

HEALTH AND RETIREMENT STUDY
2016 Core
Early, Version 2.0
July 2018

Data Description and Usage

TABLE OF CONTENTS

TABLE OF CONTENTS	III
1. OVERVIEW	1
1A. THE SAMPLE INTERVIEWED IN 2016.....	1
1B. 2016 QUESTIONNAIRE SECTIONS	2
1C. LEVELS OF FILES.....	3
1C1. Household Level Files	3
1C2. Respondent Level Files	3
1C3. Sibling Level File	3
1C4. Household Member and Child Level Files	3
1C5. Helper Level File	3
1C6. Transfer-to-Child-Level File	4
1C7. Transfer-from-Child-Level-File	4
1C8. Pension Level Files	4
1C9. Jobs level files	4
2. FILE NAMING CONVENTIONS	4
3. DATA FILES	5
4. IDENTIFICATION VARIABLES	6
4A. PRIMARY IDENTIFICATION VARIABLES	6
4A1. HHID - Household Identification Number	6
4A2. PSUBHH - 2016 Sub-household Identifier	6
4A3. PN - Person Number	6
4A4. OPN - Other Person Number	7
4B. PRIMARY IDENTIFICATION VARIABLES FOR DATASETS AT EACH OF THE NINE LEVELS	7
4C. SECONDARY IDENTIFICATION VARIABLES	7
4D. DATASETS INCLUDING PSUBHH AND OPN AS PRIMARY IDENTIFIERS.....	8
5. DISTRIBUTION FILES AND DIRECTORY STRUCTURE	8
5A. DISTRIBUTION FILES	8
5B. DIRECTORY STRUCTURE	9
6. PROGRAM STATEMENTS	9
6A. USING THE FILES WITH SAS	9
6B. USING THE FILES WITH SPSS	10
6C. USING THE FILES WITH STATA	10
7. LOADING HRS DATA PRODUCTS IN A NON-WINDOWS ENVIRONMENT	10
7A. WORKING IN AN OS X ENVIRONMENT	10
7A1. OS X: SPSS	11
7A2. OS X: Stata	11
7B. WORKING IN A LINUX ENVIRONMENT	12
7B1. Linux: Stata	12
7B2. Linux: SAS	12
8. DOCUMENTATION	13
8A. CODEBOOK	13
8A1. Variable Names	13
8A2. Other Types of Documentation	15
8B. MASTER CODES	15
9. ADDITIONAL NOTES	15

9A. HOUSEHOLDS WITH NO FAMILY OR FINANCIAL RESPONDENTS	15
9B. UNFOLDING BRACKET VARIABLES AND IMPUTATIONS	16
9C. MODULES FOR THE 2016 DATA COLLECTION	18
9D. EXPLANATION OF THE DIFFERENCE BETWEEN TRACKER AND CORE DATA RELEASES	18
9E. ENHANCED FACE-TO-FACE SAMPLE DESIGN AND CONTENT	18
9F. SUMMARY OF DATA MODEL (DM) CHANGES DURING 2016 PRODUCTION	19
10. OBTAINING THE DATA.....	24
10A. REGISTRATION AND DOWNLOADING THE DATA	24
10B. CONDITIONS OF USE	24
10C. PUBLICATIONS BASED ON DATA.....	24
11. IF YOU NEED TO KNOW MORE.....	25
11A. HRS INTERNET SITE	25
11B. CONTACT INFORMATION:	25
APPENDIX.....	I
A. EXAMPLES OF SUB-HOUSEHOLD AND RESPONDENT PERSON NUMBER AND OTHER PERSON NUMBER ASSIGNMENTS.....	I
A1. MARRIED COUPLE STAYS MARRIED	I
A2. COUPLE DIVORCES.....	II
A3. ONE OR BOTH RESPONDENTS DIE.....	II
A4. SINGLE RESPONDENT MARRIES	II
A5. COUPLE DIVORCES, ONE RESPONDENT REMARRIES AND DIVORCES	III
A6. COUPLE DIVORCES AND MARRIES AGAIN	IV
A7. MARRIED COUPLE WITH CHILDREN AND SIBLINGS	IV
A8. COUPLE WITH CHILDREN AND SIBLINGS DIVORCES.....	V
A9. COUPLE DIVORCES, ONE RESPONDENT REMARRIES, BOTH SPLIT-OFF HOUSEHOLDS HAVE NEW MEMBERS ..	V

Data Description and Usage

1. Overview

The 2016 HRS Core (Early, Version 2.0) data release consists of data obtained as part of the Health and Retirement Study (HRS), a national longitudinal study of the economic, health, marital, and family status, as well as public and private support systems, of older Americans. The National Institute on Aging provided funding (NIA U01 AG009740), with supplemental support from the Social Security Administration. The Institute for Social Research (ISR) Survey Research Center (SRC) at the University of Michigan conducted the survey.

The 2016 HRS Core (Early, Version 2.0) is an early public release that includes cases from the six continuing HRS cohorts (HRS, AHEAD, CODA, WAR BABY, EARLY BABY BOOMER and MID BABY BOOMER) and the newly screened LATE BABY BOOMER cohort. HRS early data releases are provided for users who wish to begin analyzing data prior to a final public release being available. Early release files have undergone a fair amount of data cleaning, but errors may be present in the data files and/or the associated codebooks. The Early release data files have been structured in a format consistent with Final release datasets.

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

1A. The Sample Interviewed in 2016

The data collection period for the 2016 interview was April 2016 through April 2018. The HRS sample is comprised of seven sub-samples (HRS, AHEAD, CODA, WAR BABY, EARLY BABY BOOMER, MIDDLE BABY BOOMER and LATE BABY BOOMER).

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

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

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

The War Baby (WB) sub-sample consists of people who were born in 1942 through 1947, were household residents of the conterminous U.S. in the spring 1992, who, at that time, did not have a spouse or partner born before 1924 or between 1931 and 1941, and were still household residents at the time of the first interview in 1998, and their spouses or partners at the time of the initial interview or

at the time of any subsequent interview. The War Baby sub-sample was interviewed in 1998 and every two years thereafter.

The Early Baby Boomer (EBB) sub-sample consists of people who were born in 1948 through 1953, were household residents of the U.S. when first interviewed in 2004, and who, at that time, did not have a spouse or partner who was born before 1948, and their spouses or partners at the time of the initial interview or at the time of any subsequent interview. The Early Baby Boomer sub-sample was interviewed in 2004 and every two years thereafter.

The Mid Baby Boomer (MBB) sub-sample consists of people who were born between 1954 and 1959, were household residents of the U.S. when first interviewed in 2010/2011, and who, at that time, did not have a spouse or partner who was born before 1954, along with their spouses or partners at the time of the initial interview or at the time of any subsequent interview. The Middle Baby Boomer sub-sample was interviewed in 2010/2011 and every two years thereafter.

Both the EBB and MBB sub-samples were supplemented in the 2010 wave with a sample of individuals residing in areas with 10% or higher concentrations of Black and/or Hispanic populations in order to boost the size of the minority samples in those cohorts.

The Late Baby Boomer (LBB) sub-sample consists of people who were born in 1960-1965, were household residents of the U.S. when first interviewed in 2016, and who, at that time, did not have a spouse or partner who was born before 1959, along with their spouses or partners at the time of the initial interview or at the time of any subsequent interview. The Late Baby Boomer sub-sample was interviewed in 2016 and will be interviewed every two years thereafter.

Original sample members are those selected as described above and their spouses or partners at the time of the initial interview in 1992 (HRS), 1993 (AHEAD), 1998 (CODA or WB), 2004 (EBB), 2010/2011 (MBB) or 2016 (LBB). For more details about the sample, see our [Web site](#).

