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

Prepared by Marita A. Servais

Survey Research Center
Institute for Social Research
University of Michigan, Ann Arbor

June 2004
(Updated June 2010)

Table of Contents

ii

1 Introduction

The University of Michigan Health and Retirement Study (HRS) surveys more than 22,000 Americans over the age of 50 every two years. The study paints an emerging portrait of an aging America's physical and mental health, insurance coverage, financial status, family support systems, labor market status, and retirement planning. This document provides an overview of the full scope of data files that the study has released as it has evolved since the original data collection in 1992.

The Health and Retirement Study (HRS) and Asset and HEAlth Dynamics Among the Oldest Old (AHEAD) studies were created as separate but related surveys. In its original conceptualization, the HRS survey was designed to follow age-eligible individuals and their spouses or partners as they made the transition from active worker into retirement; the AHEAD survey was designed to examine the dynamic interactions between health, family, and economic variables, in the post-retirement period at the end of life. Both studies obtain detailed information in a number of domains: demographics, health status, housing, family structure, employment of respondent, work history and current employment, disability, retirement plans, net worth, income, and health and life insurance.

The HRS collected data in 1992, 1994, and 1996. AHEAD collected data in 1993 and 1995. In 1998, 2000, and 2002, the HRS, AHEAD and two new sub-samples, War Baby (WB) and Children of the Depression Age (CODA), were interviewed. In 2004, 2006, and 2008 HRS, AHEAD, WB, CODA, and another new sub-sample, Early Baby Boomer (EBB), were interviewed. A new sub-sample, Mid Baby Boomer (MBB) will be added in 2010. For more details about the sample, see the Sample Design. See "Data Collection Path" under "Quick links" at our Web site for visual overview of the sample and data collection efforts. Click on any cell for more information about a specific data product.

Sub-sample Abbreviation	Sub-sample Name	Birth Year of Age-Eligible Respondents	First Wave	Data Collected
AHEAD	Aging & HEAlth Dynamics	1890-1923	1993	1993 1995 1998 2000 2002 2004 2006 2008 2010
CODA	Children Of the Depression Age	1924-1930	1998	1998 2000 2002 2004 2006 2008 2010
HRS	Health & Retirement Study	1931-1941	1992	1992 1994 1996 1998 2000 2002 2004 2006 2008 2010
WB	War Baby	1942-1947	1998	1998 2000 2002 2004 2006 2008 2010
EBB	Early Baby Boomer	1948-1953	2004	2004 2006 2008 2010
MBB	Mid Baby Boomer	1954-1959	2010	2010

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

HRS data are available free to researchers and analysts at the HRS Web site. By receiving the data, which have been freely provided, you must agree to use them for research and statistical purposes only and make no effort to identify the respondents therein. In addition, you must agree to send us a copy of any publications you produce based on the data.

Most HRS data are public data available after a simple registration process has been completed. HRS special access data require registration through a supplemental registration system. HRS restricted data are only available under very restrictive conditions. See Obtaining the Data for additional details.

2   HRS Data Files

The HRS public data files consist of five types, 1) Core Files, 2) Exit Files, 3) Post-Exit Files, 4) Imputation Files, 5) Cross-Wave Files, 6) Off-Year Studies, and 7) Sensitive Health Data Products. Also available at the public data download page on our Web site, but not further discussed here, are

- Researcher Contributions[1],
- Restricted File Documentation,
- RAND Contributed Files[2], and a
- Software Program.

The data are provided in ASCII format, with fixed-length records. You'll want to use associated SAS, SPSS or Stata program statements to read the data into the analysis package of your choice. See Program Statements for information about creating datasets with the program statements provided for your particular software.

2.1   List of Current Public Data Products

All HRS public data products (as of June 2010) are listed below. Clicking on the Quick link "Data Product List" at our Web site will take you to "What's Available (Public)" where you will find a year-by-year listing of public datasets and files that are currently available to registered users. To see metadata information -- including data descriptions, codebooks, and questionnaires -- click on the Dataset Name. Be sure to review "Alerts" prior to using any datasets in order to learn of any post-release changes.

---

[1] For more information about these files, see "What's Available (Public)" at the "Data Products" section at our Web site.

[2] The RAND HRS Data file is a respondent-level data file containing derived variables covering a broad range of measures that have been constructed and named consistently across waves. The RAND HRS Data file (Version J) used data from 1992, 1993, 1994, 1995, 1996, 1998, 2000, 2002, 2004, 2006, and 2008 early release. For further information, see "Rand Contributions" at "What's Available (Public)" at the "Data Products" section at our Web site.

HRS Data Products

- 1992 HRS (Final) (v.2.0)
- 1992 HRS Imputations (Final) (v.3.0)
-
- 1993 AHEAD (Final) (v.2.10)
-
- 1994 HRS (Final) (v.2.0)
- 1994 HRS Imputations (Final) (v.3.0)
-
- 1995 AHEAD Core (Final) (v.2.0)
- 1995 AHEAD Core Imputations (Final) (v.3.0)
- 1995 AHEAD Exit (Final) (v.2.0)
- 1995 AHEAD Exit Imputations (Final) (v.1.0)
-
- 1996 HRS Core (Final) (v.4.0)
- 1996 HRS Core Imputations (Final) (v.3.0)
- 1996 HRS Core Supplement (Final) (v.1.0)
- 1996 HRS Exit (Final) (v.1.0)
- 1996 HRS Exit Imputations (Final) (v.1.0)
-
- 1998 HRS Core (Final) (v.2.3)
- 1998 HRS Core Imputations (Final) (v.3.0)
- 1998 HRS Exit (Final) (v.1.0)
- 1998 HRS Exit Imputations (Early) (v.2.0)
- 1998 HRS Post-Exit (Final) (v.1.0)
-
- 1999 HRS Mailout (v.1.0)
-
- 2000 HRS Core (Final) (v.1.0)
- 2000 HRS Core Imputations (Final) (v.2.0)
- 2000 HRS Exit (Final) (v.1.0)
- 2000 HRS Exit Imputations (Early) (v.1.0)
- 2000 HRS Post-Exit (Final) (v.1.0)
-
- 2001 HRS Mailout CAMS (v.2.0)
- 2001 HRS Mailout HUMS (v.1.0)

- 2001 HRS Mailout HUMS Imputations (Early) (v.1.0)
-
- 2002 HRS Core (Final) (v.2.0)
- 2002 HRS Core Imputations (Final) (v.1.0)
- 2002 HRS Exit (Final) (v.1.0)
- 2002 HRS Post-Exit (Final) (v.1.0)
-
- 2003 HRS Mailout CAMS (v.1.0)
- 2003 Internet Survey (v.1.0)
-
- 2004 HRS Core (Final) (v.1.0)
- 2004 HRS Core Imputations (Early) (v.1.0)
- 2004 HRS Exit (Final) (v.1.0)
- 2004 HRS Post-Exit (Final) (v.1.0)
-
- 2005 HRS Mailout CAMS (v.1.0)
-
- 2006 HRS Core (Final) (v.2.0)
- 2006 HRS Exit (Final) (v.1.0)
- 2006 HRS Post-Exit (Final) (v.1.0)
- 2006 HRS Core Imputations (Final) (v.A)
- 2006 Internet Survey (v.1.0)
-
- 2007 HRS Mailout DVS (v.1.0)
- 2007 HRS Mailout CAMS (v.1.0)
- 2007 Internet Survey (v.1.0)
-
- 2008 HRS Core (Early) (v.2.0)
- 2008 HRS Exit (Early) (v.1.0)
- 2008 HRS Core Imputations (Early) (v.A)
- 2008 HRS Post-Exit (Final) (v.1.0)
-
- 2009 HRS Mailout CAMS (v.1.0)
-
- Cross-Wave: Tracker 2008 File (v.1.0)
- Cross-Wave: Region and Mobility File (v.2.0)
- Cross-Wave: LOPN 2002 (v.1.1)
- Cross-Wave: Labor Section Carry Forward Variables (v.2.0)

- Cross-Wave: Imputations for Pension-Related Variables (v.1.0)
- Cross-Wave: Employer Pension Tracker File (v.1.0)
- Cross-Wave: Imputations for Pension Wealth 1992 and 1998 (v.2.0)
- Cross-Wave: Respondent Pension Tracker File (v.1.0)
- Cross-Wave: Imputations for Employer-Sponsored Pension Wealth from Current Jobs in 2004 (v.1.0)
- Cross-Wave: Imputation of Cognitive Functioning Measures 1992-2006 (v.1.0)
- Cross-Wave: Master ID File (v.1.0)
- Cross-Wave: SS Wealth File (v.3.1)
- Cross-Wave: Child Proximity (v.1.0)

Researcher Contributions

- 1992 HRS 401K Flag (v.1.0)
- 1992 HRS Pension (Level 1) Present Value Database (v.1.0)
- 1992 HRS Self-Reported Pension Wealth (v.1.0)
- 1994 HRS Imputed Medical Expenses (v.1.0)
- Second Home Ownership and Equity Corrections (v.1.0)
- Tax Calculations for HRS 2000 and 2002 (v.A)
- Table Data: Pensions in the Health and Retirement Study (v.1.0)

Restricted File Documentation

- National Death Index (v.3.1 [Codebook])

Geographic Information (v.2.0 [Codebook])
- Detailed Industry-Occupation (v.7.0 [Codebook])
- SSA Respondent Cross-Year Benefits (v.2.0 [Codebook])
- SSA Respondent Cross-Year Summary Earnings (v.1.1 [Codebook])
- SSA Respondent Cross-Year Detail Earnings (v.1.0 [Codebook])

RAND Contributed Files

- RAND HRS Data File (v.J)
- RAND HRS Distribution CD Key (v.1.0)

HRS Special Access Files

- *Various, see [HRS Sensitive Health Data](#)*

Requesting HRS Special Access Files

- Application for 2001 ADAMS Data (v.1.0)
- Application for 2003 Diabetes Data (v.1.0)
- Application for 2005 Prescription Drug Study Data (v.1.0)
- Application for 2006 Biomarker Data (v.1.0)

Software

- Pension Estimation Program (v.2.0)

2.1.1 Core Files

The first waves of the study, 1992 for the HRS sub-sample, 1993 for the AHEAD sub-sample, 1998 for the CODA and WB sub-samples, 2004 for the EBB sub-sample, and 2010 for the MBB sub-sample, contain data from the original interview. In subsequent years, "core" interviews were taken with living respondents, and "exit" interviews were taken on behalf of deceased respondents.

