HEALTH AND RETIREMENT STUDY
1994 Core and Exit
Final, Version 2.0
August 2004
(revised October 2004)
Data Description and Usage:
What's New in Version 2.0

TABLE OF CONTENTS

OVERVIEW OF VERSION 2.00 RELEASE

This document is intended to serve as an overview of what's new in the version 2.0 release of the 1994 HRS Core (Final) data. The primary change is the provision of standard identification variables in character (string), instead of numeric format, to facilitate merging with the tracker file and later years' data files. Additional changes are listed below.

- The combined variables, HHIDPN, HHIDPRN, and HHIDIPN have been dropped. Other identification variables and CVERSION have been added. Other than these changes, no variables have been deleted or added in the version 2.00 release.
- Data alerts that necessitated changes to the data have been incorporated. As a result, both the number of records in some files and the values of some variables, including identification variables, have changed.
- CSUBHH has been changed for two cases.
- A listing of codebook files, the variables contained therein, and the corresponding data file is provided. The codebook files have not changed since 1998; the frequencies in the codebook files are based on data available at that time.
- The questionnaire files are now distributed in Adobe .PDF format.
- For historical, contextual, purposes, two documents, Data Description and Usage, Version 1.00 (datadesc.txt) and Difference between the HRS Wave 2 Partial Public Release and the HRS Wave 2 Full Public Release (partrel.txt), are included, unedited, at the end of this document.

DATA AND PROGRAM STATEMENT FILES

Number of Records and Number of Variables in Data Files

The table below lists the number of records in the version 2.0 file and number of variables in each of the version 1.0 and version 2.0 files.

- The number of variables changes because of the change in identification variables; see CHANGES TO VARIABLES section below for details.
- The number of records changes because of the incorporation of data alerts that necessitated changes to the data; see the CHANGES TO DATA VALUES section below for details.

File	Number of Records V1.0	Number of Records V2.0	Records Deleted	Number of Variables V1.0	Number of Variables V2.0	Net number of Variables Added
W2KIDS	22,930	22,861	(69)	36	43	7
W2HHLIST	21,635	21,626	(9)	17	24	7
W2SIBS	17,880	17,800	(80)	54	61	7
W2CS	13,006	13,006	-	56	62	6
W2A	11,596	11,596	-	75	81	6
W2B	11,596	11,596	-	176	183	7
W2C	11,596	11,596	-	82	88	6
W2E	11,596	11,596	-	72	78	6
W2FA	11,596	11,596	-	568	574	6
W2FB	11,596	11,596	-	460	466	6
W2FC	11,596	11,596	-	228	234	6
W2G	11,596	11,596	-	74	81	7
W2H	11,596	11,596	-	200	207	7

File	Number of Records V1.0	Number of Records V2.0	Records Deleted	Number of Variables V1.0	Number of Variables V2.0	Net number of Variables Added
W2J	11,596	11,596	-	141	148	7
W2R	11,596	11,596	-	77	84	7
W2S	11,596	11,596	-	78	85	7
W2PARS	8,902	8,886	(16)	42	49	7
W2D	6,979	6,979	-	140	148	8
W2K	6,979	6,979	-	108	116	8
W2N	6,979	6,979	-	577	584	7
W2V	6,979	6,979	-	68	75	7
W2MOD4	1,561	1,561	-	52	60	8
W2MOD7	827	827	-	8	16	8
W2MOD8	822	822	-	40	48	8
W2MOD2	817	817	-	9	17	8
W2MOD1	815	815	-	16	24	8
W2MOD5	801	801	-	7	15	8
W2MOD3	771	771	-	19	27	8
W2MOD0	222	222	-	57	65	8
W2MOD6	203	203	-	32	40	8
W2MOD9	179	179	-	28	36	8

CHANGES TO VARIABLES

A summary of variables dropped, converted, renamed, or added appears below. Please be aware that the codebook has not been updated to reflect these changes. See Codeframes for Identifiers and Other Added Variables section below for more detailed variable information.

Respondent-level and W2PARS Files Variable Changes

Data Files	Dropped	Converted to character (string)	Renamed	Added
W2A W2C W2CS W2E W2FA W2FB W2FC	HHIDPN	HHID PN	Was: W2SUBHH Now: CSUBHH*	CPN_SP COPN_CS CFAMR CFINR CCSR ASUBHH CVERSION
W2B W2G W2H W2J W2R W2S		HHID PN	Was: W2SUBHH Now: CSUBHH*	CPN_SP COPN_CS CFAMR CFINR CCSR ASUBHH CVERSION

Data Files	Dropped	Converted to character (string)	Renamed	Added
W2MOD0 W2MOD1 W2MOD2 W2MOD3 W2MOD4 W2MOD5 W2MOD6 W2MOD7 W2MOD8 W2MOD9		HHID PN		CSUBHH CPN_SP COPN_CS CFAMR CFINR CCSR ASUBHH CVERSION
W2PARS	HHIDRPN	HHID PN	Was: W2SUBHH Now: CSUBHH*	CPARCODE CPN_SP COPN_CS CFAMR CFINR CCSR ASUBHH CVERSION

* Converted to character (string)

Household-level and W2KIDS, W2SIBS, and W2HHLIST Files Variable Changes

Data Files	Dropped	Converted to character (string)	Renamed	Added
W2N W2V	HHIDPN	HHID	Was: W2SUBHH Now: CSUBHH* Was: PN Now: CPN_FIN*	CPN_FAM CPN_NFAM CPN_NFIN CPN_CS CPN_NCS COPN_CS ASUBHH CVERSION
W2D W2K		HHID	Was: W2SUBHH Now: CSUBHH* Was: PN Now: CPN_FIN*	CPN_FAM CPN_NFAM CPN_NFIN CPN_CS CPN_NCS COPN_CS CVERSION

Data Files	Dropped	Converted to character (string)	Renamed	Added
W2KIDS W2SIBS	HHIDPN HHIDRPN	HHID	Was: W2SUBHH Now: CSUBHH* Was: PN Now: OPN*	CPN_FAM CPN_NFAM CPN_FIN CPN_NFIN CPN_CS CPN_NCS COPN_CS ASUBHH CVERSION
W2HHLIST	HHIDPN HHIDIPN	HHID	Was: W2SUBHH Now: CSUBHH* Was: PN Now: OPN*	CPN_FAM CPN_NFAM CPN_FIN CPN_NFIN CPN_CS CPN_NCS COPN_CS ASUBHH CVERSION

* Converted to character (string)

CHANGES TO DATA

Data Alerts

Data alerts that necessitated changes to the data have been incorporated in version 2.0 release. Details are provided below.

(1998-08-24) User Warning: HRS Wave 2 Income Imputation Data in W2N

Prior to the release of the 1994 imputation files, changes were made to the imputed variables included in the core file W2N -- W5907, W5913, W5919, W5925, W5949, W5955, W5961, W5967, W5973, W5979, W5986, W6003, W6022, W6083, W6095, W6103, W6118, W6134, W6204, W6284, W6285, W6296, W6302, W6304, W6376, W6383, W6390, W6404, W6501, and W16164. The corrections may be found at our Web site on the Retired or Superseded Data Alerts page. A total of 81 cells were changed for these variables.

While these changes have been incorporated the version 2.0 release, we suggest you use the 1994 core imputation file H94I_N available at our Web site for imputed variables.

(1998-08-25) W3617: Values Outside of Range in W2FA

Values of greater than 95 were changed to 95 for W3617 (FA44). Details may be found at our Web site on the Retired or Superseded Data Alerts page. A total of seven cells were changed for this variable.

(1998-10-12) Duplicate Person Numbers in W2KIDS [UPDATE]

Twenty-three households had duplicate, non-000 and non-999, child person numbers. Suggestions for resolution were superseded by data alert issued 2003-11-13, see below. Details may be found at our Web site on the Retired or Superseded Data Alerts page.

(1998-11-24) Problematic Age Variable (W103)

This data alert alluded to problems with the age variable, actually W104, not W103, but included no correction statements. We suggest you use BIRTHMO and BIRTHYR from the tracker file instead of W104 - HHCS7d. Age or W53 - Year of Birth or W214 - A11. R:Birthdate:Year. The retired data alert may be found at our Web site on the Retired or Superseded Data Alerts page.

(2001-08-07) Fix for W2S (Section S) SPSS statements

New program SPSS program statements were provided for W6867, W6868, W6869, and W6870. Details may be found at our Web site on the Retired or Superseded Data Alerts page.

(2003-02-25 and 2002-12-09) W2PARS (Parents) File Replaced

The W2PARS.DA file was re-issued. The value of PN was changed for 194 records. Ninety-five records had a value for PN when previously they had none. Ninety-nine records had a different value for PN than they did previously.

The 12/9/2002 replacement file, W2PARS.DA contained 8,902 records. The previous file contained 12,247 records. The deleted records were parent (or parent-in-law) records that contained mother (or mother-in-law) data that also appeared, properly, on the mother (or mother-in-law) records. In addition to extraneous parent records, the old file also contained extraneous parent (or parent-in-law) data in mother (or mother-in-law) records for four variables,

W8210 - Parent(s) own home?
W8213 - Parent(s) live with whom?
W8217 - Parent(s) live w/in 10 miles
W8219 - Parent(s): contact with

Minor changes were also made to two additional variables, W8220 (Parent(s): received fin assist?) and W8233 (Parent(s) fin assist r?).

The codebook file, 21_W2PAR.TXT, was not re-issued and contains frequencies based on the old, 12,247 record, file.

Additional details may be found at our Web site on the Retired or Superseded Data Alerts page.

PN (PERSON NUMBER) was added to facilitate merging parent information with 1994 respondent files. PN had a value for all records even if the spouse/partner was not a respondent in 1994. For parent records, PARCODE is 1-7, PN was equal to RPN - PERSON NUMBER (OF FAMILY R). For parent-in-law records, PARCODE is 8-14, PN was not equal to RPN and was assigned as follows.

- For spouse/partners who provided an interview in 1994, PN was obtained from a 1994 respondent file.
- For 1994 non-respondent spouse/partners who provided an interview in 1992, PN was obtained from the1994 W2HHLIST file.
- For parent-in-law information provided in exit interviews, PN was obtained from the 1994 W2HHLIST file.
- For 1994 new spouse/partner non-respondents, PN was obtained from the TRACK30 file.
- For spouse/partners not part of the sample or not listed, PN was assigned a value of 990.

The file contained records for mothers, fathers, parents, mothers and stepfathers, fathers and stepmothers, stepfathers, and stepmothers of family respondents and analogous records for in-laws.

The type of record, PARCODE, determined which variables in the W2PAR file had valid data. Listed below are the variables and the record types that contained valid data for the variables.

