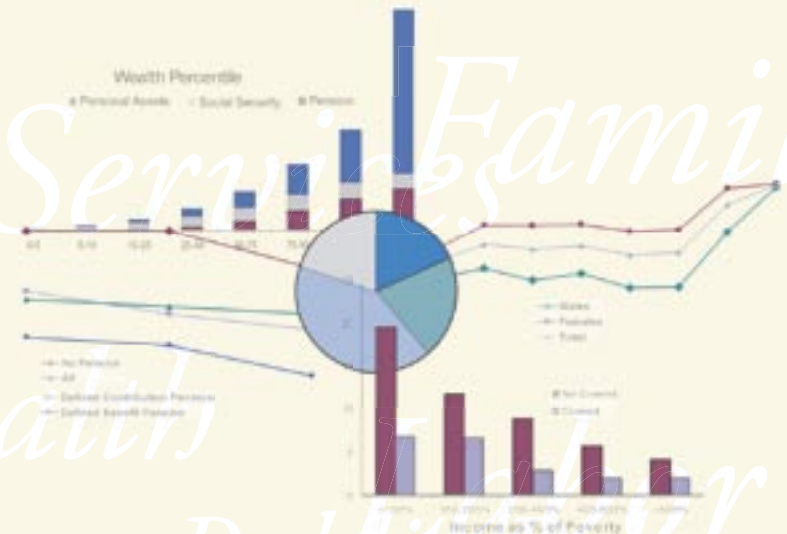
Getting Started with the Health and Retirement Study



Version 1.0

Getting Started with the Health and Retirement Study

Spring 2006

Version 1.0

Compiled and Edited by Cindi Leacock

Excerpts written by HRS Co-Investigators and Staff

Editorial input by: Cathy Liebowitz, Michael Nolte, Janet Keller, Marita Servais, Stacey Bruestle, Patricia St. Clair at RAND and HRS Co-Is

Cover art by Melinda Trombley

How to Contact the Health and Retirement Study

Mail:
Health and Retirement Study
Survey Research Center
Institute for Social Research
University of Michigan
426 Thompson Street
Ann Arbor, MI 48104

Phone: 734-936-0314 Fax: 734-647-1186

Questions? e-mail:
hrsquest@isr.umich.edu

Web site
http:\\hrsonline.isr.umich.edu

Our Primary Sponsor
National Institute on Aging
Richard Suzman, Program Officer
Phone: 301-496-3138 Fax: 301-402-0051

The Health and Retirement Study Co-Investigators

Robert J. Willis
Co-Director,
University of Michigan

David R. Weir
Co-Director,
University of Michigan

F. Thomas Juster
Director Emeritus,
University of Michigan

Charles C. Brown
University of Michigan

Alan L. Gustman
Dartmouth College

John C. Henretta
University of Florida

Daniel H. Hill
University of Michigan

Michael D. Hurd
RAND

Kenneth M. Langa
University of Michigan

John J. McArdle
University of Southern California

Olivia S. Mitchell
University of Pennsylvania

Mary Beth Ofstedal
University of Michigan

Willard L. Rodgers
University of Michigan

Beth J. Soldo
University of Pennsylvania

Thomas L. Steinmeier
Texas Tech University

Robert B. Wallace
University of Iowa

Table of Contents (Hyperlinked to Page)

Introduction

Chapter 1 A Quick Intro to the HRS Web Site

Chapter 2 The HRS Web Site for Data Users

Chapter 3 Using the Web Site to Prepare for Data Analysis

Chapter 4 Other HRS Data Products and Tools

Chapter 5 HRS Glossary & Acronyms

Part Two: HRS RAND DATA using STATA Tutorial

Chapter 6 Using the HRS RAND Data with STATA

Chapter 7 Classical Ordinary Least Squares Regression

Chapter 8 HRS RAND Data Weights, Clustering, and Strata

Introduction

An Overview of the Health and Retirement Study

A Quick Look

Since 1992, the University of Michigan Health and Retirement Study (HRS) has conducted biennial surveys of Americans over the age of 50. In 2004, more than 26,000 individuals were contacted in 2004. Supported by the National Institute on Aging (U01 AG00 9740), this panel study provides longitudinal data for an array of social science disciplines, including these topics:

- Health and cognitive conditions and status
- Retirement plans and perspectives
- Expectations for the future
- Family structure and transfers
- Employment status and job history
- Job demands and requirements
- Disability
- Demographic background
- Housing
- Income and net worth
- Health insurance and pension plans
- Experimental modules

OUR SPONSORS

National Institute on Aging

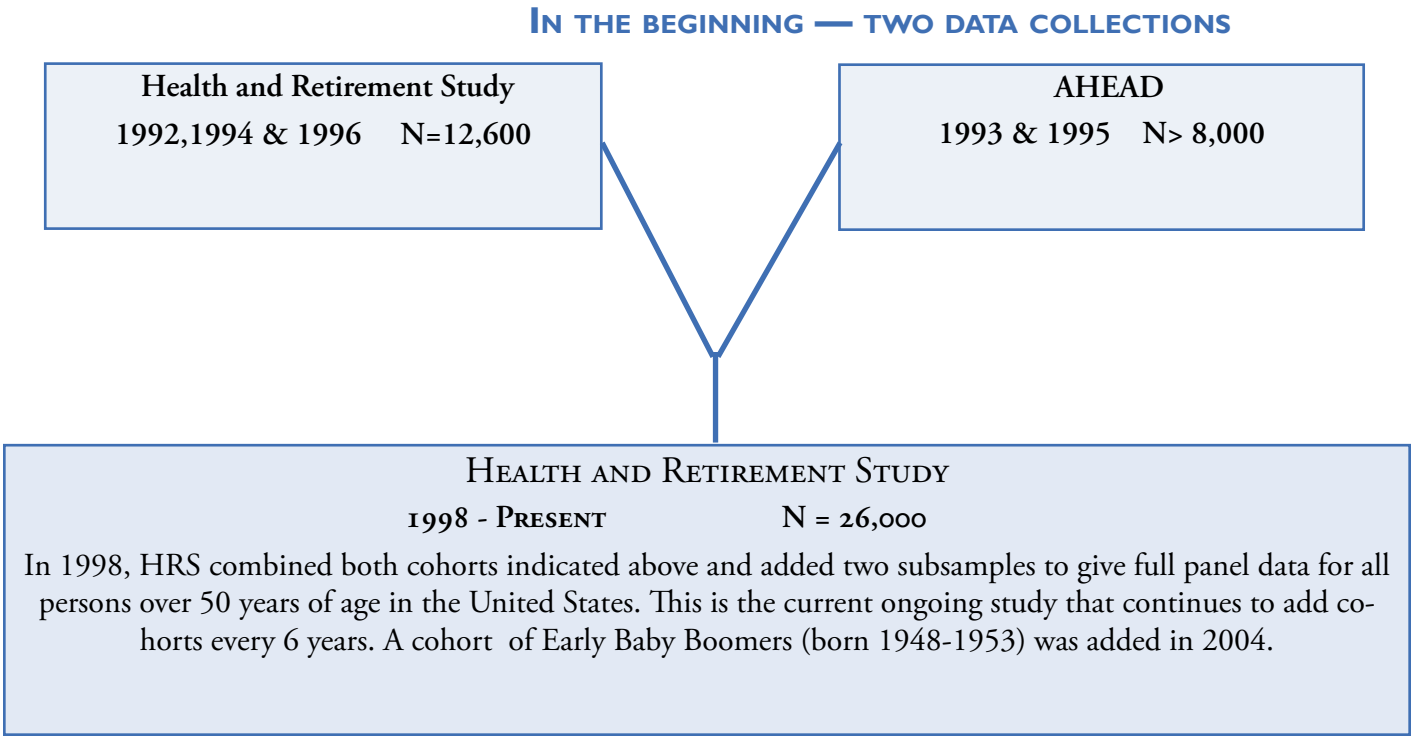
Social Security Administration

Health and Retirement Study Components

Two Studies Became One

In order to understand the data and what is available for research, you will find it is important to know that at on time the current Health and Retirement Study was composed of two distinct data collections. The following diagram outlines the melding of the two original data collections, Health and Retirement Study (HRS) and the Study of Asset and Health Dynamics among the Oldest Old (AHEAD), into the current study. In addition to the merge of HRS and AHEAD, sub-samples were added to create a complete panel of respondents representing all persons over the age of 50 years in the United States. New respondents are now added every six years to continue to represent those over 50. For example a third sub-sample of Early Baby Boomers (born 1948-1953) was added in 2004.

The original study designs had different objectives, questionnaire topics, and data availability. These studies were redesigned and joined in 1998 to provide a common longitudinal format. Longitudinal data from HRS 1992, 1994, and 1996 and AHEAD 1993 and 1995 can be merged with the current HRS study.



The Span of Social Science Disciplines Using the HRS

Over 1000 publications including books, book chapters, journal articles, working papers, newspaper articles have been produced using the Health and Retirement Study data. The study is a multidisciplinary effort designed by a broad range of researchers in such fields as:

- economics
- survey methodology
- sociology
- demography
- psychology
- medicine
- public policy

The social science data base from the HRS study serves a function analogous to the data collections produced by large-instrumentation (e.g. satellite) studies used in the natural sciences.

Figure 1 demonstrates the growth of research and the type of output using HRS data over the period 1993-2005.

From its inception in 1992, the Health and Retirement Study (HRS) has been unique from a social science standpoint. It represents a major investment by the National Institute on Aging to provide data for the research community to examine issues facing an aging society in the 21st century.

Now we have included the landmark Aging, Demography, and Memory Study (ADAMS). This in-depth study of dementia is unlike any other in comprehensive data collection.

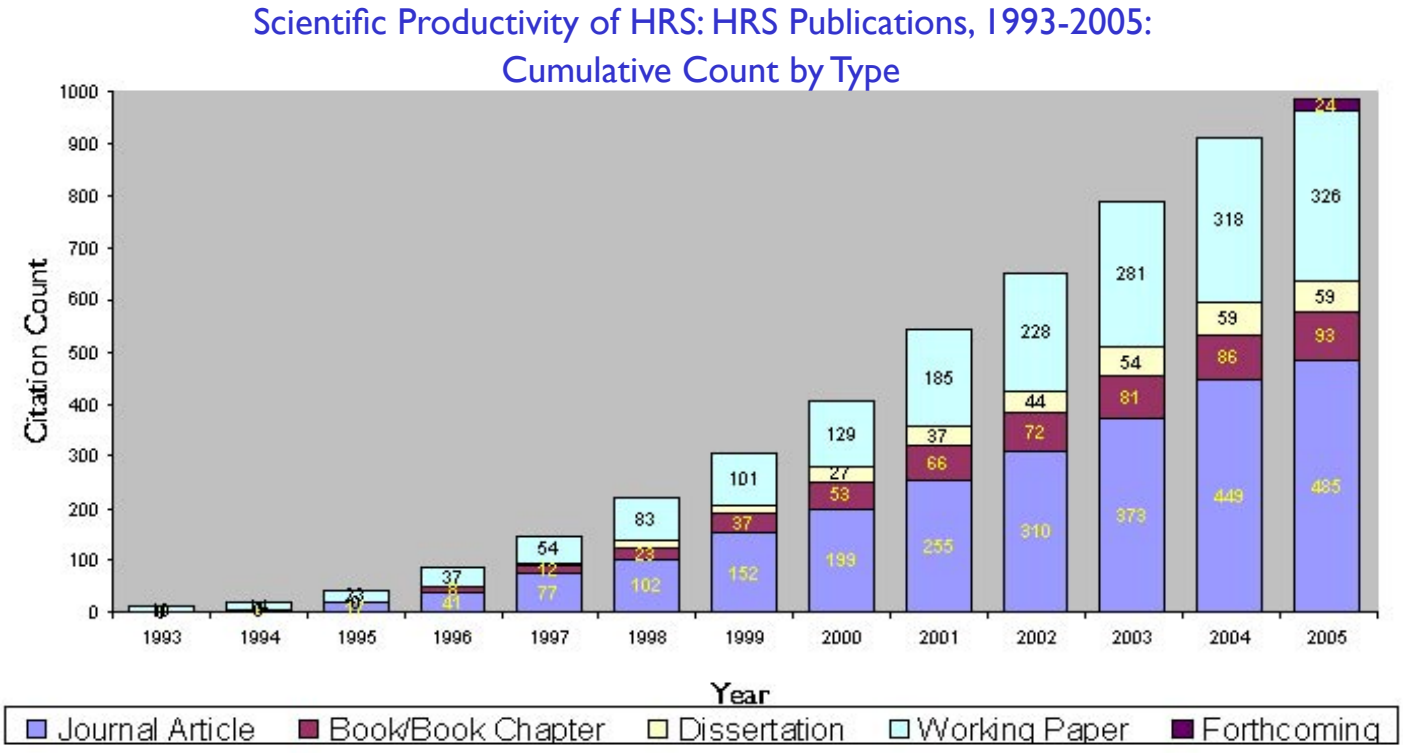


Figure 1. Since 1993, social scientists such as economists, physicians, demographers, and sociologists, have collaborated to research areas vital to the improvement of life for people over 50 in the United States.

The Original Health and Retirement Study (HRS)

Conducted 1992, 1994, and 1996

In 1990, the National Institute on Aging (NIA) awarded F. Thomas Juster of the University of Michigan (U of M), Institute for Social Research a five-year grant to design and begin a study to answer pressing questions about the needs of retirees. The sample included people making retirement decisions still in the labor force, as well as those already retired. The objectives were to gather information about the antecendants and consequences of retirement across multiple disciplines of study.

The data would allow researchers to examine relationships between health, income, wealth, and family over time. This created an opportunity to study the patterns of wealth and consumption, the effects of disability, widowhood, and other life changes on wealth and health.