Core Survey Content

While there is some variation over the years, the description provided for HRS 2008 gives a feeling for the type of information available. For more detailed information, see "Content" or "Questionnaires" or "Codebooks" at the "Documentation" section at [our Web site](#).

o HRS 2008 Core

o Section A: Coverscreen. Household information; interview information; living arrangements; marital status; respondent information.

o Section B: Demographics. Birthplace; education; family history; language; marital history; military history; race; religion; residence.

o Section C: Physical Health. Alcohol; arthritis; cancer; daily activities; depression; diabetes; eyes; fall; fractures; hearing; heart problems; height/weight; hypertension; incontinence; insomnia; lung disease; medical conditions; memory-related disease; pain; preventive procedures; psychiatric; self-rated health; stroke; symptoms; tobacco; height; weight.

o Section D: Cognition. Proxy cognition; self cognition.

o Section E: Family Structure. Child transfers; children assistance; children demographics; financial/time assistance; grandchildren; grandchildren assistance; household information; household members; neighbors; transfers.

o Section F: Parents and Sibling/Couple Decisions. Marriage; neighbors; parent assistance; parent demographics; sibling demographics.

o Section G: Functional Limitations and Helpers. Activities of daily living; family structure; Instrumental Activities of Daily Living; physical activities; volunteer work/help others.

o Section H: Housing. Condominimum/Coop/Association information; deed; equity line of credit; farm or ranch; home facilities/modifications; home ownership; mobile home; mortgage; neighborhood; other loans; purchase; rent; retirement services; second home; second mortgage; taxes; type of residence; value.

o Section J: Employment (Core). Main job (benefits, characteristics, dates, description, earnings, hours, training); self-employment dates and earnings; other jobs (earnings, hours, reason for leaving); formeremployer (dates, description, earnings, reason for leaving); job requirements; income (annuities, pension, social security); early retirement window; pension and retirement plans.

o Section J: Employment (Exit). Main job (benefits, characteristics, dates, description, earnings, hours, training); self-employment dates and earnings; other jobs (earnings,

hours, reason for leaving); formeremployer (dates, description, earnings, reason for leaving); job requirements; income (annuities, pension, social security); early retirement window; pension and retirement plans.

o Section K: Last Job. Job characteristics; dates; description; earnings; hours; pension and retirement plans.

o Section L: Job History. Earnings, hours; pension and retirement plans; early retirement window.

o Section M: Disability. Benefits (Social Security/SSDI/SSI, Veterans Administration, Workers Compensation, other programs); impairment history; injuries at work.

o Section N: Health Services and Insurance. Health providers (dentist, doctor); drugs; financial assistance; government health insurance; health insurance; hospitalization; in-home care/special facilities; long-term care insurance; Medicaid; Medicare; nursing home information; outpatient surgery.

o Section P: Expectations. Probability; risk tolerance.

o Section Q: Assets and Income. Assets (bonds, business or farm, CD, T-Bill, checking/savings/money market, IRA, other, pension, real estate, stocks, transportation, trusts); expenses (charity, medical, debts, food); federal tax return; income (employment, annuities, bonds, CD, T-Bill, checking/savings/money market, self-employment, food stamps, pension, profession/trade, rental, social security, stocks, supplemental security (SSI), tips, bonus, unemployment, veterans benefits, welfare, workers compensation, other employment, other sources); lump sum payments.

o Section R: Asset Change. Business purchased; business sold; household member addition assets/debts; own home; real estate-purchased; real estate-sold; residence-bought or sold; major home improvement; stocks.

o Section S: Widowhood and Divorce. Death; divorce; earnings/work change; health insurance change; life insurance; pension; Social Security.

o Section T: Wills, Insurance and Trusts. Will provisions; life insurance; beneficiaries; asset disposal (Exit)

o Section U: Asset Verification. Verification of substantial changes in net worth or asset value between 2006 and 2008.

o Section W: Internet Use and Social Security Permission. Internet use; social security permissions.

o Section LB: Psycho-Social (Participant Lifestyle Leave-Behind Questionnaire). Psycho-social; participant lifestyle

Structure of Core Questionnaires

The structure of the core questionnaires has evolved over the years. There were two major redesigns. First in anticipation of a combined data collection effort in 1998, the 1995 AHEAD and 1996 HRS questionnaires were redesigned. The advent of new data collection software in 2002 marked the second major redesign. The table below provides an overview.

Table 1: Comparison of Sections 2008 - 1992

Content	2008 2006 2004 2002	2000 1998 1996 1995	1994 1992	1993
Preload	PR	PR		
Coverscreen	A	CS		
Demographics	B	A	A	A
Physical Health / Health Status	C	B	B	B
Cognition / Expectations	D	C	C 1994 L 1992	C
Proxy Cognition		PC		
Family Structure and Transfers	E	D	E	D
Parents, Siblings and Transfers	F			
Functional Limitations and Helpers / Health Care Utilization & ADLs / Health Care Costs	G	E		E
Housing	H	F	D	F
Physical Measures/Biomarkers 2006	I			
Employment	J	G	F	G
Last Job	K		G	
Job History	L		H	
Disability	M		J	
Health Services and Insurance	N			
Expectations / Retirement Plans	P	H	K 1992	H
Assets and Income	Q	J	N	J
Net Worth			K 1994 M 1992	
Asset Change / Asset Change & Widowhood / Capital Gains / Net Worth	R	N	V 1994	K
Widowhood and Divorce	S		S	
Wills and Life Insurance / Health Insurance & Event History	T	R	R	R
Asset Reconciliation	U			
Modules	V	M		
Event History, Internet Use & Social Security Permission	W	EV 1998		
Time Calculations, Thumbnail	Y	T	X	

Core Modules

A brief overview of the contents of the modules section in each interviewing year is included below. For more detailed information, see "Modules" at the "Documentation" section at our Web site.

□ HRS 2008 Modules: Annuities and lump sum payments; breast and prostate cancer screening; weight history; quality of care; coping strategies; transfers; financial sophistication and investment decision making; retirement behavior; dental health, access to care and utilization.

□ HRS 2006 Modules: Financial preparedness; risk aversion; parents' and siblings' health and long-term care use; valuation of health insurance benefits; Medicare valuation; characteristics of the home environment; subjective probabilities of health-related events; informed consent with Alzheimer's disease patients; cognition - number series; cognition - retrieval fluency.

□ HRS 2004 Modules: Annuities; arthritis and health behaviors; asset ownership; cognition; disability; loneliness; norms on transfer behavior; occupational health; pension characteristics; pension documents; probability alternative to bracketing; retirement planning; risk aversion.

□ HRS 2002 Modules: Self-assessed health utilities; willingness to pay for disease prevention; restless leg syndrome, night leg cramps, and neck and shoulder pain; risk aversion; Internet use; loneliness, stress and social support/social burden; ELSA health questions; numeracy; positive well-being; later life education; subjective uncertainty about stock market returns.

□ HRS 2000 Modules: Medicare knowledge; alternative medicine; planning and expectations for retirement; social and economic altruism; benevolence and obligation; health plan booklet; health utilities index; risk tolerance; alcohol consumption and IADL measures; proxy validation; valuing health.

□ HRS 1998 Modules: There were limited modules in HRS 1998, due to the addition of two new cohorts, CODA and WB, and merger with the two original HRS and AHEAD cohorts into one study. The 1998 modules were asked primarily of the AHEAD sample members who were asked ADL and cognition questions corresponding to similar modules in previous waves.

□ HRS 1996 Modules: Consumption and anchoring, health during childhood, health pedigree, personality inventory, Medicare attitudes and preferences, volunteerism and time use, preference parameters for consumption, saving and labor supply, advance directives, attitudes toward inter-familial transfers, retirement planning, saving for retirement

□ AHEAD 1995 Modules: Unfolding brackets with different entry points, Wave 1 ADL questions, LSOA2 ADL questions, security and safety, sleep, living wills, in-depth ADLs

□ HRS 1994 Modules: CES-D depression scale, crystallized intelligence, functional health, long-run income elasticity of labor supply, risk aversion, social support, parent-child transfers, ADLs, activities and time allocation, nutrition

- □ AHEAD 1993 Modules: Resilience, time use, alternative ADLs, WAIS Similarities, quality of life, in-depth ADLs, financial pressure

- □ HRS 1992 Modules: Physiological health measures, ADL measures from NLTCS and NHIS, meta-memory, process benefits, employment alternatives, parental wealth, occupational injuries, health risks, substitution elasticity of consumption

2.1.2 Exit Files

The sample design provides for 'exit interviews' with a surviving spouse, child or other informant. Questions are asked concerning medical expenditures and family interactions with the deceased during the final stages of life and information about the disposition of assets following death. Exit interviews were attempted in each interview wave since the initial interview. The 1994 data from both core and exit interviews are contained in one set of files; for subsequent years, core and exit files are distributed separately.

2.1.3 Post-Exit Files

Post-exit interviews are designed to elicit information about unresolved exit interview issues such as the settlement of a deceased respondent's estate. These interviews are typically administered in the wave that follows the exit interview. The post-exit interview is triggered when the exit proxy respondent reports that a deceased respondent's estate has not been settled or for some other reason. We attempt to locate, if possible, the proxy who gave the exit interview and ask the estate distribution question sequence that was not answered in the exit interview. The first post-exit interviews were administered in 1998.

2.1.4 Imputation Files

Imputations for most "unfolding bracket" variables have been constructed and are provided in separate files. The number of records in the imputation files corresponds to the number of records in the files for which the imputations have been made. The imputation files may be matched with the original files by its primary identifiers.

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 questions asked -- if the amount was less, more, or about a specified amount. Most often the series consisted of two to four unfolding bracket questions. For some unfolding bracket series, a preload variable assigned an entry point, low, medium or high, which was used to determine the asking order of the unfolding bracket questions.

Imputations are provided for the variables satisfying our ad hoc selection rules listed below.

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

For each amount variable, there are six imputation-related variables: the "X," "C," "S," "D," "E," and "T" variables to accommodate various needs of perspective users. For more detailed information, see "Imputations" at the "Documentation" section at our Web site.

Those who just want to have imputed values may keep only the "X" variable with the imputed amount. For instance, the imputed variable JH020X, "HOME PRESENT VALUE" will be found in the 2004 imputation file H04I_HH corresponding to the original amount variable, JH020 in the 2004 public release data H04H_H.

Also included are five aggregate variables -- main home equity, and second home equity, total assets, total net worth, and household income. The variables listed below are from the 2004 files.

> JHHINC – household income in the last calendar year
> JASSETS – household assets at the time of the interview
> JHOME1 – primary home equity
> JHOME2 – second home equity
> JNETWRTH – sum of JASSETS, JHOME1, and JHOME2

For detailed information about the imputation process, see "IMPUTE: A SAS Application System for Missing Value Imputations--With Special Reference to HRS Income/Assets" HRS Documentation Report DR-007 (2001) available in "User Guides" at the "Documentation" section at our Web site.

The 2006 final-release and 2008 early-release income and wealth imputations files were prepared by RAND and follows different conventions.

2.1.5 Cross-wave Files

A brief description of HRS Longitudinal and Cross-Wave Data Products is included below. For more information about these files, see "What's Available (Public)" at the "Data Products" section at our Web site.

Tracker 2008 File

The tracker file contains a record for each individual eligible to be interviewed, whether actually interviewed or not, in any wave from 1992 through the most recent year. The tracker file contains some basic demographic information as well as some basic information obtained from the National Death Index (NDI). In addition it contains basic information about the interview year such as whether a person gave an interview in a particular wave so that it can be used to understand the interview history of respondents. The tracker file also contains cross-sectional weights for each interview year 1992 through the most recent year for both respondent and household level analysis. The primary identification variables for this file are HHID and PN.

Region and Mobility File

The Region and Mobility File: 1992-2008 contains information on inter-wave changes in respondent's location of residence. Distance, in miles, was calculated for all respondents who were interviewed in a residence location differing from that of the previous wave and for whom geographic information was available. Also included are two one-time variables, region in which respondent was born and region in which respondent attended school, and, for each interview year, region lived in at the time of interview and urbanicity. Version 3.0, released June 2010, contains corrected mobility and rural-urban information. Previous versions should no longer be used. The primary identification variables for these files are HHID and PN.