Variables	Parent Record Types	Parent-in-law Record Types
W8201 Parent still living? W8202 Parent age W8203 Parent needs help? W8204 Parent can be left alone? W8205 Parent year of death W8206 Parent age at death W8207 Parent have illness lasting 3+ mo W8208 Parent ever live in nursing home	1 Mother 2 Father	8 Mother-in-law 9 Father-in-law
W8209 Parents married?	1 Mother 2 Father 3 Parents	8 Mother-in-law 9 Father-in-law 10 Parents-in-law
W8210 Parent(s) own home? W8211 Parent(s) financial situation	1 Mother (and stepfather) 2 Father (and stepmother) 3 Parents	8 Mother-in-law (and stepfather-in-law) 9 Father-in-law (and stepmother-in-law) 10 Parents-in-law
W8212 Parent(s) live together?	3 Parents	10 Parents-in-law
W8213 Parent(s) live with whom? W8214 Parent(s): move in w/whom W8215 Parent(s): year of move W8216 Parent(s) move helped whom W8217 Parent(s) live w/in 10 miles W8218 Parent(s) state of residence W8219 Parent(s): contact with	1 Mother (and stepfather) 2 Father (and stepmother) 3 Parents	8 Mother-in-law (and stepfather-in-law) 9 Father-in-law (and stepmother-in-law) 10 Parents-in-law

Variables	Parent Record Types	Parent-in-law Record Types
W8220 Parent(s): received fin assist? W8221 Parent(s): how much fin assist W8222 Parent(s): fin assist brackets W8223 Parent(s): received per assist? W8224 Parent(s): hours fam r per assist W8225 Parent(s): 100+ hrs fam r per assist W8226 Parent(s): hours spouse per assist W8227 Parent(s): 100+ hrs spouse per assist W8228 Parent(s): received oth assist? W8229 Parent(s): hours fam r oth assist W8230 Parent(s): 100+ hrs fam r oth assist W8231 Parent(s): hours spouse oth assist W8232 Parent(s): 100+ hrs spouse oth assist W8233 Parent(s) fin assist r? W8234 Parent(s) amount fin assist r W8235 Parent(s) amount fin assist r brackets	1 Mother 2 Father 3 Parents 4 Mother and Stepfather 5 Father and Stepmother 6 Stepmother 7 Stepfather	8 Mother-in-law 9 Father-in-law 10 Parents-in-law 11 Mother-in-law and Stepfather-in-law 12 Father-in-law and Stepmother-in-law 13 Stepfather-in-law 14 Stepmother-in-law

(2003-11-13) Correction Statements for W2KIDS, W2SIBS, and W2HHLIST

In the 1994 W2KIDS, W2SIBS, and W2HHLIST files a number of problems were corrected. The problems included values of 000 or 999 for (O)PN for new household members, reports from two family respondents for ten households, see More Than One Family Respondent section below, and other duplicate records.

- For W2KIDS, a total of 218 cells were changed for (O)PN, a total of 104 cells were changed for W8001, W8002, W8003, and FAMFLAG, and a total of 69 records were deleted.
- For W2SIBS, total of 270 cells were changed for (O)PN, a total of 23 cells were changed for W8101, W8102, W8103, W8106, W8017, and FAMFLAG and a total of 80 records were deleted.
- For W2HHLIST, total of 1,251 cells were changed for (O)PN, two cells were changed for W8301, and a total of nine records were deleted.

The corrections may be found at our Web site on the Retired or Superseded Data Alerts page.

(2004-10-01) Data File W2PARS Replaced

One data file, W2PARS.DA, was replaced. One record was deleted and CSUBHH was changed for 47 records.

Other Changes

More Than One Family Respondent

One respondent was supposed be designated as the Family Respondent and to provide family information for the household. However, in error, ten households had family information provided by both respondents.

In version 1.0, records with information from each respondent were included in the W2KIDS, W2SIBS, and W2PARS files. Records from households with two family respondents were flagged by the FAMFLAG variable, Family Respondent Problem Flag, where a code value of

> 1. [Record] is associated with a household that contains two Family Respondents (will require special attention by analyst).

In version 2.0, we have deleted the records with information provided by the second Family Respondent. Records with information provided from the second Family Respondent in W2KIDS and W2SIBS were deleted in correction statements provided for the data alert issued 2003-11-13, see above. In addition, 15 records with information provided from the second Family Respondent were deleted from W2PARS.

The FAMFLAG was changed to a value of 2 for records in W2KIDS, W2SIBS, and W2PARS retained from these ten households.

> 2. Record is associated with a household that contained two Family Respondents; information from one respondent has been retained; no special attention required by analyst.

In the respondent-level file W2E for 20 respondents, FAMFLAG was changed to 3 or 4.

> 3. Respondent is in a household that contained two Family Respondents; information from this respondent has been retained; no special attention required by analyst.

> 4. Respondent is in a household that contained two Family Respondents; information from this respondent has been deleted; no special attention required by analyst.

CSUBHH Changed for Two Cases

As a result of later information, CSUBHH has changed for two cases. These changes will synchronize the version 2.0 1994 files with the current tracker file, TRK2002.

HHID	PN or RPN	Old CSUBHH	New CSUBHH
015730	020	3	0
057447	020	3	4

CSUBHH in W2HHLIST

CSUBHH in W2HHLIST has values of 0, 1, and 2, but not of 3. Records for two respondent households in 1992 that split in 1994 were duplicated and have values of 1 and 2 for CSUBHH. Records for two respondent households in 1992 in which one, or both, of the respondents died by 1994 were not duplicated and have values of 0 for CSUBHH. Records for one respondent households in 1992 in which the respondent died by 1994 have values of CSUBHH of 0.

PRIMARY AND SECONDARY IDENTIFIERS

Primary identifiers uniquely identify records in a file. Secondary identifiers are provided to facilitate merging with other files. The primary and secondary identifiers for each of the data sets are listed in the tables below.

Respondent-level and W2PARS Files Identifiers

Data Files		Primary Identifiers	Secondary Identifiers
W2A	W2MOD0	HHID PN	CSUBHH ASUBHH CPN_SP COPN_CS CFAMR CFINR CCSR
W2B	W2MOD1		
W2C	W2MOD2		
W2CS	W2MOD3		
W2E	W2MOD4		
W2FA	W2MOD5		
W2FB	W2MOD6		
W2FC	W2MOD7		
W2G	W2MOD8		
W2H	W2MOD9		
W2J			
W2R			
W2S			
W2PARS		HHID PN CPARCODE CSUBHH	

Household-level and W2KIDS, W2SIBS, and W2HHLIST Files Identifiers

Data Files	Primary Identifiers	Secondary Identifiers
W2D	HHID CSUBHH	ASUBHH CPN_FAM CPN_NFAM
W2K		
W2N		
W2V		

Data Files	Primary Identifiers	Secondary Identifiers
W2KIDS W2SIBS W2HHLIST	HHID CSUBHH OPN	CPN_FIN CPN_NFIN CPN_CS CPN_NCS COPN_CS

CODEBOOK INFORMATION

Codebook Files

In 1994 codebook files were distributed in separate files for each section of the interview. Since there is not a one-to-one correspondence between codebook files and data files, the contents of the original codebook files and the data files in which the variables appear are noted in the table below. Variables with imputed versions (W1nnnn) appear in the codebook next to original variable (Wnnnn); only the original variable is listed in the table below.

For documentation of identification variables, see the Codeframes for Identifiers and Other Added Variables section of this document. For more recently calculated imputations, see the imputation files, 1994 HRS Imputations (Final) (v.3.0), available at our Web site.

Codebook file	Contents	Variable range in codebook	Data files	Level*
01_W2MAS.TXT	Master Codes			---
02_W2CS.TXT	Household and Individual Coversheet Data	W100-W150	W2CS	R
03_W2A.TXT	Section A: Demographics and Miscellaneous	W45-W62, W200-W237	W2A	R
04_W2B.TXT	Section B: Health	W301- W473	W2B	R
05_W2D.TXT	Section D: Housing	W600-W723	W2D	H
06_W2E.TXT	Section E: Family Structure	W901- W963	W2E	R
07_W2FA.TXT	Section FA: Employment (Employees)	W3300-W4115	W2FA	R
08_W2FB.TXT	Section FB: Employment (Self-Employed)	W4200-W4768	W2FB	R
09_W2FC.TXT	Section FC: Employment (Unemployed)	W4800-W5154	W2FC	R
10_W2J.TXT	Section J: Disability	W5200-W5337	W2J	R
11_W2K.TXT	Section K: Net Worth	W5500-W5585	W2K	H
12_W2V.TXT	Section V: Capital Gains	W5700-W5763	W2V	H
13_W2C.TXT	Section C: Cognition	W5800-W5877	W2C	R
14_W2N.TXT	Section N: Income	W5900-W6527	W2N	H
15_W2R.TXT	Section R: Health Insurance	W6700-W6773	W2R	R
16_W2S.TXT	Section S: Widowhood	W6800-W6874	W2S	R

Codebook file	Contents	Variable range in codebook	Data files	Level*
17_W2G.TXT	Section G: Last Job R Not Working Now	W7000-W7070	W2G	R
18_W2H.TXT	Section H: Job History	W7100-W7296	W2H	R
19_W2KID.TXT	Section E: Children file	W8001-W8028	W2KIDS	C
20_W2SIB.TXT	Section E: Siblings file	W8101-W8144	W2SIBS	S
21_W2PAR.TXT	Section E: Parents file	W8201-W8235	W2PARS	P
22_W2HHL.TXT		W8301-W8310	W2HHLIST	M
23_W2MD0.TXT	Experimental Module 0: Activities and Nutrition	W9201-W9255	W2MOD0	R
24_W2MD1.TXT	Experimental Module 1: Depression Scale	W9256-W9269	W2MOD1	R
25_W2MD2.TXT	Experimental Module 2: Similarities	W9270-W9276	W2MOD2	R
26_W2MD3.TXT	Experimental Module 3: Physical Functioning	W9277-W9293	W2MOD3	R
27_W2MD4.TXT	Experimental Module 4: Spending and Saving	W9294-W9343	W2MOD4	R
28_W2MD5.TXT	Experimental Module 5: Risk Aversion	W9344-W9348	W2MOD5	R
29_W2MD6.TXT	Experimental Module 6: Social Support	W9349-W9378	W2MOD6	R
30_W2MD7.TXT	Experimental Module 7: Transfers	W9379-W9384	W2MOD7	R
31_W2MD8.TXT	Experimental Module 8: Help with ADLs	W9385-W9422	W2MOD8	R

* Where

> R = respondent
> H = household
> C = child
> S = sibling
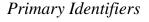> P = parent information
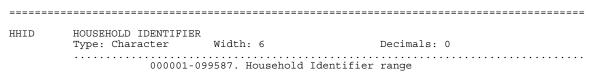> M = household member

Codeframes for Identifiers and Other Added Variables

Since the changed or added identifiers and variables are not documented in the original codebooks, the following information is provided.

