit (Final, Version 1.0) data release. These include the codebook, the 2002 box and arrow questionnaire, and the data description.

7A. Codebook

The HRS 2002 Exit Codebook is provided as a series of 17 ASCII text files, as well as a single file containing all sections. There is a codebook file corresponding to each data file, and a complete codebook that includes all sections. Each variable has its own codebook entry. The format of the codebook is, for the most part, consistent with the codebooks of 1995, 1996, 1998, and 2000. The most notable differences are the naming of variables in the 2002 data, and the presentation of ranges and means for continuous variables.

7A1. Variable Names

Variable names begin with a letter designating the wave of data collection (S for 2002 Exit), followed by the section letter, and numbers after the section letter. For example, for the variable SC189, S indicates 2002 Exit, C designates Section C (physical health), and 189 is the variable number. Variables from the preload section of the instrument will have either SX or SZ as prefix letters. The "X" indicates a variable that is updated by data collected in later sections of the questionnaire, whereas the "Z" indicates preloaded data that were not changed by subsequent answers to questions. For example:

```
SX007      RESP FAM/FIN TYPE - UPDATED
SZ007      PREV WAVE R FIN/FAM TYPE

SX065_R    COUPLENESS STATUS - UPDATED
SZ066_R    PREVIOUS WAVE COUPLENESS - INDIVIDUAL
```

7A1a. Multiple Response and Looped Variables

There are two types of variables with multiple mention indicators, simple multiple mentions, and multiple mention questions within a loop.

Simple multiple mention variables take the form: (wave prefix) + (section letter) + (variable number) + (mention number). For example, SC021M1 through SC021M6 are 2002 Exit variables from section C with one to six mentions.

Variable names for multiple mentions to questions within a loop take the form: (wave prefix) + (section letter) + (variable number) + (underscore) + (loop iteration) + (letter designating mention number). For example, SJ140_1A, is a 2002 Exit variable from section J, variable number 140 in the first iteration of the loop, and the first mention.

Simple loop variables (without a multiple mention) have an underscore (_) in their name and a suffix that designates the loop, e.g., SJ119_1.

Null multiple mention variables and variables from null loops beyond the first mention or first loop are not included in the data. It is generally the case that one null multiple mention variable and one null loop is retained.

7A1b. Masked Variables

To protect the confidentiality of the information that respondents provide, a number of variables have been masked or are simply not included in the 2002 Exit Final (Version 1.0) 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.

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.

The names of variables that were masked for confidentiality end in the letter "M"; for example, variable SX026M (1ST ADDRESS STATE - MASKED) and variable SA068M (REGION FACILITY LOCATED - MASKED). In the 2002 Exit release the following variables are masked:

| <u>VARIABLE</u> | <u>LABEL</u> |
|-----------------|---|
| SX026M | 1ST ADDRESS STATE - MASKED |
| SX030M | 2ND ADDRESS STATE - MASKED |
| SZ031M | PREV WAVE 2ND RESIDENCE STATE - MASKED |
| SA126M | R DIED- STATE - MASKED |
| SA129M | R DEATH CERTIFICATE- STATE - MASKED |
| SA133M1M | CAUSE OF DEATH-MASKED- 1 |
| SA133M2M | CAUSE OF DEATH-MASKED- 2 |
| SA068M | REGION FACILITY LOCATED - MASKED |
| SA076M | CURRENT RESIDENCE STATE - MASKED |
| SA081M | OTHER RESIDENCE STATE - MASKED |
| SB003M | STATE BORN - MASKED |
| SB005M | COUNTRY BORN - MASKED |
| SB017M | R HIGHEST DEGREE - MASKED |
| SB024M | FATHERS USUAL OCC-MASKED |
| SB029M1M | TYPE HISPANIC - MASKED-1 |
| SB031M | R RACE - MASKED |
| SB047M | ST/COUNTRY LIVED DURING SCHOOL - MASKED |
| SB052M | DENOMINATION- MASKED |
| SC209M1M | DIAGNOSIS OF MEMORY PROBLEM-MASKED-1 |
| SC209M2M | DIAGNOSIS OF MEMORY PROBLEM-MASKED-2 |
| SC108M1M | OTHER HEALTH-SPECIFY- MASKED-1 |
| SC108M2M | OTHER HEALTH-SPECIFY- MASKED-2 |
| SN314M1M | WHY ADMITTED - FINAL- 1- MASKED |
| SN314M2M | WHY ADMITTED - FINAL- 2- MASKED |
| ST158M | WHICH STATE WILL PROBATED - MASKED |

7B. Other Types of Documentation

In addition to this document and the codebook, three additional types of documentation are available.