1B. 2016 Questionnaire Sections

2016 Section	Content
PR	Preload
A	Coverscreen
B	Demographics
C	Physical Health
D	Cognition
E	Family Structure
F	Parents, Siblings and Transfers
G	Functional Limitations and Helpers
H	Housing
I	Physical Measures and Biomarkers
J	Employment
J2	Pensions
J3	Retirement and Social Security
M	Disability
N	Health Services and Insurance
P	Expectations
Q	Assets and Income
R	Asset Change
S	Widowhood and Divorce
T	Wills and Life Insurance
U	Asset Verification

V	Modules
W	Event History, Internet Use and Social Security
Y	Time Calculations
IO	Interviewer Observations
LB	Leave-Behind Questionnaires
TN	Thumbnails

NOTE: As of the 2016 data collection wave, sections K and L have been incorporated into Section J. Some items from Section L have been incorporated into Section J3.

1C. Levels of Files

In the 2016 data collection instrument, most questions were asked of all respondents. Some questions were asked about the household. For two-respondent households, household level questions were asked of one respondent who was designated as the financial respondent, family respondent, or coverscreen respondent (the first respondent interviewed) on behalf of the entire household.

In addition to the familiar household-level and respondent-level files, the 2016 HRS Core (Early, Version 2.0), contains files at seven other levels: household-member-and-child, sibling, helper, transfer-to-child, transfer-from-child, jobs, and pension.

1C1. Household Level Files

Household-level files contain questions that were asked about the household of a designated coverscreen, financial, or family respondent. A coverscreen respondent answered family questions (section A) on behalf of the entire household; the coverscreen respondent may or may not be the family respondent. A family respondent answered family questions (section E) on behalf of the entire household, and a financial respondent answered household-level financial questions (sections H, Q, R and U) on behalf of the entire household. The household-level files contain one record for each household in which at least one interview was obtained in 2016.

1C2. Respondent Level Files

Respondent-level files contain questions that were asked of all respondents about themselves (or asked of a proxy about the respondent if the respondent was not able to give an interview). The files contain one record for each respondent or proxy who gave an interview in 2016.

1C3. Sibling Level File

The sibling-level file consists of characteristics of the respondent's siblings. If a respondent had at least one living parent, he/she was asked a variety of questions about his/her siblings (Section F). The sibling file contains one record for each sibling of a respondent. Each respondent reports on his/her own parents and siblings. Sibling data are also stored in the preload section, H16PR_SB.

1C4. Household Member and Child Level Files

Household-member-and-child-level-files contain characteristics about household members and children. This information can come from Section A. (coverscreen) or the family respondent in Section E and also in the preload section.

1C5. Helper Level File

The helper-level file contains information provided by each respondent about helpers. A helper may be a person or organization that was reported by the respondent as providing help with Activities of Daily Living (ADLs) or

Instrumental Activities of Daily Living (IADLs). The file contains one record for each helper. If a child helped both respondents in a two respondent household, the helper file will contain two records - one of mother's report of the child's helping her and one for father's report of the child's helping him.

1C6. Transfer-to-Child-Level File

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

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

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

1C8. Pension Level Files

Pension-level files contain information about pensions that the respondent has reported over the years during their involvement with the HRS. Each file contains one record for each pension that has been reported and followed up on in the current data collection.

1C9. Jobs level files

Job-level files contain information on jobs the respondent has reported over time. Current jobs are represented, as well as past employment with unresolved pensions from past waves. The files have one record for each job represented.

2. File Naming Conventions

Files are named beginning with "H16" for HRS 2016, followed by a letter (or two) designating the questionnaire section. A separator, "_" and then one or two letters designating the level follows the section letter designator.

- H for household-level
- R for respondent-level
- MC for household-member-and-child-level
- SB for sibling-level
- HP for helper-level
- TC for transfer-to-child-level
- FC for transfer-from-child-level
- JB for jobs-level
- P for pension-level

For example, the file H16A_R includes variables from Section A (coverscreen) at the respondent level, whereas, H16A_H contains variables from Section A (coverscreen) at the household level.

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

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

One of each of these file types is provided for each of the 40 data files for the 2016 HRS Core (Early, Version 2.0) data release. For example,

H16A_R.DA contains respondent data from Section A,
H16A_R.SAS contains corresponding SAS program statements,
H16A_R.SPS contains corresponding SPSS program statements,
H16A_R.DO contains corresponding STATA DO statements,
H16A_R.DCT contains corresponding STATA dictionary statements, and
H16A_R.TXT contains the ASCII codebook.

3. Data Files

The 2016 HRS Core (Early, Version 2.0) data are distributed in 40 data files. The files are listed below along with the number of cases (N), number of variables (NV), and the primary identifiers (IDS). The records in the data files are sorted in order by these primary identifiers.

The 2016 HRS Core data are provided in ASCII format, with fixed-length records. Use associated SAS, SPSS or STATA program statements to read the data into the analysis package of your choice.

Household level files, IDS=HHID OSUBHH

H16PR_H	N=15027	NV= 44
H16A_H	N=15027	NV= 47
H16E_H	N=15027	NV= 137
H16H_H	N=15027	NV= 196
H16Q_H	N=15027	NV= 679
H16R_H	N=15027	NV= 111
H16U_H	N=15027	NV= 246
H16IO_H	N=15027	NV= 56

Respondent level files, IDS=HHID PN

H16PR_R	N=20918	NV= 175
H16A_R	N=20918	NV= 36
H16B_R	N=20918	NV= 164
H16C_R	N=20918	NV= 240
H16D_R	N=20918	NV= 232
H16F_R	N=20918	NV= 186
H16G_R	N=20918	NV= 160
H16I_R	N=20918	NV= 209
H16IO_R	N=20918	NV= 72
H16J_R	N=20918	NV= 614
H16J3_R	N=20918	NV= 195
H16LB_R	N=20918	NV= 402
H16M1_R	N=20918	NV= 261
H16M2_R	N=20918	NV= 382
H16N_R	N=20918	NV= 345
H16P_R	N=20918	NV= 54
H16S_R	N=20918	NV= 114
H16T_R	N=20918	NV= 106
H16TN_R	N=20918	NV= 18
H16V_R	N=20918	NV= 316
H16W_R	N=20918	NV= 19
H16Y_R	N=20918	NV= 34

Household member and child level files, IDS=HHID OSUBHH OPN
H16PR_MC N=80508 NV= 26
H16E_MC N=66980 NV= 28

Sibling level files, IDS=HHID PN OPN
H16PR_SB N=14912 NV= 16
H16F_SB N=27405 NV= 24

Transfer-to-child-level-file, IDS=HHID OSUBHH
H16E_TC N= 7022 NV= 23

Transfer-from-child-level-file, IDS=HHID OSUBHH
H16E_FC N= 1185 NV= 23

Helper level file, IDS=HHID PN OPN
H16G_HP N= 6252 NV= 30

Jobs level file, IDS=HHID PN JOBID
H16PR_JB N=16460 NV= 16

Pension level files, IDS=HHID PN Pension ID
H16PR_P N= 8667 NV= 17
H16J2_P N=14996 NV= 190

At a later date, household and respondent level weight variables for the 2016 sample will be available in the HRS cross-wave Tracker file.

4. Identification Variables

Identification variables for HRS 2016 are stored in character format.

4A. Primary Identification Variables

Several variables, HHID, PSUBHH, PN, OPN, are used in various combinations to uniquely identify the nine different level datasets that comprise this data release.

4A1. HHID - Household Identification Number

In the initial wave of data collection (in 1992 for the HRS sub-sample, in 1993 for the AHEAD sub-sample, 1998 for the WB and CODA sub-samples, 2004 for the EBB sub-sample, 2010/2011 for the MBB sub-sample and 2016 for the LBB sub-sample), each sample household was assigned a Household Identifier. HHID is stable across waves of data collection and uniquely identifies the original household and any households derived from that household in subsequent waves of data collection. HHID has six-digits.