The questionnaire contained subjects such as people's health, retirement plans and their understanding of their needs for retirement. In addition, it addressed family structure; time and money transfers; as well as demographics and housing. Also, the study addressed health insurance and pension plans.

The study was conducted on a national panel with an initial sample of 12,654. Blacks and Hispanics were over-sampled to provide an accurate overview of the cohort born between 1931 and 1941.

Other features of interest include links to administrative agencies data files such as the employers' pensions, the earning and benefits from Social Security, and the National Death Index. These files have restricted use and researchers must meet a strict criteria to access these files.

The study was conducted over three waves before it was merged with AHEAD; the cohort is still intact.

The Study of Asset, Health Dynamics Among the Oldest Old (AHEAD)

Conducted 1993 (Wave 1) and 1995 (Wave 2)

During the same time period as the HRS, a group of reserachers at ISR led by Willard Rogers proposed a competitive supplement to the files designed to answer a broad range of questions about resources and later life health transitions in people born in or before 1923. The rationale for this study was that older individuals have differing levels and mixes of resources and consume them at different rates as they age. Their health and functioning changes and affects these resources.

The questionnaire dealt with aspects of life related to those resources and health. Among these issues measured were:

- cost of illness, what was borne by family
- different resources used to offset mental, physical, and functional losses
- effectiveness of various care arrangements in preserving function and delaying institutionalization
- time and money transfers received to help slow or prevent impoverishment
- extent of dissaving and Medicaid spend down

This national panel study oversampled Blacks and Hispanics. The initial sample was 8,222 of those born in 1923 and before.

Researchers have used these data to research transitions during advanced old age in physical, cognitive, and functional health. They have examined the relationships of these transitions to changes in wealth and income, as well as other economic resources and intergenerational transfers. Analysts have studied the mix of effects of economics, family, and programs on institutionalization, dissaving, and health declines.

This cohort is still surveyed and has been since it merged with HRS in 1998.

The Present Health and Retirement Study

Merged HRS and AHEAD Wave 1998, Waves 2000, 2002 and 2004, and beyond — A Panel Study Representing all Persons Over 50 in the United States

In 1998, the HRS and AHEAD studies were merged into a single data collection effort and instrument. Since then cohorts have been added to create a sample of respondents that includes all ages over the age of 50. The steady-state design means that new cohort of respsondents ages 50-56 will be added to the sample every 6 years.

The first merged survey included four cohorts:

1. Wave 4 of the HRS cohort
2. Wave 3 of the AHEAD cohort
3. Wave 1 of an added Children of the Depression—CODA (born from 1924-30) cohort
4. Wave 1 of an added War Babies—WB (born from 1942-47) cohort

The steady-state design called for a fifth cohort which was added in 2004: Wave 1 of the newly added Early Baby Boomers —EBB (born from 1948-1953) cohort. Future plans to continue to populate the study with 50-56 year olds include adding the Mid Boomers (born 1953-60) in 2010.

The steady-state design is diagrammed in a visual overview on page XI of the Introduction — The Data Collection Path. The HRS is now a panel study completed every two years, representing all persons over 50 in the United States. This merger and addition of cohorts is important to keep in mind in all data analysis, particularly with longitudinal data. The original HRS cohort can still be studied longitudinally back to 1992, the original AHEAD cohort can be studied back to 1993.

The problem of assessing the well-being of the population over 50 in the United States will become increasingly complex as that population's size increases with the addition of the Baby Boom generation and the continued reduction in mortality rates. Since 1998, the HRS has surveyed the many aspects of life concerning this population to provide the social sciences with a large data base from which to study the inter-relationships of health, wealth, demographics, retirement, and a multitude of other areas.

The questionnaire was redesigned in 1998 by merging the questions from the two studies and adding new ones. The wording of some questions was changed and some questions were eliminated. The combination of the purposes and objectives created a powerful data collection tool that is updated each year.

From 1998 on topics include:

- Retirement plans and perspectives
- Retirement attitudes, preferences, expectations, and subjective probabilities
- Claims on transfer programs and contingent claims; dissaving and Medicaid eligibility
- Cognitive performance-based testing
- Demographic characteristics
- Employment status, job history, and disability
- Family structure and transfers
- Health insurance and pension plans
- Housing (including access to services) and services use (community and nursing home)
- Out-of-pocket costs for all services
- Physical and functional health

The data can be analyzed by researchers in an exponentially larger number of combinations using variables from the growing and dynamic multi-disciplinary data set. Interestingly, researchers are performing cross-disciplinary analysis, combing fields such as economics and medicine.

The Restricted Data Files

These data can be linked to HRS participants' data.

Social Security Administrative Data
Income and benefit information for HRS participants derived from SSA earnings and benefits records

Pension Data
Comprehensive pension and projected pension information derived from employers' records

Health and Retirement Medicare Linkage File Data
Information on health care costs, use of services, and diagnosis

National Death Index Data
Death information on deceased HRS participants: month/year and cause

Industry/Occupation Data
Participant occupation and industry details

Geographical Data
Participant location information: state, county, zip code, and census tract

Other Special Files: Exact Date of Interview, Exact Date of Birth

Experimental Modules and Mail Surveys

Experimental questions, typically different for each Wave, are asked for a wide variety of purposes at the end of each questionnaire. For example, one year the questions were regarding the extent of use of alternative medicine.

The sample for the HRS has been also involved in additional surveys designed as supplemental information to the current questionnaire. Details can be obtained by clicking the link on this page or by putting http://hrsonline.isr.umich.edu/data/avail URL in your browser and clicking on the survey.

The Consumption and Activities Mail Survey (CAMS)

This longitudinal survey has questions about specific use of time and money. Data collection occurs in off-year mail outs from the core.

Human Capital and Educational Expenses Mail Survey

This one-time survey has questions about respondents' capital and spending on education.

Diabetes Mail Survey

This one-time survey was sent only to respondents to study self-care and knowleged of diabetes in a sub-population with a diagnosis of diabetes.

The Aging and Demographics Memory Study (ADAMS)

ADAMS was a supplement to the HRS with the specific aim of conducting a population-based study of dementia. ADAMS fieldwork began in August, 2001 continued through March, 2005. Our projection is for data to be available for researchers by late 2005. At the beginning, HRS formed a partnership with a research team at Duke University to conduct in-person assessments for dementia on selected HRS respondents. The purpose was to gather information on respondents' cognitive status and assign a diagnosis related to dementia. 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 was 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 primary goal of the ADAMS study was to collect data that will allow researchers to estimate the prevalence of dementia in the U.S. elderly population. This research will facilitate our understanding of the natural history of of dementia in the changing the health and social function of older Americans.

For more information refer to this journal article: Langa, K., et al., **The Aging, Demographics, and Memory Study: Study Design and Methods.** Neuroepidemiology, 25, 181-191, (2005).

Wide-Spread use of Hrs Data

Researchers have published over 1,000 articles using HRS data in 25 types of journals in fields such as these:

- Business and Economics
- Medical Sciences
- Gerontology and Geriatrics
- Sociology
- Public Health and Safety
- Psychology
- Population Studies
- Insurance
- Statistics
- Women
- Social Sciences (Comprehensive)
- Health Administration
- Disability
- Drug Abuse and Alcoholism
- Social Services and Welfare
- Political Science
- Philosophy
- Pharmacology
- Occupational Health and Safety
- Housing and Urban Planning
- Education
- Real Estate

Today the HRS is used internationally as a template for similar national studies and NIA has funded research to compare data between these studies. Many national studies are being designed, such as those in Israel, Korea, and Japan. Three are in progress and have data available:

–The Mexican Health and Aging Study (MHAS) is a prospective panel study of health and aging in Mexico. Data and documentation are available.

–The English Longitudinal Study of Ageing (ELSA). data set is available through the Economic and Social Data Service.

–The Survey of Health, Aging and Retirement in Europe (SHARE): The initial release of SHARE data is now available.

Links to Cross-National Studies
Find under Intro/Guide>Links

All the uses for HRS data are only beginning to be tapped. Researchers continue developing new models for studying the data both intra- and inter-disciplinarily.

HRS is positioned to address the need for data to analyze implications of changes in economy, society, demography, technology, and policy.
As the Baby Boom approaches retirement, the growth in longevity continues, and the 20th century divorce revolution changes family structures, new insights are needed.

Researchers are challenged to address the changing landscape of work and retirement that includes a continued decline in defined benefit pensions, solvency issues, an increase in 401(k), other defined contribution plans along with a decrease in employer health insurance and disappearing retiree health insurance. With increasing complexity of decisions about work, savings, health, and retirement, the Health and Retirement Study data is ready to meet researchers' needs.

Data Collection Path

The following diagram from our Web site is a visual overview of the preceding study descriptions. It shows all the cohorts, the years they entered, the survey year, and type of study: either AHEAD, HRS, mailouts and other types of surveys.

Enter http://hrsonline.isr.umich.edu/intro/dataflow.html into your Internet browser to access detailed information such as:

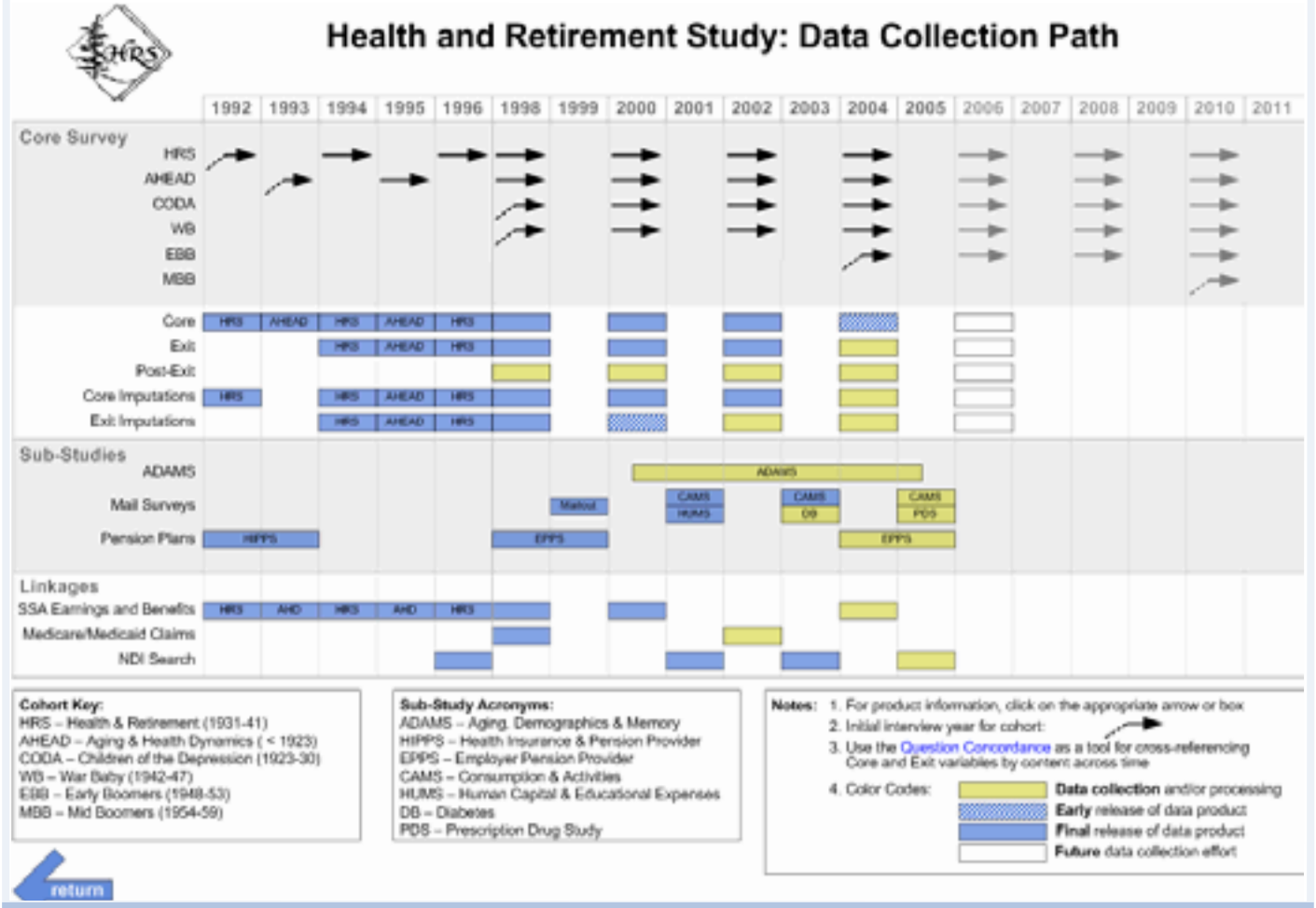
–Sample size
–Response rate
–Entry cohort
–Latest Release
–Data Release Date
–Data alerts

Each arrow and box are linked to information about the item indicated, cohort or survey.