7B1. 2000 HRS Final Release Data Description

The 2000 HRS Final Release Data Description document contains useful file merge examples and additional information about HRS data files in general.

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 2002 HRS Exit 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

The following are miscellaneous additional notes regarding the HRS 2002 Exit Final Data Release (Version 1.0). If we become aware of additional issues, they will be posted on our Web site in the Data Alerts section.

8A. Changes in Data Resulting from Respondent Comments

Interviewers recorded comments made by respondents during the course of the interview. HRS staff reviewed these comments for selected questions. The original 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 sentence case.

When a change was made as a result of the comment review process, inconsistencies with subsequent variables may result. For consistency, a limited number of changes 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.

(What does this sentence mean? "The INAP text in the codebook does not include codeframes added in the comment review process.")

8B. Estate Callback

In the months of September through December 2003, Exit Proxy phone interviews were conducted with 2002 Exit proxy respondents. These interviews are referred to as the Estate Call Back (ECB). Of the 874 proxy respondents that were part of the ECB sample, 695 completed the ECB. The ECB interviews were designed to capture information missing from the HRS 2000 Exit Proxy (Section N) and 2002 Exit Proxy (Section T) interviews regarding home ownership and/or the value of an estate. More specifically, respondents who owned a home at the time of their initial Exit Proxy interview missed questions about their homes. Additionally, changes in the question text made it unclear as to whether the total value of their estate included the value of the home or not. Both of these issues form the basis for the ECB interviews.

Interviewers were instructed to conduct the ECB with the same proxy who completed the original Exit proxy interview in 2000 or 2002. In the event that the same proxy was not available, the ECB interview was not administered. The data from the ECB are contained in the data file x02T_r. The data file contains information both for respondents who were and for respondents who were not a part of the ECB. The variable ECBFLAG identifies cases selected for the ECB, and also designates which of the selected cases completed an ECB interview. The variable ECBPATH (also in x02T_r) identifies which questions in Section T a particular proxy respondent was asked. Please check the documentation area for the ECB box and arrow questionnaire, which is included with this release as file x00ecb.pdf.

8C. Unfolding Bracket Variables and Imputations

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 of "unfolding" questions were asked. The manner in which the unfolding questions were programmed (Blaise) is different for 2002 data compared to the CAI (SurveyCraft) software used for 1993 through 2000. This change was transparent to the respondents, since exactly the same questions were asked with the new software as would have been asked with the old software; but it did have an implication for the data that were actually stored and also for the data that

are released.

Instead of storing the response to each unfolding question, three summary variables were generated: the minimum, maximum, and result values for the amount, given the answers to the unfolding questions. If the Respondent said "more than" to the unfolding question with the highest value, then the maximum value was stored as ten times that value. If the last answer a respondent gave in an unfolding sequence was either "Don't Know" or "Refused," that is stored in the result variable.

For most analysts, those three variables (and in particular, the minimum and maximum of the possible range) will be sufficient for analyses. For any analyst who needs the more detailed information, it should be noted that the three variables, combined with the information about the unfolding questions provided in the box-and-arrow and codebook, are sufficient to allow the analyst to reconstruct the sequence of questions asked of any respondent, and the answers to each of those questions in many of the unfolding sequences.

For other sequences, those in which respondents were randomly assigned to one of three "entry" points for the first unfolding question, the analyst will also need to take into account a fourth variable (located in the preload sections) that specifies the entry point for each respondent. The following example shows the preload variable (SZ083) and the unfolding sequence that uses the random entry point from SZ083.