4A2. PSUBHH - 2016 Sub-household Identifier

In combination with HHID, PSUBHH uniquely identifies a household at the time of the 2016 data collection. Sub-household identifiers can be different at each wave. PSUBHH has one-digit. For more information, see Examples of [Sub-Household and Respondent Person Number and Other Person Number Assignments](#).

4A3. PN - Person Number

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

4A4. OPN – Other Person Number.

In the 2016 data collection HHID, PSUBHH and OPN uniquely identify another person in the household member and child files; HHID, PN, and OPN uniquely identify another person in the helper or sibling files. OPN has three-digits.

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

Two identifiers uniquely identify records in the

- o respondent-level datasets:
 - 1) HHID HOUSEHOLD IDENTIFICATION NUMBER
 - 2) PN RESPONDENT PERSON IDENTIFICATION NUMBER

Two identifiers uniquely identify records in the

- o household-level datasets:
 - 1) HHID HOUSEHOLD IDENTIFICATION NUMBER
 - 2) PSUBHH 2016 SUB HOUSEHOLD IDENTIFICATION NUMBER

Three identifiers uniquely identify records in the

- o helper-level datasets:
 - 1) HHID HOUSEHOLD IDENTIFICATION NUMBER
 - 2) PN RESPONDENT PERSON IDENTIFICATION NUMBER
 - 3) OPN OTHER PERSON NUMBER

Three identifiers uniquely identify records in the

- o household-member-and-child-level:
 - 1) HHID HOUSEHOLD IDENTIFICATION NUMBER
 - 2) PSUBHH 2016 SUB HOUSEHOLD IDENTIFICATION NUMBER
 - 3) OPN OTHER PERSON NUMBER

Three identifiers uniquely identify records in the

- o sibling-level datasets:
 - 1) HHID HOUSEHOLD IDENTIFICATION NUMBER
 - 2) PN RESPONDENT PERSON IDENTIFICATION NUMBER
 - 3) OPN OTHER PERSON NUMBER

Three identifiers uniquely identify records in the

- o transfer-to-child-level, and
- o transfer-from-child-level datasets:
 - 1) HHID HOUSEHOLD IDENTIFICATION NUMBER
 - 2) PSUBHH 2016 SUB HOUSEHOLD IDENTIFICATION NUMBER
 - 3) OPN OTHER PERSON NUMBER

Three identifiers uniquely identify records in the

- o Pension Level Data sets
 - 1) HHID HOUSEHOLD IDENTIFICATION NUMBER
 - 2) PN RESPONDENT PERSON IDENTIFICATION NUMBER
 - 3) Z505 PENSION TRACKING NUMBER

Three identifiers uniquely identify records in the

- o Jobs level datasets:
 - 1) HHID HOUSEHOLD IDENTIFICATION NUMBER
 - 2) PN RESPONDENT PERSON IDENTIFICATION NUMBER
 - 3) Z504 JOB TRACKING NUMBER

4C. Secondary Identification Variables

In addition to the primary identification variables that uniquely identify records in a dataset, secondary identification variables that allow links to

other datasets are provided. Two sets of secondary identification variables are provided. They are listed below.

Secondary Identification Variables for

- o respondent-level datasets
- o sibling level datasets
- o helper-level datasets
- o job-level datasets and
- o Pension-level datasets

PSUBHH	2016	SUB	HOUSEHOLD	IDENTIFICATION	NUMBER
OSUBHH	2014	SUB	HOUSEHOLD	IDENTIFICATION	NUMBER
PPN_SP	2016	SPOUSE/PARTNER	PERSON	NUMBER	
PCSR	2016	WHETHER	COVERSHEET	RESPONDENT	
PFAMR	2016	WHETHER	FAMILY	RESPONDENT	
PFINR	2016	WHETHER	FINANCIAL	RESPONDENT	

Secondary Identification Variables for

- o household level datasets
- o household member or child level datasets
- o transfer-from-child-level datasets and
- o transfer-to-child-level datasets

PSUBHH	2016	SUB	HOUSEHOLD	IDENTIFICATION	NUMBER		
OSUBHH	2014	SUB	HOUSEHOLD	IDENTIFICATION	NUMBER		
OPN_CS		PERSON	NUMBER	OF	COVER	SCREEN	RESPONDENT
PPN_FAM	2016	FAMILY	RESP	PERSON	NUMBER		
PPN_FIN	2016	FINANCIAL	RESP	PERSON	NUMBER		
PPN_NCS	2016	NON-COVERSCREEN	RESP	PERSON	NUMBER		
PPN_NFAM	2016	NON-FAMILY	RESP	PERSON	NUMBER		
PPN_NFIN	2016	NON-FINANCIAL	RESP	PERSON	NUMBER		

4D. Datasets Including PSUBHH and OPN as Primary Identifiers

When working with datasets including PSUBHH and OPN as primary identifiers -- household-member-and-child-level, it is essential to use PSUBHH to distinguish other persons.

Sometimes records with the same HHID and OPN will be separate reports about the same person. For example, [Couple with Children and Siblings Divorces](#), the record where HHID is 890121, PSUBHH is 1, and OPN is 101 is the report about the child from the mother, and the record where HHID is 890121, PSUBHH is 2, and OPN is 101 is the report about the same child from the father.

At other times records with the same HHID and OPN will be reports about different persons. For example, a couple divorces, one respondent remarries, and both split-off households have new members. When a couple divorces and both split-off households have new members, the record where HHID is 871231, PSUBHH is 2, and OPN is 151 is for the ex-wife's mother, while the record where HHID is 871231, PSUBHH is 1, and OPN is 151 is for the ex-husband's new stepchild.

5. Distribution Files and Directory Structure

5A. Distribution Files

The files are packaged for download from our Web site in two different ways - as one large .zip file that contains six smaller .zip files, one .pdf file, and one .txt file, or the eight smaller files available individually for separate download. The combined file is H16core.zip.

The individual .zip files for separate download are:

Data file

H16da.zip contains data files.

Program statement files

H16sas.zip contains SAS program statements.

H16sps.zip contains SPSS program statements.

H16sta.zip contains STATA program statements.

Documentation files

H16cb.zip contains the codebook.

H16qn.zip contains the questionnaire.

H16dd.pdf - this document.

5B. Directory Structure

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

Directory	Contents
c:\hrs2016	Files downloaded from Web site
c:\hrs2016\codebook	Unzipped files from H16cb.zip
c:\hrs2016\data	Unzipped files from H16da.zip
c:\hrs2016\qnaire	Unzipped files from H16qn.zip
c:\hrs2016\sas	Unzipped files from H16sas.zip
c:\hrs2016\spss	Unzipped files from H16sps.zip
c:\hrs2016\stata	Unzipped files from H16sta.zip

Decompress the selected .zip files into the appropriate subdirectories. You will need about 450 MB of free space on your storage device to store the 40 .DA files.

6. Program Statements

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

The statement files are named beginning with the same prefix as the corresponding data file. For example, SAS statements in the file H16A_R.SAS go with the H16A_R.DA data file.

6A. Using the Files with SAS

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

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

If the *.SAS file is located in "c:\hrs2016\sas" and the data file is located in "c:\hrs2016\data", you can run the file as is. A SAS system file (*.SD2 or *.SAS7BDAT) will be saved to directory "c:\hrs2016\sas".

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

6B. Using the Files with SPSS

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

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

If the *.SPS file is located in "c:\hrs2016\spss" and the data file is located in "c:\hrs2016\data", you can run the file as is. An SPSS system file (*.SAV) will be saved to directory "c:\hrs2016\spss".