Longitudinal Other Person Number (LOPN) Files

The Longitudinal Other Person Number files have been prepared to facilitate the merging of information obtained about children and other household members in different waves of the HRS study. Two variables, HHID and LOPN, may be used to uniquely identify children and other household members across all waves of the study from 1992 through 2002. The LOPN files contain no substantive analytic variables, only identification variables. Eighteen data files are provided. Three are cross-wave files; the primary identification variables for these files are HHID, LOPN and STORY. Fifteen are single-wave files; the primary identification variables for these files are HHID, cySUBHH and OPN. The Version 1 LOPN files were constructed using files publicly available in January 2004.[3]

Labor Section Carry Forward Variables

The Labor Section Carry Forward data set consists of respondent-level, cross-sectional files constructed from the employment sections of HRS 1994 (Wave 2), HRS 1996 (Wave 3), HRS 1998 (Wave 4), HRS 2000 (Wave 5), HRS 2002 (Wave 6), and HRS 2004 (Wave 7). When a respondent reported no change in employment status since the previous interview, certain questions relating to current employment were not asked. Additional questions are skipped if the respondent holds the same position or title in addition to the same employer. As a service to users, the files in this data release carry forward the information from each prior wave whenever certain employment variables are "skipped" in the current wave. N: 1994=11596; 1996=10964; 1998=13113; 2000=12455; 2002=18167; 2004=20129. The primary identification variables for the file are HHID and PN.

Imputations for Pension-Related Variables

The Imputations for Pension-Related Variables files contain imputations for don't know, refused, and missing responses for key pension related variables from the employment sections of the 1992, 1994, 1996, 1998, 2000 and 2002 interviews. The imputations were performed for the HRS cohort in 1992, 1994, and 1996 and for the HRS and War Baby cohorts in 1998, 2000 and 2002; AHEAD and CODA cohorts were not included in these imputations. The imputations for each wave are cross-sectional. N: 1992: 12652; 1994:

---

[3] Contact HRS for information about how to create LOPN using all currently available data.

11596; 1996: 10964; 1998: 21384; 2000: 19580; 2002: 18167. The primary identifica-
tion variables for the files are HHID and PN.

Employer Pension Tracker File

The Employer Pension Data Tracker file is constructed to help analysts in identifying
employees with self-reported pension coverage and employer pension plan descriptions
data from the 1993 and/or 1999 employer survey(s). This file is a respondent level file
that contains one record for every respondent among HRS and War Baby cohorts who
participated in the core survey in 1992 and/or 1998. There are 15615 observations in the
data file. The data file contains variables indicating whether respondents had pension
coverage in 1992 and/or 1998 from current, last, and/or previous jobs, whether respon-
dents worked for an employer with 25 or more employees at all locations in the1998 sur-
vey, whether respondents' employer contact information was available in the 1998
survey, and whether respondents had matched plan descriptions from the 1993 and/or
1999 employer survey(s). For the 1998 survey, two additional variables are included:
whether or not the respondents' employer was interviewed, and whether or not matched
plans were obtained through an interview and are available for the respondent. N: 15615.
The primary identification variables for the file are HHID and PN.

Imputations for Pension Wealth 1992 and 1998

The Imputations for Pension Wealth file contains imputations for pension wealth data
from current jobs on both the self-reported and the employer data. DB values are calcu-
lated or imputed from the 1993 and 1999 employer data which were obtained based on
employer contact information provided by the respondent in the 1992 and 1998 core sur-
veys. DC value imputations are based on the self-reported data. The data file is a respon-
dent level file containing 15,879 records, one record for every respondent in the HRS and
War Baby cohorts who participated in the core survey in 1992 to 2000. N: 15879. The
primary identification variables for the file are HHID and PN.

Respondent Pension Tracker Files

The Respondent Pension Tracker Files are designed to enhance the user's ability to use
pension information from various jobs reported by respondents in up to seven waves of
the Health and Retirement Study, for the HRS cohort, 1992, 1994, and 1996 and for all
cohorts for 1998, 2000, 2002, and 2004. These data files are especially useful to users
who are interested in analyzing detailed information on respondents' pension history,
current coverage, and pension wealth at each of the survey years. Each file includes vari-
ables indicating pension coverage from each job, identified by the job on which the pen-
sion is or was held, and the number of pension plans from current and previous
employments. Each data file also includes an index identifying each of the previous pen-
sion plans that are dormant (i.e. not in pay status, not cashed out, rolled over, converted
to an annuity, or lost), number of pension plans from current job, total number of previ-
ous pension plans through a current interview date, and the number of dormant pension
plans that a respondent is entitled to (has active claim on) as of that interview date. In
each wave, pension indices from previous pensions are brought forward to the current
wave. In addition, the information collected about changes in older pension plans are
used to adjust reports in earlier waves about which pension plans were dormant at the

time of a previous survey. There is one respondent-level data file for each of the waves 1992, 1994, 1996, 1998, 2000, 2002, and 2004. The primary identification variables for the files are HHID and PN.

Imputations for Employer-Sponsored Pension Wealth from Current Jobs in 2004

This data set contains imputations for pension data from current jobs based on both self-reported and employer data. For this version of pension wealth imputations, we use the respondent's self-reported pension plan type to determine which method of wealth estimation to use. If a respondent reported being covered by one or more defined contribution (DC) plans on his/her current job in 2004, DC wealth for each such plan is taken from the 2004 survey self-reported account balances. If a respondent reported being covered by one or more defined benefit (DB) plans in 2004, the DB pension wealth value for each such plan is calculated by the HRS Pension Estimation Program using a combination of self-reported data from the employment section (J) of the 2004 survey and pension plan rules obtained from the plan's Summary Plan Description (SPD). The primary identification variables for the files are HHID and PN.

Imputation of Cognitive Functioning Measures 1992-2006

The Imputation of Cognitive Functioning Measures 1992-2006 (V1.0) release contains imputations for cognitive functioning data for HRS 1992 through 2006. The objective was to perform imputations for respondents with missing cognition data using a multi-variate, regression-based procedure using Imputation and Variance Estimation (IVEware) software (http://www.isr.umich.edu/src/smp/ive/). We used a combination of relevant demographic, health, and economic variables, as well as prior and current wave cognitive variables to perform the imputations. Prior wave cognitive scores were used to perform the imputations, except for the baseline waves for each of the cohorts where subsequent wave scores were used instead. Additional information about our imputation strategy is provided in the data description for this data product. The primary identification variables for the files are HHID and PN.

Master ID File

The HRS Master ID file can be used to merge biannual HRS files (see also Linking Respondents across Time). The HRS uses two identification variables in combination to uniquely identify individuals: HHID and PN, however, a few[4] individuals have had changes in their HHID or PN variables in different waves of the study. It contains wave-specific HHID, PN, and SUBHH identification variables for every wave of HRS in Final Core release data, for every observation appearing in that wave which match exactly the identification variables as they appear in the core release of each wave.

---

[4] This affects 108 respondents in 1992, two respondents in 1993 and 1995, and one respondent in 1998 and 2000. See the "Cross-Wave Tracker File Data Description" for more information about these "overlap" cases, at "Data Descriptions" page at the "Documentation" section at our Web site .

Prospective Social Security Wealth Measures of Pre-Retirees

The Prospective Social Security Wealth Measures of Pre-Retirees data set consists of respondent-level, cross-sectional files constructed from the employment sections of the HRS 1992 (wave 1), HRS 1998 (wave 4), HRS 2004 (wave 7) and the restricted SSA summary and detailed earnings and benefits files. From these data, we calculate the value of Social Security wealth at different ages (at early retirement age, normal retirement age and age 70). Version 2 of these data offered an improvement in the number of respondents needing SS wealth imputations and the overall imputation strategy. Version 3 offered an improvement in the imputation strategy for household wealth values. Version 3.1 corrects undocumented code frame values for these variables: R1IMPUTE, R4IMPUTE, R7IMPUTE, S1IMPUTE, S4IMPUTE, S7IMPUTE, R1CLAIMED, S1CLAIMED, R4CLAIMED, S4CLAIMED, R7CLAIMED, and S7CLAIMED. The primary identification variables for the files are HHID and PN.

Child Proximity

These data sets contain proximity measures for HRS family respondents and children and step-children for 2004, 2006 and 2008. Proximity measures were calculated for children who lived more than 10 miles from the family respondent. The primary identification variables for the files are HHID, SUBHH, PN, and OPN.

2.1.6 Off-Year Studies

These off-year surveys include mail surveys and will also include internet and other types of surveys. A brief description of HRS Off-Year Studies is included below. For more information about these files, see "What's Available (Public)" at the "Data Products" section at our Web site.

2009, 2007, 2005, 2003, and 2001 Consumption and Activities Mail Survey (CAMS)

The surveys included questions about individual activities, household patterns of consumption, and, in 2001, use of prescription drugs. In 2003 three additional questions were added to the activities section, and many more to the consumption section.

- In 2001, questionnaires were mailed to 5,000 households interviewed in the HRS 2000 core survey. If a selected household had two panel members in it, one panel member was randomly selected to receive the questionnaire.
- In 2003, questionnaires were mailed to 4,156 of the respondents who were in the 2001 CAMS. The remaining 843 respondents were lost due to death (n=372), loss to follow-up (n=173), and exclusion from the 2003 CAMS because they were participating in other HRS supplemental studies (n=298).
- In 2005, in coupled or partnered households, both individuals were included in the sample. They are referred to as R and SP, where R is the respondent who was selected for prior waves of CAMS and SP is the spouse or partner of the R. In the 2005 CAMS there were two different versions of questionnaires. The first consisted of activities and consumption, or the "full" questionnaire (as was the case in 2001 and 2003), and the second "partial" questionnaire contained only the questions about activities. The "full" questionnaire was mailed to Rs (as defined above) and the "partial" questionnaire was mailed to their spouse or partner (SP). In the fall of 2005, a total of 8,124

questionnaires were mailed. Of the 8,124 questionnaires, 5,440 were "full" and 2,684 were "partial" questionnaires. A total of 5,815 questionnaires were returned, 3,880 were the "full" version and 1,935 were the "partial" version.

□ In the fall of 2007, a total of 7,741 questionnaires were mailed. Of the 7,741 questionnaires, 5,209 were respondent questionnaires and 2,532 were spouse questionnaires. A total of 5,612 questionnaires were returned, 3,738 were the respondent version and 1,874 were the spouse version.

□ In the fall of 2009, a total of 7,231 questionnaires were mailed. Of the 7,231 questionnaires, 4,954 were the "full" version, and 2,277 were the "partial" version. The sample for the 2009 CAMS mirrored that of 2005 and 2007.

The primary identification variables for these files are HHID and PN.

2009 Internet Survey

A sub-sample (N=5,742) of the Health and Retirement Study (HRS) participated in the 2009 Internet-based survey developed jointly by the HRS, Survey Research Center (SRC), and the Institute for Social Research (ISR) at the University of Michigan, and the RAND Corporation. The 2009 Internet Survey Final Release (Version 1.0) collected information on a number of topical areas including: health (physical and mental, health behaviors); psychosocial items; economics (income, assets, expectations, and consumption); and retirement. The 2009 Internet Survey is the fourth in a series of surveys conducted on the Internet. Completed interviews were obtained from 4,433 HRS respondents. The primary identification variables for these files are HHID and PN.

2007 Disability Vignette Study (DVS)

The 2007 Disability Vignette Study (DVS) includes a short sequence of questions about the respondents' own health and disability status, followed by a set of anchoring vignettes. The vignettes provide short descriptions of people in different states of health, which respondents are asked to rate on the same dimensions and scales as they rated their own health. Comparison of respondents' ratings of their own health and their vignette ratings allows researchers to evaluate whether respondents exhibit different response styles. Two versions of the DVS questionnaire (A and B) were administered to assess question ordering and gender effects. The DVS Sample consisted of respondents who had completed a self-interview in the HRS 2006 Core and who (prior to the start of the DVS field period) had not died or requested removal from the sample, and who were not in the HRS 2007 Consumption and Activities Mail Survey (CAMS) or the HRS 2007 Prescription Drug Study (PDS). Of the 5,678 questionnaires mailed in the fall of 2007, 4,639 were returned. The primary identification variables for these files are HHID and PN.