Example Random Entry Assignment Variable from Preload:

Preload Variable from the data file H02xpr_r:

```
=====
SZ083      UNFOLD VALUE - TOTAL COST MEDICAL
           Section: PR      Level: Respondent  Type: Numeric      Width: 1
           Decimals: 0
           CAI Reference: B_Rvars.Z083_UnfMedCost_V
.....
           468              1. RANDOM ASSIGNMENT
           478              2. RANDOM ASSIGNMENT
           554              3. RANDOM ASSIGNMENT
           1                Blank. Data from prior wave not available
=====
```

Unfolding Series from section N that uses SZ083 to assign respondents an entry point:

```
=====
SN222      TOTAL MEDICAL COSTS - MIN
           Section: N      Level: Respondent  Type: Numeric      Width: 6
           Decimals: 0
           CAI Reference: BN.N222_

           N222_ - N224_  Unfolding Sequence
           Did it amount to less than $____ , more than $____ , or what?

           BREAKPOINTS:   1000    5000    25000    100000    500000

           PROCEDURE(S):  ENTRY POINT(S):
=====
```

UNFM_1up3down 100000
UNFM_2up2down 25000
UNFM_3Up1down 5000

User Note: The entry point for the SN222 - SN224 unfolding sequence is determined by the random assignment variable SZ083.

.....
190 0. Value of Breakpoint
7 1000. Value of Breakpoint
51 1001. Value of Breakpoint
45 5000. Value of Breakpoint
209 5001. Value of Breakpoint
74 25000. Value of Breakpoint
513 25001. Value of Breakpoint
95 100000. Value of Breakpoint
236 100001. Value of Breakpoint
27 500000. Value of Breakpoint
39 500001. Value of Breakpoint
15 Blank. INAP (Inapplicable)

=====
SN223 TOTAL MEDICAL COSTS - MAX
Section: N Level: Respondent Type: Numeric Width: 7
Decimals: 0
CAI Reference: BN.N223_

.....
36 999. Value of Breakpoint
7 1000. Value of Breakpoint
55 4999. Value of Breakpoint
45 5000. Value of Breakpoint
206 24999. Value of Breakpoint
74 25000. Value of Breakpoint
404 99999. Value of Breakpoint
95 100000. Value of Breakpoint
221 499999. Value of Breakpoint
27 500000. Value of Breakpoint
316 5000000. Value of Breakpoint
15 Blank. INAP (Inapplicable)

=====
SN224 TOTAL MEDICAL COSTS - RESULT
Section: N Level: Respondent Type: Numeric Width: 2
Decimals: 0
CAI Reference: BN.N224_

.....
97. Data Not Available
306 98. DK (Don't Know); NA (Not Ascertained)
5 99. RF (Refused)
1190 Blank. INAP (Inapplicable)

8D. Improper Household Splitting Between Core and Exit Instruments

Ideally, when one person in a household dies between waves of HRS data collection, the living spouse who is being interviewed in the core instrument should answer 5=NO to question A020.1ST R SAME SP/P, which would trigger follow-up questions used to determine if the respondent's spouse had died. Consequently, the household would be split correctly, thereby making the deceased spouse the coverscreen, family, and financial respondent within his or her own household. As such, the exit proxy reporter would receive questions designated for the first, family and financial respondent.

In the 2002 data collection period, however, there were 17 core respondents who answered A020 with 1=YES, AND SP/PARTNER IS LIVING, when their spouse was, in fact, deceased. A 'YES' at A020 caused the household not to split properly. The respondent was designated for an exit proxy interview, but some were not designated as the first, family, and/or financial R. This resulted in the exit proxy reporter getting skipped past questions in the exit interview that should have been answered if the split had occurred properly.

The following households in the 2002 Exit have no Financial Respondent (SX007 = 2 or 4):

| <u>HHID</u> | <u>SSUBHH</u> |
|-------------|---------------|
| 034096 | 3 |
| 038692 | 3 |
| 043131 | 3 |
| 087465 | 3 |
| 200168 | 3 |
| 200398 | 3 |
| 200516 | 3 |
| 203927 | 3 |
| 210130 | 3 |
| 210405 | 3 |

The following households in the 2002 Exit have no Family Respondent (SX007 = 1 or 4):

| <u>HHID</u> | <u>SSUBHH</u> |
|-------------|---------------|
| 034096 | 3 |
| 038692 | 3 |
| 043131 | 3 |
| 043587 | 3 |
| 049511 | 3 |
| 059920 | 3 |
| 087465 | 3 |
| 200168 | 3 |
| 200398 | 3 |
| 200516 | 3 |
| 203927 | 3 |
| 204258 | 3 |
| 206151 | 3 |
| 210405 | 3 |

The following households in the 2002 Exit have no Coverscreen Respondent (SCSR = 5):

| <u>HHID</u> | <u>SSUBHH</u> |
|-------------|---------------|
| 034096 | 3 |
| 038692 | 3 |
| 043131 | 3 |
| 043587 | 3 |
| 049511 | 3 |
| 052716 | 3 |
| 059920 | 3 |
| 087465 | 3 |
| 200168 | 3 |
| 200398 | 3 |
| 200516 | 3 |
| 203927 | 3 |
| 204258 | 3 |
| 204763 | 3 |
| 206151 | 3 |
| 210130 | 3 |
| 210405 | 3 |

8E. Dropped Preload in the 2002 Exit Instrument

For the respondent HHID = 034975 PN = 020, some preload was inadvertently deleted during the interview. The deleted preload resulted in erroneous skip patterns in the data.