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

6C. Using the Files with STATA

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

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

If the *.DO and *.DCT files are located in "c:\hrs2016\STATA" and the data file is located in "c:\hrs2016\data", you can run the .DO file as is.

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

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

7. Loading HRS Data Products in a Non-Windows Environment

All files released by The Health and Retirement Study are created in a Windows environment. This means that non-Microsoft users will need to modify the default Windows file structure syntax to match that of their own operating system. The instructions in this section provide advice on how to make these changes.

7A. Working in an OS X Environment

The following examples should work for both Macintosh OS X and any Unix/Linux distribution. Open the SAS program file(s), SPSS syntax file(s) or the Stata do/dct files in an ASCII editor and make the changes indicated below.

Important items for Mac OS X users to consider:

1. Use **Stuffit** instead of **Archive Utility** to decompress the downloaded zip file into a desktop folder.
2. When using the Apple ASCII text editor (**TextEdit**), select *Preferences* and make these changes...
 - Under Format, select *Plain text*

- Under Options, uncheck
 - *Smart quotes*
 - *Smart dashes*
- Better yet, install and use **TextWrangler**

7A1. OS X: SPSS

In this example, we assume that the user has downloaded the 2014 tracker dataset and placed the files in a **Desktop** folder called **trk2014** with the ASCII data file (.da) stored in subfolder **data** and the syntax file (.sps) in subfolder **spss**. The commands in the syntax file(s) should be modified to look like this:

```
FILE HANDLE trk2014TR_R /name='~/Desktop/trk2014/data/TRK2014TR_R.da'
LRECL=221.
DATA LIST FILE= trk2014TR_R /
HHID 1-6(A)
[rest of syntax file goes here]
.
execute.
SAVE /outfile '~/Desktop/trk2014/spss/TRK2014TR_R.sav'.
Execute.
```

7A2. OS X: Stata

In the following example we assume that the user has downloaded the 2014 tracker dataset and placed the files in a **Desktop** folder called **trk2014** with the ASCII data file (.da) stored in subfolder **data** and the syntax files (.do/.dct) in subfolder **stata**. Next step is to modify the .do and .dct files in the stata folder:

File TRK2014TR_R.do:

```
Change from...
    infile using c:\trk2014\stata\TRK2014TR_R.dct
To...
    infile using ~/Desktop/trk2014/stata/TRK2014TR_R.dct
```

```
Change from...
    save c:\trk2014\stata\TRK2014TR_R.dta
To...
    Save ~/Desktop/trk2014/stata/TRK2015TR_R.dta
```

File TRK2014TR_R.dct:

```
Change from...
dictionary using c:\trk2014\data\TRK2014TR_R.da {
column(1) str6 HHID %6s "HOUSEHOLD IDENTIFIER"
column(7) str3 PN %3s "PERSON NUMBER"

    [additional dictionary statements]
}
To...
dictionary using ~/Desktop/trk2014/data/TRK2014TR_R.da {

column(1) str6 HHID %6s "HOUSEHOLD IDENTIFIER"
column(7) str3 PN %3s "PERSON NUMBER"

    [additional dictionary statements]
```

7B. Working in a Linux Environment

7B1. Linux: Stata

In the following system file build example we assume that the user (user1) has downloaded the 2014 tracker dataset and placed the files in a folder called **trk2014** (/home/user1/trk2014), with the ASCII data file (.da) stored in subfolder **data** (/home/user1/trk2014/data) and the syntax files (.do/.dct) in subfolder **stata** (/home/user1/trk2014/stata).

This process is very similar to the OS X example shown above. The main difference is that the file/folder structure will be somewhat different when he .do and .dct files in the stata folder are modified.

File TRK2014TR_R.do:

Change from...

```
infile using c:\trk2014\stata\trk2014tr_r.dct
```

To...

```
infile using /home/user1/trk2014/stata/trk2014tr_r.dct
```

Change...

```
save c:\trk2014\stata\TRK2014TR_R.dta
```

To...

```
Save /home/user1/trk2014/stata/trk2014tr_r.dta
```

File TRK2014TR_R.dct:

Change from...

```
dictionary using c:\trk2014\data\trk2014tr_r.da {  
column(1) str6 HHID %6s "HOUSEHOLD IDENTIFIER"  
column(7) str3 PN %3s "PERSON NUMBER"
```

```
    [additional dictionary statements]
```

```
}
```

To...

```
dictionary using /home/user1/trk2014/data/trk2014tr_r.da {
```

```
column(1) str6 HHID %6s "HOUSEHOLD IDENTIFIER"  
column(7) str3 PN %3s "PERSON NUMBER"
```

```
    [additional dictionary statements]
```

```
}
```

7B2. Linux: SAS

In the following examples we assume that the user (user1) is working in a folder named **saswork** (/home/user1/saswork), and is creating a SAS version of HRS products in a separate sub-folders (e.g.m, /home/user1/saswork/2014/sas).

Building SAS system files (One Section)

In the following example we assume that the user has downloaded the 2014 tracker dataset and placed the files in a folder called **trk2014** (/home/user1/saswork/trk2014) with the ASCII data file (.da) stored in subfolder **data** (/home/user1/saswork/trk2014) and the syntax file (.sas) in subfolder **sas** (/home/user1/saswork/trk2014/sas). Next step is to modify trk2014tr_r.sas to match the folder structure:

In file trk2014tr_r.sas:

Change from...

```
LIBNAME trk2014 'c:\trk2014\sas\';  
DATA trk2014.trk2014tr_r;  
INFILE 'c:\trk2014\data\trk2014tr_r.da' LRECL = 1028;
```

To...

```
LIBNAME trk2014 '/home/user1/saswork/trk2014/sas/';  
DATA trk2014.trk2014tr_r;  
INFILE '/home/user1/saswork/trk2014/trk2014tr_r.da' LRECL = 1028;
```

When the SAS file is run, it will create file trk2014tr_r.sas7bdat in folder /home/user1/saswork/trk2014/sas/.

Building SAS system files (Multiple Sections)

In the following example we assume that the user has downloaded the HRS 2016 Core dataset and placed the files in a folder called hrs2016 (/home/user1/saswork/hrs2016) with the ASCII data files (H16{sec}_{lvl}.da) stored in subfolder **data** (/home/user1/saswork/hrs2016/data) and the syntax files (H16{sec}_{lvl}.sas) in subfolder **sas** (/home/user1/saswork/hrs2016/sas).

Important: Begin by modifying each syntax file as follows:

Change all occurrences of...

```
LIBNAME H2016 'c:\hrs2016\sas\';
```

To...

```
LIBNAME H2016 '/home/user1/saswork/hrs2016/sas/';
```

Change the INFILE statement in each .sas file from...

```
INFILE 'c:\hrs2016\data\H16{sec}_{lvl}.da' LRECL = nnn;
```

To...

```
INFILE '/home/user1/saswork/hrs2016/data/H16{sec}_{lvl}.da' LRECL=nnn;
```

Example

```
INFILE '/home/user1/saswork/hrs2016/data/H16A_H.da' LRECL=87;
```

8. Documentation

There are several types of documentation available for use with the 2016 HRS Core (Early, Version 2.0) data release. These include a codebook and the 2016 box-and-arrow questionnaire. In addition, 2016 variables have been added to the [Online Concordance](#).

8A. Codebook

The HRS 2016 Codebook is provided as a series of 40 ASCII text files, as well as a file containing all sections. There is a codebook file corresponding to each data file. Each variable has its own codebook entry. The format of the codebook is, for the most part, consistent with all previous releases.