Visit our Website for more information:

http://hrsonline.isr.umich.edu

EASY TIP


Health and Retirement Study: Data Collection Path
1992 1993 1994 1995 1996 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
Core Survey
HRS
AHEAD
CODA
WB
EBB
MBB
Core
Exit
Post-Exit
Core Imputations
Exit Imputations
Sub-Studies
ADAMS
Mail Surveys
Pension Plans
Linkages
SSA Earnings and Benefits
Medicare/Medicaid Claims
NDI Search
Cohort Key:
HRS – Health & Retirement (1931-41)
AHEAD – Aging & Health Dynamics (< 1923)
CODA – Children of the Depression (1923-30)
WB – War Baby (1942-47)
EBB – Early Boomers (1948-53)
MBB – Mid Boomers (1954-59)
Sub-Study Acronyms:
ADAMS – Aging, Demographics & Memory
HIPPS – Health Insurance & Pension Provider
EPPS – Employer Pension Provider
CAMS – Consumption & Activities
HUMS – Human Capital & Educational Expenses
DB – Diabetes
PDS – Prescription Drug Study
Notes:
1. For product information, click on the appropriate arrow or box
2. Initial interview year for cohort:
3. Use the Question Concordance as a tool for cross-referencing Core and Exit variables by content across time
4. Color Codes:
Data collection and/or processing
Early release of data product
Final release of data product
Future data collection effort
return

How to Use Our Guidebook

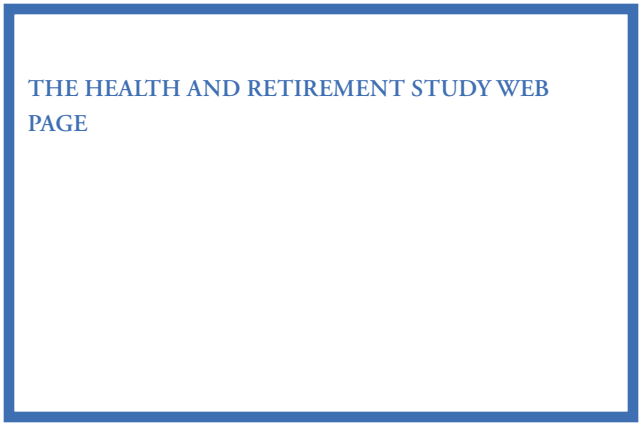
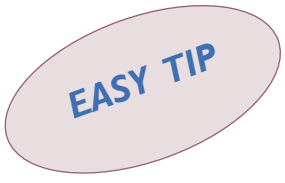


Figure number. Title of Figure — Hyperlinked to URL.

How this page is helpful to you.

Web site illustrations are intended for demonstration and may not represent the most current version of a Web page.

Information with emphasis



Our Web site http://hrsonline.isr.umich.edu

Designed for Two Types of Readers

Reader One — Just Looking

If you are interested in finding out basic information about HRS and learning what research is available:

- The Introduction and Chapter 1 should meet your needs.

Reader Two — Planning a Research Project or Using HRS Data

If you are exploring the Health and Retirement Study Data and thinking of using it for research or are currently using it:

- Chapters 2-5 are written for you.
  - Chapter 2 will take you through the basic steps to get started with registration and exploration of our data files.
  - Chapter 3 & 4 describe the Toolbox for using the HRS data, what the tools are and how they are used.
  - Chapter 5 will introduce you to the language of HRS and give you insight into past and current glossary entries and acronyms you may find while using our documentation.

Hyperlinked Text

This text is either linked to information within the guidebook itself or to our Web site pages. This link will take you to the Home Page of our Web site.

A Quick Intro to the HRS Web site

1

Introduction to Basic HRS Website Information

Home

Intro/Guide

Publications

Search

&

HRS Participant Web site

The Health and Retirement Study Web Site

The HRS Home Page

Our home page contains the latest news in product announcements, news items, general HRS information, data alerts (corrections to data already released), the latest HRS Data User Newsletter, and a link to the HRS Participant Web site.

Methods

1. To view the home page of the HRS Web site (Figure 2) insert: http://hrsonline.isr.umich.edu into your internet browser address box.

   Please bookmark this page, we refer to it often.

2. Notice that on our Web site each page has two navigation bars that indicate where you are in the site page hierarchy. (Figure 3)

   – The top (primary) navigation bar has links to the main sections: Home, Intro/Guide, Documentation, Data, and Publications. For convenience, a hot link is also provided to the Search function (powered by Google).

   – The name of the section you are currently visiting is highlighted in burgundy on a blue background.

   – The secondary navigation bar below the top bar is specific to each section, providing links to the main subdivisions of that section.

HRS Products include all waves of data, documentation, and data links.

Figure 2. HRS Home Page
This Web page is continually updated with our latest releases. *Please bookmark this page.*

Figure 3. HRS Home Page
Use the *Top Navigation Bar to find Web page categories and the *Secondary navigation bar for specific pages in that section.

Figure 4. HRS Home Page — User Registration
Click here to register to obtain HRS data.

User Registration

To register to use the HRS data start from the Home Page, find User Registration on the secondary navigation bar (Figure 4) and click the hot link. These data are free for public use. (For more information refer to Chapter 2, pg. 14 for a complete explanation of how to register.)

EASY TIP

Intro/Guide Section

The Site Map – *Use the Web site the Easy Way*

The quickest overview of the Web site is on the Site Map. You can quickly move to the Web page you are looking for by clicking the links on this page. The site map displays a comprehensive overview of the contents of the Web site. You can click on hyperlinked text to find information easily.

Figure 5. Site Map
Each bullet point is hyperlinked so you can access any page on the Web site easily.

Methods

1. URL: http://hrsonline.isr.umich.edu/intro/sho_intro.php?hfyle=sitemap.

or

2. From the HRS home page.
   - Click Intro/Guide on the primary navigation bar.
   - Click on Site Map on the secondary navigation bar (Figure 5).

Intro/Guide Section

Co-Investigators

Each HRS Co-Investigator (Co-I) has a Web page with contact information and an outline of their primary research interests. If you would like information on a specific research area in the HRS, search the contact information on the Co-Investigator pages.

Methods

1. URL: http://hrsonline.isr.umich.edu/intro/sho_intro.php?hfyle=copis.

   – As you move your cursor over their pictures the photos turn from black and white to color because the pictures are hyperlinked to that individual's Web page. Click on the pictures or the names of the Co-Is of interest.

   – This page has hyperlinked photos to each Co-I, as well as hyperlinked names under the photos.

2. From the HRS home page:

   – Click on Intro/Guide on the primary navigation bar.

   – Click Co-I on the secondary navigation bar (Figure 6).

   – This page has hyperlinked photos to each Co-I, as well as hyperlinked names under the photos.

Figure 6. HRS Co-Investigators and Committees
Click on either the picture or the name of the Co-I to open their Web page.

How to Contact the HRS Staff

Methods

1. URL: http://hrsonline.isr.umich.edu/intro/sho_intro.php?hfyle=contact
or

2. Click on Intro/Guide on the top navigation bar and Contact on the secondary navigation bar.

Figure 7. How to Contact Us in INTRO/GUIDE Section

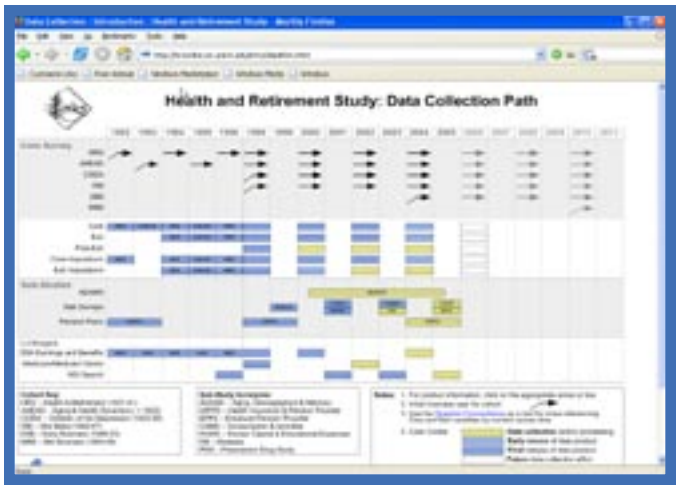


Figure 8. Image Map of the Data Collection Path
You can see the history of HRS and AHEAD, the merger in 1998 with the 2 sub-samples, and the current HRS steady state design on the top row. The sub-sample of Early Baby Boomers was added in 2004 with Middle Boomers to be added in 2010.

Figure 9. What's Available
Core data are from the surveys taken by respondents. Exit data are from surveys taken by proxy after a respondent has died. Mailout surveys have a specific topic (e.g., Internet Use).

Intro/Guide

Overview of Waves, Cohorts, Subsamples of HRS — The Available Data

To get a quick idea of the available HRS datasets, we have two formats to present what is available for you — an image map or a table. The image map displays a comprehensive overview of the contents of the Web site. You can click on hyperlinked text to find information easily. Both methods link to a Metadata Summary with information specific to that sample.

1. Image Map (Figure 8.)
   - URL: http://hrsonline.isr.umich.edu/intro/data-flow.html.
   - From the HRS home page:
     – Click on Intro/Guide on the primary navigation bar.
     – Click on Data Flow on the secondary navigation bar.

   Each box/arrow **is** hyperlinked to a Metadata summary, the contents are listed in the box below.

2. What's Available Table (Figure 9.)
   - URL: http://hrsonline.isr.umich.edu/data/avail.html
   - From the HRS home page:
     – Click on Intro/Guide on the primary navigation bar.
     – Click on What's Available on the secondary navigation bar
   - Data products are displayed by year (Figure 9). Click on the appropriate hyperlink to view a page containing a the same metadata summary as found with the links to the data collection path links.

The Metadata summary contains the:

a. number in the sample
b. entry cohort information
c. latest version of data release
d. field dates
e. response rate
f. data alerts
g. hyperlinks to questionnaire, data description, codebook, and modules.

EASY TIP

Publications

To find papers, journal articles, and other reference materials click publications on the top navigation bar.

The Online Bibliography

If you would like to find research using the HRS data by author, title, or keyword the searchable *Online Bibliography is the easiest method.*
Use the Online Bibliography to search for books, research papers, articles, and working papers on the topics of interest to you.

Methods
1a. URL: http://hrsonline.isr.umich.edu/biblio/index.html (Figure 10).

or

1b. From the HRS home page:

- Click Publications on the primary navigation bar.
- Click Online Bibliography, in the descriptive paragraph, click on the hyperlinked text:

2. Select the type of publication you would like to view.
3. Select one of more content areas (Figure 11., e.g., Consumption and Savings), or search by author, title, or years.
4. Adjust the retrieval threshold between 1 and 100 depending on the desired number of papers. A value of 1 will retrieve more records, 100 will result in fewer records.
5. Select a sort rule (e.g., author, title, or date).
6. Or use the Freetext Search.
7. Click Submit.

Figure 10. The Online Bibliography
This searchable HRS bibliography contains the publications using HRS data. It is updated continually.

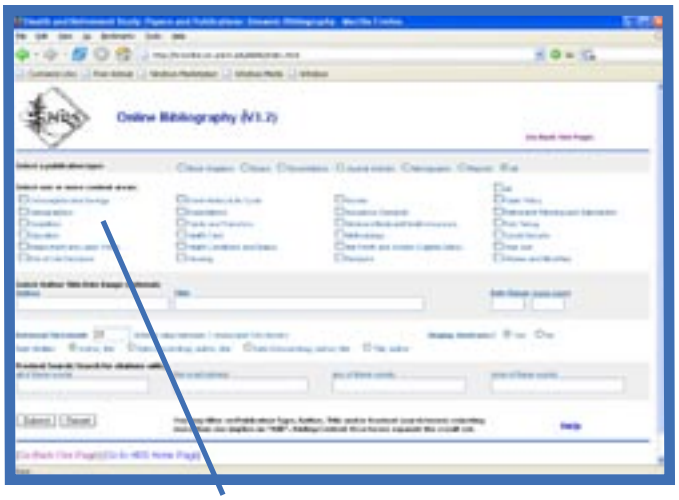


Figure 11. Select a Content Area
Example for selecting articles and papers about Consumption and Saving

Figure 12. The Populations Studies Center Resources link can be found in the Publications section.

Figure 13. Early HRS Working Papers Archive
The purpose of these papers was to analyze data quality, to inform future improvements needed and stimulate research using these new data..

Population Studies Center Access

To find journal articles or abstracts using the Population Studies Center (PSC) resources.

Methods

1. URL: http://hrsonline.isr.umich/papers/shopapers.php?hfyle=pscpubs (Figure 12)

2. Clink the link just below Online Bibliography on the Publications page.

Population Studies Center Resources

Access to Resources on Aging

The Michigan Center on the Demography of Aging (MiCDA), through the University of Michigan's Population Studies Center (PSC) has been active in collecting a variety of papers on aging research. Other PSC resources include:

* The Ronald and Deborah Freedman Library holds over 45,000 publications, government and U.N. documents, dissertations, working papers, and monographs.

* The PSC Publications page facilitates communication among population researchers by making working papers, research reports, and reprints authored by Center researchers accessible online. Search for a specific author or title by using the PSC Publications Search Engine.

Early HRS Working Papers Archive

At the bottom of the PSC Access page are the papers in the HRS/AHEAD Working Paper Series the results from the analysis of early release datasets created from the first waves of data collection on the studies. The series includes papers from four early results workshops (HRS Wave 1, HRS Minority Perspectives, AHEAD Wave 1, and HRS Wave 2). See the List of HRS Working Papers (Figure 13).

Please note that the HRS staff have not attempted to validate either the data or the statistical analyses contained in the working papers.

Additional HRS Publication Links

Methods

1. Insert URL: http://hrsonline.isr.umich.edu/papers/index.html.

2. From the HRS Home Page click Publications on the top navigation bar.

This is the list of links:

Register Your Paper!

Please inform us of your papers and publications based on HRS data.

Conference Proceedings

Papers and presentations associated with conferences sponsored by the Health and Retirement Study.

User Guides

These documents provide additional documentation about the concepts, measures and questions in the HRS surveys. They expand upon the information found in codebooks, questionnaires and data descriptions.(PDF format).

Documentation Report Series

Contributed reports with technical information about file construction.

HRS in the News

Newspaper stories, magazine articles, and press releases concerning HRS data and findings (Figure 15).

Figure 14. Papers and Publications sponsored by the HRS.

Figure 15. An example of HRS in the News — Press releases, newspaper stories, magazine articles can be selected from the publications page.