Identification Variables

For Respondent-level Files and for W2PARS

Primary Identifiers

```
=========================================================================================
HHID        HOUSEHOLD IDENTIFIER
            Type: Character        Width: 6              Decimals: 0
            ....................................................................
                    000001-099587. Household Identifier range
```

```
========================================================================================
PN         PERSON NUMBER
           Type: Character      Width: 3                   Decimals: 0
           ....................................................................................

                          010. Person Number
                          011. Person Number
                          020. Person Number
                          021. Person Number
                          030. Person Number
                          031. Person Number
                          040. Person Number
                          990. Non-sample spouse/partner (in W2PARS)

========================================================================================
For W2PARS only:

CPARCODE   1994 TYPE OF PARENT RECORD
           Type: Character      Width: 2                   Decimals: 0

           User note:  Denotes relationship to respondent.  See PARCODE for additional
           information.
           ....................................................................................

                            1.  Mother*
                            2.  Father*
                            3.  Parents
                            4.  Mother and Stepfather
                            5.  Father and Stepmother
                            6.  Stepfather
                            7.  Stepmother


CSUBHH     1994 SUB-HOUSEHOLD IDENTIFIER
           Type: Character      Width: 1                   Decimals: 0
           ....................................................................................

                          0. Original sample household - no split from divorce or
                             separation of spouses or partners
                          1. Split household - one half of couple from SUBHH 0 and new
                             spouse or partner, if any
                          2. Split household - other half of couple from SUBHH 0 and new
                             spouse or partner, if any
                          3. Deceased respondent household

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

Secondary Identifiers

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

CSUBHH     1994 SUB-HOUSEHOLD IDENTIFIER
           Type: Character      Width: 1                   Decimals: 0
           ....................................................................................

                          0. Original sample household - no split from divorce or
                             separation of spouses or partners
                          1. Split household - one half of couple from SUBHH 0 and new
                             spouse or partner, if any
                          2. Split household - other half of couple from SUBHH 0 and new
                             spouse or partner, if any
                          3. Deceased respondent household

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

ASUBHH     1992 SUB-HOUSEHOLD IDENTIFIER
           Type: Character      Width: 1                   Decimals: 0
           ....................................................................................

                            0. Original sample household

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

CPN_SP     1994 SPOUSE/PARTNER PERSON NUMBER
           Type: Character      Width: 3                   Decimals: 0
           ....................................................................................
```

```
                010. Person Number
                011. Person Number
                020. Person Number
                021. Person Number
                030. Person Number
                031. Person Number
                040. Person Number
         Blank. INAP (Inapplicable) - No spouse interview; no interview

================================================================================
COPN_CS  1994 ACTUAL COVERSHEET PERSON NUMBER
         Type: Character      Width: 3              Decimals: 0

         User note:  Coversheet respondent may have been an informant other than either
         of the respondents.  Character (string) version of IPN in W2HHLIST and W101 in
         W2CS.
         ......................................................................
                010. Respondent Person Number
                011. Respondent Person Number
                020. Respondent Person Number
                021. Respondent Person Number
                030. Respondent Person Number
                031. Respondent Person Number
            033-034. Other Person Number
                040. Respondent Person Number
            041-083. Other Person Number
            101-399. Other Person Number
            401-499. Other Person Number
            501-699. Other Person Number
            701-799. Other Person Number
            901-998. Other Person Number
                999. Other person without assigned person number
              Blank. No interview

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

CFAMR    1994 WHETHER FAMILY RESPONDENT
         Type: Numeric        Width: 1              Decimals: 0
         ......................................................................
                  1. Yes
                  2. Yes - household had two family respondents and the family
                     information given by this respondent has been retained.
                  3. Yes, exit interview
                  4. No - household had no family respondents but the person number
                     of this respondent appears as the CPN_FAM in the household
                     record.

                  5. No
                  6. No - household had two family respondents and the family
                     information given by this respondent has not been retained.
                  8. No - household had no family respondents and the person number
                     of this respondent appears as the CPN_NFAM in the household
                     record..
              Blank. INAP (Inapplicable) - No interview

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

CFINR    1994 WHETHER FINANCIAL RESPONDENT
         Type: Numeric        Width: 1              Decimals: 0
         ......................................................................
                  1. Yes
                  3. Yes, exit interview
                  4. No - household had no financial respondents but the person
                     number of this respondent appears as the CPN_FIN in the
                     household record.

                  5. No
                  8. No - household had no financial respondents and the person
                     number of this respondent appears as the CPN_NFIN in the
                     household record
              Blank. INAP (Inapplicable) - No interview

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

CCSR     1994 WHETHER COVERSHEET RESPONDENT
         Type: Numeric        Width: 1              Decimals: 0

         ......................................................................
```

```
                1. Yes
                3. Yes, exit interview
                4. No – neither respondent was the coversheet respondent, but the
                   person number of this respondent appears as the CPN_CS in the
                   household record.

                5. No
                8. No – neither respondent was the coversheet respondent, but the
                   person number of this respondent appears as the CPN_NCS in the
                   household record.

           Blank. INAP (Inapplicable) - No interview
```

===

For Household-level Files and for W2KIDS, W2SIBS, W2HHLIST

Primary Identifiers

===

```
HHID     HOUSEHOLD IDENTIFIER
         Type: Character      Width: 6           Decimals: 0
.............................................................................
                    000001-099587. Household Identifier range
```

===

```
CSUBHH   1994 SUB-HOUSEHOLD IDENTIFIER
         Type: Character      Width: 1           Decimals: 0
.............................................................................

             0. Original sample household - no split from divorce or
                separation of spouses or partners
             1. Split household - one half of couple from SUBHH 0 and new
                spouse or partner, if any
             2. Split household - other half of couple from SUBHH 0 and new
                spouse or partner, if any
```

===
For W2KIDS, W2SIBS, and W2HHLIST only:

```
OPN      OTHER PERSON NUMBER
         Type: Character      Width: 3           Decimals: 0
.............................................................................

                010. Respondent Person Number
                011. Respondent Person Number
                020. Respondent Person Number
                021. Respondent Person Number
                030. Respondent Person Number
                031. Respondent Person Number
            033-034. Other Person Number
                040. Respondent Person Number
            041-083. Other Person Number
            101-399. Other Person Number
            401-499. Other Person Number
            501-699. Other Person Number
            701-799. Other Person Number
            901-998. Other Person Number
```

===

Secondary Identifiers

===

```
ASUBHH   1992 SUB-HOUSEHOLD IDENTIFIER
         Type: Character      Width: 1           Decimals: 0
.............................................................................
                    0. Original sample household
```

===

```
CPN_FAM    1994 FAMILY RESP PERSON NUMBER
           Type: Character       Width: 3              Decimals: 0

           User note:  Character (string) version of RPN for W2KIDS and W2SIBS
           ......................................................................................

                         010. Person Number
                         011. Person Number
                         020. Person Number
                         021. Person Number
                         030. Person Number
                         031. Person Number
                         040. Person Number
                    Blank. No family respondent; no interview

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

CPN_NFAM   1994 NON-FAMILY RESP PERSON NUMBER
           Type: Character       Width: 3              Decimals: 0
           ......................................................................................

                         010. Person Number
                         011. Person Number
                         020. Person Number
                         021. Person Number
                         030. Person Number
                         031. Person Number
                         040. Person Number
                    Blank. No spouse/partner; no non-family respondent; no interview

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

CPN_FIN    1994 FINANCIAL RESP PERSON NUMBER
           Type: Character       Width: 3              Decimals: 0
           ......................................................................................

                         010. Person Number
                         011. Person Number
                         020. Person Number
                         021. Person Number
                         030. Person Number
                         031. Person Number
                         040. Person Number
                    Blank. No financial respondent; no interview

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

CPN_NFIN   1994 NON-FINANCIAL RESP PERSON NUMBER
           Type: Character       Width: 3              Decimals: 0
           ......................................................................................

                         010. Person Number
                         011. Person Number
                         020. Person Number
                         021. Person Number
                         030. Person Number
                         031. Person Number
                         040. Person Number
                    Blank. No spouse/partner; no non-financial respondent; no interview

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

CPN_CS     1994 COVERSHEET RESP PERSON NUMBER
           Type: Character       Width: 3              Decimals: 0

           User note:  Character (string) version of RPN for W2KIDS and W2SIBS
           ......................................................................................

                         010. Person Number
                         011. Person Number
                         020. Person Number
                         021. Person Number
                         030. Person Number
                         031. Person Number
                         040. Person Number
                    Blank. No coversheet respondent; no interview

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

CPN_NCS  1994 NON- COVERSHEET RESP PERSON NUMBER
```

```
           Type: Character          Width: 3                 Decimals: 0
           ...............................................................................
                      010. Person Number
                      011. Person Number
                      020. Person Number
                      021. Person Number
                      030. Person Number
                      031. Person Number
                      040. Person Number
                    Blank. No spouse/partner; no non-coversheet respondent; no interview

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

COPN_CS  1994 ACTUAL COVERSHEET PERSON NUMBER
           Type: Character          Width: 3                 Decimals: 0

           User note:  Coversheet respondent may have been an informant other than either
           of the respondents.  Character (string) version of IPN in W2HHLIST and W101 in
           W2CS.
           ...............................................................................
                      010. Respondent Person Number
                      011. Respondent Person Number
                      020. Respondent Person Number
                      021. Respondent Person Number
                      030. Respondent Person Number
                      031. Respondent Person Number
                  033-034. Other Person Number
                      040. Respondent Person Number
                  041-083. Other Person Number
                  101-399. Other Person Number
                  401-499. Other Person Number
                  501-699. Other Person Number
                  701-799. Other Person Number
                  901-998. Other Person Number
                      999. Other person without assigned person number
                    Blank. No interview

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

Other Variables

For All Files

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

CVERSION  1994 DATA RELEASE VERSION
           Type: Numeric           Width: 3                 Decimals: 0
           ...............................................................................
                      2. Version 2, July 2004

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

For W2E

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

FAMFLAG   Family Respondent Problem Flag
           Type: Numeric           Width: 1                 Decimals: 0
           ...............................................................................
                      0.      Respondent is in a household that contains one Family
                              Respondent (no special attention needed)

                      2.      Respondent is in a household that contains no Family
                              Respondent (no family information was given)

                      3.      Respondent is in a household that contained two Family
                              Respondents; information from this respondent has been
                              retained; no special attention required by analyst.

                      4.      Respondent is in a household that contained two Family
                              Respondents; information from this respondent has been
                              deleted; no special attention required by analyst.

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

For W2KIDS, W2SIBS, and W2PARS

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

FAMFLAG   Family Respondent Problem Flag
          Type: Numeric          Width: 1                    Decimals: 0
          ...............................................................................

                        0.        [Record] is associated with a household that contains one
                                  Family Respondent (no special attention needed)

                        2.        Record is associated with a household that contained two
                                  Family Respondents; information from one respondent has
                                  been retained; no special attention required by analyst.

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

IF YOU NEED TO KNOW MORE

This document is intended to serve as an overview of what's new in the version 2.0 release of the 1994 HRS Core and Exit (Final) 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.

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

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

PREVIOUS DOCUMENTATION

Two documents, Data Description and Usage, Version 1.00 (datadesc.txt) and Difference between the HRS Wave 2 Partial Public Release and the HRS Wave 2 Full Public Release (partrel.txt) are included, unedited, in this document. **Much, but not all, of the material in these two documents is currently obsolete.** Please disregard the following information when it is in conflict with the proceeding information or with other, more recent, information at our Web site. As always, please contact us if you have additional questions.

Difference between the HRS Wave 2 Partial Public Release and the HRS Wave 2 Full Public Release (partrel.txt)

June 18, 1998

To follow is a list of changes that users of the HRS Wave 2 Partial Public Release may need to be aware of when switching to the HRS Wave 2 Full Public Release. Users who obtained their first HRS Wave 2 Public Release on June 18, 1998 or later need not read this document.

Changes between the partial and full releases of HRS Wave 2:

- SECTION CS (COVERSHEET - HOUSEHOLD AND INDIVIDUAL)

  Section CS, containing coversheet information, was added.

- SECTION A (MISCELLANEOUS, DEMOGRAPHIC BACKGROUND)

  A number of new variables were added to the beginning of Section A.

- SECTION C (COGNITION)

  Variables in Section C were renamed so that the interview sections were in the same order as they appeared in the interview. The original variable names were W7500-W7577.   The new variable names are W5800-W5877, obtained by subtracting 1700 from each original variable name.

- SECTION E (FAMILY)

   A number of new variables were added to the end of Section E.

- SECTION IDV

  The file W2IDV was eliminated from the HRS Wave 2 Full Public Release. It is replaced by the HRS/AHEAD Tracker File, available from the same area of the Web page as other HRS/AHEAD public release files.

- SECTION S (WIDOWHOOD AND DIVORCE)

  Variables in Section S were renamed so that the interview sections were in the same order as they appeared in the interview. The original variable names were W6600-W6674.   The new variable names are W6800-W6874, obtained by adding 200 to each original variable name.

_____