8A1. Variable Names

Variable names begin with a letter designating the wave of data collection (P for 2016), followed by the section letter, and numbers after the section letter. For example, PC001 where P=2016, C=section C (physical health), and 001 is the variable number. Variables from the preload section of the instrument will have either PX or PZ as prefix letters. The X indicates a variable that is updated by data collected in later sections of the questionnaire, whereas the Z indicates preloaded data that were not changed by subsequent answers to questions.

For example:

```
PX007_R    RESP FAM/FIN TYPE - UPDATED
PZ077_R    PREV WAVE R FIN/FAM TYPE
```

PX007_R indicates that the preloaded Financial or Family respondent was changed, or updated, later in the instrument from what had been preloaded (or assigned) prior to the start of the 2016 interview.

8A1a. Multiple-response and Looped Variables

There are two types of variables with multiple mention indicators. First are simple multiple mentions and second are multiple mentions within loops.

Simple multiple mention variables take the form: (wave prefix) + (section letter) + (variable number) + (mention number). For example, PN219M1 through PN219M5 are 2016 variables from section N with one through 5 mentions.

Variable names for multiple mentions to questions within a loop take the form: (wave prefix) + (section letter) + (variable number) + (underscore) + (loop iteration) + (letter designating mention number). For example, PN049_1M1, is a 2016 variable from section N, variable number 049 in the first iteration of the loop, and the first mention.

Simple loop variables (not a multiple mention) have an underscore (_) in their name and a suffix that designates the loop, e.g., PN025_1.

For variables that have a "W" right after the section designator, the variable names are slightly different. Variable names for multiple mentions to questions within a W-loop take the form: (wave prefix) + (section letter) + ("W") + (variable number) + (letter designating loop iteration) + (mention number). For example, PKW097A1, is a 2016 variable from section K, variable number 097 in the first iteration of the loop, and the first mention. Other non-multiple mention variables within this type of loop are named with the letter designating the loop iteration. For example, PKW002A, is variable number 002 in the first iteration of the loop.

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

8A1b. Masked Variables

To protect the confidentiality of the information that respondents provide, a number of variables have been masked or are simply not included in the Early release public dataset. Some of these variables may be made available to analysts as restricted data. See our Web site for details.

Names, addresses, days of birth, information on geographical relocation and similar variables are not included in publicly released files.

Geographical locations are recoded to a level no more detailed than U.S. Census Region and Division. Data on the highest educational degree earned have been further grouped together to increase cell sizes. The 2010 U.S. Census codes have been used to code the occupation and industry data in 2016.

The names of variables that were masked for confidentiality end in the capital letter "M"; for example, variable PX026M (1ST ADDRESS STATE - MASKED) and variable PB024M (FATHER USUAL OCCUPATION - MASKED).

8A2. Other Types of Documentation

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

8A2a. Overview of HRS Public Data Files for Cross-sectional and Longitudinal Analysis

This is a thorough document that provides a description of structure and content for all HRS public data files and a detailed explanation for handling the longitudinal nature of the survey. It is located here:

<http://hrsonline.isr.umich.edu/sitedocs/dmgt/OverviewofHRSPublicData.pdf>

8A2b. Box and Arrow Questionnaire

The research community has referred to the type of documentation that describes the questions asked in the interview as a "questionnaire". Since the 2016 HRS data were collected using a CAI program, a traditional hard-copy questionnaire was not produced as part of the data collection phase. Therefore, in order to document question sequencing, we have provided a current-wave version of the traditional box and arrow questionnaire.

8A2c. Cross-Wave Tracker File

The cross-wave tracker file contains basic demographic information, interview status, and if, when and how an interview was conducted during a specific data collection period. The cross-wave tracker file also contains weight variables. The tracker file contains one record for every person who was ever eligible to be interviewed in any wave. Please see the cross-wave tracker data description for a more detailed description.

8B. Master Codes

A master code file contains detailed codeframes used in several sections of the codebook. The master codes include health conditions, occupation codes, industry codes, and state and country codeframes. The master codes appear in [Appendix B](#).

9. Additional Notes

The following are miscellaneous additional notes regarding the 2016 HRS Core (Early, Version 2.0). If we become aware of additional issues, they will be posted on our Web site in the [Data Alerts](#) section.

9A. Households with No Family or Financial Respondents

As noted earlier in this document, the data collection design was to have asked most questions of all respondents and some questions of just a designated coverscreen, or family, or financial respondent on behalf of the household. However, occasionally that is not what happened. For some households we did not obtain an interview from a family or financial respondent. There were 461 households that had no family respondent; 256 households had no financial respondent. There are 4 missing coverscreen respondents in this wave. The household records for these households contain null values for the missing information. Households missing a family or financial respondent can be identified, respectively, by values of "Blank. No family/financial respondent" (in the household record) in the following variables:

PPN_FAM - 2016 FAMILY RESP PERSON NUMBER

PPN_FIN - 2016 FINANCIAL RESP PERSON NUMBER

9B. Unfolding Bracket Variables and Imputations

Typically, a series of unfolding bracket questions followed a lead-in question asking for an amount. If an actual amount was not given, a series of "unfolding" questions were asked. The manner in which the unfolding questions were programmed (Blaise) is different for the 2002 through 2010 data compared to the CAI (SurveyCraft) software used for 1993 through 2000. This change was transparent to the respondents, since exactly the same questions were asked with the new software as would have been asked with the old software; but it did have an implication for the data that were actually stored and also for the data that are released.

Instead of storing the response to each unfolding question, three summary variables were generated: the minimum and maximum values for the amount, given the answers to the unfolding questions, and if the last answer a respondent gave in an unfolding sequence was either "Don't Know" or "Refused," what that answer was. In 2002, if the Respondent said "more than" to the unfolding question with the highest value, then the maximum value was stored as ten times that value. However, in 2004 and 2010, if the Respondent said "more than" to the unfolding question with the highest value, then the maximum value was stored as 99999996.

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

For other sequences -- those in which respondents were randomly assigned to one of three "entry" points for the first unfolding question -- the analyst will also need to take into account a fourth variable (located in the preload sections) that specifies the entry point for each respondent. The following example shows the preload variable (PZ041) and the unfolding sequence that uses the random entry point from PZ041.

Example Random Entry Assignment Variable from Preload:

```

Preload Variable from the data file H16pr_h:
.....
PZ041          PREASSIGNED UNFOLD RANDOM VALUE - SELF EMPLOYMENT
Section: PR    Level: Household      Type: Numeric   Width: 1   Decimals: 0
Ref: HH.X041_UnfSEmpInc_V

UNFOLD ASSIGN - SELF EMPLOYMENT INCOME
.....

3897          1.  RANDOM ASSIGNMENT 1
3871          2.  RANDOM ASSIGNMENT 2
3895          3.  RANDOM ASSIGNMENT 3
7            Blank. Data Missing
.....

```

.....
Unfolding Series from section Q that uses OZ041 to assign respondents an entry point:
.....

PQ016 R INCOME FROM SELF EMPLOYMENT - MIN
 Section: Q Level: Household Type: Numeric Width: 6 Decimals: 0
 Ref: SecQ.RIncome.Q016_

Did it amount to less than \$____ , more than \$____ , or what?