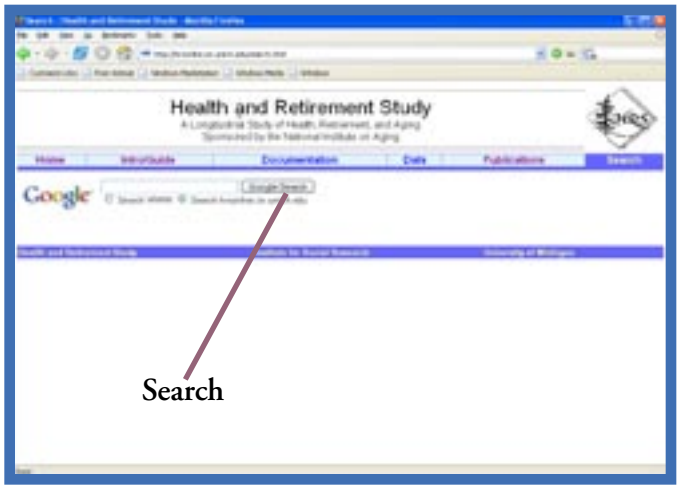


Figure 16. Search
Search for information on the HRS Web page using a Google search engine.

Search

Looking for something in particular? To Search the HRS Web site use keywords in a Google search engine. Click on the Search button on the top navigation bar.

Methods

1. URL: http://hrsonline.isr.umich.edu/search.html

or

2. From the HRS home page a Search hot link is provided on the top navigation bar at the far right (powered by Google). Click Search. Enter keywords in search engine box (Figure 16). Hit Enter.

Figure 17. The HRS Participant Web Site
Participants commonly request the types of information found on the Web site page. You may find the Participant Newsletters helpful for a number of uses. Summaries of HRS research are included in each newsletter.

HRS Participant Website

The participants (respondents) have their own Web site that contains their annual newsletters. These have easy-to-read summaries of findings based on peer-reviewed journal articles using HRS data.

Methods

1. URL: http://hrsparticipants.isr.umich.edu

You may want to bookmark this page since a newsletter is added every year.

2. Choose either the English or Spanish page.

3. Find hot links to other organizations helpful to those over 50.

Participants are interviewed either in English or Spanish. The households where a Spanish interview was conducted receive both an English and Spanish newsletter. Each newsletter title on the Web site is hyperlinked to a PDF file containing the entire newsletter available for reprinting. There are no copyright concerns because this is public information funded by the government; however, we appreciate bibliographic citations.

The HRS Web site for Data Users

2

Home/ Data User Registration

Obtaining HRS Public Data on Our Web site

The primary method of communication between researchers and the Health and Retirement Study staff is through the Web site. The Web site has hundreds of pages of information. Understanding what information is available and how to obtain it is key to an easier analytic experience using the data.

The next section outlines the HRS Web site. HRS data products are available without cost to researchers and analysts. User Registration is required to download files. See details on our Web site.

Methods

1. URL: https://ssl.isr.umich.edu/hrs/.

or

2. Click on Data on the primary navigation bar and then User Registration (Figure 18)

- We go to great lengths to protect the confidentiality of HRS Respondents. See Conditions of Use (pg. 16) for further understanding of our policies.
- See special instructions for obtaining access to sensitive health data files which will be posted.
- See special instructions for obtaining access to restricted data files at: http://hrsonline.isr.umich.edu/rda/. Restricted data include linkages to Social Security Administrative data, Employer Pension data, National Death Index, Geographic Information, Industry/Occupation, Imputated College Tuition, and Health Care Information.

Note:
The information you provide will not be used for any commercial use, and will not be redistributed to third parties.

Figure 18. User Registration
From the HRS Home Page click User Registration on the secondary navigation bar.

Figure 19. User Registration
This is the gateway page to the data. You will click User Registation not only the first time you register, but everytime you access the data after you have already registered.

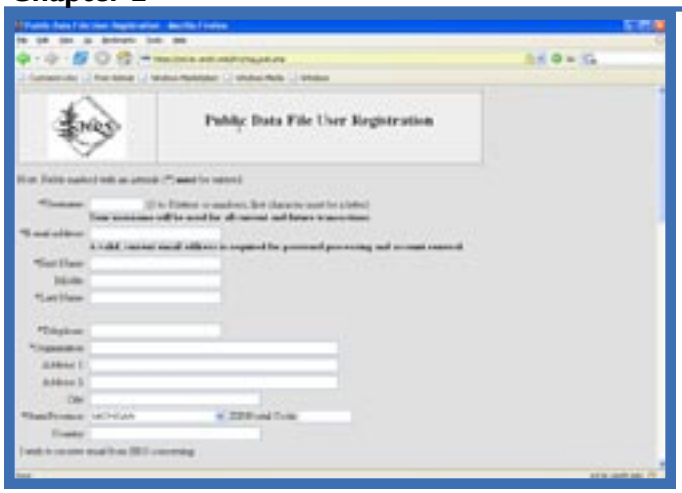


Figure 20. Public Data File User Registration
Fill in the form to register with HRS, then you will be e-mailed a user ID and password.

When You Register to Use HRS Data

You will receive your user id and password by e-mail to access the data shortly after you register. Return to this page, fill in the information and click on the button – you have already registered. The file directory will open. You will always access the data through this page (Figure 19).

By registering all users, we are able to document for our sponsors the size and diversity of our user community, allowing us to continue to collect these important data.

Registration and Downloading the Data

Methods

1. URL: https://ssl.isr.umich.edu/hrs/login.php.

or

2. From the HRS Home Page click on User Registration. Read the Conditions of Use Agreement and the Privacy and Security Notice.
3. Fill in the form on the Web site to register with HRS.
   - There are two check boxes.

     I wish to receive e-mail from HRS concerning:

     ☑ Data Alerts (notification of dataset changes and upgrades)

     ☑ Announcements of newsletters, conferences and other items of interest to analysts

   - If you check these boxes, registered users receive user support in the forms of information data alerts; data user newsletters; future data and data product releases; workshops to learn to use the HRS data; and publication lists.
4. After completing the form, click the register button at the bottom of the page.
   - After HRS validates your e-mail, you will be e-mailed a user ID and password.
   - Your username and password are required to gain access to the HRS Data Distribution Center. (See Figure 23, Page 18.)

TIP: If you do not hear back from HRS shortly after registering, there may be a problem with the e-mail address you entered. Please contact HRS Quest e-mail with the information that you have registered and not heard back from HRS.

Conditions of Use

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

You must agree to:

* Make no attempts to identify study participants.

* Not to transfer HRS Public Release data to any third party other than staff or students for whom you are directly responsible except as indicated below.

* Not to allow others to use your username and password to access this site.

* To certify the destruction of any downloaded Public Release data file as well as any data files derived from the downloaded file when requested to do so by the Health and Retirement Study.

* Report immediately to the Health and Retirement Study at hrsquest@isr.umich.edu any disclosure of study participant identity as well as any discovery of flaws or errors in the data or documentation files.

* Notify the Health and Retirement Study through use of the update function provided at this site or by electronic mail directed to hrsquest@isr.umich.edu of changes in your electronic mail address, postal address, telephone number, organizational affiliation or organizational status.

Figure 21. Conditions of Use for Public Data Files
Print the page to read it and keep in your HRS file.

Three Conditions of Use For Publications

Include the following citation in any research reports, papers, or publications based on Public Release data:

In text:
"The HRS (Health and Retirement Study) is sponsored by the National Institute of Aging (grant number NIA U01AG009740) and conducted by the University of Michigan."

In references:
"Health and Retirement Study, (Wave [x]/Year [yyyy]) public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG09740). Ann Arbor, MI, (year)."

Include the following citation in any research reports, papers, or publications based on any Public Release data file tagged as "Early" or "Preliminary":

"This analysis uses Early Release data from the Health and Retirement Study, (Wave [x]/Year [yyyy]), sponsored by the National Institute of Aging (grant number NIA U01AG009740) and conducted by the Unversity of Michigan. These data have not been cleaned and may contain errors that will be corrected in the Final Public Release version of the dataset."

Provide information regarding publications based on data obtained from the Health and Retirement Study by sending a copy of any papers or publications using HRS public files or datasets to:
Health and Retirement Study
Room 3050 ISR
P.O. Box 1248
Ann Arbor, MI 48106-1248.

Figure 22. How to Contact Us
The HRS e-mail address is hrsquest@isr.umich.edu.

Home/ Help Desk

You may contact a member of the HRS staff through the form on the Help Desk page or by e-mail. Your question is routed to the staff specializing in that area of expertise.

Methods

1. To view the home page of the HRS Web site (Figure 2) insert: http://hrsonline.isr.umich.edu into your internet browser address box.
2. Click on Help Desk at the far right on the secondary navigation page.

This page is designed to assist users who:

- have questions about the HRS study and its procedures
- are seeking solutions to problems with HRS data products
- wish to contact the study concerning other issues.

Using this form, provide as much detail as possible so that we may best assist you. Your question will be passed on to the HRS Help Desk and routed to the HRS staff member who is most familiar with the data product or area of study.

3. Or send e-mail to hrsquest@isr.umich.edu.

Search for HRS Citations

Use this page to find an HRS citation by article or by Co-Investigator in various data bases such as JSTOR, Web of Knowlege, and many others.

Method

Go to:
http://hrsonline.isr.umich.edu/papers/sho_papers.php?hfyle=citesearch

Using Your Password

From the HRS Data Distribution Center, you have access to the following types of Health and Retirement Study available data products:

Final Release Datasets and Files

Final Release files have been cleaned, documented, and made easier to use. They are aligned with the Tracker file (explained in Chapter 4, pg. 45), and supplemented by imputation and data description files. Codebooks (including frequencies) in ASCII and HTML format, and questionnaires in PDF and Word format accompany all Public Release files.

Early Release Datasets and Files

Early release files are made available a short time after the end of the field period, and have undergone minimal processing. A Codebook in ASCII format is provided that matches the format of the released dataset.

Cross-Wave Files

Longitudinal datasets such as the Tracker File which assist researchers in linking respondents across time.

Supplemental Files

Additional releases, usually of the same data in new forms that may be helpful to analysts.

Researcher Contributions

These files are provided by fellow researchers interested in sharing their work as a service to the research community. The Health and Retirement Study is not responsible for the support, content, or usability of these files.

Documentation for Restricted Files

Users who plan to apply for access to HRS restricted data may wish to review file documentation (Chapter 5).

Figure 23. Data Sets and Files (You must register to access this Web page.) Click on the hot link to the Data Set you would like to use.

An HRS Data Product Example

Figure 24. 2002 HRS Core Final (v.1.0) (You must register to access this Web page.) These are the zip files of documents and data you will find when you click 2002 HRS Core Final (v.1.0).

An HRS Data Product Example

2002 HRS Core Final (v.1.0)

Codebook files
Distribution set
Data files
User information for distribution data file
Questionnaire
SAS data descriptions
SPSS data descriptions
STATA data descriptions
HRS 2002 to HRS 2000 cross-reference file

Figure 25. The Elements of the Data Files in 2002 HRS Core Final (v.1.0) as seen in Figure 24.

Health and Retirement Study

Three Types of Data

- **Public Use**
- **HRS Special Access Files (available in 2006)**
- **Restricted Data**

Terminology Describing A Wave of Data

Product terms

• **Wave core data** are the responses of the total living respondents and their proxies in one survey.

• **Wave exit data** are the responses of proxies for deceased respondents in that survey.

• **Lower-case v.** refers to the version of that set of data. When a data set is updated the version number is changed.

• **Early and final** refer to the stages of release of data. Early refers to the first stage of release for the purpose of providing users access to data as soon as possible, and is equivalent to a final draft. This is a documented and analyzable data set. The second stage, Final, which is provided after analysts have worked further to clean the data this is a fully documented and analyzabable finished product.

HRS Data Files are Created for Use With These Statistical Packages

- SAS

- SPSS

- STATA (Further tutorials are provided in Chapter 6-8).

Examples of how to use each statistical package with HRS data are included in the data description so designated.

What's Available

Find Data Set Products from HRS in an Overview of Waves, Cohorts, Subsamples of HRS

To find the available HRS datasets, we have two different formats of What's Available for you in a table or an image map. The site map displays a comprehensive overview of the contents of the Web site. You can click on hyperlinked text to find information easily.

1. What's Available Table (Figure 26.)
   - URL: http://hrsonline.isr.umich.edu/data/avail.html
   - From the HRS home page:
     - Click on Intro/Guide on the primary navigation bar.
     - Click on What's Available on the secondary navigation bar
   - Data products are displayed by year. Click on the appropriate hyperlink to view a page containing a metadata summary – data about the data. (See below for specific examples).

2. Image Map (Figure 8.)
   - URL: http://hrsonline.isr.umich.edu/intro/dataflow.html.
   - From the HRS home page:
     - Click on Intro/Guide on the primary navigation bar.
     - Click on Data Flow on the secondary navigation bar.
   - Data products are displayed by year. Click on the appropriate hyperlink to view a page containing a metadata summary

Figure 26. What's Available
An overview of waves, cohorts and subsamples for available data.

A Metadata Summary

Each Metadata summary (Figure 27.) includes the following information the data set: N, Entry Cohort Info, Field Dates, and Response Rates, with links to data tools. Click on the hyperlinked text to find information easily.

Figure 27. Metadata Summary Example 2004
Provides data user with N, entry cohort info, field dates, response rates and links to all the necessary data tools for each data set. (See Chapters 3 & 4)

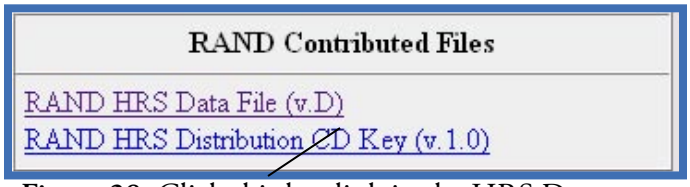


Figure 28. Click this hot link in the HRS Data Distribution Center to access the page in Figure 29 below. The data files are listed beneath the text on this Web page.