```
              HEALTH AND RETIREMENT STUDY, WAVE 2
                Full Public Release, Version 1.00
                         June 18, 1998


                    Data Description and Usage
```

_____

David Howell of the HRS/AHEAD staff authored this document.
It is intended to serve as an outline and approach to using the
data, but not to be a comprehensive guide. Because of the nature
of the process, the data description is often created under time
constraints and compiled from various sources in a hasty manner.
If errors are present, we apologize - please send notification to
David by electronic mail to "dahowell@isr.umich.edu".

NOTE: There has been much progress made recently on an HRS/AHEAD
User Guide that will provide much of the information contained
within, but in a more polished manner. When it becomes available,
overlapping content will be removed from this document.

_____

This release of HRS Wave 2 data set is intended for use by the
general public. By receiving these data which have been freely
provided, you are agreeing to use them for research and statistical
purposes only and make no effort to identify the respondents
therein. In addition, you are in good faith agreeing to send a copy
of any publications you produce based on the data to the address
below.

HRS/AHEAD Papers and Publications
Institute for Social Research, Room 3250
The University of Michigan
P.O. Box 1248
Ann Arbor, MI (USA) 48106-1248

_____

TABLE OF CONTENTS

```
 Sections
 --------


 ...1  Acknowledgments


 ...2  Contact Information


 ...3  Obtaining the Data
 ......3-a  Registration/Conditions of Use
 ......3-b  Internet Site


 ...4  Files Description
 ......4-a  Files List
 ......4-b  File Types
 .........  4-b-1  Documentation
 ............  4-b-1-a  Data Description
 ............  4-b-1-b  Interview/Questionnaire/Box-and-Arrow
 ............  4-b-1-d  Codebook
 ............  4-b-1-d  Questionnaire and Codebook Files List
 .........  4-b-2  Raw Data
 .........  4-b-3  Descriptor Statements
 ......4-c  Weights
 ......4-d  Imputations
 ......4-e  Identification Variables
 ......4-f  Structure
 .........  4-f-1  HRS/AHEAD Tracker File
 .........  4-f-2  Individual-level Files
 .........  4-f-3  Household-level Financial Files
 .........  4-f-4  Family and Household Listing Files
 ......4-g  Merging
 .........  4-g-1  Individual (Respondent) Level File Creation
 .........  4-g-2  Individual (Family/Helper/Household Member)
                  Level File Creation
 .........  4-g-3  Household level file creation.
 .........  4-g-4  Merging with HRS Wave 1


 ...5  Using the Files
 ......5-a  Setup
 ......5-b  Decompressing the Files
 .........  5-b-1  Decompressing Files Using HRS2.BAT
 .........  5-b-2  Decompressing Files Yourself
 ......5-c  Subdirectory Structure
 ......5-d  Using the Files with SAS
 ......5-e  Using the Files with SPSS
 ......5-f  Using the Files with STATA
```

_____

...1 Acknowledgments

The Health and Retirement Study (HRS) is a national longitudinal
study that focuses on persons born between the years 1931 and 1941
and their health, retirement, and economic status. It is a
cooperative agreement between the Institute for Social Research at
the University of Michigan and the National Institute on Aging.

Funding has been provided by the National Institute on Aging at NIH,
the Social Security Administration, the Department of Labor Pension
and Welfare Benefits Administration, the Office of the Assistant
Secretary for Planning and Evaluation at DHHS, the State of Florida
Department of Elder Affairs, the NIH Office of Research on Minority
Health, and the NIH Office of Research on Women's Health.

The data, with appropriate masking for purposes of respondent
confidentiality, is being made available to the public via the
Internet in hopes that a broad group of persons will make use of
this very important collection of data.

...2 Contact Information

If you have questions, concerns, or comments that are not adequately
addressed here or on our web page (http://www.umich.edu/~hrswww),
please feel free to contact us.

        E-Mail: hrsquest@isr.umich.edu

        Postal service: Health and Retirement Study
                        Institute for Social Research, Room 1248
                        P.O. Box 1248
                        Ann Arbor, MI 48106-1248

        FAX: (734) 647-1186

        Phone: (734) 647-1186

...3 Obtaining the Data

......3-a Registration/Conditions of Use

Before working with HRS or AHEAD data, you must first register.
Through your registration, we are able to convey to our sponsors
the size and diversity of our user community, allowing us to
continue to collect this important data. Registered users will
receive user support, as well as information related to errors in
the data, future releases, workshops, and publication lists. The
information provided will _not_ be for commercial use, and will
not be redistributed to third parties.

If you have already registered, thank you; you need not register
again unless the information submitted has changed.

If you have not yet registered, you may register your use of
HRS/AHEAD data by one of the following three methods:

- by completing the online registration form on the HRS/AHEAD
  Web Page (http://www.umich.edu/~hrswww/datafile/register.html)
- by returning the form below by E-Mail to dahowell@isr.umich.edu
- by returning the form below by postal service to David Howell,
  Institute for Social Research, Room 3114, The University of
  Michigan, P.O. Box 1248, Ann Arbor, MI (USA) 48106-1248

HRS/AHEAD PUBLIC RELEASE DATA
REGISTRATION AND CONDITIONS OF USE

By receiving these data which have been freely provided, you are
agreeing to use them for research and statistical purposes only and
make no effort to identify the respondents therein. In addition, you
are in good faith agreeing to send a copy of any publications you
produce based on the data to the address below.

HRS/AHEAD Papers and Publications
Institute for Social Research, Room 3250
The University of Michigan
P.O. Box 1248
Ann Arbor, MI (USA) 48106-1248

If you agree to the HRS/AHEAD Public Release Data Conditions of
Use, please fill out the information below.

Name:
E-Mail address:

```
Organization name:
Address:
City:
State:
Zip code:
Phone number:
FAX number:


Thank you for your assistance in this matter!


......3-b  Internet Site


Health and Retirement Study public release data is available
through the World-Wide-Web.


To access the data and other relevant information, you may point the
World-Wide-Web browser of your choice to the HRS/AHEAD Web Page at:


                    http://www.umich.edu/~hrswww


Choose "Datasets and Files" under the header of the same name.


...4  Files Description


The descriptions that follow deal only with files included with
and specific to the HRS Wave 2 Full Public Release.


Files associated with the same data set generally have the same
prefix.  For instance, SAS file "W2A.SAS" and EXTRACT file
"W2A.EDI" go with data file "W2A.DA".  Questionnaire and
codebook files have a slightly different prefix in that they
are preceded by a two digit number (and underscore) that indicates
the order in which the files should be printed to create a
properly ordered, complete copy of the documentation.


In addition to the files provided in the HRS Wave 2 Full Public
Release, there are two other HRS/AHEAD public release files users
will probably want to obtain.  The first is the HRS/AHEAD Tracker
File, which provides weights, tracking, demographic, and other
information for the entire sample in a single data set.  The second
is the HRS/AHEAD Concordance, a database that allows users to track
content and identify similar questions longitudinally.  Both files
are available from the HRS/AHEAD Web Page, in the same area of
Datasets and Files as the HRS Wave 2 Full Public Release.


......4-a  Files List
```

The number of files contained in the HRS Wave 2 Full Public
Release seem daunting at first. It eases the mind, however, to
realize that there is no need to access every one of the files.
Some files are specific for the codebook, others are for SAS users,
and so on. Indeed, it is unlikely that persons will use every
data set, or even every variable within a data set. Rather, it
is best to determine what content areas are of interest, and focus
in on just the files containing the variables of interest.

Data sets included in the HRS Wave 2 Public Release are:

```
Data File      Contains
---------      --------
W2A            Section A: Demographics, and Miscellaneous
W2B            Section B: Health
W2C            Section C: Cognition
W2CS           Household and Individual Coversheet Data
W2D            Section D: Housing
W2E            Section E: Family Structure
W2FA           Section FA: Employment (Employees)
W2FB           Section FB: Employment (Self-Employed)
W2FC           Section FC: Employment (Unemployed)
W2G            Section G: Last Job, R Not Working Now
W2H            Section H: Job History
W2HHLIST       Coversheet: Household Listing
W2J            Section J: Disability
W2K            Section K: Net Worth
W2KIDS         Section E: Children file
W2MOD0         Experimental Module 0: Activities and Nutrition
W2MOD1         Experimental Module 1: Depression Scale
W2MOD2         Experimental Module 2: Similarities
W2MOD3         Experimental Module 3: Physical Functioning
W2MOD4         Experimental Module 4: Spending and Saving
W2MOD5         Experimental Module 5: Risk Aversion
W2MOD6         Experimental Module 6: Social Support
W2MOD7         Experimental Module 7: Transfers
W2MOD8         Experimental Module 8: Help with ADLs
W2MOD9         Experimental Module 9: Activities and Time Use
W2N            Section N: Income
W2PARS         Section E: Parents file
W2R            Section R: Health Insurance
W2S            Section S: Widowhood
W2SIBS         Section E: Siblings file
W2V            Section V: Capital Gains
```

```
......4-b  File Types


.........  4-b-1  Documentation


There are three types of documentation available for use
specifically with the HRS Wave 2 Full Public Release.  They are
the Data Description, Interview/Questionnaire/Box-and-Arrow, and
Codebook.  Users of the data will want to become familiar with
all three and reference them often.


............  4-b-1-a Data Description


Subdirectory: (any)
Suffix:       TXT


The Data Description, which you are currently reading, gives
a rough overview of the data set.  It is stored as an ASCII text
file, and should be looked over prior to working with the data.


............  4-b-1-b Interview/Questionnaire/Box-and-Arrow


Subdirectory: C:\HRS\WAVE2\IVIEW
Suffix:       WP5


There are three names the research community uses that all refer
to basically the same piece of documentation: interview,
questionnaire, and box-and-arrow.  For purposes of this document,
we will refer to the document as a questionnaire.


The HRS Wave 2 Questionnaire is stored as  a set of WordPerfect
Version 5.0 files.  The questionnaire is the only part of the HRS
Wave 2 Public Release that is not in ASCII text form.  Because of
the graphical nature of the questionnaire, adequate conversion to
ASCII text format was not feasible.


The questionnaire is helpful when used in tandem with the codebook,
as the questionnaire graphically depicts skip patterns and the
flow of the interview, which some users find very helpful.


For a list of all HRS Wave 2 Questionnaire files, see Part 4-b-1-d
of this document.