2007 Internet Survey

Beginning in June 2007, a sub sample (N=3,721) of the Health and Retirement Study (HRS) participated in an Internet based survey developed jointly by the HRS, Survey Research Center (SRC), Institute for Social Research (ISR), at the University of Michigan, and the RAND Corporation. Completed interviews were obtained from 2,665 HRS respondents. The 2007 Internet Survey Final Release (Version 1.0) collected information on a number of topical areas, including Internet/computers, health and emotional prob-

lems, Social Security, numeracy items, psychosocial items, household composition, expectations, questions about housing/checking accounts, and stocks and prescription drug usage and insurance. The 2007 Internet Survey is the third in a series of surveys that will be conducted on the Internet. The primary identification variables for these files are HHID and PN.

2006 Internet Survey

The 2006 Internet Survey is the second in a series of surveys that will be conducted on the Internet. In March 2006, a sub sample (N=1,920) of the Health and Retirement Study (HRS) participated in an Internet based survey developed jointly by the HRS, Survey Research Center, Institute for Social Research (ISR), at the University of Michigan, and the RAND Corporation. The National Institute on Aging at NIH (R01 AG020638) provided funding for the 2006 Internet Survey. The 2006 Internet Survey Final Release (Version 1.0) questionnaire contained several different topical areas, including Internet and computers, health and emotional problems, Social Security, numeracy items, psychosocial items, expectations, questions about housing, checking accounts, stocks, prescription drug usage, and insurance. Completed interviews were obtained from 1,352 HRS respondents. The primary identification variables for these files are HHID and PN.

2003 Internet Survey

In December 2002, a sub-sample of the Health and Retirement Study (HRS) participated in an Internet based survey developed jointly by the HRS, Survey Research Center, Institute for Social Research (ISR), at the University of Michigan and the RAND Corporation. The National Institute on Aging at NIH (R01 AG020638) provided funding for the 2003 Internet Survey. The 2003 Internet Survey Final Release (Version 1.0) questionnaire contained several different topical areas including Internet/computers, health problem, disability and work limitations, numeracy items, psychosocial items, expectations, and questions about housing/checking accounts, and stocks. Many of the questions were taken from the HRS survey, the National Health Interview Survey (NHIS) and from the Current Population Survey (CPS). Completed interviews were obtained from 2,197 HRS respondents. The primary identification variables for these files are HHID and PN.

2001 Human Capital Mail Survey (HUMS)

In the fall of 2001, HUMS questionnaires were mailed to 3,862 households interviewed in the HRS 2000 core survey, and who were likely to have had at least one child (ever) 18 years of age or older. If a selected household had two panel members in it, one panel member was randomly selected to receive the questionnaire. The survey included questions about parental economic investment in the education of children and about children's educational attainment and the costs associated with attending college. The data include one record per child in a given household for which responses were obtained. The primary identification variables are HHID, PN, and OPN.

2001 HUMS College Tuition Imputations

The 2001 HRS HUMS contained one section, referred to as Section H: Child Educational Attainment and Expenses. This section contains questions about high school attendance, college attendance (number, duration, type), and the proportion of college expenses (tui-

tion, food, housing) paid for by the parent of the child. Respondents answered the questions for each child (up to and including 10 children) N: 10,437 (3,031 questionnaires) The primary identification variables are HHID, PN, and OPN.

1999 HRS Mail Survey

An experimental data collection was conducted in 1999 with a sample of 2,998 respondents from the 1998 interviews. Questionnaires were returned by 2,454 of these sample members. Most of the questions replicate those asked in the core interviews. In addition, there are questions about preferences for different possible spending patterns over time. The primary identification variables for this file are HHID and PN.

2.1.7 Sensitive Health Data Products

The Sensitive Health Data Product files are available under terms of a supplemental registration system that requires special download procedures. See HRS Special Access Data for additional information. A brief description of HRS Sensitive Health Data Products is included below. For more information about these files, see "What's Available (Public)" at the "Data Products" section at our Web site.

2006 Biomarker Data

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 preselected for the enhanced face-to-face interview in 2006. Selected respondents who completed an in person self-interview (as opposed an interview done with a proxy), 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 file contains data pertaining to the administration and analysis of the blood spots. The test results available to date include Hemoglobin A1c, total cholesterol and HDL cholesterol. Analyses for C-reactive protein and cystatin C are forthcoming, and a new version of this file will be issued when those analyses have been completed. The file contains 8,392 records. The primary identification variables for the file are HHID and PN.

2005 Prescription Drug Survey (PDS)

The HRS 2005 Prescription Drug Study is the first wave of a two-wave mail survey designed to track changes in prescription drug utilization as Medicare Part D, the prescription drug benefit, is phased in. The baseline wave, administered in 2005, was intended to capture prescription drug use, coverage, and satisfaction prior to the implementation of Medicare Part D, as well as awareness of the new drug benefit and available subsidies. In addition, questions about sources of information on Medicare Part D and expectations of the impact of Part D on prescription drug cost, coverage, and health were asked. The second wave, scheduled for 2007, will capture similar information post-implementation. The sample for the Prescription Drug Study (PDS) was drawn from respondents to HRS 2004. The study sample included HRS respondents born in 1942 or earlier (65th birthday in 2007), or already covered by Medicare or Medicaid at some time between 2002 and

2004. Respondents interviewed by proxy, and those in nursing homes were eligible for inclusion. A sample of 5,654 persons was drawn from the eligible respondents, with oversamples of persons lacking prescription drug coverage or having low income and wealth. After attempting contact with the sample, and based on partial information from HRS 2006 interviews, we determined that 340 persons died prior to the October 2005 start of the first wave of the Prescription Drug Study and were determined to be ineligible for inclusion in the sample. Of the 5,314 remaining eligible cases, 4,684 returned questionnaires or completed a telephone interview. The primary identification variables for these files are HHID and PN.

2003 Diabetes Study

The 2003 Diabetes Study is a supplemental study on diabetes conducted in 2003 and 2004. The study was conducted by mail with a sample of persons reporting diabetes in the 2002 core wave. The study was administered in two stages: first, a self-administered questionnaire about diabetes care, self-management, and health care utilization, and second, a mail-in kit with a finger-stick dried blood spot sample to measure levels of hemoglobin A1c. Of the 3,194 respondents who reported a diagnosis of diabetes in the 2002 HRS, 2,381 were eligible to participate in the study after exclusions for death or random assignment to another study. Of these eligible cases, 1,901 returned a questionnaire and 1,233 had valid laboratory data. The primary identification variables for these files are HHID and PN.

The Aging, Demographics, and Memory Study (ADAMS) Tracker File

The Aging, Demographics, and Memory Study (ADAMS), a supplement to the Health and Retirement Study, was funded by the National Institute on Aging with the specific aim of conducting a population-based study of dementia. HRS formed a partnership with a research team led by Brenda L. Plassman, Ph.D., director of the Epidemiology of Dementia Program at the Duke University Medical Center. The purpose of this collaboration was to conduct in-person clinical assessments for dementia on selected HRS respondents in order to gather information on their cognitive status. A diagnosis of dementia, cognitive impairment but not demented (CIND), or non-case was assigned on the basis of this assessment. Prior community-based studies of dementia have focused on a particular geographical area or have been based on nationally distributed samples that are not representative of the population. This study is the first of its kind to conduct in-home assessments of dementia in a national sample that is representative of the U. S. elderly population. The ADAMS Tracker file contains demographic variables, sample weights, and field outcomes for Waves A, B, and C for the entire ADAMS sample of 1,700 respondents.

The Aging, Demographics, and Memory Study (ADAMS) Wave A

Initial ADAMS assessments (Wave A) were completed for 856 subjects between August 2001 and December 2003. The primary identification variables for most of these files are HHID and PN. All sections of Wave A are currently available.

       Section AB: Assessment Data Part I - Initial Visit (Respondent)
       Section AC: Clinical History - Initial Visit (Respondent)

Section AD: Assessment Data Part II: Dementia Diagnoses/Summary Scores - Initial Visit (Respondent)
Section AE: Medications - Initial Visit (Respondent)
Section AE: Medications - Initial Visit (Drug)
Section AF: Family History - Initial Visit (Respondent)
Section AF: Familyi History - Initial Visit (Sibling)
Section AF: Family History - Initial Visit (Child)
Section AG: Informant Caregiving Questionnaire - Initial Visit (Respondent)
Section AH: Medical Conditions - Initial Visit (Respondent)
Section AH: Medical Conditions - Initial Visit (Condition)
Section AJ: Dementia Checklist & Neurological Exam - Initial Visit (Respondent)
Section AM: Medical History - Initial Visit (Respondent)
Section AN: Assessment Data Part III: Neuropsychological Exam - Initial Visit (Respondent)

The Aging, Demographics, and Memory Study (ADAMS) Wave B

Wave B contains follow-up assessments completed for 252 Wave A respondents between November 2002 and March 2005. Wave B assessments were attempted for 333 respondents - those with a diagnosis of "cognitive impairment, not demented," mild dementia, or borderline normal cognition and for whom longitudinal information would likely clarify the diagnosis. The primary identification variables for most of these files are HHID and PN. All sections of Wave B, analogous to those included in Wave A, are currently available.

The Aging, Demographics, and Memory Study (ADAMS) Wave C

Wave C contains follow-up visits for all Wave A respondents who were still alive in 2006 and not previously diagnosed with dementia. A total of 466 respondents were targeted for these follow-up visits. A Wave C assessment was completed for 315 respondents between June 2006 and May 2008. The primary identification variables for these files are HHID and PN. Two sections of Wave C are currently available.
Section CC: Clinical History - Follow-up Visit (Respondent)
Section CD: Assessment Data Part II: Dementia Diagnoses/Summary Scores - Follow-up Visit (Respondent)

2.2 Levels of Core and Exit Files

Most questions were asked of all respondents. Some questions were asked about the household -- for two-respondent households, these questions were asked of a designated financial respondent, family respondent, or coversheet respondent (the first respondent interviewed) on behalf of the entire household.

In addition to household-level and respondent-level files, the 1995, 1996, 1998, 2000, 2002, 2004, 2006, and 2008 Core Data Releases contain files at five other levels: household-member-and-child, sibling, helper, transfer-to-child and transfer-from-child. The 2002 Core Data Release contains files at two additional levels: assets and pensions. The file structures for 1992, 1993 Final Core Data Releases and 1994 Final Core/Exit Data Release were slightly different. The files are described below.

▫ Household-Level Files

Household-level files contain questions that were asked about the household of a designated coversheet, financial or family respondent. A coversheet respondent answered family questions on behalf of the entire household. Similarly, a family respondent answered family questions on behalf of the entire household, and a financial respondent answered household-level financial questions on behalf of the entire household. The household-level files contain one record for each household in which at least one interview was obtained. These files are available for core files for all years.

▫ 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. These files are available for core and exit files for all years.

▫ Household-Member-and-Child-Level Files

These files contain information provided by the family respondent or financial respondent about each household member or child of the respondent or of the respondent's spouse or partner. The files contain one record for each household member (other than the respondent or the respondent's spouse or partner) or child. These files are available for each year's core and exit files, except 1992 and 1994.