EASY TIP

Figure 29. RAND HRS Data File (v. F)
(You must register on the HRS Registration Page to access this Web page.) These files are produced by the RAND Corporation from HRS public data releases.

Supported by NIA and SSA

Popular — RAND Data Produced from HRS Data

This section contains data descriptions and documentation for contributed files produced by RAND Corporation from public versions of Health and Retirement Study data.

RAND HRS Data File (v. F based on 2004 Early Release)

The RAND HRS Data file is a cleaned and easy-to-use version of key variables from eight waves of the Health and Retirement Study data, including four entry cohorts: the original 1992 Health and Retirement Study (HRS) cohort; the 1993 Study of Asset and Health Dynamics (AHEAD) cohort; and the Children of Depression (CODA) and War Baby (WB) cohorts entering in 1998.

As of March 2006, it incorporates data from 1992, 1993, 1994, 1995, 1996, 1998, 2000, 2002 Final Release versions, and 2004 Early Release version of HRS data.

All data products have a version letter, not number from RAND. Please check the version letter for the latest release.

Note:

The Health and Retirement Study is not responsible for the support, content, or usability of these files. HRS cannot respond to any queries concerning these files. Questions or problems concerning the RAND HRS Data should be addressed to RANDHRShelp@rand.org.

Using the Web site to Prepare for Data Analysis

3

Introduction to Data Analysis Tools

HRS Documentation Products

Data Descriptions
Questionnaires
Concordance
Codebooks
Module Descriptions
Imputation Datasets
Data Alerts
User Guides

HRS Documentation Products

This is your "toolbox" for HRS data analysis. New products are being added every few months so continue checking the HRS Web site for products that you may find useful in the future.

Data Descriptions (pg. 25)
: Text description of each wave with details of level of files, naming conventions, identification variables, directory structure, codebook documentation, merging code for various waves of data, and longitudinal issues

Questionnaires (pg. 28)
: Representations of interviews in "box-and-arrow" format some also document the question flow or skip patterns in the survey

Concordance (pg. 30)
: The question concordance is a tool for cross-referencing interview questions by content across time (1992 through 2004 interview years). For interview years 1995 and beyond, each record in a retrieval set contains a link to the appropriate codebook table. To see details for a variable (jump instructions, question text, user notes, and codeframe) follow this link.

Codebooks (pg. 32)
: Codebooks for all data products are available for download in ASCII and HTML format

Module Descriptions (pg.36)
: Details on the contents of experimental modules of questions on the survey for each interviewing year

Imputation Datasets (pg.37)
: Information on imputation products currently available to the public

Data Alerts (pg. 38)
: Data release errors, omissions, notes, and corrections

User Guides (pg. 39)
: These documents provide additional documentation about the concepts, measures, and questions in the HRS surveys. They expand upon the information found in codebooks, questionnaires, and data descriptions (PDF format).

Figure 30. Documentation Products
Each heading is hyperlinked to the documentation product named.

Data Analysis Tools

Documentation Products (Figure 30) is a major heading on the primary navigation bar on the HRS Home Page and can be found at: http://hrsonline.isr.umich.edu/meta/index.html.

Figure 31. Data Descriptions for HRS Public Files
Data Descriptions are most helpful in understanding a data set in a particular wave. Information for how to begin data analysis is often included in the more recent descriptions.

Data Description Table of Contents
Example Early Release Core 2002

a. The Sample
b. Questionnaire Sections
c. Levels of Files
d. File Naming Conventions
e. Data Files
f. Identification Variables
g. Distribution Files and Directory Structure Program Statements: Using The Files With SAS, SPSS, and STATA
h. Codebook
i. Basic Demographic Variables
j. Selecting The Correct Variable For The Respondent From Preload Information
k. Variables Combined in Household Member Child Data

Figure 32. Example of a Table of Contents for a Data Description for a Wave of HRS Data

Data Descriptions

We strive for the satisfaction of all researchers using HRS data by writing clear, specific and detailed data descriptions for using a particular wave of data. Data Descriptions have improved over time as we receive feedback from researchers. For this reason, sometimes reading a method in a recent Data Description (e.g., 2002) may help with an analysis of an earlier dataset.

The Data Description is your best source of information for understanding and using the data.

The data description (Figure 31) contains:

- the sample — a detailed text description of the total population – the N, including the sub-populations. The populations change in waves (Remember the study merges and other changes that have taken place over the years.)
- the questionnaire sections
- what files are available (referred to by **level**)
- explanations for other elements of that data wave

Levels of Files in the Data Description

Each Data Description for a wave has a detailed explanation of the levels of files available in the data. Levels refers to the categories of questions asked, as well as to whom they were asked. For example, the Household-Level Files contain questions that were asked about the household of a designated coversheet, financial, or family respondent. There are household; respondent; household member and child; sibling; helper; transfer-to-child; and transfer-from-child level files. The data description for each wave is the source for understanding levels of files.

TIP — SAVE HOURS

Read the Data Description before doing anything else with the data. This gives you an overview of the HRS syntax in naming conventions and identification of variables. There is a clear description of the levels of the files. And an overview of the sample and demographics is provided.

QUESTIONNAIRES

The "box and arrow" (B & A) representation of the survey questionnaire, while very technical, can help you understand the question flow for a particular study. It can be used to identify the subset of respondents who were asked a particular question.

Uses for the Questionnaires:

- Variable labels can be cross-referenced to actual questions.
- Some questions are nested in other questions; the B & A format shows how they are handled.
- Some information is preloaded into the survey from previous interviews.
- Some questions are skipped for certain respondents due to 1.) inapplicability, 2.) no change from the previous wave, 3.) preloaded information answering the question, or other circumstances.

Some questions are nested in other questions; the Box & Arrow format shows how they are handled.

NOTE: Letters are also used to designate Year /Wave of an interview, for both Core and Exit this may be the first character in a variable. The Section of the questionnaire (or data set) is the second character in the variable (Chapter 3, pg. 35).

Questionnaire Sections & Categories from the 2002 Core Data.

Section	Category
PR	Preload
A	Cover Screen
B	Demographics
C	Health
D	Cognition
E	Family Structure
F	Parents/Sibs/Decisions
G	Functional Limitations
H	Housing
J	Employment
K	Last Job
L	Job History
M	Disability
N	Health Services and Insurance
P	Expectations
Q	Assets and Income
R	Asset Change
S	Widowhood & Divorce
T	Wills and Life Insurance
V	Modules

Figure 33. Example Questionnaire Section List HRS 2002 – Questionnaire Sections and Categories vary from wave to wave. Use the Concordance to find the section of a category and the label for each wave (Figure 37, pg. 30).

Quick Finder — Section Content by Wave

Content	1992 HRS	1993	1994 HRS	1996 HRS	1998 HRS	2000 HRS	2002 HRS	2004 HRS
Preload Coversheet Demographics	— CS A	— CS A	PRE CS A	PR CS A	PR CS A	PR CS A	PR A B	PR A B
Health Status Health Services OOP Spending ADLs	B B — B	B E E E	B B B B	B E E E	B E E E	B E E E	C N N G	C N N G
Cognition Proxy Cognition Expectations	L — L	C PC H	C — C	C PC H	C PC H	C PC H	D D P	D D P
Family Structure Transfers	CS E	D —	E EE	D[2] (n/a)	D	D	E F	E F
Housing Assets, Debts Income Capital Gains	D M N —	F K J —	D K N V	F J J N	F J J N	F J J N	H Q Q R	H Q Q R
Employment Retirement Pensions	F G H	G — —	FA FB FC G,H	G GG GH	G GG GH	G GG GH	J K L	J K L
Retirement Plans	K		—	—	—	—	—	—
Disability	J	—	J	GD —	GD GJ	GD GJ	M	M
Health Insurance Life Insurance	R R	R R	R R	R R	R R	R R	N T	N T
Widowhood	S,N	—	S	N	N	N		
Divorce Event History	— —	— —	S —	— —	N EV	N EV	S W	S W
Internet Use	—	—	—	—	—	—		
Iwer Obs Thumbnail	X —	— —	— —	— —	— TN	— TN	— —	— —
Physical Measures	—	—	—	—	—	—	—	I

Figure 34. The Health and Retirement Study began as two distinct data collections that merged in 1998. Due to changes in coding sections and adding questions to the questionnaire the letter designations have changed with each wave.

The Health and Retirement Study

A Longitudinal Study of Health, Retirement, and Aging
Sponsored by the National Institute on Aging

Home | Intro/Guide | Documentation | Data | Publications | Search
Descriptions | Questionnaires | Concordance | Codebooks | Modules | Imputations | Data Alerts | User Guides

Questionnaires

Biennial Interview Questionnaires

The following table cross-references Core interview questionnaire content by section across wave. Questionnaires are stored in Adobe PDF format. For year 2002 and thereafter, box and arrow representations of both Core and Exit interviews are included in each questionnaire set; prior to 2002 box and arrow questionnaires were only provided for Core interviews.

On this page...

- Biennial Questionnaires (Online)
- Biennial Questionnaire Sets (Download)
- Off-Year Questionnaires (Online)
- Notes

View Biennial Interview Questionnaire by Section [1] [3] [4]

Content	1992 HRS (Wave 1)	1993 AHEAD (Wave 1)	1994 HRS (Wave 2)	1995 AHEAD (Wave 2)	1996 HRS (Wave 3)	1998 HRS	2000 HRS	2002 HRS	2004 HRS
Table of Contents Preface	- -	Contents Preface	- -	Contents Summary	- -	Contents Preface	Contents Preface	- -	- -
Preload Coversheet Demographics	- Section CS Section A	- Section CS Section A	Section PRE Section CS Section A	Section PRE Section CS Section A	- Section CS Section A	- Section CS Section A	- Section CS Section A	- Section A Section B	- Section A Section B
Health Status Health Care Utilization Health Care Costs ADLs	Section B Section B - Section B	Section B Section E Section E Section E	Section B Section B Section B Section B	Section B Section E Section E Section E	Section B Section E Section E Section E	Section B Section E Section E Section E	Section B Section E Section E Section E	Section C Section N Section N Section G	Section C Section N Section N Section G
Cognition Proxy Cognition Expectations	Section L - Section L	Section C Section PC Section H	Section C - Section C	Section C Section PC Section H	Section C Section PC Section H	Section C Section PC Section H	Section C Section PC Section H	Section D Section D Section P	Section D Section D Section P
Family Structure,	Coversheet Section E	Section D	Section E	Section D	Section D[2]	Section D	Section D	Section E	Section E Section F

Figure 35. Questionnaires Web page
Look on this page for the location of data content by wave section.

Questionnaires

While HRS attempts to maintain longitudinal integrity the questionnaires are different for every study and every wave. Some questions have been reworded or added. Section designations have changed.

See Figure 35. for the location of the section content by wave.

Methods

1a. URL: http://hrsonline.isr.umich.edu/meta/sho meta.php?hfyle=qnaires

1b. From the HRS Home Page click Documentation on the top navigation bar. Click Questionnaires from the list on the documentation page.

2. Locate the Content in the first column and the Year in the top row to find the Section letter for that content on our Web site.

Note: To locate a variable over waves use the

Example From a Questionnaire Page

A26a. Have you always lived in or around MAIN RESIDENCE [CITY/CITY & STATE]?

1. YES 5. NO 8. DK 9. RF

A26b BRANCHPOINT: IF R HAS ALWAYS LIVED IN AREA OF CURRENT MAIN RESIDENCE(A26a=1) or IF MAIN RESIDENCE IS IN SAME AREA AS A RESIDENCE IN R's LAST IW (CS36={1 or 3}), GO TO A36b

A26b. IF KNOW YEAR R MOVED TO CURRENT RESIDENCE (A25):
Before you moved to MAIN RESIDENCE [CITY/CITY & STATE] in YEAR (A25),had you ever lived there before?

IF KNOW HOW MANY YEARS R HAS LIVED AT CURRENT RESIDENCE (CS33A):
Before you moved to MAIN RESIDENCE [CITY/CITY & STATE] YEARS AT CURRENT RESIDENCE (CS33A) years ago, had you ever lived there before?

OTHERWISE:
Before you moved to MAIN RESIDENCE [CITY/CITY & STATE], had you ever lived there before?
1. YES 5. NO 8. DK 9. RF

GO TO A27
BRANCHPOINT

Figure 36. Question A26a. An example question HRS 2002.

Questionnaire Page Example Contents

- The question
- Question number
- The answer codes
- Instructions to the interviewer on how to proceed with the interview considering each respondent's circumstances or previous answers.

Variable Notes

The interview question becomes a variable in the dataset. Each variable is given a unique code.

Since 1995:

- The first letter in a data variable designates the year/ wave of an interview
- The second letter designates the Section in that data file or questionnaire.

Respondent Variables

Numeric and Character Variables

– Early respondent identifier variables in HRS 1992 and 1994 are numeric

– Respondent identifier variables in later waves are characters

– 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

– Most software packages will require that the variables be of the same type. HRS analysts recommend converting numeric identifiers to character type.

Concordance

How to Locate a Variable in a Different Wave/Year

Locate a particular HRS research variable or topic using the Concordance to quickly search our data for the locations of a particular variable.

- The concordance is a tool for cross-referencing interview questions by content across time.
- It contains one record for each of the questions represented by one or more variables in a specified public file. – See User Notes.
- Concordance records are linked over time only in that they have Interest Areas, keywords in common, or *appear in similar sections in different years.*
- Keyword groups are supplemented by the cross-reference feature which links identical or similar questions across waves.