............  4-b-1-c  Codebook


Subdirectory: C:\HRS\WAVE2\CODEBOOK
Suffix:       TXT
```

The HRS Wave 2 Codebook is stored as ASCII text files. There
should be a codebook file that corresponds to each dataset.
When accessing the codebook, it is sometimes also useful to
reference the associated questionnaire files.

The codebook conveys variable names, labels, question text,
code values, code labels. It also conveys some skip logic
in a non-graphical format. In addition, frequencies or means
are presented for each variable. Please note that the frequencies
and means are UNWEIGHTED. In addition, the means include missing
data values, and are intended only to be used to check that your
data read in correctly. The means and associated univariates
should not be used to examine the data analytically.

When accessing the codebook, it is sometimes also useful to
reference the associated questionnaire files. While it is possible
to work with the data at some level without the questionnaire, it
is nearly impossible to use the data without the codebook.

For a list of all HRS Wave 2 Codebook files, see Part 4-b-1-d
of this document.

.... 4-b-1-d Questionnaire and Codebook Files List

For users that wish to print out the entire HRS Wave 2
Questionnaire, or the entire HRS Wave 2 Codebook, the first
two digits of each file name indicate the order in which they
should be printed.

Questionnaire and codebook files included in the HRS Wave 2
Full Public Release are:

```
Q'Naire        Codebook
Files          Files          Contains
------------   ------------   --------
*1             01_W2MAS.TXT   Master Codes
01_W2CS.WP5    02_W2CS.TXT    Household and Individual Coversheet Data
02_W2A.WP5     03_W2A.TXT     Section A: Demographics, and Miscellaneous
03_W2B.WP5     04_W2B.TXT     Section B: Health
04_W2D.WP5     05_W2D.TXT     Section D: Housing
05_W2E.WP5     06_W2E.TXT     Section E: Family Structure
06_W2EE.WP5    *2             Section EE: Family Structure
07_W2FA.WP5    07_W2FA.TXT    Section FA: Employment (Employees)
08_W2FB.WP5    08_W2FB.TXT    Section FB: Employment (Self-Employed)
09_W2FC.WP5    09_W2FC.TXT    Section FC: Employment (Unemployed)
```

```
10_W2J.WP5    10_W2J.TXT    Section J: Disability
11_W2K.WP5    11_W2K.TXT    Section K: Net Worth
12_W2V.WP5    12_W2V.TXT    Section V: Capital Gains
13_W2C.WP5    13_W2C.TXT    Section C: Cognition
14_W2N.WP5    14_W2N.TXT    Section N: Income
15_W2R.WP5    15_W2R.TXT    Section R: Health Insurance
16_W2S.WP5    16_W2S.TXT    Section S: Widowhood
17_W2G.WP5    17_W2G.TXT    Section G: Last Job, R Not Working Now
18_W2H.WP5    18_W2H.TXT    Section H: Job History
*3            19_W2KID.TXT  Children Information from Section E
*3            20_W2SIB.TXT  Sibling Information from Section E
*3            21_W2PAR.TXT  Parent Information from Section E
*3            22_W2HHL.TXT  Household Listing from Section E
19_W2MD0.WP5  23_W2MD0.TXT  Module 0: Activities and Nutrition
20_W2MD1.WP5  24_W2MD1.TXT  Module 1: Depression Scale
21_W2MD2.WP5  25_W2MD2.TXT  Module 2: Similarities
22_W2MD3.WP5  26_W2MD3.TXT  Module 3: Physical Functioning
23_W2MD4.WP5  27_W2MD4.TXT  Module 4: Spending and Saving
24_W2MD5.WP5  28_W2MD5.TXT  Module 5: Risk Aversion
25_W2MD6.WP5  29_W2MD6.TXT  Module 6: Social Support
26_W2MD7.WP5  30_W2MD7.TXT  Module 7: Transfers
27_W2MD8.WP5  31_W2MD8.TXT  Module 8: Help with ADLs
28_W2MD9.WP5  32_W2MD9.TXT  Module 9: Activities and Time Use
```

*1 - The Master Codes file contains meanings for select
     large code sets, so that they only need to be listed out
     once. The Master Codes list is then referred to when
     needed in other codebooks.
*2 - Section EE has its own questionnaire entry, but is
     combined into the other Section E portions of the codebook.
*3 - Certain family and household listing information has been
   *3 - The coversheet's household listing and select parts of
        Section E (Family) have been broken out into separate
        files after collection. The questionnaire represents
        how they were actually collected, and the codebook indicates
        how they ended up.

.... 4-b-2 Raw Data

Subdirectory: C:\HRS\WAVE2\DATA
Suffix: DA

Files with the extension "DA" are raw data files. HRS Wave 2
data are stored in ASCII text format, with fixed-length records.
All HRS Wave 2 data should be numeric.

```
.........  4-b-3  Descriptor Statements


Subdirectory:          Suffixes:   Software:
C:\HRS\WAVE2\EXTRACT   EDI         EXTRACT
C:\HRS\WAVE2\OSIRIS    DI          OSIRIS
C:\HRS\WAVE2\SAS       SAS, SAI    SAS
C:\HRS\WAVE2\SPSS      SPS, SPI    SPSS
C:\HRS\WAVE2\STATA     DO, DCT     STATA
```

For software packages to understand the content of raw data
files (with the "DA" suffix), descriptor statements are
required. Because of the proprietary nature of software
packages, descriptor statements specific to the software
package are required. Please reference Part 5 of this
document for information on using the descriptor statements
provided with the software package of your choice.

......4-c Weights

Household and person-level weights for the HRS Wave 2 Public
Release are present in the HRS/AHEAD Tracker File. The
HRS/AHEAD Tracker File can be obtained from the same area of
the HRS/AHEAD Web Site as the HRS Wave 2 Public Release.

For the time being, the weights are not included in with the
HRS Wave 2 Public Release data files. In future versions of the
data, we will include them in each dataset; until that time, we
apologize for the inconvenience.

......4-d Imputations

A large number of variables were imputed in the HRS Wave 2 Public
Release data set. Those analysts who wish to use the imputations
need do nothing out of the ordinary. For those who do not care
to use the imputed values, or perhaps just wish to know what the
original value was, we have created imputation indicators.

The codebook indicates whether a particular variable was imputed.
Variables which are imputed should have the tag "[IMPUTED]"
below the variable label in the codebook. A variables which is
imputed should have an associated variable that is its imputation
indicator. The variable name for the imputation indicator is always
the variable number plus an additional 10,000. For example, if W100
were imputed, its imputation indicator would be W10100; for W9999
the imputation indicator would be W19999, et cetera. The presence
of an imputation indicator for a variable is further evidence that

a variable has been imputed.

The imputation indicators are one digit codes reflecting the
original value of the data. The meaning of the imputation
indicators are shown in the table below.

```
        Indicator
        Value      Meaning (original value of variable)
        ---------  -----------------------------------
        1          Partially missing data: brackets were used
                   to obtain the value
        2          Original value was termed Inap.
        3          Original value was not missing, and was inside
                   the valid range of codes (most probably, a
                   prior variable was imputed and changed the
                   skip pattern of this variable)
        4          Original value was not missing, but was outside
                   the valid range of codes
        5          Missing data: refused
        6          Partially missing data: the range card was
                   used to obtain the value
        7          Missing data: loss/negative, DK/NA; ]
                   "Q not relevant to R", "other"
        8          Missing data: DK/don't know
        9          Missing data: NA/Not ascertained
```

Analysts who so choose should be able to use the imputation
indicator in combination with the imputed variable to restore
the original values for the variable.

......4-e Identification Variables

Identification variables are distinguishable from other variables
in that they identify a record in a data set for a particular
level of analysis.

Household level. Upon being interviewed, each sample household was
assigned a Household Identifier (HHID). The HHID is stable, and
uniquely identifies the original household across time. At each
cross-section, however, the status of a household may change due to
the severance of a partnership or the death of a respondent. The
Sub-Household Identifier (SUBHH) for each cross-section, when used
in combination with the HHID, uniquely identifies a household as of
a particular cross-section. All households are assigned a SUBHH of
O in the first year of collection. Thereafter, a SUBHH of O
indicates that the original household remains intact. A SUBHH of

1 or 2 recognizes households that have broken off from the original
household due to the severance of a partnership. A SUBHH of 3
indicates a deceased respondent, who for practical reasons is
considered to now be in a household of their own.

In summary, to identify an original household use the HHID by
itself, but to identify a household as of a particular cross-section
use the HHID in combination with the SUBHH.

  Example 11-1. Two respondents in a sample household are married
  as of the first cross-section. Each respondent is assigned a
  HHID of 12345 and a SUBHH of 0. As of the second cross-section
  the two respondents are still married, and each retains their HHID
  of 12345 and their SUBHH of 0.

  Example 11-2. Two respondents in a sample household are married
  as of the first cross-section. Each respondent is assigned a
  HHID of 23456 and a SUBHH of 0. As of the second cross-section,
  the married couple divorces. At the second cross-section, both
  respondents retain their HHID of 23456, but each is assigned a
  SUBHH of 1 and 2, respectively.

  Example 11-3. Two respondents in a sample household are married
  as of the first cross-section. Each respondent is assigned a
  HHID of 34567 and a SUBHH of 0. One respondent dies before the
  next wave. At the next wave, both respondents retain their HHID
  of 34567; the living respondent retains their SUBHH of 0, but
  the deceased respondent is assigned a SUBHH of 3.

  Example 11-4. A respondent who has never been married is in the
  first cross-section. The respondent is assigned a HHID of 45678
  and a SUBHH of 0. As of the second cross-section, the
  respondent marries. Both the respondent and their new spouse are
  assigned a HHID of 45678 and a SUBHH of 0 as of the second
  cross-section. (The household was not divided or otherwise
  changed; it was added to.)

Individual level. Individuals, whether they be respondents,
children, siblings, or otherwise, are at their root persons
associated with a sample household. For that reason, they are able
to share a single identifier, the Person Number (PN). Person
numbers are unique within a household, meaning no two persons
associated with a household should ever have the same PN. In
addition, the PN assigned to a person never changes. When used
together, the HHID of the original household and the PN form an
identifier for the person that is unique across time. Because

HHID and PN do form a unique person identifier, a single combined
variable called HHIDPN has been included in many files for the
convenience of the analyst (though HHID and PN appear separately
as well).

  Example 11-5. A sample household with a HHID of 56789 contains
  two respondents assigned PNs of 010 and 020, respectively.
  Associated with the household are three children with PNs of
  101, 102, and 201, and two siblings with PNs of 051 and 052.
  A friend who lives with the respondents is assigned a PN of 80.
  All eight persons will keep those same PNs across time.

When dealing with individual level family, helper, and household
member files, be aware that households broken off from the
original household due to the severance of a partnership can each
contain a separate report on the same person.

  Example 11-6. A sample household at the first cross-section
  contains a respondent (who has a PN of 010), their spouse who
  is also a respondent (and has a PN of 020), and their mutual
  child (who has a PN of 201). As of the first cross-section, the
  household has a HHID of 67891 and a SUBHH of 0. Prior to the
  second cross-section, the respondents divorce. Thus, as of the
  second cross-section there are two sub-households. The first
  sub-household has a HHID of 67891 and a SUBHH of 1; it contains
  the first respondent (PN 010) and a report on their mutual child
  (PN 040). The second sub-household has a HHID of 67891 and a
  SUBHH of 2; it contains the other respondent (PN 020) and a
  report on their mutual child (PN 040). Note that the child is
  reported on by both sub-households and thus the information
  appears twice.

Because the current identification variable scheme was not
implemented until after multiple waves of data were collected,
some exceptions to the PN identification method remain in the
data. Those exceptions will be made consistent in the near
future, but until that time, the exceptions are listed below.

Special case: HRS Wave 1 Person Identifiers. As of the public
release of HRS Wave 1, we had not yet decided on using the HHIDPN
as a unique person identifier for respondents. For this reason,
in HRS Wave 1 only, the Case ID uniquely identifies each
respondent, not the HHIDPN. Analysts who plan to merge HRS Wave 1
to other HRS/AHEAD data sets will find that both the Case ID and
HHIDPN are present in the HRS/AHEAD Tracker File, available via
the web page in the same location as other HRS/AHEAD public release

data sets. Because both identifiers are present, the HRS/AHEAD
Tracker File can be used as a bridge to merge the HHIDPN and other
variables on to HRS Wave 1 files. There are plans to re-release
HRS Wave 1 with the new identification scheme sometime in late
summer, 1998.

Special case: HRS Wave 1 Parents. Parents at HRS Wave 1 were
assigned PNs in the range 31-49. As of HRS Wave 2, it was decided
that PNs in the range 10-49 would be reserved for respondents.
Because the parent PNs overlap with the respondent PNs, the parent
PNs are not unique. To solve this problem, it is suggested that
analysts add 40 to the parents' HRS Wave 1 PNs prior to merging
with other files. The new HRS Wave 1 parent PNs will then be in
the range 71-89. HRS/AHEAD staff will fix the problem in this
manner in the HRS Wave 1 public data set as of the re-release in
late summer, 1998.

Special case: HRS Wave 2 Parents. For some reason, PNs were not
a part of the HRS Wave 2 parents data file. As a result, for now
there are no PNs in the data set for parents at HRS Wave 2. Merging
HRS Wave 1 and HRS Wave 2 parent data thus becomes fairly difficult.
HRS/AHEAD processing staff will attempt to add PNs to the HRS Wave 2
parents file in the near future. In the meantime, analysts will
need to employ an alternative means of merging. One suggestion is
to merge based on the relationship of the parent to each respondent.

In summary, identification variables you will find in HRS Wave 2
are...

HHID   HRS Household Identifier
        [five digits]
        Uniquely identifies the original HRS Wave 1 household.

W2SUBHH HRS Wave 2 Sub-household Identifier
        [one digit]
        When used along with the HHID, uniquely identifies the
        household as of the HRS Wave 2 cross-section. Household
        composition at each wave can change due to the death of a
        respondent in a household, or the splitting of a partnered
        pair (as an example, perhaps due to divorce).

        A "0" in this variable indicates the original Wave 1
        household is still intact.
        A "1" or "2" indicates a Wave 1 household that has split
        as of this wave.
        A "3" indicates a deceased person who, because of their

deceased status, is considered for practical reasons to be
in a sub-household of their own.

PN   Person Number
        [three digits]
        Uniquely identifies a person associated with a household;
        this person may be a respondent, spouse/partner, child,
        sibling, parent, or other household member.
        Variants of this identifier are RPN (Respondent Person
        Number, in the family sections) and IPN (Informant
        Person Number, in the household listing).

HHIDPN HRS Household Identifier + Person Number
        [eight digits]
        A convenient combination of HHID (the first five digits)
        and PN (the last three digits). Uniquely identifies a
        person both cross-sectionally and longitudinally.
        Variants of this identifier are HHIDRPN and HHIDIPN.

......4-f Structure

The file structure for the HRS is most easily understood once the
method for collecting the data is understood.

First, at each cross-section, there are questions asked of all
respondents, questions asked of a designated Financial Respondent
on behalf of the entire household, and questions asked of a
designated Family Respondent on behalf of the entire household.

Second, most questions are also asked in other waves, introducing
a longitudinal aspect to the file structure.

We like to refer to the way our data is collected as being at
different "levels". One example of a level of collection is the
household level; another is the individual/respondent level.
Data at these different levels are associated by the identification
variables in Section 4-e. The HRS can thus be thought of as one
large relational database.

.... 4-f-1 HRS/AHEAD Tracker File

We attempt to aid users in tracking the HRS sample longitudinally
through use of our "tracker file". The version of the HRS/AHEAD
Tracker File updated as of HRS Wave 2 (file "W2TRACK") is available
from the same area of the web page as the HRS Wave 2 Public
Release.

Each record in the tracker file corresponds to one of the
13142 respondents that are a part of the sample as of HRS Wave 1
and HRS Wave 2.

Information in the tracker file includes select identification
variables, demographic information, weights, whether a person gave
an interview in a particular wave, and whether a person was the
Family or Financial Respondent in a particular wave, among other
things.

.... 4-f-2 Individual-level Files

When we say that a file was collected at the individual level,
we mean that all respondents were to be asked the questions in
these sections. Thus, there should be a record in each file for
each of the 11596 respondents that gave information in HRS Wave 2.

Asked of all respondents:

```
W2A            Section A: Demographics
W2B            Section B: Health
W2C            Section C: Cognition
W2E            Section E/EE: Family Structure *
W2FA           Section FA: Employment (Employees)
W2FB           Section FB: Employment (Self-Employed)
W2FC           Section FC: Employment (Unemployed)
W2G            Section G: Last Job, R Not Working Now
W2H            Section H: Job History
W2J            Section J: Disability
W2R            Section R: Health Insurance
W2S            Section S: Widowhood
```

There were also a set of experimental modules that were not
asked of all respondents, but rather a subset of available
respondents.

Asked of a subset of respondents:

```
W2E            Section E/EE: Family Structure *
W2MOD0         Experimental Module 0: Activities and Nutrition
W2MOD1         Experimental Module 1: Depression Scale
W2MOD2         Experimental Module 2: Similarities
W2MOD3         Experimental Module 3: Physical Functioning
W2MOD4         Experimental Module 4: Spending and Saving
W2MOD5         Experimental Module 5: Risk Aversion
```

```
W2MOD6          Experimental Module 6: Social Support
W2MOD7          Experimental Module 7: Transfers
W2MOD8          Experimental Module 8: Help with ADLs
W2MOD9          Experimental Module 9: Activities and Time Use
```

* In Section E, some questions were asked of all respondents,
  and others of just Family Respondents.

.... 4-f-3 Household-level Financial Files

Each household was to have a person designated to be the
"Financial Respondent"

Examples of household-level financial files are...

```
W2D             Section D: Housing
W2K             Section K: Net Worth
W2N             Section N: Income
W2V             Section V: Capital Gains
```

.... 4-f-4 Family and Household Listing Files

In reality, family files are still individual level files.
In this case, however, there is not one respondent per
individual line, but one household member, child, sibling,
or parent per individual line.

Files of this nature include...