In 2002, 2004, 2006, and 2008 the household-member/child files contain a separate record for each child, child's spouse/partner, and other household members. All records in the 2002, 2004, 2006, and 2008 household-member/child files are *individual* records.

In prior years, 1993, 1995, 1996, 1998, 2000, information about a non-resident child's spouse/partner is contained in the non-resident child's record while each resident, whether a child, spouse/partner of child, or other resident, has a separate record. In other words, for non-resident children, the records in these files are *couple* records while for residents, they are *individual* records.

▫ Household Members-Level File

These files contain one record for each respondent, resident child, or other household member. These files are available for 1992 and 1994.

▫ Children-Level File

These files contain one record for each resident or non-resident child. These files are available for 1992 and 1994.

▫ Sibling-Level Files

These files contain information about some of the respondent's siblings. Detailed questions about siblings were only asked if the respondent's mother or father was alive this wave or was alive the previous wave.[5]

In 2000 and earlier years, the designated family respondent reported on his/her own siblings and also on the siblings of his/her spouse or partner.[6] The 2002 and subsequent years each respondent reported on his/her own siblings.

For most years[7], the files contain one record for each sibling *about whom detailed questions were asked*. These files are available for 1992, 1994, 1996, 1998, 2000, 2002, 2004, 2006, and 2008 core files.

▫ Helper-Level File

This file contains information provided by each respondent about helpers other than the respondent's spouse or partner. A helper is a person or organization reported by the respondent as providing help with ADLs or IADLs. If the helper was a child, the record contains information about the help provided by the child and the child's spouse or partner, if any. The file contains one record for each helper (or, if the helper is a married or partnered child, the helping couple) for each respondent reporting help. If a child helped both mother and father, the 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. These files are available for each year's core and exit files, except 1992 and 1994.

▫ Transfer-to-Child-Level File

This file contains information provided by the family respondent about transfers of money to a child or grandchild. The file contains one record for each transfer to a child or grandchild. These files are available for each year's core and exit files, except 1992, 1993, and 1994.

▫ Transfer-from-Child-Level-File

This file contains information provided by the family respondent about transfers of money from children or grandchildren. The file contains one record for each transfer from a child or grandchild. These files are available for each year's core and exit files, except 1992, 1993, and 1994.

▫ Parents-Level File

The 1992 Parents file contains two records per respondent, one for the respondent's mother and one for the respondent's father. The 1994 Parents file is more complex. It may contain records for mothers, fathers, parents, mothers and stepfathers, fathers

---

[5] Detailed questions were also asked about siblings of a spouse or partner who died during the course of the study, if the respondent did not have a new spouse or partner, and if the deceased spouse's or partner's mother or father was alive this wave or was alive the previous wave.

[6] If the designated family respondent did not provide an interview, information from the non-family respondent was used for PR sibling file.

[7] For 2002 and 2004, the number of records in the sibling files, e.g., H02PR_SB and H02F_SB differ. Also in the file H02F_SB, one sibling may appear in two records, once as reported by the respondent and once as reported by the respondent's spouse/partner. See biannual data descriptions for additional information.

and stepmothers, stepfathers, or stepmothers of family respondents and analogous re-cords for in-laws. See also Parent Data across Time. These files are available for 1992 and 1994.

▫ Pension-Level File

This file contains one record for each pension reported by the respondent. In prior years, these data would have been part of Section GG (Last Job) and Section GH (Job History). These files are available for 2002 core.

▫ Asset-Level File

This file contains one record for each asset eligible for "Asset Reconciliation". It contains the answers respondents give for assets that appear to be discrepant on their preload and/or asset and income counterparts. These files are available for 2002 core.

2.3 Identification Variables for Core and Exit Files

The primary identification variables, those used to uniquely identify a record in the file, for core and exit files are listed in the tables below.

	Table 2: Unique Identifiers for Core Files										

Level	1992	1993	1994[8]	1995	1996	1998	2000	2002	2004	2006	2008
Respondent	HHID PN	HHID PN	HHID PN	HHID PN	HHID PN	HHID PN	HHID PN	HHID PN	HHID PN	HHID PN	HHID PN
Household	HHID ASUBHH	HHID BSUBHH	HHID CSUBHH	HHID DSUBHH	HHID ESUBHH	HHID FSUBHH	HHID GSUBHH	HHID HSUBHH	HHID JSUBHH	HHID KSUBHH	HHID LSUBHH
Household Member and Child		HHID BSUBHH OPN		HHID DSUBHH OPN	HHID ESUBHH OPN	HHID FSUBHH OPN	HHID GSUBHH OPN	HHID HSUBHH OPN	HHID JSUBHH OPN	HHID KSUBHH OPN	HHID LSUBHH OPN[9]
Sibling	HHID ASUBHH OPN		HHID CSUBHH OPN		HHID ESUBHH OPN	HHID FSUBHH OPN	HHID GSUBHH OPN	HHID PN OPN	HHID PN[10] OPN	HHID PN OPN	HHID PN OPN[11]
Helper		HHID PN OPN HLPTYPE		HHID PN OPN DHLP-TYPE	HHID PN OPN	HHID PN OPN	HHID PN OPN	HHID PN OPN	HHID PN OPN	HHID PN OPN	HHID PN OPN[12]
Transfer to				HHID DSUBHH OPN DTRAN-NUM	HHID ESUBHH OPN ETYPTRN	HHID FSUBHH OPN GTRAN-NUM	HHID GSUBHH OPN FTRAN-NUM	HHID HSUBHH OPN HTC_NDX*	HHID JSUBHH JTC_NDX*	HHID KSUBHH OPN[13]	HHID LSUBHH OPN[14]

---

[8] The 1994 files include both core and exit interviews.
[9] There are 53 duplicates in H08E_MC.
[10] PN not included in the final release version of H04PR_SB. See April 02, 2007 data alert for information to add PN.
[11] There are 3 duplicates in H08F_SB.
[12] There are 2 duplicates in H08G_HP.
[13] There are 250 duplicates in H06E_TC.
[14] There are 200 duplicates in H08E_TC.

Level	1992	1993	1994[8]	1995	1996	1998	2000	2002	2004	2006	2008
Transfer from				HHID DSUBHH OPN DTRAN-NUM	HHID ESUBHH OPN	HHID FSUBHH OPN GTRAN-NUM	HHID GSUBHH OPN FTRAN-NUM	HHID HSUBHH OPN HFC_NDX*	HHID JSUBHH JFC_NDX*	HHID KSUBHH OPN[15]	HHID LSUBHH OPN[16]
Children	HHID ASUBHH OPN		HHID CSUBHH OPN								
Household Members	HHID ASUBHH OPN		HHID CSUBHH OPN								
Parents	HHID PN ATYPE-PAR		HHID PN CPAR-CODE CSUBHH								
Pension								HHID PN HZ139*			
Asset								HHID HSUBHH TYP-ASST*			

* Numeric variable

---

[15] There are 5 duplicates in H06E _FC.
[16] There are 7 duplicates in H08E_FC.

Level	1995	1996	1998	2000	2002	2004	2006	2008
Respondent	HHID PN	HHID PN	HHID PN	HHID PN	HHID PN	HHID PN	HHID PN	HHID PN
Helper	HHID PN OPN	HHID PN OPN	HHID PN OPN	HHID PN OPN	HHID PN OPN	HHID PN OPN	HHID PN OPN	HHID PN OPN[17]
Transfer to	HHID NSUBHH N1522* NTYP-TRAN*	HHID PSUBHH OPN PTYP-TRAN*	HHID QSUBHH OPN QTRAN-NUM*	HHID RSUBHH OPN RTRAN-NUM*	HHID SSUBHH OPN STC_NDX*	HHID TSUBHH TTC_NDX *	HHID UTUBHH OPN[18]	HHID VSUBHH OPN[19]
Transfer from	HHID NSUBHH N1475* NTYP-TRAN*[20]	HHID PSUBHH OPN PTYP-TRAN*	HHID QSUBHH OPN QTRAN-NUM*	HHID RSUBHH OPN RTRAN-NUM*	HHID SSUBHH OPN SFC_NDX*	HHID TSUBHH TFC_NDX *	HHID USUBHH OPN[21]	HHID VSUBHH OPN[22]
Household Members and Children	HHID NSUBHH OPN	HHID PSUBHH OPN	HHID QSUBHH OPN	HHID RSUBHH OPN	HHID SSUBHH OPN	HHID TSUBHH OPN	HHID USUBHH OPN	HHID VSUBHH OPN

* Numeric variable

---

[17] There are 2 duplicates in X08G_HP.

[18] There are 21 duplicates in X06E_TC.

[19] There are 24 duplicates in X08E_TC.

[20] The 1995 exit final release file X95D_FC contains one duplicate record.

[21] There are 4 duplicates in X06E_FC.

[22] There are 2 duplicates in X08E_FC.

2.4 Merging HRS Files

Many analyses require variables that appear in separate files. Sometimes you will need to obtain variables from files at different levels that contain different numbers of records. Before you can do your analysis work, the files will need to be merged in an appropriate manner. For an introduction to merging HRS files, with SPSS, SAS and Stata examples, see the document "An Elementary Cookbook of Data Management using HRS Data with SPSS, SAS and Stata Examples" (June 2004) on the "User Guides" page at the "Documentation" section at our Web site.

2.5 File Naming Conventions

2.5.1 File Prefixes

For the years 1995, 1996, 1998, 2000, 2002, 2004, 2006, and 2008, Core files are named beginning with a prefix indicating the study and the year, e.g., "H02" for HRS 2002. Next a letter or two designating the questionnaire section. Next a separator, "_" and then one or two letters designating the level:

> H for household-level
> R for respondent-level
> MC for household-member-and-child-level
> SB for sibling-level
> HP for helper-level
> TC for transfer-to-child-level
> FC for transfer-from-child-level
> A for asset-level (2002 Core)
> P for pension-level (2002 Core)

Putting it all together, files named H02A_R include variables from 2002 section A at the respondent-level. Other years and other products follow different naming conventions. A complete listing of all public data files is provided at the end of this document.

2.5.2 File Extensions

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
> .TXT for codebook files

One of each of these file types is provided for each of parts of the data release. For example:

> H02A_R.DA contains 2002 respondent data from section A
> H02A_R.SAS contains corresponding SAS program statements

H02A_R.SPS contains corresponding SPSS program statements
H02A_R.DO contains corresponding Stata do statements
H02A_R.DCT contains corresponding Stata dictionary statements
H02A_R.TXT contains the ASCII codebook

3 Program Statements

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

The statement files are named beginning with the same prefix as the corresponding data file. For example, SAS statements in the file H02A_R.SAS go with the H02A_R.DA data file.

3.1 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.

o   If the .DA files are located in the directory specified in the .SAS files and you wish to write the SAS system file to the directory specified, you can run the file as is.
o   If you wish to specify other directories, you will need to edit the .SAS file to reflect the proper path names prior to running the file.

A SAS system file (SAS7BDAT) will be saved to the specified directory.

3.2 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.

o   If the .DA files are located in the directory specified in the .SPSS files and you wish to write the SPSS system file to the directory specified, you can run the file as is.
o   If you wish to specify other directories, you will need to edit the .SPSS file to reflect the proper path names prior to running the file.

A SPSS system file (.SAV) will be saved to the specified directory.

3.3 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.

o If the .DA files are located in the directory specified in the .DCT and .DO files and you wish to write the Stata system file to the directory specified, you can run the file as is.
o If you wish to specify other directories, you will need to edit the .DCT and .DO files to reflect the proper path names prior to running the file.