Methods

1a. URL: http://hrsonline.isr.umich.edu/concord/index.html.

1b. From the HRS home page

– Click Documentation on the primary navigation bar

– Click on the hyperlinked word **Concordance**

2. Create a question subset by filtering on wave, interest area, section, short-labeled range, and/or release type. A free-text search option is also available.
3. Choose the variable you would like to view (Figure 38. Example: Time Use).

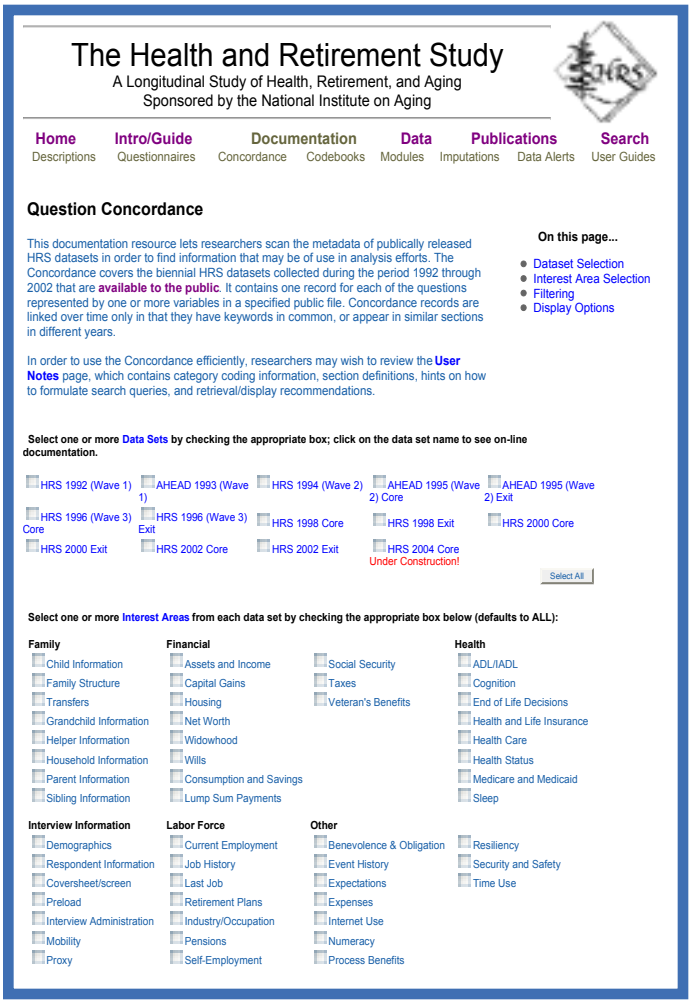


Figure 37. Question Concordance
Check the boxes for one or more data sets and check the boxes for interest areas or select all. Press Enter on your keyboard or click the submit button.

Time Use Variables from HRS Concordance

Wave	Variable Name	Short Label	Long Label	Category	Subcategory	Detail	Xref
AHD 1993 (W1)	V1984	M2-1	MODULE 2: TIME USE	Module	Time Use	MODULE 2: TIME USE	
AHD 1993 (W1)	V1985	M2-2	HOME IMPROVEMENTS LAST 12 MOS	Module	Time Use	HOME IMPROVEMENTS LAST 12 MOS	
AHD 1993 (W1)	V1986	M2-3	YARDWORK LAST 12 MOS	Module	Time Use	YARDWORK LAST 12 MOS	
AHD 1993 (W1)	V1987	M2-4	GROW/STORE/CAN FOOD LAST 12 MOS	Module	Time Use	GROW/STORE/CAN FOOD LAST 12 MOS	
AHD 1993 (W1)	V1988	M2-5	REPAIR OWN AUTO LAST 12 MOS	Module	Time Use	REPAIR OWN AUTO LAST 12 MOS	
AHD 1993 (W1)	V1989	M2-6	#HOURS PREV ACTIVITIES LAST 12 MOS	Module	Time Use	#HOURS PREV ACTIVITIES LAST 12 MOS	
AHD 1993 (W1)	V1990	M2-7	VOLUNTEER FOR CHURCH LAST 12 MOS	Module	Time Use	VOLUNTEER FOR CHURCH LAST 12 MOS	
AHD 1993 (W1)	V1991	M2-8	VOLUNTEER FOR SCHOOL LAST 12 MOS	Module	Time Use	VOLUNTEER FOR SCHOOL LAST 12 MOS	
AHD 1993 (W1)	V1992	M2-9	VOLUNTEER FOR SENIORS LAST 12 MOS	Module	Time Use	VOLUNTEER FOR SENIORS LAST 12 MOS	
AHD 1993 (W1)	V1993	M2-10	VOLUNTEER OTHER LAST 12 MOS	Module	Time Use	VOLUNTEER OTHER LAST 12 MOS	
AHD 1993 (W1)	V1994	M2-11	# HOURS VOLUNTEER LAST 12 MOS	Module	Time Use	# HOURS VOLUNTEER LAST 12 MOS	
AHD 1993 (W1)	V1995	M2-12	HELP OTHERS TRANSPORT LAST 12 MOS	Module	Time Use	HELP OTHERS TRANSPORT LAST 12 MOS	
AHD 1993 (W1)	V1996	M2-13	HELP OTHERS HOUSEWORK LAST 12 MOS	Module	Time Use	HELP OTHERS HOUSEWORK LAST 12 MOS	
AHD 1993 (W1)	V1997	M2-14	HELP OTHERS CHILDCARE LAST 12 MOS	Module	Time Use	HELP OTHERS CHILDCARE LAST 12 MOS	
AHD 1993 (W1)	V1998	M2-15	HELP OTHER THINGS LAST 12 MOS	Module	Time Use	HELP OTHER THINGS LAST 12 MOS	
AHD 1993 (W1)	V1999	M2-16	# HOURS HELP LAST 12 MONTHS	Module	Time Use	# HOURS HELP LAST 12 MONTHS	
HRS 2000 Core	G6954	M9-5	TRADE YEARS FOR HEALTH	Module	Time Tradeoff	YEARS FOR HEALTH	
HRS 2000 Core	G6955	M9-6	NUMBER OF YEARS	Module	Time Tradeoff	YEARS FOR HEALTH-NUMBER YEARS	
HRS 2000 Core	G6956	M9-6A	NUMBER OF MONTHS	Module	Time Tradeoff	NUMBER OF MONTHS	
HRS 2000 Core	G6958	M9-7A	GIVE MORE THAN ONE YEAR	Module	Time Tradeoff	YEARS FOR HEALTH-1 YEAR	
HRS 2000 Core	G6959	M9-7B	GIVE MORE THAN THREE YEARS	Module	Time Tradeoff	YEARS FOR HEALTH-3 YEARS	
HRS 2000 Core	G6960	M9-7C	GIVE MORE THAN FIVE YEARS	Module	Time Tradeoff	YEARS FOR HEALTH-5 YEARS	
HRS 2000 Core	G6961	M9-7D	GIVE MORE THAN SIX MONTHS	Module	Time Tradeoff	YEARS FOR HEALTH-6 MONTHS	

Figure 38. Example: Concordance Search Result–HRS 2000 Core/Time Use
Variables are identified by a variable name, the initial letter indicates the Wave year (e.g., G=2000), a short label, long label closer to the survey question asked, category type of data (core, exit, and module). A module is from the series of experimental questions asked with each survey; subcategory reveals the topic, and detail for the variable. Xref" links to the cross-reference table. Clicking on a link in that column will bring up a window that lists variables from 1995 on that have same or similar content.

Codebooks — Interview Qs

A codebook documenting all variables in each dataset is created for each data product.

Methods

1. URL: http://hrsonline.isr.umich.edu/meta/sho_meta.php?hfyle=codebks.
2. From the HRS Home Page
   - Open Documentation
   - Open Codebooks and click on the year of data applicable

Codebook Contents

The codebook is an explanation of:

- implicitly the wave code indicates the year a wave was done
- the name of the variable
- in what section of the questionnaire the question from which the variable derived can be found
- the level of the question (the question refers only to the respondent i.e., the respondent level, or to the entire household i.e., the household level, or other levels defined by the Data Description and Usage Documentation)
- computer assisted interview (CAI) reference number
- storage mode — variable type is numeric or character (string)
- width refers to the size or the number of storage location in the ASCII data file, i.e. the number of characters in the variable
- decimals indicate the number of implicit decimals, if any
- the section letter and the question number
- the question in the survey
- the distribution (frequencies) of responses for each answer

Figure 39. Codebooks for HRS Data Files
The codebook documents all the variables of a data set.

Online Codebooks for HRS Public Files

Biennial Data Products

2004 Core (Early, August 2005)
Core Imputations (Early, October 2005)
Exit (Early, January 2006)

2002 Core (Final v. 1.0, September 2004)
Core Imputations (Final v1.0, June 2005)
Exit (V1.0, February 2004)

2000 Core (Ver 1.0, Sep 2002)
Core Imputations (Ver 2.0, Dec 2002)
Exit (Final v1.0, March 2005)
Exit Imputations (Early v1.0, Sep 2002)

1998 Core (Ver 2.3, November 2003)
Core Imputations (Ver 3.0, Dec 2002)
Exit (Final v1.0, January 2005)
Exit Imputations (v2.0, June 2004)
Post-Exit (Final V1.0, February 2006)

1996 Core (Ver 4.00, September 2003)
Core Imputations (Ver 3.0, December 2002)
Exit (Ver 1.00, September 2003)
Exit Imputations (Ver 1.0, June 2004)

1995 AHEAD Core (Ver 2.0, May 2002)
AHEAD Exit (Ver 2.0, June 2003)
AHEAD Core Imputations (Ver 3.0, December 2002)
AHEAD Exit Imputations (Ver 1.0, June 2004)

1994 Core (Ver 2.0, September 2004)
Core Imputations (Ver 3.0, November 2003)

1993 AHEAD Core (Ver 2.10, March 1998)

1992 Core (Ver 2.0, September 2004)
Core Imputations (Ver 3.0, November 2003)

Longitudinal and Cross-Wave Data Products

Labor Section Carry Forward Variables (Ver 1.0, June 2005)

Imputations for Pension Related Variables (Ver 1.0, June 2005)

Employer Pension Data Tracker File (Ver 1.0, June 2005)

Imputations for Pension Wealth (Ver 1.0, June 2005)

Cross-Wave Tracker File (2002 ver 2.0, October 2005)

Respondent Region File (Ver 2.0, February 2004)

Longitudinal Other Person Number (LOPN) (Ver 1.0, February 2004)

Off-Year Studies

2003 Consumption and Activities Mail Survey (CAMS) (Early v. 1.0, June 2004)

2001 Consumption and Activities Mail Survey (CAMS) (Ver 1.1, Apr 2004)

Human Capital and Educational Expens es Mail Survey (HUMS) (Early, July 2004)

HUMS College Tuition Imputations (Ver 1.0, June 2005)

1999 1999 HRS Mailout Survey - No Code book Available.

Figure 40. Enlargement of Codebook Web page

Finding Variables in the Codebook

Variables are denoted differently in each wave. The codebook is key to understanding the wave/s of data of interest.

The codebook lists:

- The variable code & name
- Section, Level, Type, Width, Decimals
- CAI Reference (a computer reference)
- Hyperlinks to same variable/s in other waves
- The Interview question on the questionnaire
- The frequencies of response type
- The code frame: the potential response types

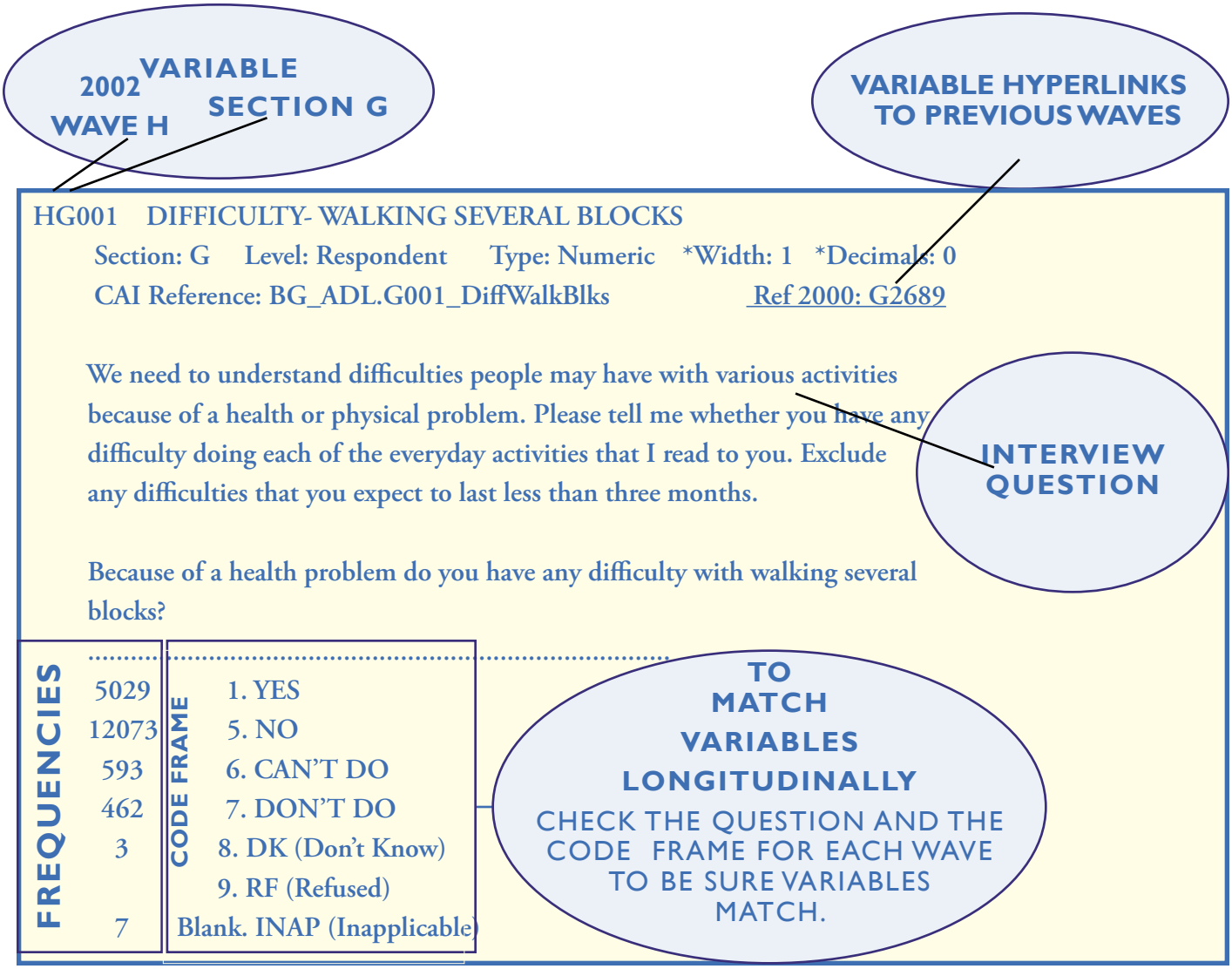


Figure 41. 2002 Core Release Codebook
Variable: DIFFICULTY-WALKING SEVERAL BLOCKS
Section G of Wave H (2002)