8F. Exit Re-interviews

The following cases were interviewed in the 1998 Exit, and were re-interviewed in the 2002 Exit. The analyst should decide whether these cases should be dropped from their analyses.

HHID = 011863 PN = 010
 HHID = 203802 PN = 010

8G. Dropped Case

In November 2004, we issued a data alert to drop the respondent with HHID=041212 and PN=010 from the 2002 Exit Early release, because the respondent was not deceased. In the 2002 Exit Final release, this case has been deleted from all data files.

8H. Summary of Data Model (DM) Changes during 2002 Exit Production

New Data Models are periodically released to interviewers during the data collection, or field period. New Data Models consist of changes in the Blaise code in the HRS CAPI/CATI survey instrument. Feedback from interviewers and quality control checks of the production data are used to identify necessary changes. Every time a new Data Model is released to interviewers, it is given a "version number" and date. In all of the 2002 Exit Final release data files, there is a variable called, SVDATE, which tells what version of the Data Model was used to interview a given household. The following is a brief summary of the programming changes associated with each new Data Model issued during the 2002 Exit field period.

Data Model 2

Section C:

SC020 (PAST CANCER TREATED) was skipped in error when SZ103 (PREV WAVE R HAS CANCER)=5 and SC018(CANCER OF ANY KIND EXCLUDING SKIN)=1. Additionally, SC140 (WEIGHT GAIN/LOSS 10 LBS. SINCE PREV WAVE) and SC198 (DIFFICULTY BREATHING) through SC208 (INCONTINENCE) were skipped in error. These problems were corrected in data model 2.

Section T:

The program erroneously only allowed one mention at ST207M1 (EOL WHO AUTHORITY - 1); it was programmed as a multiple mention in data model 2.

Data Model 3

Section N:

Exit proxies for the respondent who was in a nursing home when he/she died were asked at SN114 (EVER PATIENT OVERNIGHT IN NURSING HOME) if the respondent had any nursing home stays besides the one when he/she died, and, if not, skipped questions SN115 (# TIMES SPENT OVERNIGHT IN NURSING HOME) - SN117 (NUM MOS R SPENT OVERNIGHT IN NH). Therefore, they were skipped over the SN123 (MONTH R MOVED TO NURSING HOME) loop and missed questions SN127 (ELIGIBLE FOR MEDICAID START NH STAY) - SN133 (LIVE WITH WHICH CHILD AFTER NH STAY). This problem was corrected in data model 3.

Data Model 4

No changes documented.

9. Obtaining the Data

9A. *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 on 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.

9B. *Conditions of Use*

By registering, you agree to the Conditions of Use governing access to Health and Retirement Study 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.

9C. 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.

10. If You Need to Know More

This document is intended to serve as a brief overview and to provide guidelines to using the 2002 HRS Exit (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.

10A. 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 URL is <http://hrsonline.isr.umich.edu/>.

10B. Contact Information

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

Internet: Help Desk on 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

11. Appendix

11A. 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 reunited².

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

11A1. 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

11A2. 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

11A3. 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

One Respondent Dies:

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

Both Respondents Die in same wave:

Household records

HHID=034567 CSUBHH=3 (in exit interview, first respondent to die)

HHID=034567 CSUBHH=4 (in exit interview, second respondent to die)
Respondent records
HHID=034567 PN=010 CSUBHH=3 (in exit interview, first respondent to die)
HHID=034567 PN=020 CSUBHH=4 (in exit interview, second respondent to die)

11A4. 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

11A5. 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, the 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