A Stata system file (.DTA) will be saved to the specified directory.

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.

4 Longitudinal Issues

4.1 Sample Design

As of 2008, the HRS sample was comprised of five sub-samples or cohorts.

o 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.

o 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.

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

o The CODA (Children of the Depression Age) sub-sample consists of people who were born 1924 through 1930, were household residents of the conterminous U.S.

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

o   The EBB (Early Baby Boomer) sub-sample consists of people who were born in 1948 through 1953, were household residents of the conterminous 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 will be interviewed every two years thereafter.

o   The MBB (Middle Baby Boomer) sub-sample consists of people who were born in 1954 through 1959, were household residents of the conterminous U.S. when first interviewed in 2010, and who, at that time, did not have a spouse or partner who was born before 1954, and 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 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), and 2010 (MBB). For more details about the sample, see "Survey Design" at the "Documentation" section at our Web site.

4.2 Linking Respondents across Time

Respondent records from all waves and from the tracker file may be linked by HHID and PN. You should determine whether you wish the resulting dataset to contain the union of records (the output file contains a record for a respondent with a record in any of the input files) or intersection of records (the output files contains a record only for respondents with records in specified input files). When merging with the tracker file, some records, called overlap cases, from the 1992 biannual files will not be matched.[23]

---

[23] Each respondent or potential respondent is in the tracker file only once and is uniquely identified by HHID and PN. However a few respondents had different HHIDs and PNs in one or more earlier waves of the study. These are called overlap cases. They may require special handling in constructing longitudinal files and in merging tracker to wave-specific files. The current HHID and PN reflect the current status of the case. Overlap cases also have a former HHID and PN in some previous wave and these are given in the variables OVHHID and OVPN. You should be aware of these situations when merging the biannual files with the tracker file. There are two basic types of overlaps.

➤   First, there were a number of original HRS 1992 (Wave 1) households eligible to be either an HRS or AHEAD household. An interview was attempted for all of them in HRS 1992. Afterwards, a random sub-sampling was performed, with 60% of the cases staying in HRS and the remaining going to AHEAD. We refer to the 134 cases transferred to AHEAD as HRS "inter-study overlap". Among those 134 cases assigned to AHEAD, 109 were actually interviewed in AHEAD, and 25 were not. These respondents were eligible to be interviewed with a base-line interview both in 1992 and in 1993. From 1993 onward they were considered part of the AHEAD sub-sample.

The sub-household identifiers can be used to link cross-sectional household data with cross-sectional respondent data prior to linking respondent data longitudinally. The sub-household variables for each of the years for core and exit are listed below.

Year	SUBHH - Core	SUBHH – Exit	Sub-samples	Wave
1992	ASUBHH	--	HRS	Wave 1
1993	BSUBHH	--	AHEAD	Wave 1
1994	CSUBHH	--	HRS	Wave 2
1995	DSUBHH	NSUBHH	AHEAD	Wave 2
1996	ESUBHH	PSUBHH	HRS	Wave 3
1998	FSUBHH	QSUBHH	HRS AHEAD CODA WB	Various
2000	GSUBHH	RSUBHH	HRS AHEAD CODA WB	Various
2002	HSUBHH	SSUBHH	HRS AHEAD CODA WB	Various
2004	JSUBHH	TSUBHH	HRS AHEAD CODA WB EBB	Various
2006	KSUBHH	USUBHH	HRS AHEAD CODA WB EBB	Various
2008	LSUBHH	VSUBHH	HRS AHEAD CODA WB EBB	Various
2010	MSUBHH	WSUBHH	HRS AHEAD CODA WB EBB MBB	Various

4.3 Parent Data across Time

Over the course of HRS and AHEAD data collection and distribution, variables about parents have been asked and distributed in a variety of different ways.
- o For 1992 (HRS sub-sample) the distribution format is one record per parent.
- o For 1994 (HRS sub-sample) the distribution format includes many types of records.
- o For 1993 and 1995 (AHEAD sub-sample) and for 2002, 2004, 2006 and 2008 (combined sub-samples) the distribution format is one record per respondent with information about two parents.
- o For 1996 (HRS sub-sample) and for 1998 and 2000 (combined sub-samples) the distribution format is one record per household with information about four parents.

➢ Second, there are two cases (as of the 2008 version of the tracker) belonging to what we call "household merge overlaps" which result from intermarriage between respondents who entered the study in separate households.

While it is possible to retrieve the information from the earlier waves by using the values of OVHHID and OVPN, since the number of cases is small -- 109 respondents in 1992, two respondents in 1993 and 1995, one respondent in 1998 and 2000 -- for most types of analysis it is reasonable to retain only the information from the respondent's most recent "incarnation". If you choose this option, you should be aware that some 1992 biannual records (or for any of the years prior to "intermarriage" for the two overlaps resulting from intermarriage) will not be matched by a record in the tracker file when matching by HHID and PN. If you wish to obtain the 1992 (or prior to intermarriage) data, see " Merging the Tracker with other HRS Data " in the "Cross-Wave Tracker File Data Description" available at the "Data Descriptions and Release Notes " page in the "Documentation" section at our Web site.

To facilitate longitudinal analysis, the analyst may wish to create files using a standard format for parent information for all waves of data. For detailed information about creating files with one record per respondent for each year, see "Creating Parent Files with One Record per Respondent" at "Family Data" at the "Documentation" section at our Web site. The tasks required to create files with a standardized format of one record per respondent including both mother and father information from the distribution files are described, in brief, in this document. Detailed descriptions are provided in companion documents.

4.4 Linking Other Persons across Time

Linking other person records across time is more complex. Other person records are uniquely identified by HHID, nSUBHH, and OPN where nSUBHH is the current wave's sub-household identifier. In the event of split-households, there may be more than one report about a single other person in a given wave. The Longitudinal Other Person Number files have been prepared to facilitate the merging of information obtained about children and other household members in different waves of the HRS study. Two variables, HHID and LOPN, may be used to uniquely identify children and other household-members across all waves of the study.[24] For more detailed information, see "Longitudinal Other Person Number (LOPN) Data Description" at "Data Descriptions and Release Notes" at the "Documentation" section at our Web site.

Alternatively, depending on your analytic needs, you may link other person records longitudinally through a series of sequential steps linking HHID, xSUBHH (the current-year's sub-household), and OPN to HHID, ySUBHH (the previous-year's sub-household), and OPN[25]. Contact us for additional information.

4.5 Summary of Changes from Early Waves

Over the course of the study there have been changes in identification variable storage mode and in INAP codes.

4.5.1 Character Type Identification Variables

In some prior releases of HRS, all variables were stored in NUMERIC format. Releasing identification variables in numeric format caused some problems. Thus starting with 1995 data collections, we released all primary and secondary identification variables in character format and are no longer creating combined identification variables[26]. The HRS Wave 1 (1992) and Wave 2 (1994) and AHEAD Wave 1 (1993) files have been re-

---

[24] The Version 1 LOPN files were constructed using files publicly available in January 2004.

[25] Be aware that this technique of matching OPN records by HHID, previous wave SUBHH and OPN to track children and household-members across waves of the study does not work well for persons who assumed the OPN number of their deceased spouse or partner during the 1993 to 2000 waves, for spouses or partners assigned a new OPN in 2002, and for persons with more than one OPN or for OPNs used by more than one person

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

released with character identification variables. As of this writing, no releases still include numeric identification variables.[27]

4.5.2 INAP Codes Stored as Blanks

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

5 Obtaining the Data

HRS data are available free to researchers and analysts at the HRS Web site. Most HRS data are released as public data and are available after a simple registration process, described below, has been completed. The additional requirements to obtain other data collections are also described.

5.1 HRS Public Data

In order to obtain public release data, you must first register at the "Access to Public Data" page at the "Data Products" section 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 public data files. The information you provide will not be used for any commercial use, and will not be redistributed to third parties.

---

[27] When merging datasets, most software packages require that the identification variables in each dataset be of the same type. We suggest that analysts of older HRS or AHEAD datasets convert the old, numeric identifiers to character type.   Each software package has different ways of converting variable types in an existing dataset.

 In SAS, for example, to convert numeric variable HHID to character variable HHID, you can use code such as this.
```
        * numeric to character;
        CHHID=put(HHID,z6.0);
        Attrib CHHID label=' HOUSEHOLD IDENTIFIER' length=$6 format=$char6.;
        drop HHID;
        rename CHHID=HHID;
```

In STATA to convert numeric variable HHID to string variable HHID, you can use code such as this.
```
        * numeric to string
        gen str6 chhid=substr(string(hhid+1000000),2,6)
        drop hhid
        rename chhid hhid
```

In SPSS, to convert a numeric variable to a string variable, use the STRING() function.
```
        string chhid (a6).
        compute chhid = string(hhid,n6).
        delete variables hhid.
        rename variable (chhid = hhid).
```

5.2 HRS Sensitive Health Data

The Health and Retirement Study strives to provide high quality data without compromising respondent confidentiality. Since respondent health data records contain particularly sensitive information, such data products are released to researchers who qualify for access only through a supplemental registration system. See details about the application process under "Sensitive Health Data" at the "Data Products" section at our Web site. To apply:

- o You must be a registered HRS user (see preceding section).
- o You must download an application and mail it to the address provided.
- o Once you receive your approval notification, login at the "HRS Public File Download Area," go to the "Download Data Products" page where you will then see links to the data products that you have requested.

5.3 Conditions of Use for HRS Public and Sensitive Health Data

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

- o Make no attempts to identify study participants
- o Not to transfer data to any third party except as specified
- o Not to allow others to use your username and password
- o To include specified citations in work based on HRS data
- o To provide us information about publications based on HRS data
- o To report apparent errors in the HRS data or documentation files
- o To 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 "Access to Public Data" page at the "Data Products" section at our Web site.

5.4 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
> Room 3050 ISR
> P.O. Box 1248
> Ann Arbor, MI 48106-1248

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

5.5 HRS Restricted Data

HRS Restricted data files contain sensitive respondent information. Restricted data products include: SSA (Social Security Administration) data, geographic information, detailed industry/occupation information, pension estimation data, health care information including information from Medicare claims and the National Death Index data, and other miscellaneous data.

These files are only available under terms of a formal agreement negotiated between the researcher and HRS. The two crucial requirements that you must meet in order to be granted access to HRS restricted data are 1) you must be affiliated with an institution that has an NIH-approved Human Subjects Review Process via a Multiple Projects Assurance and 2) you must be a Principal Investigator or Co-Principal Investigator on a project funded by a current United States Government research grant or contract.   See details about the application process under "Restricted Data" at the "Data Products" section at our Web site. Alternatively, the Michigan Center on the Demography of Aging (MiCDA) Data Enclave located in Ann Arbor may be used by researchers who cannot meet the usual restricted data requirements.

6 If You Need to Know More

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

6.1 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/

The Quick links provide direct access to:

> Genetics
> Bibliography
> Question Concordance
> Data Collection Path
> Data Product List
> Study Content User Guides
> Document Finder
> News and Events
> Data Alerts
>
> Site Map
> Help Desk
> How to contact us
> Other Research Sites

Details and live links to Documentation, Data Products, and Papers and Publications sections of our Web site are provided below.