Variable: HG001
Level of File: Respondent
Type: Variable type is numeric rather than a character

System of Designation Years/Waves	Code Letter Core/Exit
2004: 2004 HRS Core (Early)	
2002: 2002 HRS Core (Final); 2002 HRS Exit (Early)	H/S
2000: 2000 HRS Core (Final) ; 2000 HRS Exit (Early)	G/R
1998: 1998 HRS Core (Final) ; 1998 HRS Exit (Final)	F/Q
1996: 1996 HRS [Wave 3] Core (Final) ; 1996 HRS [Wave 3] Exit (Final)	E/P
1995: 1995 AHEAD [Wave 2] Core (Final) 1995 AHEAD [Wave 2] Exit (Final)	D/N
A Different System of Designation 1994: 1994 HRS [Wave 2] Core (Final)	W
1993: 1993 AHEAD [Wave 1] Core (Final)	V
1992: 1992 HRS [Wave 1] Core (Final)	V

Figure 42. Wave Designation by Letter Code Core -Living respondents and Exit - Deceased respondents.

Variable Code Letters

Since 1995 code letters fall into three categories:

1. wave core as the first character in a variable
2. wave exit also as the first character in a variable
3. section code as the second character in a variable.

In Figure 42, you see that in a data variable the first letter is the Year/Wave either Core (living respondents) or Exit (proxies for deceased respondents) and the second letter is the questionnaire section.

Module Descriptions

Variables in Our Experimental Modules

In each wave the survey contains some experimental questions, contained in modules asked at the end of the interview. On average, there are about 10 modules each wave. In general modules are:

- To test new topics/questions
- About 3 minutes long
- Administered to a random "nth" of the sample

You can see a description of the experimental module at each wave on our Web site.

Methods

1. URL: http://hrsonline.isr.umich.edu/meta/sho meta.php?hfyle=modules.

or

2. From the HRS home page.
   - Click on Documentation on the navigation bar.
   - On the secondary navigation bar click on Modules.

Figure 43. Descriptions of Experimental Modules
For each wave the topics for the experimental survey questions are listed. Typically, these are only asked in one wave and therefore they cannot be merged longitudinally.

Figure 44. Imputed Datasets
Datasets with high rates on *non-response* are imputed.

An Imputation Example – The Holdings of Stocks and Mutual Funds*

First, the interviewer asks whether the respondent (or his/her spouse) own any shares of stock or stock mutual funds. If affirmative, the interviewer asks the value of these stock holdings. If the respondent is unable or unwilling to provide an exact amount, the interviewer asks whether it is more than $25,000. Depending on the response, additional, smaller brackets are explored so that in the end the range is narrowed down to $0-2,500; $2,500-25,000; $25,000-125,000; $125,000-400,000; $400,000 or more.

These ranges are known as "brackets;" the sequence of probes into increasingly narrow ranges is known as "unfolding brackets" questions. The brackets vary by asset and income category. For example, the cut-off values for dividend and interest income are $200, $500, $2,500, and $10,000.

* RAND HRS Data Documentation, Version C Patricia St.Clair, Delia Bugliari, Sandy Chien, Steven Haider, Orla Hayden, Michael Hurd, Serhii Ilchuk, Gabriela Lopez, David Loughran, Constantijn Panis, Philip Pantoja, Monika Reti, Julie Zissimopoulos. July 2003

Imputation Datasets

The Missing Data Imputation Process

HRS/AHEAD researchers have developed special methods to collect information on variables that can be subject to high rates of non-response, e.g., income, financial assets, housing equity, and medical expenditures.

A key feature of the process is the use of special questionnaire designs to elicit "bracket" information from respondents who are unable or unwilling to answer amount questions. Bracketed data are used to help impute missing data.

Methods

1. URL: http://hrsonline.isr.umich.edu/meta/sho_meta.php?hfyle=impute_files

2. Click Documentation on the primary navigation bar, then Imputations on the secondary navigation bar

Note: For detailed information on the techniques involved in the imputation process, please consult IMPUTE: A SAS Application System for Missing Value Imputations — with Special Reference to HRS Income/Assets Imputations.

Or

3. Open the Home Page.

   – Click on Documentation on the navigation bar

   – Click User Guide on the second navigation bar to find the PDF file

Data Alerts

Data alerts are notices of errors, corrections, or problems in HRS early and final public data releases and associated documentation.

The date associated with each alert reflects the most recent version of that alert. For data alerts which have been superseded, or which apply to past versions of data releases, see the Superseded Data Alerts page. Please report any data alert problems to the HRS Help Desk at (hrsquest@isr.umich.edu)

Methods

1. Check the latest Data Alerts on the HRS home page
2. Click Data Alerts on the secondary navigation bar
3. Find the year of the data you are using

or

1. URL: http://hrsonline.isr.umich.edu/meta/sho_meta.php?hfyle=alerts/index
2. Check the home page for up-to-date data alerts.

Figure 45. Data Alerts
Check here for updates on data alerts while doing analysis.

Documentation Report Series

This Web page contains links to PDF documents from contributors with technical information about survey design or file construction.

Methods

1a. URL: http://hrsonline.isr.umich.edu/papers/sho_papers.php?hfyle=dreptbib

or

1b. From the HRS home page

2. Click Documentation on the top navigation bar.
3. Click Documentation Report Series to the far right on the secondary navigation bar

Figure 46. Documentation Report Series
Researcher contributions of technical information about file construction.

Figure 47. User Guides
An overview of the sample design with hyperlinks to a more comprehensive User Guide.

User Guides

This Web page contains links to PDF documents designed to provide additional information to "hands-on" users of HRS data products.

Methods

1a. URL: http://hrsonline.isr.umich.edu/docs/sho_refs.php?hfyle=index&xtyp=3

or

1b. From the HRS home page

2. Click Documentation on the top navigation bar

3. Click User Guides to the far right on the secondary navigation bar

Current User Guides

User Guides are designed to provide analysts with documentation about the concepts, measures and questions in the HRS surveys. These reports expand upon the information found in codebooks, questionnaires and data descriptions. They also provide comprehensive descriptions of created measures (including their origin), changes made across waves, variable distributions, and results from data quality analysis.

Honggao Cao, John Henretta, Theresa Norgard, Beth Soldo and David Weir, HRS 2001 HUMS College Tuition Imputations (June 2005).

Gwenith G. Fisher, Jessica D. Faul, David R. Weir, and Robert B. Wallace, Documentation of Chronic Disease Measures in the Health and Retirement Study (HRS/AHEAD). HRS Documentation Report DR-009 (February 2005).

Stephanie Fonda and A. Regula Herzog, Documentation of Physical Functioning Measures in the Health and Retirement Study and the Asset and Health Dynamics among the Oldest Old Study. HRS Documentation Report (December 2004)

Mary Beth Ofstedal, Gwenith G. Fisher and A. Regula Herzog, Documentation of Cognitive Functioning Measures in the Health and Retirement Study. HRS Documentation Report (March 2005). Please note that this replaces earlier (2001 and 2002) versions of this report.

Diane E. Steffick, Documentation of Affective Functioning Measures in the Health and Retirement Study HRS Documentation Report DR-005 (2000).

Data Use Guides

Data Use Guides are designed to provide analysts, research assistants and other support personnel responsible for data management with additional documentation about the data structure of individual HRS surveys as well as information on special data manipulation topics.

Michael A. Nolte and Marita A. Servais, Occupation and Industry Coding in HRS/AHEAD. (March 2003).

Marita A. Servais, Overview of HRS Public Data Files for Cross-sectional and Longitudinal Analysis. (May 2005). Please note that this replaces the June 2004 version of this report.

Marita A. Servais, An Elementary Cookbook of Data Management using HRS Data with SPSS, SAS and STATA Examples. (June 2004).

Links to Other Special Documentation Pages

Resources for Analysis of Family Data

Background Information on study design and sampling techniques.

EASY TIP

Other HRS Data Products and Tools

4

What's Available?

Metadata Summary

Experimental Modules

Tracker file

Restricted Data:

Links to

Administrative Data

Pension User Guides and

Estimator Guide

What's Available

Finding All the types of Datasets

Overview of Waves, Cohorts, Subsamples of HRS — The Available Data

To get a quick idea of available HRS data sets, we have two different formats of "what's available" for you on an image map or in a table.

Methods

1. Image Map (Figure 48.)

Click on the first Web page link entitled —Health and Retirement Study: Data Collection Flow.

- URL: http://hrsonline.isr.umich.edu/intro/dataflow.html.

- From the HRS home page:
  - Click on Intro/Guide on the primary navigation bar
  - Click on Data Flow on the secondary navigation bar

Each box/arrow is hyperlinked to a Metadata Summary (Metadata Summary is explained in on the next page).

2. Using the What's Available table (Figure 49.)

- URL: http://hrsonline.isr.umich.edu/data/avail.html

- From the HRS home page:
  - Click on Intro/Guide on the primary navigation bar
  - Click on What's Available on the secondary navigation bar
- Data products are displayed by year (Figure 49). Click on the appropriate hyperlink to view a page containing a summary of metadata resources. The same type of information is in the example in Figure 48. Each entry is hyperlinked to a Metadata Summary (Example Figure 50).

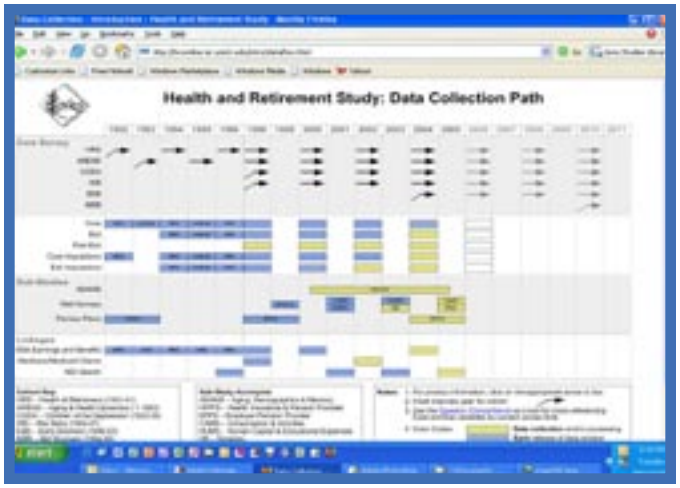


Figure 48. Image Map of the Data Flow or Data Collection Path
You can easily see the history of HRS, AHEAD, the 1998 merger and the current HRS steady state design.

Figure 49. What's Available
Core data are from the surveys taken by respondents. Exit data are from surveys taken by proxies after a respondent has died. Mailout surveys have specific topic (e.g., Time Use).

The Metadata Summary

A detailed summary of the dataset

The previous two methods explained how to navigate to the metadata summary for any given data set.

The Metadata Summary contains the:

a. sample size

b. entry cohort information

c. latest version of data release

d. field dates

e. response rate

f. data alerts

g. hyperlinks to the questionnaire, the data description, codebook, and modules.

See Figure 50. for an example of the content you will obtain.

Example

Metadata Summary for HRS 1998 Core

N: 21,380 (HRS=10,865; AHEAD=6,947; CODA=2,320; WB=2,529)

Field Dates: Feb 1998-Mar 1999

Entry Cohort Info:
HRS: Original R born 1931-1941
AHEAD: Original R born before 1923
CODA (Children of Depression): Original R born 1924-1930
WB (War Baby): Original R born 1942-1947

Response Rate:
(HRS=84.4%; AHEAD=90.5%; CODA=72.4%; WB=70.1%)

Latest Release: November 2003 (Ver 2.3)

Basic Information:

Online Questionnaire	Core Data Description	Codebook	Modules

Alerts:	(2003-11-04)	Respondent-level Data Files Replaced
	(2003-09-30)	Three Replaced or New Data Files
	(2003-09-05)	Data File H98CS_H Replaced
	(2003-05-27)	Detailed Comparison of Variables in Versions 1 and 2 of HRS
	(2003-05-27)	1998 Core
	(2003-05-27)	Corrections to Household Member and Child-level Files (1998)
		New Release of HRS 1998 Core (Final, Version 2.0)

Figure 50. Example: Metadata Summary for HRS 1998 Core
Clicking on either the Data Collection Flow arrows or What's Available table cells will take you to this type of summary information for each data set.