PROCEDURES: 3Up, 2Up1Down, 1Up2Down
 BREAKPOINTS: \$5,000, \$10,000, \$25,000, \$100,000
 RANDOM ENTRY POINT ASSIGNMENT [1 (\$5,000)] or [2 (\$10,000)] or
 [{NOT 1 and NOT 2} (\$25,000)] AT X041

.....
 124 0. Value of Breakpoint
 4 5000. Value of Breakpoint
 16 5001. Value of Breakpoint
 6 10000. Value of Breakpoint
 20 10001. Value of Breakpoint
 1 25000. Value of Breakpoint
 17 25001. Value of Breakpoint
 1 100000. Value of Breakpoint
 5 100001. Value of Breakpoint
 11476 Blank. INAP (Inapplicable); Partial Interview

.....
 PQ017 R INCOME FROM SELF EMPLOYMENT - MAX
 Section: Q Level: Household Type: Numeric Width: 8 Decimals: 0
 Ref: SecQ.RIncome.Q017_

*

.....
 38 4999. Value of Breakpoint
 4 5000. Value of Breakpoint
 16 9999. Value of Breakpoint
 6 10000. Value of Breakpoint
 18 24999. Value of Breakpoint
 1 25000. Value of Breakpoint
 15 99999. Value of Breakpoint
 1 100000. Value of Breakpoint
 95 99999996. Greater than Maximum Breakpoint
 11476 Blank. INAP (Inapplicable); Partial Interview

.....
 PQ018 R INCOME FROM SELF EMPLOYMENT - RESULT
 Section: Q Level: Household Type: Numeric Width: 2 Decimals: 0
 Ref: SecQ.RIncome.Q018_

*

.....
 24 98. DK (Don't Know); NA (Not Ascertained)
 70 99. RF (Refused)
 11576 Blank. INAP (Inapplicable); Partial Interview

9C. Modules for the 2016 Data Collection

There are 10 modules for 2016 HRS. The module data can be found in the file H16V_R. Topical areas of the modules include:

Module 1: Symptom Prevalence of Adult Attention Deficit Hyperactivity Disorder (ADHD)
Module 2: Financial Mismanagement at Older Ages
Module 3: Financial Advice and Financial Capacity at Older Ages
Module 4: Material Hardship
Module 5: Subjective Conditional Probabilities of Working (R's < age 65)
Module 6: Long-term Care Expectations
Module 7: Medication Non-Adherence
Module 8: Religious Life Histories
Module 9: Affordable Care Act
Module 10: HRS Coordinated Care

9D. Explanation of the Difference between Tracker and Core Data Releases

There are some areas where the Tracker File and the core data are different. Below are some examples. The examples rely on data from the HRS 2008 Core.

LSUBHH - The tracker file assigns a '9' for all new spouses in 2008 for the previous wave's SUBHH number. The 2008 Final Release data, however, lists the SUBHH number of the household that the new spouse married into, since the new spouse will 'inherit' many of the relationships from that household.

LPN_SP - Even if a spouse did not give an interview in 2008, a spouse PN (LPN_SP) will exist in the core data. Additionally, for a non-original R's new spouse, we assign a spouse PN (LPN_SP) and track the spouse's influence on the household with that number, even though we will never interview that person. In contrast, the tracker file will not reference these spouses.

Family and Financial R Assignments - As we keep a record of a non-interviewed spouse in the core data, we also release their assigned role in the household (i.e. family or financial R) with a listing for LPN_FIN, and LPN_FAM in the household sections. The tracker file will simply list the assignments of the people interviewed. Therefore, for example, the 2008 Core Final data has a blank line in Section H for a household where '020' was the non-interviewed financial R. The tracker file would show the same household as not having a financial R.

Marital Status - It is important to note that marital status, as it was assigned in the 2008 data (LB063) may be different from the marital status variable (xMARST) in the Tracker file. See the [Cross-wave Tracker File Data Description](#) for a more complete explanation of how marital status was assigned and the criteria used to reclassify respondents. The marital status variable in the Tracker file was constructed by looking at several different sources of data, whereas the marital status variable in the 2008 Core Final data was not changed using outside sources, in order to preserve the flow through the instrument for respondents.

9E. Enhanced Face-to-Face Sample Design and Content

In 2006, HRS initiated what is referred to as an Enhanced Face-to-Face Interview. In addition to the core interview, the Enhanced Face-to-Face Interview includes a set of physical performance measures, collection of biomarkers, and a Leave-Behind Questionnaire on psychosocial topics. A random one-half of households were pre-selected for the enhanced face-to-face interview

in 2006, with the other half of the sample selected for 2008, and the design is repeated in each subsequent wave. In coupled households, both members of the couple are selected. Selected respondents who completed a self (as opposed to proxy) interview, in-person (at least through Section I - the physical measures and biomarkers section), and who were non-institutionalized at the time of the interview were eligible for the physical measures and biomarkers components. This same group of respondents, plus those for whom an in-person proxy interview was completed, were also eligible for the Leave-Behind Questionnaire.

Psychosocial Leave-Behind Questionnaire (Section LB_R)

In 2004, HRS added a new feature for data collection in the form of self-administered questionnaires that were left with respondents upon the completion of an in-person Core Interview, referred to as the Leave-Behind Questionnaire. The purpose of the Leave-Behind Questionnaire is to collect additional information from respondents without adding to the interview length. In 2016, as in previous data collections since 2006, the Leave-Behind Questionnaire was incorporated into the Enhanced Face-to-Face Interview. The Leave-Behind Questionnaire administered in 2010 is called the Participant Lifestyle Questionnaire, and it includes questions on participation in general activities, relationships with others, and views on their life in general as well as specific aspects of their life.

Physical Performance Measures and Biomarkers (Section I_R)

Blood pressure
Breathing test (peak flow)
Grip strength
Timed walk (8 ft.)
Balance tests (semi-tandem, side-by-side, full tandem)
Height
Weight
Waist circumference
Saliva (for which DNA was extracted and stored)
Dry blood spots (analyzed for Hemoglobin Alc, total cholesterol and HDL cholesterol)

With the exception of the timed walk, which was administered only to respondents 65 years of age or older, the physical measures and biomarkers were conducted on the full enhanced face-to-face sample.

Three separate consents were obtained to cover: 1) all of the physical measures, plus blood pressure; 2) saliva sample; and 3) dry blood spots. The physical measures booklet, which the interviewers used to administer the physical measures and biomarkers, is available on the HRS website [Physical Measures 2016](#). The booklet contains the consent forms, as well as instructions and protocols for all of the measurements.

9F. Summary of Data Model (DM) Changes during 2016 Production

--

Data Model 2 Changes:

Section N - Health Insurance

N291_Placeofcare: Problem: The response options aren't read to the R and there's no on screen instruction if the R just keeps answering "yes". Fix: Reverted back to the 2014 codeframe.

--

Data Model 3 Changes:

Section A - Coverscreen

A071_SameMainRes: Problem: Per Co-Pi, "we do not want to count a move between first and second homes as a move in Section A for new cohort Rs." Fix: Restricted the assignment of X033 = 1 to reinterview HH only.

Section A2

X072AMoChgResStat: Problem: Daughter who was preloaded as PW no contact (X056=6 and X057=6) moved in but the fills are now missing. Fix: Added no contact to the code. Fills now are working.

Section B - Demographics

B134_: Problem: New Cohort Rs are the only Rs who should be asked the new 'orientation' questions B134 and B135. We should be using X024 instead of Z076 to help control flow at the B133BP so that New Spouses of reinterview Rs do not get asked. Fix: Changed to use X024 not 1 (new cohort).

Section H - Housing

H066_: Problem: Analyst indicates that a max value of 9,999,999 is too low a threshold for housing values. Please change the fixed length/max value of H066 to match H020. Fix: The max value will now equal 99,999,999.

H100_HousingAssoc: Problem: This is a New Cohort case and the flow through H086BP seems incorrect. Since I'm a new R and I don't own or rent (H014=7) I shouldn't jump to H100 as the instrument is currently doing. Fix: Adjusted the code to correct the flow.

Section N - Health Insurance and Service Utilization

N260_, N261_ and N262_: Problem: Asks for main reason but the response options allow for multiple responses." This was supposed to have been changed in 2012 but we missed it. Fix: Per Co-Pi: Changed the code to single entry instead of multi.