6.1.1 Documentation

Document Finder	Keyword search for documentation items on this site.
Questionnaires	Representations of interviews in "box-and-arrow" format.
Concordance	The question concordance is a tool for cross-referencing survey questions by content across time for all biennial interview years. Each record in a retrieval set contains a link to the appropriate codebook section, allowing the user to see data element details (jump instructions, variable description, question text, user notes, and codeframe). For years 1995 and beyond, cross-reference information is available for core and exit interview variables that are appear in multiple waves.
Data Descriptions	Information on how to use public release datasets, including details on content, data manipulation techniques, cross-wave merging and other items of interest.
Codebooks	Links to online HTML codebooks for biennial and off-year data products.
Content	Overview of content by section for each interviewing year (linked to questionnaire).
Modules	Module Descriptions: Overview of experimental module content for each interviewing year (linked to questionnaire).
Imputations	Information on imputation products currently available to the public.
Survey Design	Overview of HRS design and methodology.
Data Alerts	Data release errors, omissions, notes, and corrections.
User Guides	These documents explain the concepts, measures, and questions in the HRS surveys. They expand upon the information found in codebooks, questionnaires and data descriptions.
Family Data	Resources for Analysis of Family Data: Reference materials pertaining to HRS data content related to family issues. Includes direct links to questionnaire areas, codebook content and bibliographic materials, as well as other relevant material.

6.1.2 Data Products

What's Available (Public)	A comprehensive listing of HRS Public and Sensitive Health data products. • Biennial Datasets • Longitudinal Datasets • Off-Year Studies • Sensitive Health data
Access to Public Data *(Registration required)*:	Register and download data products from these categories. • HRS Public Data • Sensitive Health Data (with approved application) • HRS Restricted Data (documentation only) • Researcher Contributions • **RAND Contributions**
Sensitive Health Data	The Health and Retirement Study strives to provide high quality data without compromising respondent confidentiality. Since respondent health data records contain particularly sensitive information, such data products are released to researchers who qualify for access only through a supplemental registration system.
Restricted Data	HRS Restricted data files contain sensitive respondent information. They are only available under terms of a formal agreement negotiated between the researcher and HRS. For a **list of restricted data products** and **application materials**, visit the **HRS Restricted Data** Web site.
Family Data	Resources for Analysis of Family Data: Reference materials pertaining to HRS data content related to family issues. Includes direct links to questionnaire areas, codebook content and bibliographic materials, as well as other relevant material.
Redistribution	Third Party Redistribution Policy: Information for organizations interested in redistributing or archiving HRS data products.

6.1.3 Publications

Online Bibliography	A dynamic interface to the HRS bibliographic database that allows a variety of searches and queries. A **static list** of all papers and publications based on HRS data, sorted by category and first-author, is available as a supplement to the online database. You may also wish to **search for HRS citations** through Google Scholar and other online indexes. Summary tables of bibliography database contents by **workform and content area** are also available.
User Guides	These documents explain the concepts, measures, and questions in the HRS surveys. They expand upon the information found in codebooks, questionnaires and data descriptions.
PSC Publications	Access to Population Studies Center resources on aging.
Conference Proceedings	Papers and presentations associated with conferences sponsored by the Health and Retirement Study.
Midterm Reviews	Papers produced by the NIA Data Monitoring Committee reviewing the study at the midpoint of the grant cycle (2002).
HRS in the News	**Newspaper stories**, **magazine articles**, and **press releases** concerning HRS data and findings.
Register Your Paper!	Please inform us of your papers and publications based on HRS data.

6.2 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-mailto: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

7 Listing of All Files Contained in Each of the HRS Public and Special Release Data Products

The file prefixes (file names) and the number of records (Ns) in each of the data files included in the HRS public and special access data products, as of June 2010, is provided below.

Biannual Files

2008HRS EARLY CORE
- H08PR_MC (N=69523)
- H08E_MC (N=66128)
- H08A_R H08B_R H08C_R H08D_R H08F_R H08G_R H08IO_R H08I_R H08J_R H08K_R H08LB_R H08L_R H08M1_R H08M2_R H08N_R H08PR_R H08P_R H08RC_R H08S_R H08TN_R H08T_R H08V_R H08W_R H08Y_R (N=17217)
- H08PR_SB (N=15599)
- H08A_H H08E_H H08H_H H08IO_H H08PR_H H08Q_H H08R_H H08U_H (N=11897)
- H08F_SB (N=6400)
- H08E_TC (N=5521)
- H08G_HP (N=4678)
- H08E_FC (N=926)

2008HRS EARLY COREIMP
- INCWLTH08E1A (N=17217)

2008HRS EARLY EXIT
- X08PR_MC (N=7595)
- X08G_HP (N=2744)
- X08A_R X08B_R X08C_R X08D_R X08E_R X08G_R X08IO_R X08J_R X08N_R X08PR_R X08T_R X08Y_R (N=1333)
- X08E_TC (N=415)
- X08E_FC (N=207)

2008HRS FINAL POSTEXIT
- PX08PR_MC (N=664)
- PX08A_R PX08PR_R PX08T_R (N=118)

2006HRS FINAL CORE
- H06PR_MC (N=72080)
- H06E_MC (N=63958)
- H06A_R H06B_R H06C_R H06D_R H06F_R H06G_R H06IO_R H06I_R H06J_R H06K_R H06LB_R H06L_R H06M0_R H06M1_R H06M2_R H06N_R H06PR_R H06P_R H06RC_R H06S_R H06TN_R H06T_R H06V_R H06W_R H06Y_R (N=18469)
- H06PR_SB (N=17622)
- H06A_H H06E_H H06H_H H06IO_H H06PR_H H06Q_H H06R_H H06U_H (N=12605)
- H06F_SB (N=7473)
- H06E_TC (N=6072)
- H06G_HP (N=4926)
- H06E_FC (N=893)

2006HRS FINAL COREIMP
- INCWLTH06F2A (N=18469)

2006HRS FINAL EXIT
- X06PR_MC (N=7784)
- X06G_HP (N=2707)
- X06A_R X06B_R X06C_R X06D_R X06E_R X06G_R X06IO_R X06J_R X06N_R X06PR_R X06T_R X06Y_R (N=1310)
- X06E_TC (N=432)
- X06E_FC (N=165)

2006HRS FINAL POSTEXIT
- PX06PR_MC (N=822)
- PX06A_R PX06PR_R PX06T_R (N=147)

2004HRS FINAL CORE
- H04PR_MC (N=76284)
- H04E_MC (N=68796)
- H04PR_SB (N=20588)
- H04A_R H04B_R H04C_R H04D_R
  H04F_R H04G_R H04IO_R H04I_R
  H04J_R H04K_R H04LB_R H04L_R
  H04M1_R H04M2_R H04N_R
  H04PR_R H04P_R H04RC_R
  H04S_R H04T_R H04V_R H04W_R
  H04Y_R (N=20129)
- H04A_H H04E_H H04H_H
  H04PR_H H04Q_H H04R_H
  H04U_H (N=13645)
- H04F_SB (N=10873)
- H04E_TC (N=6206)
- H04G_HP (N=5026)
- H04E_FC (N=921)

2004HRS EARLY COREIMP
- H04I_CR H04I_FR H04I_JR
  H04I_KR H04I_LR H04I_M1R
  H04I_M2R H04I_NR H04I_PR
  H04I_SR H04I_TR (N=20139)
- H04I_EH H04I_HH H04I_QH
  H04I_RH H04I_UHH H04I_UHJ
  (N=13651)
- H04I_ETC (N=6196)
- H04I_GHP (N=5026)
- H04I_EFC (N=921)

2004HRS FINAL EXIT
- X04PR_MC (N=6776)
- X04G_HP (N=2493)
- X04A_R X04B_R X04C_R X04D_R
  X04E_R X04G_R X04IO_R X04J_R
  X04N_R X04PR_R X04T_R
  X04W_R X04Y_R (N=1227)
- X04E_TC (N=283)
- X04E_FC (N=153)

2004HRS FINAL POSTEXIT
- PX04PR_MC (N=1870)
- PX04A_R PX04PR_R PX04T_R
  (N=465)

2002HRS FINAL CORE
- H02PR_MC (N=70127)
- H02E_MC (N=65668)
- H02A_R H02B_R H02C_R
  H02D_R H02F_R H02G_R
  H02IO_R H02J_R H02K_R
  H02L_R H02M1_R H02M2_R
  H02N_R H02PR_R H02P_R
  H02RC_R H02S_R H02T_R
  H02V_R H02W_R H02Y_R
  (N=18167)
- H02PR_SB (N=14129)
- H02A_H H02E_H H02H_H
  H02PR_H H02Q_H H02R_H
  (N=12350)
- H02F_SB (N=6196)
- H02E_TC (N=5133)
- H02G_HP (N=4686)
- H02U_A (N=1478)
- H02E_FC (N=921)
- H02PR_P (N=565)

2002HRS FINAL COREIMP
- H02I_CR H02I_FR H02I_GR
  H02I_JR H02I_KR H02I_LR
  H02I_M1R H02I_M2R H02I_NR
  H02I_PR H02I_SR H02I_TR
  (N=18167)
- H02I_EH H02I_HH H02I_QH
  H02I_RH (N=12350)
- H02I_ETC (N=5133)
- H02I_GHP (N=4684)
- H02I_EFC (N=921)
- H02I_UAG H02I_UAH (N=711)

2002HRS FINAL EXIT
- X02PR_MC (N=8292)
- X02G_HP (N=2736)
- X02A_R X02B_R X02C_R
  X02D_R X02E_R X02G_R
  X02IO_R X02J_R X02N_R
  X02PR_R X02T_R X02W_R
  X02Y_R (N=1501)
- X02E_TC (N=306)
- X02E_FC (N=141)

□ H2000XI (N=1348)

2002HRS FINAL POSTEXIT
□ PX02PR_MC (N=2195)
□ PX02A_R PX02PR_R PX02T_R
(N=486)

2000HRS FINAL POSTEXIT
PX00PR_MC (N=644)
PX00CS_R PX00N_R PX00PR_R
(N=359)

2000HRS FINAL CORE
□ H00CS_MC H00D_MC H00J_MC
H00PR_MC (N=46023)
□ H00A_R H00B_R H00CS_R H00C_R
H00D_R H00E_R H00G_R H00H_R
H00J_R H00M_R H00N_R H00PC_R
H00PR_R H00R_R H00T_R
(N=19580)
□ H00D_SB H00PR_SB (N=17590)
□ H00CS_H H00D_H H00F_H H00J_H
H00N_H (N=13214)
□ H00D_TC (N=5839)
□ H00E_HP (N=4156)
□ H00D_FC (N=852)

1998HRS FINAL CORE
□ H98CS_MC H98D_MC H98J_MC
H98PR_MC (N=49013)
□ H98D_SB H98PR_SB (N=21409)
□ H98A_R H98B_R H98CS_R
H98C_R H98D_R H98E_R
H98G_R H98H_R H98J_R
H98M_R H98N_R H98PC_R
H98PR_R H98R_R H98T_R
(N=21384)
□ H98CS_H H98D_H H98F_H
H98J_H H98N_H (N=14395)
□ H98D_TC (N=6176)
□ H98E_HP (N=3455)
□ H98D_FC (N=797)

2000HRS FINAL COREIMP
□ H00I_JMC (N=46023)
□ H00I_ER H00I_NR H00I_RR
(N=19580)
□ H00I_DH H00I_FH H00I_JH
H00I_NH (N=13214)
□ H00I_DTC (N=5839)
□ H00I_EHP (N=4156)
□ H00I_DFC (N=852)

1998HRS FINAL COREIMP
□ H98I_ER H98I_RR (N=21384)
□ H98I_DH H98I_FH H98I_JH
H98I_NH (N=14395)
□ H98I_DTC (N=6176)
□ H98I_EHP (N=3455)
□ H98I_DFC (N=797)