```
W2HHLIST        Coversheet: Household Listing

W2KIDS          Section E: Children file
W2PARS          Section E: Parents file
W2SIBS          Section E: Siblings file
```

The information in the household listing was not given by
each household member, but instead by a single informant.
The information in the family files was not given by each
family member, but rather by a single person designated
as the "Family Respondent" in each household.

......4-g Merging

Merging would not be a particularly difficult task if all
datasets were structured the same. Unfortunately, because of the
hierarchical nature of the data many datasets are not structured

alike, and so merging becomes one of the more difficult data
management tasks facing the analyst.

Many analyses require variables that appear in separate files.
Before doing analysis work, the files will need to be merged in an
appropriate manner. Prior to doing any data management, however,
analysts should ask themselves two questions:

First, what variables are of interest? Predetermining what
variables are needed for an analysis allows the analyst to subset
their files to include only the necessary variables, weights, and
identification variables. The smaller files are, the more
manageable they are to work with.

Second, what should the final analysis file look like? Knowing
beforehand whether the intended analysis requires one household per
data set record, one respondent per data set record, or some other
configuration makes planning the merging of the files much easier.

After these two questions are answered, there are three main types
of final analysis files that analysts create. Descriptions of the
three file types are below, followed by instructions on how to
construct them.

Individual (respondent) level files contain information about one
respondent on each record. If there is more than one respondent
in a household, each will have their own record in the file.
Examples of this sort of file are the demographics, health, and
experimental modules files.

Individual (family/helper/household member) level files have
information about one person (who is not a respondent) on each
record. If there is more than one person of that type associated
with the household, each will have their own record in the file.
Examples of this sort of file are the children, parents, siblings,
helper, and household listing files.

Household level files have information about one household on
each record. Examples include household financial data, and
family data that is not specific to a single person associated
with the household.

.... 4-g-1 Individual (Respondent) Level File Creation

This set of instructions can be used to create a final analysis
file with one respondent per line.

1.   Subset your original files to include only the necessary
     weight, identification, and analysis variables.

2.   Identify subsetted files that are already at the individual
     (respondent) level and merge them together.
2a. Sort them by HHIDPN.
2b. Merge them by HHIDPN.
     The result should be a individual (respondent) level file with
     all individual (respondent) level variables in it.

3.   Identify subsetted files that are at the household level and
     merge them together.
     3a. Sort them by HHID and SUBHH.
     3b. Merge them by HHID and SUBHH.
     The result should be a household level file with all household
     level variables in it.

4.   Identify subsetted files that are at the individual (family/
     helper/household member) level, make them into household level
     files, and merge them together.
     4a. Sort them by HHID and SUBHH.
     4b. Determine the maximum number of persons per household in
          each subsetted file. In step 4c, you will create this
          many sub-files from each subsetted file.
     4c. For each subsetted file, create a number of sub-files,
          the first of which contains the first person in the
          household, the second of which contains the second person
          in the household, and so on until the nth file contains
          the nth person in the household.
     4d. Uniquely rename variables in each sub-file so that they
          do not write over each other when merging. You will not
          want to rename HHID and SUBHH and weight variables. You
          will want to uniquely rename the PN of the person as well
          as other variables that specifically refer to that person,
          such as age and education.
     4e. Sort all of the sub-files by HHID and SUBHH, if this is
          not already the case.
     4f. Merge all of the sub-files together by HHID and SUBHH.
     The result should be a household level file with all individual
     (family/helper/household member) variables strung out on each
     line.

5.   Merge the resultant files from steps 2, 3, and 4 together.
     5a. Sort the resultant files by HHID and SUBHH.
     5b. Merge the household level files from steps 3 and 4 to the

respondent level file from step 2 by HHID and SUBHH. Be
sure to have your merging routine allow for multiple
matches in the respondent level file. Because there can
be multiple respondents per household, you need to allow
the household level data to be matched to each.

The result should be an individual (respondent) level file, with
household and individual (family/helper/household member) data
present on each respondent's record. Account in your analyses
for the fact that information not originally at the individual
(respondent) level is duplicated in the file for households with
more than one respondent.

Example 11-7a. Sample households with HHID 78912 and 78193 have
one and two respondents in them, respectively. Household 78912
has a household income of $10,000 a year, and household 78913
$40,000 a year. When an individual (respondent) level data file
is created from these two households, there are three records that
result, one for each respondent. If we now run a mean household
income on the individual (respondent) file we get an average
household income of $30,000, which is incorrect because we have
counted the income of household 78913 twice (they have two records
in the file).

Example 11-7b. Sample households with HHID 78912 and 78913 have
one and two respondents in them, respectively. Household 78912
has a household income of $10,000 a year, and household 78913
$40,000 a year. When an individual (respondent) level data file
is created from these two households, there are three records that
result, one for each respondent. Before running a mean household
income, we keep the first record for each HHID and SUBHH, and
discard all but the first record for households with a duplicate
HHID and SUBHH. If we now run a mean household income on the
revised individual (respondent) file we get an average household
income of $25,000, which is correct as each household is counted
only once.

.... 4-g-2 Individual (Family/Helper/Household Member)
                  Level File Creation

This set of instructions can be used to create a final analysis
file with one family member, household member, or helper per line.
Before creating this particularly complex type of file, consider
whether a final analysis file at the household or individual
(respondent) level might be an acceptable alternative.

1.   Subset your original files to include only the necessary

weight, identification, and analysis variables.

2.   Identify subsetted files that are at the individual
     (respondent) level, merge them together, and make them into
     household level files.
     2a.   Sort them by HHIDPN.
     2b.   Merge them by HHIDPN.
     2c.   Take the resultant file and create two sub-files, the
           first of which contains the first person in the household,
           the second of which contains the second person in the
           household, if present.
     2d.   Uniquely rename variables in each sub-file so that they
           do not write over each other when merging. You will not
           want to rename HHID and SUBHH and weight variables. You
           will want to uniquely rename the PN of the respondent as
           well as other variables that specifically refer to that
           respondent, such as age and education.
     2e.   Sort the two sub-files by HHID and SUBHH.
     2f.   Merge the two sub-files together by HHID and SUBHH.
     The result should be a household level file with all
     individual (respondent) level variables in it.

3.   Identify subsetted files that are at the household level and
     merge them together.
     3a.   Sort them by HHID and SUBHH.
     3b.   Merge them by HHID and SUBHH.
     The result should be a household level file with all household
     level variables in it.

4.   Identify the child, parent, sibling, helper, or household
     member individual level file that you want to merge all other
     files to. We will call this the "core" file.   For instance,
     you may want to attach information about household members
     and respondents' parents to each of their children. In this
     case, the child file is the "core" file, and the household
     member and parent files are going to be merged to it. In
     other words, the final analysis file will have one child per
     line, not one parent or household member per line. After you
     have identified the core file, leave it alone until step 6.

5.   Identify subsetted files other than the core file that are at
     the individual (family/helper/household member) level, make
     them into household level files, and merge them together.
     5a.   Sort them by HHID and SUBHH.
     5b.   Determine the maximum number of persons per household in
           each subsetted file. In step 4c, you will create this

many sub-files from each subsetted file.

5c. For each subsetted file, create a number of sub-files, the first of which contains the first person in the household, the second of which contains the second person in the household, and so on until the nth file contains the nth person in the household.

5d. Uniquely rename variables in each sub-file so that they do not write over each other when merging. You will not want to rename HHID and SUBHH and weight variables. You will want to uniquely rename the PN of the person as well as other variables that specifically refer to that person, such as age and education.

5e. Sort all of the sub-files by HHID and SUBHH, if this is not already the case.

5f. Merge all of the sub-files together by HHID and SUBHH.

The result should be a household level file with all individual (family/helper/household member) variables, except those in the core file, strung out on each line.

6. Merge the resultant files from steps 2, 3, 4, and 5 together.

6a. Sort the resultant files by HHID and SUBHH.

6b. Merge the household level files from steps 2, 3, and 5 to the core file from step 4 by HHID and SUBHH. Be sure to have your merging routine allow for multiple matches in the core file. Because there can be multiple persons per household in the core file, you need to allow the household level data to be matched to each.

The result should be an individual (core) level file, with individual (respondent), household, and individual (family/ helper/household member) level data present on each respondent's record. Account in your analyses for the fact that household and individual (family/helper/household member) information is duplicated in the file for households with more than one core individual.

Example 11-8a. Sample households with HHID 89123 and 89124 have one and two resident children in them, respectively. Household 78912 has a household income of $10,000 a year, and household 89123 $40,000 a year. When an individual (child) level data file is created from these two households, there are three records that result, one for each child. If we now run a mean household income on the individual (child) file we get an average household income of $30,000, which is incorrect because we have counted the income of household 89124 twice (the household has two children, and each has a record in the file).

Example 11-8b. Sample households with HHID 89123 and 89124 have
one and two resident children in them, respectively. Household
89123 has a household income of $10,000 a year, and household
89124 $40,000 a year. When an individual (child) level data file
is created from these two households, there are three records that
result, one for each child. Before running a mean household
income, we keep the first record for each HHID and SUBHH, and
discard all but the first record for households with a duplicate
HHID and SUBHH. If we now run a mean household income on the
revised individual (child) file we get an average household income
of $25,000, which is correct as each household is counted only
once.

.... 4-g-3 Household level file creation.

This set of instructions can be used to create a final analysis file
with one household per line.

1. Subset your original files to include only the necessary
    weight, identification, and analysis variables.

2. Identify subsetted files that are at the individual
    (respondent) level, merge them together, and make them into
    household level files.
    2a. Sort them by HHIDPN.
    2b. Merge them by HHIDPN.
    2c. Take the resultant file and create two sub-files, the
         first of which contains the first person in the household,
         the second of which contains the second person in the
         household, if present.
    2d. Uniquely rename variables in each sub-file so that they do
         not write over each other when merging. You will not want
         to rename HHID and SUBHH and weight variables. You will
         want to uniquely rename the PN of the respondent as well
         as other variables that specifically refer to that
         respondent, such as age and education.
    2e. Sort the two sub-files by HHID and SUBHH.
    2f. Merge the two sub-files together by HHID and SUBHH.
    The result should be a household level file with all individual
    (respondent) level variables in it.

3. Identify subsetted files that are already at the household
    level and merge them together.
    3a. Sort them by HHID and SUBHH.
    3b. Merge them by HHID and SUBHH.
    The result should be a household level file with all household

level variables in it.

4.   Identify subsetted files that are at the individual (family/
      helper/household member) level, make them into household level
      files, and merge them together.
      4a.   Sort them by HHID and SUBHH.
      4b.   Determine the maximum number of persons per household in
            each subsetted file. In step 4c, you will create this
            many sub-files from each subsetted file.
      4c.   For each subsetted file, create a number of sub-files, the
            first of which contains the first person in the household,
            the second of which contains the second person in the
            household, and so on until the nth file contains the nth
            person in the household.
      4d.   Uniquely rename variables in each sub-file so that they do
            not write over each other when merging. You will not want
            to rename HHID and SUBHH and weight variables. You will
            want to uniquely rename the PN of the person as well as
            other variables that specifically refer to that person,
            such as age and education.
      4e.   Sort all of the sub-files by HHID and SUBHH, if this is
            not already the case.
      4f.   Merge all of the sub-files together by HHID and SUBHH.
      The result should be a household level file with all individual
      (family/helper/household member) variables strung out on each
      line.

5.   Merge the resultant files from steps 2, 3, and 4 together.
      5a.   Sort the resultant files by HHID and SUBHH.
      5b.   Merge the resultant files together by HHID and SUBHH.
      The result should be a household level file, with individual
      (respondent) and individual (family/helper/household member)
      data strung out on each household's record.

.... 4-g-2 Merging with HRS Wave 1

Longitudinal merges are basically the same as cross-sectional
merges except that you need to make sure your variable names do
not overlap and thus overwrite each other. That should not
generally be a problem in the HRS, as Wave 1 variables are
preceded by a "V" and most Wave 2 variables are preceded by a
"W".

Because HRS Wave 1 was distributed prior to a revision in how
we think about our identification variables, it uses a different
individual-level identification variables for respondents.

The variable the uniquely identifies HRS Wave 1 respondents
is W1CASE. Fortunately, both W1CASE and HHIDPN are present in
the HRS/AHEAD Tracker File (see Part 4-f-1). Thus, the
HRS/AHEAD Tracker File can be used to add HHIDPN to HRS Wave 1,
of W1CASE to HRS Wave 2, thus solving the problem.

Otherwise, dealing with HRS Wave 1 data is mostly similar to
dealing with HRS Wave 2 data.

...5 Using the Files

......5-a Setup

While a particular setup is not required for using the HRS files,
we do recommend the following.

Create the directory C:\HRS\WAVE2 on your hard drive.

Copy all of the HRS Wave 2 files you retrieved from the HRS/AHEAD
Web Page to C:\HRS\WAVE2.

By using this directory structure, you ensure that HRS2.BAT will
work appropriately with your files and that you will not have to
change the path names in your descriptor files. This method also
well organizes your files and makes user support easier.

......5-b Decompressing the Files

First, verify that the files you copied from the HRS/AHEAD Web
Page are in directory C:\HRS\WAVE2.

At this point, you may choose to decompress the files by one
of two methods: using HRS2.BAT, or by doing it yourself.

.... 5-b-1 Decompressing Files Using HRS2.BAT

We have provided a DOS batch utility called HRS2.BAT to aid in
decompression of your files. You may run the utility from DOS,
or launch it using the Run command in Windows.

The first screen identifies the program and announces that
you may press the [CTRL] and [BREAK] key at any time to halt
the program. A brief reminder of responsible use of the data is
included. You are then prompted to press a key to continue.

The second screen reminds you that you need to have created the

directory C:\HRS\WAVE2 on your hard drive. It also lists the
compressed (self-extracting) files that you may decompress with
HRS2.BAT and what they contain. If you have completed the
instructions as stated, you may press 'Y' (yes) to go on. Pressing
'N' (for no) will result in the program halting execution so you
may make the appropriate changes and run it again.

The third screen of HRS2.BAT is the one in which you choose
which of the files you wish to decompress. The files,
along with their compressed size, decompressed size, and number of
decompressed files are in the table below. Make sure
you have enough space on your hard drive to decompress the
files you need.

When the computer prompts you to extract (decompress) a particular
file type, answer 'Y' (yes) if you wish to do so and 'N' (no) if you
do not. You do not need to decompress file types you do not plan on
using. For example, if you do not plan to use the Wave 2 Interview,
you need not decompress the associated file "IVIEW.EXE". However,
while you can choose not to decompress files of a particular type,
once you choose a type of file to decompress, with HRS2.BAT you must
decompress all files of that type.

```
                            Size
                 #  ------------------------
File          Files    Compressed Decompressed  Contents
------------  -----   ------------ ------------  -------------------
CODEBOOK.EXE     32       323,905    2,376,446   Codebook files
DATA.EXE         31     4,487,889   74,491,950   Data files
EXTRACT.EXE      31        71,791      304,928   EXTRACT descriptors
IVIEW.EXE        28       459,108    2,247,699   Wave 2 Interview
OSIRIS.EXE       31        69,351      299,956   OSIRIS descriptors
SAS.EXE          62        88,553      259,520   SAS descriptors
SPSS.EXE         62        88,606      256,370   SPSS descriptors
STATA.EXE        62        76,259      294,195   STATA descriptors
------------  -----   ------------ ------------  -------------------
Total           339     5,665,462   80,531,064
```

HRS2.BAT will print 'End of program.' to the screen when it is done.
Look in the appropriate subdirectories for the decompressed files;
the subdirectories are listed in Part 5-c.

.... 5-b-2 Decompressing Files Yourself

An advantage to decompressing files yourself is that you can
decompress only the files you need.

HRS Wave 2 files the have the EXE suffix after their filename are
PKZip self-extracting files. The software for decompression is
already built into the files. You may use many of the same commands
on these files as you would when running PKUnzip on a ZIP file.
You may also use other PKZip-compliant softwares such as WinZIP
to manipulate these files.

When decompressing files yourself, we still recommend that you
decompress the files to the same subdirectory structure as used
by HRS2.BAT. The recommended subdirectory structure is outlined
in Section 5-c.

......5-c Subdirectory Structure

After decompression, whether you used HRS2.BAT or decompressed them
yourself, the file directories should be as follows if you followed
our recommendations:

```
Subdirectory            Is to contain
------------            -------------
C:\HRS\WAVE2\CODEBOOK   Codebook files
C:\HRS\WAVE2\DATA       Data files
C:\HRS\WAVE2\EXTRACT    EXTRACT descriptor files
C:\HRS\WAVE2\IVIEW      Wave 2 Interview
C:\HRS\WAVE2\OSIRIS     OSIRIS descriptor files
C:\HRS\WAVE2\SAS        SAS descriptor files
C:\HRS\WAVE2\SPSS       SPSS descriptor files
C:\HRS\WAVE2\STATA      STATA descriptor files
```

......5-d Using the Files With SAS

To create a SAS system file for a particular dataset, the
following three file types must be present for that dataset:

```
        Directory:                      Files:
        C:\HRS\WAVE2\SAS\               *.SAS
        C:\HRS\WAVE2\SAS\               *.SAI
        C:\HRS\WAVE2\DATA\              *.DA
```

Files with the suffix "SAS" are short SAS programs which you may
use to make a SAS system file. Load them into SAS and submit
them as is; if you followed our recommended setup and all goes
well, the SAS system file should then appear in directory
C:\HRS\WAVE2\SAS.

Files with the suffix "SAI" are the SAS input statements used by
the SAS programs to describe the data.

Files with the suffix "DA" contain the raw data for SAS to read.

NOTE: If you do not want to read the entire dataset into SAS,
you may edit the SAI file to read in only the variables you
desire.

......5-e Using the Files With SPSS   ?

To create a SPSS system file for a particular dataset, the
following three file types must be present for that dataset:

        Directory: Files:
        C:\HRS\WAVE2\SPSS\   *.SPS
        C:\HRS\WAVE2\SPSS\   *.SPI
        C:\HRS\WAVE2\DATA\   *.DA

Files with the suffix "SPS" are short SPSS programs which you may
use to make an SPSS system file. Load them into SPSS and submit
them as is; if you followed our recommended setup and all goes
well, the SPSS system file should then appear in directory
C:\HRS\WAVE2\SPSS.

Files with the suffix "SPI" are the SPSS input statements used by
the SPSS programs to describe the data.

Files with the suffix "DA" contain the raw data for SPSS to read.

NOTE: If you do not want to read the entire dataset into SPSS,
you may edit the SPI file to read in only the variables you
desire.

......5-f Using the Files With STATA   ?

To use STATA with a particular dataset, the following three file
types must be present for that dataset:

        Directory: Files:
        C:\HRS\WAVE2\STATA\   *.DO
        C:\HRS\WAVE2\STATA\   *.DCT
        C:\HRS\WAVE2\DATA\   *.DA

Files with the suffix "DO" are short STATA programs ("do files")
which you may use to read in the data. Load them into STATA and

submit them as is; if you followed our recommended setup and all
goes well, STATA should read the data in appropriately.

Files with the suffix "DCT" are STATA dictionaries used by STATA
to describe the data.

Files with the suffix "DA" contain the raw data for STATA to read.

NOTE: Due to STATA's unique method of memory management, sometimes
STATA has trouble reading in exceptionally large files. To aid
in overcoming this problem, or if you do not want to read the
entire dataset into STATA, you may edit the DCT file to read in
only the variables you desire.

......5-g Using the Files With Other Software

Using the data with software other than SAS, SPSS, or STATA
requires that you be able to describe the raw data files located
in subdirectory C:\HRS\WAVE2\DATA. The raw data should all be
numeric, and are stored in ASCII text format with fixed-length
records.

Five basic types of descriptors are needed to describe the
variables in the HRS Wave 2 raw data files:

        1. Variable name (Var Name)
        2. Variable label (Var Label)
        3. Column (Column)
        4. Width (Width)
        5. Number of implicit decimals (D)
           [For example, if the number of implicit decimals is
           "2", a data point stored as "202" should be read by
           the software package as "2.02".]

All of these descriptors are stored in a convenient table as part
of the EXTRACT descriptor statements (the ones with the "EDI"
extension) under subdirectory C:\HRS\WAVE2\EXTRACT.

You may use the EXTRACT descriptors to create and edit a set of
data descriptor statements appropriate for reading the data into
a software package of your choice.

...6 Data Description

Respondents. There were a total of 13,006 persons that were
eligible to give an interview in HRS Wave 2. 11,596 persons

gave an interview, and 1,410 did not.

[NOTE: The 136 HRS Wave 1 respondents that were given over to
the AHEAD sample prior to HRS Wave 2 were not eligible to
be interviewed by HRS as of HRS Wave 2, and are not present
in Table 6-1.]

Table 6-1. Respondents

```
+--------------------------------------------+--------+
| Not interviewed at HRS Wave 2              | 1,375 |
| Not interviewed at HRS Wave 2; new spouse  |    35 |
| SUB-TOTAL                                  | 1,410 |
+--------------------------------------------+--------+
| Interviewed at HRS Wave 2                  | 11,522 |
| Interviewed at HRS Wave 2; new spouse      |    74 |
| SUB-TOTAL                                  | 11,596 |
+--------------------------------------------+--------+
| TOTAL                                      | 13,006 |
+--------------------------------------------+--------+
```

Households. There were 7,227 cross-sectional households at
HRS Wave 2. This was determined by calculating the number
of unique HHID+W2SUBHHs.

As shown in Table 6-2, there were 2,494 single households
in which the respondent gave an interview. Of 4,733 paired
households, in 4,369 both respondents gave an interview, and in
364 only one respondent gave an interview.

Table 6-2. Households

```
+--------------------------------------------+--------+
| Single household, gave interview           | 2,494 |
+--------------------------------------------+--------+
| Paired household, both gave interview      | 4,369 |
| Paired household, only one gave interview  |   364 |
| SUB-TOTAL                                  | 4,733 |
+--------------------------------------------+--------+
| TOTAL                                      | 7,227 |
+--------------------------------------------+--------+
```

Financial Respondents. Of the 7,227 HRS Wave 2 households,
6,979 had a Financial Respondent, and 248 did not.

Households may be missing a Financial Respondent for a variety

of reasons, including non-response, interviewer error, and
errors in the instrument.

Table 6-3. Financial Respondents

```
+------------------------------------------+--------+
| No Financial Respondent in household     |        |
| Single household, gave interview         |    156 |
| Paired household, both gave interview    |      5 |
| Paired household, only one gave interview|     87 |
| SUB-TOTAL                                |    248 |
+------------------------------------------+--------+
| One Financial Respondent in household    |        |
| Single household, gave interview         |  2,338 |
| Paired household, both gave interview    |  4,364 |
| Paired household, only one gave interview|    277 |
| SUB-TOTAL                                |  6,979 |
+------------------------------------------+--------+
| TOTAL                                    |  7,227 |
+------------------------------------------+--------+
```

Family Respondents. Of the 7,227 HRS Wave 2 households,
6,915 had a single Family Respondent, 10 had two Family
Respondents, and 302 had no Family Respondents.

Households may be missing a Family Respondent for a variety
of reasons, including non-response, interviewer error, and
errors in the instrument.

Households that have two Family Respondents are also in err,
likely due to interviewer error or errors in the instrument.
Because the responses given by the two Family Respondents was
sometimes inconsistent, the data for both was retained, and
should be addresses by analysts prior to using family data.

Table 6-4. Family Respondents

```
+------------------------------------------+--------+
| No Family Respondent in household        |        |
| Single household, gave interview         |    144 |
| Paired household, both gave interview    |      9 |
| Paired household, only one gave interview|    149 |
| SUB-TOTAL                                |    302 |
+------------------------------------------+--------+
| One Family Respondent in household       |        |
| Single household, gave interview         |  2,350 |
```

```
| Paired household, both gave interview     |  4,350 |
| Paired household, only one gave interview |    215 |
| SUB-TOTAL                                 |  6,915 |
+-------------------------------------------+--------+
| Two Family Respondents in household       |        |
| Single household, gave interview          |      0 |
| Paired household, both gave interview     |     10 |
| Paired household, only one gave interview |      0 |
| SUB-TOTAL                                 |     10 |
+-------------------------------------------+--------+
| TOTAL                                     |  7,227 |
+-------------------------------------------+--------+
```

Variables and Records Per File. Table 6-5 lists the number of
variables and records located in each file in the HRS Wave 2
data set.

From the amount of records, you can tell how many persons or
households (as appropriate) are in each file.

Table 6-5. Number of Variables and Records Per File

```
Data File       Variables   Records   Contains
-----------     ---------   -------   --------------------------------
W2A.DA                 75    11,596   Section A: Demographics
W2B.DA                176    11,596   Section B: Health
W2C.DA                 82    11,596   Section C: Cognition
W2CS.DA                56    13,006   Coversheet Data
W2D.DA                140     6,979   Section D: Housing
W2E.DA                 72    11,596   Section E: Family Structure
W2FA.DA               568    11,596   Section FA: (Employees)
W2FB.DA               460    11,596   Section FB: (Self-Employed)
W2FC.DA               228    11,596   Section FC: (Unemployed)
W2G.DA                 74    11,596   Section G: Last Job, R Not Working
W2H.DA                200    11,596   Section H: Job History
W2HHLIST.DA            17    21,635   Coversheet: Household Listing
W2J.DA                141    11,596   Section J: Disability
W2K.DA                108     6,979   Section K: Net Worth
W2KIDS.DA              36    22,930   Section E: Children file
W2MOD0.DA              57       222   Module 0: Activities and Nutrition
W2MOD1.DA              16       815   Module 1: Depression Scale
W2MOD2.DA               9       817   Module 2: Similarities
W2MOD3.DA              19       771   Module 3: Physical Functioning
W2MOD4.DA              52     1,561   Module 4: Spending and Saving
W2MOD5.DA               7       801   Module 5: Risk Aversion
W2MOD6.DA              32       203   Module 6: Social Support
```

```
W2MOD7.DA              8       827  Module 7: Transfers
W2MOD8.DA             40       822  Module 8: Help with ADLs
W2MOD9.DA             28       179  Module 9: Activities and Time Use
W2N.DA               577     6,979  Section N: Income
W2PARS.DA             41    12,247  Section E: Parents file
W2R.DA                77    11,596  Section R: Health Insurance
W2S.DA                78    11,596  Section S: Widowhood
W2SIBS.DA             54    17,880  Section E: Siblings file
W2V.DA                68     6,979  Section V: Capital Gains
-----------    ---------   -------  -------------------------------
TOTAL              3,596
```

The coversheet file has 13,006 records, one for each potential HRS
Wave 2 respondent.

Files with 11,596 records have a record for each respondent.

Files with 6,979 records are financial files, with a record
for each household that had a Financial Respondent.

Modules are only asked of a portion of all respondents.

All other files (household listing, kids, sibs, parents) have one
household or family member per line.

......6-a Masking for Confidentiality

We at the Health and Retirement Study are dedicated to maintaining
the confidentiality of our respondents. For that reason, a number
of variables have been recoded, removed, or set to zero for the
public release data set. A record of some of those changes
follows.

Names, addresses, and similar variables are not present.

Days of birth have been set to zero or removed altogether.

State level geography data is recoded in all places to a level no
more detailed than U.S. Census Region and Division.

Data on the highest educational degree earned has been further
grouped together to increase cell sizes.

Industry and occupation codes were recoded into a group of
thirteen and seventeen codes, respectively, from the original
three-digit U.S. Census code. The codes themselves may be found

as part of the codebook.

...7 If you have special Needs or Problems

If you have any special requests or needs, feel free to
contact the HRS/AHEAD Staff (for contact information, see Part 2
of this document). We will do our best to help. Suggestions
and/or questions concerning data content, codes, methodology,
and research-related topics are particularly welcome. Technical
support beyond problems with the files themselves is limited.

Good luck!