Section P - Expectations

P168_: Problem: During data review, analyst found that the rules for P170 were not correct and worked for only ½ of cases. He sent suggested code fix problem. Fix: Added analyst's suggested code to correctly create fill (P170) used in P168. Also Changed code to correct the gain/fallen fill at P168 when P166 = 0% or 100% or DK/RF.

P196: Problem: For analysis purposes, analyst has asked if we could save the actual fill text (more/less) for FLP166 as a permanent data value. This is important because the procedure used to create it is a little complicated (uses X083 and a mod function to assign values). Fix: Added new field P196 (based on FLP166), it is a string containing the word actually read to the respondent "more" or "less".

P197: Problem: For analysis purposes, analyst has asked if we could save the actual fill text (fallen/gained) for P170 as a permanent data value. Fix: Added new field P197 (based on P170 calculation), it is a string containing the word actually read to the respondent "fallen" or "gained".

Section V - Experimental Modules

V217: Problem: New wealth fields needed to keep track of the fills.

Fix: New wealth fields are added to keep values of presented fills:

V217 Health fill 1, used at: V203, V208

V218 Health fill 2, used at: V204, V209

V219 Health fill 3, used at: V205, V210

V220 Health fill 4, used at: V206, V211

V221 Health fill 5, used at: V207, V212

--

Data Model 4 Changes:

Section V3 - Financial Advice

V129: Problem: Change the wording to V129 to read: "How much do you trust financial advisors to provide you with useful information about your money decisions? Would you say that you trust them very much, somewhat, not very much, or not at all?"

V129: Problem: Add the words "bankers or other professional" so it reads: "How much do you trust bankers or other professional financial advisors to provide you" Fix: Added "bankers or other professional"

Section V8 - Religious Life Histories

V351_RELIGRAISED: Problem: Add an interviewer instruction BEFORE the one that is there that reads: "If R says Baptist, Methodist, Lutheran, Presbyterian, Pentecostal, Episcopal or Anglican, code as 1. Protestant. Fix: Added "If R says Baptist, Methodist, Lutheran, Presbyterian, Pentecostal, Episcopal or Anglican, code as 1. Protestant."

V351_RELIGRAISED: Problem: This is a field that is coded and is not Multiple Mention. 1.) Should this be MM or should the "Enter all that apply." be removed? 2.) if it is a MM should they be able to select 7 None and something else? Fix: Changed instruction to read: "If R mentions more than one religion, record the one that was mentioned first."

--

Data Model 5 Changes:

Section E - Children and Transfers

E105_FinHelpToRel: Correct the wording for E105 from: "..... did you give financial help... TO relatives such as siblings or nieces and nephews. Please do not count the help FROM children, grandchildren ..." To read: "..... did you give financial help... TO relatives such as siblings or nieces and nephews. Please do not count the help TO children, grandchildren ..."

Section N - Health Insurance and Service Utilization

N092_EmplHlthIns: Correct two flow problems: Rs who are covered by employer insurance were being asked N092 and N093 (does employer offer insurance/did they offer to R) Rs who were covered by employer insurance in previous wave were skipping N094 in error.

Section J2 - Pensions

W408_WhichProvider: W452 isn't a fully reliable field to use when determining whether a plan is provided by a current employer. Implement correction: Created a new variable which stores the original index used in W408 based on preload. If W408 and W790 don't match, the last condition is no longer valid,

which solves the problem when 97 is selected and also the possibility that they switch link from one preloaded job to another. Here's the revised condition:

```
IF piW408_WhichProvider = 11
OR (piW408_WhichProvider = 96 AND piJ021K007L007 = SLFEMPD)
OR (piZ503_JobStatus = pwjob AND piW417_SameJobPWandCW = yes AND
piW408_WhichProvider = piCurrentEmployer) THEN
W452_PensionFromNewJobOrSameJobAsPW := yes
ELSE
W452_PensionFromNewJobOrSameJobAsPW := no
ENDIF
```

Section U - Asset Verification

U013_globale1: We made a change to the global eligibility for getting into section U that used hard coded >=2014. We should have updated before going out in 2016, but did not. Please fix code for this wave replacing 2014 with 2016.

Section V8 - Religious Life Histories

V351: This question was changed from a multiple mention. After DM5 made, Karl could not migrate the data from DM4 so I copied V351, added a field V370 and made that the set. V351 will need to be merged into V371 post process.

--

Data Model 6 Changes:

Section J2 - Pensions

J338A: Problem: Spanish wording instructed interviewers to code 401K plans as defined benefits plans when they should be coded as defined contribution plans. Fix: Corrected Spanish wording to "contribuciones definidas" instead of "beneficios definidas."

Section V8 - Religious Life Histories

V351_PREFERENCE (and V353_DENOMINATION): Problem: If R named a Protestant denomination at V351, interviewers were instructed to code as Protestant, but then still required to ask for denomination at V353. Fix: V351 Iwer instruction changed to read: "If R says Baptist, Methodist, Lutheran, Presbyterian, Pentecostal, Episcopal or Anglican, code as 1. Protestant and enter the denomination on the next screen."

Section V9 - Affordable Care Act

V401_LASTVOTEYR: Problem: Adjust module text to reflect the start of 2016 general election voting. Fix: Added 2016 to Code 1, and removed interviewer instructions stating that general election voting had not yet occurred in 2016.

V402_WHOVOTEFOR: Problem: After nominating conventions, add names of presidential candidates to codeframe. Fix: Added name of 2016 Democratic candidate (Hillary Clinton) to Code15. Added name of 2016 Republican candidate (Donald Trump) to code 16.

--

Data Model 7 Changes:

Section V9 - Affordable Care Act

V402: Question deleted (WHO VOTED FOR LAST TIME R VOTED)

V403: Question deleted (WAS R NOT REGISTERED OR ELIGIBLE)

V404: Question deleted (IF WOULD HAVE VOTED DEM OR REP)

--

Data Model 8 changes:

Section D - Cognition

D157_: To accommodate the election results we need to accept either the current president or the president elect's names as correct. Add Iwer instruction: For D157 (president) change 1st IWER instruction to read: IWER: RECORD EITHER OBAMA OR TRUMP AS CORRECT

D158_: To accommodate the election results we need to accept either the current vice-president or the vice-president elect's names as correct. Add Iwer instruction: and for D158 (vice-president): IWER: RECORD EITHER BIDEN OR PENCE AS CORRECT

--

Data Model 9 Changes:

There were no significant changes to report in Data Model 9.

--

Data Model 10 Changes:

Section D - Cognition

D157 and D158: To accommodate the correct answers to D157 and D158 after the Jan 20, 2017 inauguration, we will prepare the following revisions for SRO to release on 1/23/17. Please revise D157 instruction to say: Answer is Trump and D158 instruction to say: Answer is Pence.

--

Data Model 11 Changes:

Section W - Event History, Internet Use and Social Security

W310 and W322: In an attempt to address Respondent concerns about interview length, interviewers have adopted the strategy of suggesting that the interview be completed in two parts. The first part is administered FTF and takes approximately 60/90 minutes to complete. The second part can then be completed by telephone.

The response options and question text can vary by mode. For example, the wording of questions W306 (pt 2) and W322 do vary in this manner. At W310, codes 1 and 3 are available for FTF but only code 2 shows for telephone Iws. This "variable text" causes a problem for these cases which are started FTF and completed by TEL.