11A6. 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

11B. Cross Wave Merge Using 2002 HRS Exit Data and 2000 Core data

11B1. SAS Example of R to R Merge - 2000 Exit to 1998 Core (for only those in the 2000 Exit data file)

```
/* create sas libraries; change paths to match local file locations*/
libname x02 "c:\x2002\sas" ;
libname h00 "c:\hrs2000\sas" ;
/*Primary identifiers for cross wave, respondent merging are HHID and PN; Sort
data by these variables */
data X02c_r (keep=hhid pn SC028 SC029);
set x02.x02c_r;
run;
proc sort; by hhid pn; run;

data h00b_r (keep=hhid pn G1274 G1275);
set h00.h00b_r;
run;
proc sort; by hhid pn; run;

/*perform the merge, creating the new data set called "cr" - "in" statement on
data set for which you want to keep all records*/
data cr;
merge x02c_r (in=y) h00b_r;
by hhid pn;
if y=1;
run;
```

11B2. Stata Example of R to R Merge - 2002 Exit to 2000 Core (for only those in the 2002 Exit data)

```
use HHID PN G1274 G1275 using h00b_r
sort HHID PN
save tmp1, replace
use HHID PN SC028 SC029 using x02c_r
sort HHID PN
merge HHID PN using tmp1
drop if _m==2
drop _m
sort HHID PN
save rdata, replace
```

11B3. SPSS Example of R to R Merge - 2002 Exit to 2000 Core (for only those in the 2002 Exit data file)

```
GET
FILE='C:\x2002\spss\x02C_R.sav'.
SORT CASES BY
HHID (A) PN (A) .
SAVE OUTFILE='C:\x2002\spss\c_r.sav'.
GET
FILE='C:\hrs2000\spss\H00B_R.SAV'.
SORT CASES BY
HHID (A) PN (A) .
SAVE OUTFILE='C:\hrs2000\spss\b_r.sav'.
MATCH FILES /FILE='C:\hrs2000\spss\B_R.SAV'
/FILE='C:\x2002\spss\c_r.sav'
/IN=source01
/BY hhid pn.
```

```
VARIABLE LABELS source01
'Case source is C:\x2002\spss\c_r.sav'.
SELECT IF( source01 = 1).
EXECUTE.
```

11C. Merge of 2002 Exit Interview to data of surviving spouse or partner in 2002 Core

11C1. SAS Example of 2002 Exit to 2002 Core surviving spouse Merge

```
/* create sas libraries; change paths to match local file locations*/
libname x02 "c:\x2002\sas" ;
libname c02 "c:\hrs2002\sas" ;
data xcr;
set x02.x02c_r;
where SPN_SP ne '';
run;
proc sort; by hhid SPN_SP; run;

data ccr;
set c02.h02c_r;
rename PN=SPN_SP;
run;
proc sort; by hhid SPN_SP; run;

data excore02;
merge xcr (in=y) ccr;
by hhid spn_sp;
if y=1;
run;
```

11C2. Stata Example of 2002 Exit to 2002 Core surviving spouse Merge

```
use HHID PN HC002 HC003 using c:\hrs2002\stata\h02c_r
rename PN SPN_SP
sort HHID SPN_SP
save tmp1, replace

use HHID PN SPN_SP SC028 SC029 using c:\x2002\stata\x02c_r
drop if SPN_SP == ""
sort HHID SPN_SP

merge HHID SPN_SP using tmp1
drop if _m==2
drop _m

sort HHID SPN_SP
save xcdata, replace
```

11C3. SPSS Example of 2002 Exit to 2002 Core surviving spouse Merge

```
GET
FILE='C:\hrs2002\spss\H02C_R.sav'.
RENAME VARIABLES (PN=SPN_SP).
```

```
    SORT CASES BY
      HHID (A) SPN_SP (A) .
  SAVE OUTFILE='C:\hrs2002\spss\CCR.sav' .

GET
  FILE='C:\x2002\spss\x02C_R.sav' .
  FILTER OFF.
  USE ALL.
  SELECT IF(SPN_SP > "0").
  EXECUTE .
  SORT CASES BY
    HHID (A) SPN_SP (A) .
  SAVE OUTFILE='C:\x2002\spss\xcr.sav' .

MATCH FILES /FILE='C:\hrs2002\spss\CCR.sav'
  /FILE='C:\x2002\spss\xcr.sav'
  /IN=source01
  /BY HHID SPN_SP.
VARIABLE LABELS source01
  'Case source is C:\x2002\spss\xcr.sav' .
SELECT IF( source01 = 1).
EXECUTE.
```