2000HRS EARLY COREWT
□ H00_WGT (N=19581)

1998HRS FINAL EXIT
□ X98PR_MC (N=4056)
□ X98E_HP (N=2160)
□ X98A_R X98B_R X98CS_R
X98D_R X98E_R X98G_R
X98N_R X98PC_R X98PR_R
X98R_R X98TN_R (N=1254)
□ X98D_TC (N=277)
□ X98CS_MC (N=260)
□ X98D_FC (N=84)

2000HRS FINAL EXIT
□ X00PR_MC (N=4262)
□ X00E_HP (N=1941)
□ X00A_R X00B_R X00CS_R X00D_R
X00EV_R X00E_R X00G_R X00N_R
X00PC_R X00PR_R X00R_R
X00S_R X00TN_R (N=1348)
□ X00D_TC (N=317)
□ X00CS_MC (N=280)
□ X00D_FC (N=133)

1998HRS EARLY EXITIMP
X98I_DR X98I_ER X98I_NR (N=1255)

2000HRS EARLY EXITIMP

1998HRS FINAL POSTEXIT

- □ PX98PR_MC (N=668)
- □ PX98N_R (N=241)

1996HRS FINAL CORE
- □ H96CS_MC H96D_MC H96J_MC H96PR_MC (N=25171)
- □ H96D_SB H96PR_SB (N=17974)
- □ H96A_R H96B_R H96CS_R H96C_R H96E_R H96G_R H96H_R H96J_R H96M_R H96N_R H96PC_R H96PR_R H96R_R H96T_R (N=10964)
- □ H96CS_H H96D_H H96F_H H96J_H H96N_H (N=6816)
- □ H96D_TC (N=3901)
- □ H96E_HP (N=799)
- □ H96D_FC (N=305)

1996HRS FINAL COREIMP
- □ H96I_DMC (N=25171)
- □ H96I_ER H96I_RR (N=10964)
- □ H96I_DH H96I_FH H96I_JH H96I_NH (N=6816)
- □ H96I_DTC (N=3901)
- □ H96I_DFC (N=305)

1996HRS FINAL CORESUPP
- □ H96D_R (N=10964)

1996HRS FINAL EXIT
- □ X96PR_MC (N=976)
- □ X96A_R X96B_R X96CS_R X96D_R X96EV_R X96E_R X96G_R X96MD_R X96N_R X96PC_R X96PR_R X96R_R X96T_R (N=234)
- □ X96E_HP (N=109)
- □ X96CS_MC (N=72)
- □ X96D_TC (N=9)
- □ X96D_FC (N=2)

1996HRS FINAL EXITIMP
- □ X96I_DR X96I_ER X96I_NR (N=234)
- □ X96I_EHP (N=109)
- □ X96I_DTC (N=9)
- □ X96I_DFC (N=2)

1995AHD FINAL CORE
- □ A95CS_MC A95D_MC A95J_MC A95PR_MC (N=15617)
- □ A95A_R A95B_R A95CS_R A95C_R A95D_R A95E_R A95G_R A95H_R A95J_R A95M_R A95N_R A95PC_R A95PR_R A95R_R A95T_R (N=7027)
- □ A95CS_H A95D_H A95F_H A95J_H A95N_H (N=5222)
- □ A95D_TC (N=2829)
- □ A95E_HP (N=2136)
- □ A95D_FC (N=829)

1995AHD FINAL COREIMP
- □ A95I_DMC A95I_JMC (N=15617)
- □ A95I_ER (N=7027)
- □ A95I_DH A95I_FH A95I_JH A95I_NH (N=5222)
- □ A95I_DTC (N=2829)
- □ A95I_EHP (N=2136)
- □ A95I_DFC (N=829)

1995AHD FINAL EXIT
- □ X95PR_MC (N=2306)
- □ X95A_R X95B_R X95CS_R X95D_R X95EV_R X95E_R X95G_R X95MD_R X95N_R X95PC_R X95PR_R X95R_R X95T_R (N=775)
- □ X95E_HP (N=764)
- □ X95CS_MC (N=172)
- □ X95D_TC (N=113)
- □ X95D_FC (N=78)

1995AHD FINAL EXITIMP
- □ X95I_DR X95I_ER X95I_NR (N=775)
- □ X95I_EHP (N=764)
- □ X95I_DTC (N=113)
- □ X95I_DFC (N=78)

1994HRS FINAL MAIN
- □ W2KIDS (N=22861)

- W2HHLIST (N=21626)
- W2SIBS (N=17800)
- W2CS (N=13006)
- W2A W2B W2C W2E W2FA W2FB W2FC W2G W2H W2J W2R W2S (N=11596)
- W2PARS (N=8886)
- W2D W2K W2N W2V (N=6979)
- W2MOD4 (N=1561)
- W2MOD7 (N=827)
- W2MOD8 (N=822)
- W2MOD2 (N=817)
- W2MOD1 (N=815)
- W2MOD5 (N=801)
- W2MOD3 (N=771)
- W2MOD0 (N=222)
- W2MOD6 (N=203)
- W2MOD9 (N=179)

1994HRS FINAL IMPUTE
- H94I_KID (N=22930)
- H94I_C H94I_E H94I_S (N=11596)
- H94I_PAR (N=8902)
- H94I_D H94I_K H94I_N H94I_V (N=6979)

1993AHD FINAL
- BOP21 (N=17424)
- BR21 (N=8222)
- BHH21 (N=6047)
- BHP21 (N=3160)

1992HRS FINAL MAIN
- PARENTS (N=25444)
- KIDS (N=24697)
- HHLIST (N=20268)
- SIBLINGS (N=18992)
- EMPLOYER HEALTH MSAFLAG (N=12652)
- HOUSEHLD (N=7607)
- WIDOWED (N=5045)
- MODULEG (N=1124)
- MODULED (N=820)
- MODULEF (N=788)
- MODULEJ (N=763)
- MODULEE (N=745)

- MODULEC (N=698)
- MODULEB (N=692)
- MODULEA (N=662)
- MODULEH (N=560)
- MODULEK (N=229)

1992HRS FINAL IMPUTE
- H92I_D H92I_M H92I_N (N=7607)

Cross-Wave Files

XWAVE TRACKER 2008
- TRK2008 (N=31022)

XWAVE REGIONANDMOBILITY FINAL
- HRSXREGION (N=31022)

XWAVE LOPN
- LC02_MC (N=70126)
- HRS9202 (N=54598)
- LC98_MC (N=49013)
- LC00_MC (N=46023)
- AHD9302 (N=32762)
- LC96_MC (N=25171)
- LC92_K (N=24697)
- LCX94_K (N=22861)
- LC93_MC (N=17424)
- CWB9802 (N=17391)
- LC95_MC (N=15617)
- LCX94_HL (N=8417)
- LX02_MC (N=8297)
- LC92_HL (N=7237)
- LX00_MC (N=4273)
- LX98_MC (N=4031)
- LX95_MC (N=2306)
- LX96_MC (N=976)

XWAVE LABORSECTCARRYFWD FINAL
- CF2002_J (N=18167)
- CF1998_G (N=13113)
- CF2000_G (N=12455)
- CF1994_FA CF1994_FB (N=11596)
- CF1996_G (N=10964)

XWAVE PENSIONIMPUTATION FINAL
- ▫ IP1998G IP1998GG IP1998GH (N=21384)
- ▫ IP2000G IP2000GG IP2000GH (N=19580)
- ▫ IP2002J IP2002K IP2002L (N=18167)
- ▫ IP1992F IP1992G IP1992H (N=12652)
- ▫ IP1994FA IP1994FB IP1994FC IP1994G IP1994H (N=11596)
- ▫ IP1996G IP1996GG IP1996GH (N=10964)

XWAVE PENSIONTRACKEREMPEMPLOYER FINAL
- ▫ PENTRKE (N=15613)

XWAVE PENSIONWEALTHIMPUTATION1992-1998 FINAL
- ▫ IPW (N=15879)

XWAVE PENSIONTRAKER FINAL
- ▫ PENTRKW4 (N=21384)
- ▫ PENTRKW7 (N=20129)
- ▫ PENTRKW5 (N=19579)
- ▫ PENTRKW6 (N=18167)
- ▫ PENTRKW1 (N=12652)
- ▫ PENTRKW2 (N=11420)
- ▫ PENTRKW3 (N=10964)

XWAVE PENSIONWEALTHIMPUTATION2004
- ▫ PWI04 (N=30888)

XWAVE COGNITIVEIMPUTATION
- ▫ COGIMPA_R (N=29016)

XWAVE MASTERID
- ▫ MASTERIDFILE (N=31022)

XWAVE SOCSECWEALTH
- ▫ SSWEALTHP (N=29070)

XWAVE CHILDPROXIMITY
CPROXA_R (N=20910)

CPROXB_R (N=18843)
CPROXC_R (N=17845)

Ancillary Surveys

2009MAIL FINAL 2009CAMS
- ▫ CAMS09_R (N=5330)

2007MAIL FINAL 2007CAMS
- ▫ CAMS07_R (N=5612)

2005MAIL FINAL 2005CAMS
- ▫ CAMS05_R (N=5815)

2003MAIL FINAL 2003CAMS
- ▫ CAMS03_R (N=3254)

2001MAIL FINAL 2001CAMS
- ▫ CAMS01_R (N=3866)

2009INET
- ▫ NET09_R (N=4433)

2007MAIL FINAL 2007DVS
- ▫ DVS07A_R (N=2329)
- ▫ DVS07B_R (N=2310)

2007INET
- ▫ NET07_R (N=2665)

2006INET
- ▫ NET06_R (N=1352)

2003INET
- ▫ NET03_R (N=2197)

2001MAIL EARLY 2001HUMS
- ▫ HUMS01_C (N=10437)

2001MAIL EARLY 2001HUMSIMP
HUMS2001T (N=10437)

1999MAIL EARLY
- ▫ MAILA (N=2454)

Special Access Files

2006HRS EARLY COREBIO
- BIOMKR06 (N=8392)

2005MAIL FINAL 2005PDS
- PDS05E_M (N=17325)
- PDS05S_R (N=5654)
- PDS05A_R (N=4684)

2003MAIL FINAL 2003DIAB
- DIAB2003S (N=3194)
- DIAB2003A (N=1901)

ADAMS1 TRACKER
ADAMS1TRK_R (N=1770)

ADAMS1 WAVEA
- ADAMS1AH_C (N=6060)
- ADAMS1AE_D (N=5765)
- ADAMS1AF_SB (N=4215)
- ADAMS1AF_CH (N=2443)
- ADAMS1AB_R ADAMS1AC_R
  ADAMS1AD_R ADAMS1AE_R
  ADAMS1AF_R ADAMS1AH_R
  ADAMS1AJ_R ADAMS1AM_R
  ADAMS1AN_R (N=856)
- ADAMS1AG_R (N=746)

ADAMS1 WAVEB
- ADAMS1BE_D (N=1793)
- ADAMS1BH_C (N=1358)
- ADAMS1BF_SB (N=1170)
- ADAMS1BF_CH (N=876)
- ADAMS1BB_R ADAMS1BC_R
  ADAMS1BD_R ADAMS1BE_R
  ADAMS1BF_R ADAMS1BH_R
  ADAMS1BJ_R ADAMS1BM_R
  ADAMS1BN_R (N=252)
- ADAMS1BG_R (N=205)

ADAMS1 WAVEC
- ADAMS1CC_R ADAMS1CD_R
  (N=315)