Fix:

We implemented changes to ST to ask SSA permission (in ST) for suspended cases, and then, upon resuming the Iw, ST will transmit the ST data into Blaise data fields. The code is now set to 3 so that it always has text. Text has been added to make sure IWERS know what not to select (for non-FTF cases).

```
{SAVE: 3}FL_LEFTwithR := '(FTF only) R agrees - form and letter left with R'
NOTE: Code 1 does the same thing. Make sure Code 1 has text all the time
{SAVE: 1}FL_SIGNEDFORMOBTAINED := '(FTF only)R agrees - signed form obtained
by interviewer'
```

--

Data Model 12 Changes:

Section A - Coverscreen

A020, A026 and A027: Incorrect preload values in X065 (and X066) were found. These "erroneous" values caused R to skip A020 (ReIw Rs) or A026/A027 (NC Rs) so we didn't get a coupleness setting for those interviews. This caused skip (or fill) issues in the subsequent sections of the questionnaire. In order to fix this, if X065 comes in empty, R is routed to A026. If X065 comes in with 2(remarried) or 4(repartnered), R will go to A020. The fills associated with this change have also been corrected.

--

Data Model 13 Changes:

Certain dates throughout the instrument were adjusted to accommodate the extended field period, which continued on into 2018.

10. OBTAINING THE DATA

10A. Registration and Downloading the Data

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

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

10B. Conditions of Use

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

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

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

10C. Publications Based on Data

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

Health and Retirement Study
Attn: Papers and Publications
The Institute for Social Research, Room 3450
P.O. Box 1248
Ann Arbor, MI (USA) 48106-1248

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

11. If You Need to Know More

This document is intended to serve as a brief overview and to provide guidelines to using the 2016 HRS Core (Early, Version 2.0) data. If you have questions or concerns that are not adequately covered here or on our Web site, or if you have any comments, please contact us. We will do our best to provide answers.

11A. 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](#).

11B. 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: hrsquestions@umich.edu

Postal Service:

Health and Retirement Study
The Institute for Social Research
The University of Michigan
P.O. Box 1248
Ann Arbor, MI 48106-1248

FAX: (734) 647-1186

Appendix

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

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

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

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

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

A value of 7 indicates respondents from split household reunited¹. It is important to understand these assignments when you merge records from different waves of the study.

A1. Married Couple Stays Married

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

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

Time 1

Household records

HHID=012345 ASUBHH=0

Respondent records

HHID=012345 PN=010 ASUBHH=0

HHID=012345 PN=020 ASUBHH=0

Time 2

Household records

HHID=012345 CSUBHH=0

Respondent records

HHID=012345 PN=010 CSUBHH=0

HHID=012345 PN=020 CSUBHH=0

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

A2. Couple Divorces

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

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

Time 1

Household records

HHID=023456 ASUBHH=0

Respondent records

HHID=023456 PN=010 ASUBHH=0

HHID=023456 PN=020 ASUBHH=0

Time 2

Household records

HHID=023456 CSUBHH=1

HHID=023456 CSUBHH=2

Respondent records

HHID=023456 PN=010 CSUBHH=1

HHID=023456 PN=020 CSUBHH=2

A3. One or Both Respondents Die

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

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

Time 1

Household records

HHID=034567 ASUBHH=0

Respondent records

HHID=034567 PN=010 ASUBHH=0

HHID=034567 PN=020 ASUBHH=0

Time 2

Household records

HHID=034567 CSUBHH=0

HHID=034567 CSUBHH=3 (in exit interview)

Respondent records

HHID=034567 PN=010 CSUBHH=3 (in exit interview)

HHID=034567 PN=020 CSUBHH=0

A4. Single Respondent Marries

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

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

Time 1

Household records

HHID=045678 ASUBHH=0

Respondent records

HHID=045678 PN=010 ASUBHH=0

Time 2

Household records

HHID=045678 CSUBHH=0

Respondent records

HHID=045678 PN=010 CSUBHH=0

HHID=045678 PN=011 CSUBHH=0

A5. Couple Divorces, One Respondent Remarries and Divorces

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

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

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

Time 1

Household records

HHID=056789 ASUBHH=0

Respondent records

HHID=056789 PN=010 ASUBHH=0

HHID=056789 PN=020 ASUBHH=0

Time 2

Household records

HHID=056789 CSUBHH=1

HHID=056789 CSUBHH=2

Respondent records

HHID=056789 PN=010 CSUBHH=2

HHID=056789 PN=011 CSUBHH=2

HHID=056789 PN=020 CSUBHH=1

Time 3

Household records

HHID=056789 FSUBHH=1

HHID=056789 FSUBHH=2

HHID=056789 FSUBHH=5

Respondent records

HHID=056789 PN=010 FSUBHH=2

HHID=056789 PN=011 FSUBHH=5

HHID=056789 PN=020 FSUBHH=1

A6. Couple Divorces and Marries Again

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

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

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

Time 1

Household records

HHID=067890 ASUBHH=0

Respondent records

HHID=067890 PN=010 ASUBHH=0

HHID=067890 PN=020 ASUBHH=0

Time 2

Household records

HHID=067890 CSUBHH=1

HHID=067890 CSUBHH=2

Respondent records

HHID=067890 PN=010 CSUBHH=1

HHID=067890 PN=020 CSUBHH=2

Time 3

Household records

HHID=067890 FSUBHH=7

Respondent records

HHID=067890 PN=010 FSUBHH=7

HHID=067890 PN=020 FSUBHH=7

A7. Married Couple with Children and Siblings

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

Time 1

Household records

HHID=078901 ASUBHH=0

Respondent records

HHID=078901 PN=010 ASUBHH=0

HHID=078901 PN=020 ASUBHH=0

Household member/child records

HHID=078901 ASUBHH=0 OPN=101 (child)

HHID=078901 ASUBHH=0 OPN=102 (child)

HHID=078901 ASUBHH=0 OPN=103 (child)

HHID=078901 ASUBHH=0 OPN=051 (sibling)

HHID=078901 ASUBHH=0 OPN=052 (sibling)

HHID=078901 ASUBHH=0 OPN=080 (friend)

A8. Couple with Children and Siblings Divorces

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

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

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

Time 1

Household records

HHID=089012 ASUBHH=0

Respondent records

HHID=089012 PN=010 ASUBHH=0

HHID=089012 PN=020 ASUBHH=0

Household member/child records

HHID=089012 ASUBHH=0 OPN=101 (child)

HHID=089012 ASUBHH=0 OPN=102 (child)

HHID=089012 ASUBHH=0 OPN=051 (her brother)

HHID=089012 ASUBHH=0 OPN=052 (her brother)

HHID=089012 ASUBHH=0 OPN=061 (his sister)

HHID=089012 ASUBHH=0 OPN=080 (friend)

Time 2

Household records

HHID=089012 CSUBHH=1

HHID=089012 CSUBHH=2

Respondent records

HHID=089012 PN=010 CSUBHH=2

HHID=089012 PN=020 CSUBHH=1

Household member/child records

HHID=089012 CSUBHH=1 OPN=101 (child)

HHID=089012 CSUBHH=1 OPN=102 (child)

HHID=089012 CSUBHH=1 OPN=051 (her brother)

HHID=089012 CSUBHH=1 OPN=052 (her brother)

HHID=089012 CSUBHH=2 OPN=101 (child)

HHID=089012 CSUBHH=2 OPN=102 (child)

HHID=089012 CSUBHH=2 OPN=061 (his sister)

A9. Couple Divorces, One Respondent Remarries, Both Split-off Households Have New Members

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

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

Time 1

Household records

HHID=090123 ASUBHH=0

Respondent records

HHID=090123 PN=010 ASUBHH=0

HHID=090123 PN=020 ASUBHH=0

Time 2

Household records

HHID=090123 CSUBHH=1

HHID=090123 CSUBHH=2

Respondent records

HHID=090123 PN=010 CSUBHH=1

HHID=090123 PN=011 CSUBHH=1 (new wife)

HHID=090123 PN=020 CSUBHH=2

Household member/child records

HHID=090123 CSUBHH=1 OPN=151 (his stepchild)

HHID=090123 CSUBHH=1 OPN=152 (his stepchild)

HHID=090123 CSUBHH=2 OPN=151 (her mom)