How to Merge data Cross-sectionally or Longitudinally

A Data Use Guide Describes the Process for Merging Different Statistical Packages and See Files That Merge on the Cross-Reference Table on the Web site

An Elementary Cookbook of Data Management using HRS Data with SPSS, SAS and STATA Examples, (June 2004) by Marita A. Servais. This user guide is your next step in data analysis after reading the Data Description if you are unfamiliar with HRS data.

Methods

1. URL: http://hrsonline.isr.umich.edu/docs/sho_refs.php?hfyle=index&xtyp=3

or

2. From the HRS Home Page: Click Documentation on the top navigation bar

3. Click Data Use Guides from the list.

Chapters 6-8 of Getting Started with the HRS

A Tutorial Using the HRS RAND Data with STATA

First-time users familiar with STATA might find this an easier alternative to starting with HRS data. The RAND HRS data is a way to become familiar with a standardized, respondent-level only

File Merge Cross-Reference Table

To check the files at a glance that can be merged together, some are restricted, go to our file merge table Figure 51.

Methods

1. URL: http://hrsonline.isr.umich.edu/rda/xref4.html

or

2. Go to the HRS Home Page and click Data on the primary navigation bar; then click HRS Restricted Data on the secondary navigation bar, the hyperlink is at the bottom of the first paragraph.
HRS File Merge Cross-Reference Table

EASY TIP

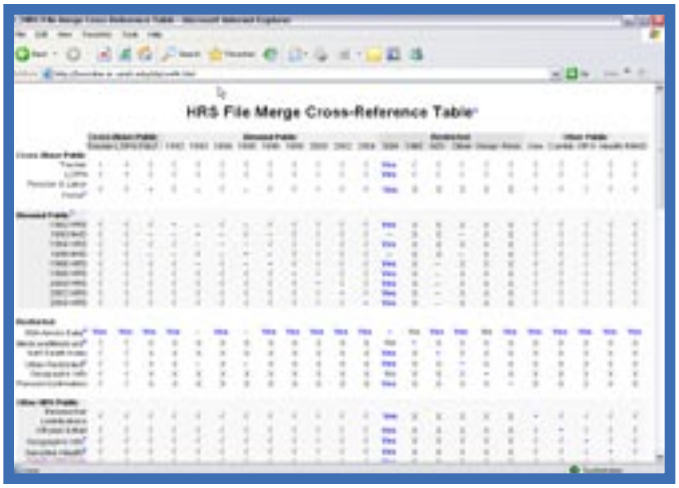


Figure 51. HRS File Merge Cross-Reference Table

Figure 52. Data Description - Tracker File

The Tracker File — How to find Respondents Across Waves

Tracker Data Documentation Products

The tracker file has a full variable listing and weights. Its purpose is to provide core data longitudinally and to faciliatate merging data sets.The data description and usage have many uses with instructions on using the data.

The Tracker file data description (Figure 52).

Methods

1. The data description for the Tracker file is at: http://hrsonline.isr.umich.edu/meta/tracker/desc/trk2002.pdf
2. You can also go to our HRS Home Page:
   - Click on Documentation
   - Click on Data Descriptions and open the hyperlinked Tracker 2002

Figure 53. Tracker 2002, Sampling Weights, LOPN.pdf

The Tracker file data description consists of sections that describe:

- Basic structure of the Tracker file; a full variable listing; and description and an explanation of naming conventions
- Issues related to the variables in the Tracker file
- General instructions on merging the Tracker file with other HRS data
- Cross-sectional weights (Figure 53) for both respondent and household analysis (http://hrsonline.isr.umich.edu/meta/tracker/desc/wghtdoc.pdf) **sample-error computation variables**, and information on inter-respondent relationships, which are vital to almost all substantive household-level analyses of the HRS/AHEAD data (Figure 53)

The Tracker File Codebook

Find the information necessary to translate the data description and documentation for tracking a respondent across waves.

Methods

1. URL: http://hrsonline.isr.umich.edu/meta/tracker/codebook/trk2002_00.html
2. From the HRS Home Page:
   - Click on Documentation. Click on Codebooks in the list
   - Tracker Codebook is at the very bottom
   - Click on the hot link to find the list shown in Figure 54

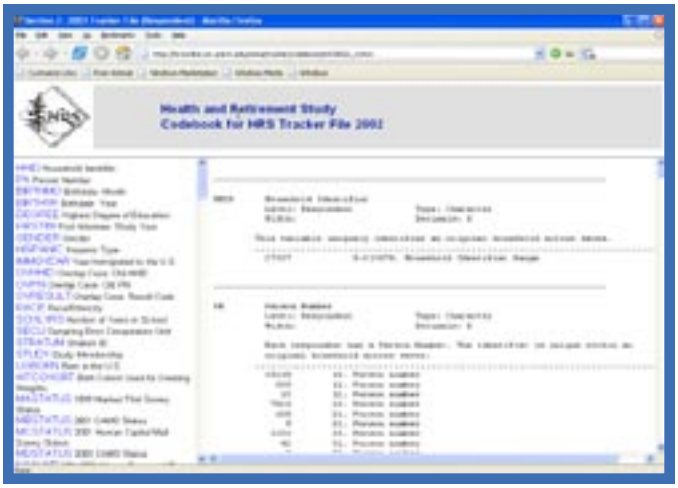


Figure 54. Codebook for HRS Tracker File 2002

The Tracker File Codebook Longitudinal Person Across Waves (LOPN)

Methods

1. URL: http://hrsonline.isr.umich.edu/meta/xyear/lopn/codebook/lopn_00.html
2. From the HRS Home Page:
   - Click on Documentation – Click on Codebooks in the list
   - Tracker Codebook LOPN is at the very bottom and the last entry
   - Click on the hot link to find the list shown in Figure 55

List of Files

AHD9302	AHEAD Sub-sample 1993-2002 (Story)
HRS9202	HRS Sub-sample 1992-2002 (Story)
CWB9802	CODA / War Baby Sub-samples 1998-2002 (Story)
LC92_HL	HRS 1992 Core (HHList)
LC92_K	HRS 1992 Core (Kid)
LC93_MC	AHEAD 1993 Core (HHMemberChild)
LCX94_HL	HRS 1994 Core/Exit (HHList)
LCX94_K	HRS 1994 Core/Exit (Kid)
LC95_MC	AHEAD 1995 Core (HHMemberChild)
LX95_MC	AHEAD 1995 Exit (HHMemberChild)
LC96_MC	HRS 1996 Core (HHMemberChild)
LX96_MC	HRS 1996 Exit (HHMemberChild)
LC98_MC	HRS 1998 Core (HHMemberChild)
LX98_MC	HRS 1998 Exit (HHMemberChild)
LC00_MC	HRS 2000 Core (HHMemberChild)
LX00_MC	HRS 2000 Exit (HHMemberChild)
LC02_MC	HRS 2002 Core (HHMemberChild)
LX02_MC	HRS 2002 Exit (HHMemberChild)

Figure 55. Codebook for HRS Tracker File 2002 LOPN List of Files

Weights, and Clustered and Stratified Sampling

The HRS includes oversamples of African American, Hispanic respondents and must be weighted to yield unbiased estimates of population parameters. Most statistical packages accommodate such sampling weights to correct for these biases. Additionally, the HRS sample is clustered and stratified and unless this is taken explicitly into account, standard errors based on the assumption of simple random sampling will be biased (generally downward). A variety of statistical packages including STATA (the "svy" command set), IVEware and som procedures in SAS.....allow for estimation of design corrected standard errors.

- This can be found on the Web site at:

  http://www.isr.umich.edu/src/smp/ive/.

Variables may be Different Across Waves

Reading the individual wave/year documentation will keep you apprised of the differences in each dataset. Remember every year is different. Every experimental module is different. Questions change and new questions are added. Respondents may participate one year, skip a year, and do an interview at a subsequent time. There are new respondents. These changes are captured in the documentation packaged with the dataset.

The HRS Tools keep you oriented to the differences and alert you as to how to accommodate them in your analysis.

Data analysis reminders

HRS Oversamples and Sampling is Clustered and Stratified

See box to the left. Sample-error computation variablesmay be found in the Tracker file.

Couples

It is important to be clear about the marital status of an R. All married respondents are considered couples, *but not all couples are married*. When assigning marital status of an R it is advisable to use the marital status found in the wave of data used in the analysis since it may change across Waves.

Age of Respondent

When determining age of an R, for an analysis confined to one wave of data, it is advisable to use the calculated age variable used in the analysis. For all other types of analysis to calculate the age refer to the Tracker file for month and year of birth and the wave year.

Working with a Published HRS Researcher

We recommend having research you are working on, for the first time with HRS data, vetted by a published HRS researcher, due to the complexity of the data and the variety of methods for weighting, merging and analyzing the data.

Send Us a Copy

Please send your working paper or published paper to: hrsquest@isr.umich.edu.

Your Own Data Tips

Write a reference of your methods for dealing with issues concerning the data. Many researchers have made contributions to our body of data user guides. If you would like to add to our list of data user guides send to:

Help Desk: Submit a question

E-mail: hrsquest@isr.umich.edu

Health and Retirement Study Special Access Files

There are three types of HRS data Public Use, Special Access Files, and Restricted Data. The Special Access Files contain sensitive medical and other data about respondents and require a special application process for use. They are only released to researchers who qualify for access through a supplemental registration system. The application and download process is described below.

Figure 56. Special Access Files Agreement

Method

1. Indentify the desired data set(s) under the HRS Special Access File Request Information heading.

- Aging, Demographics, and Memory Study (ADAMS) (to be released 2006)
- Diabetes (to be released 2006)
- Prescription Drug Study (to be released 2006)

2. Download and complete the Data Use Agreement in hardcopy format. Using the

   ADAMS data as an example: click on

http://hrsonline.isr.umich.edu/adams/DUA/HRS_ADAMS_DUA.pdf

4. Send a signed copy of the agreement via surface mail to:

   Health and Retirement Study

   DUA Review Committee

   426 Thompson Street, Room 3050 ISR

   Ann Arbor, Michigan 48104-2321

The Approval Process

The Health and Retirement Study will review your request, and based on the information that you provided when you registered at the HRS User Registration/File Download Web site, verify your identity and institutional affiliation. Once this authentication process has been completed to our satisfaction, we will authorized access to the desired data set(s). You should receive a decision within 5 working days. We will communicate with you at the email address that you provided when you registered at the File Download Web site.

How to Download the Data

1. Once you receive your approval notification, login at the HRS User Registration/File Download Web site

2. In the HRS Special Access Files box on the right hand side of the page you will now see links to the data set(s) that you have requested.

3. Click on the appropriate link and follow instructions to download the file(s) you need.

NEW in 2006!

The Aging, Demographics, and Memory Study (ADAMS) Data Access

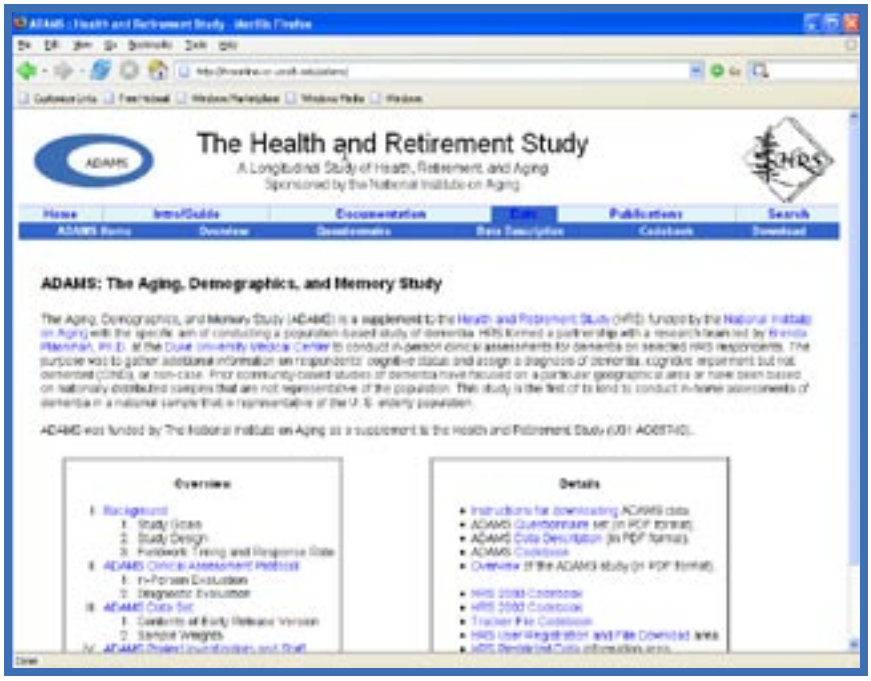


Figure 57. The Aging, Demographics, and Memory Study (ADAMS)

The ADAMS data were briefly described on pg. vii of the Introduction and are considered SpecialAccess Files. See pg. 48 to review the process for accessing the data.

Method

1. To find out more information about the study goals and design; the clinical assessment; and the available data on our Web site click on: http://hrsonline.isr.umich.edu/adams/

or

2. On the primary navigation bar click Data and on the secondary navigation bar click ADAMS Home.

What is ADAMS?

The Aging, Demographics, and Memory Study (ADAMS) is a supplement to the Health and Retirement Study (HRS) 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 Plassman, Ph.D. at the Duke University Medical Center to conduct in-person clinical assessments for dementia on selected HRS respondents. The purpose was to gather additional information on respondents' cognitive status and assign a diagnosis of dementia, cognitive impairment but not demented (CIND), or non-case. 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.

ADAMS was funded by The National Institute on Aging as a supplement to the Health and Retirement Study (U01 AG09740).

The Restricted Data Files

Some files are detailed data provided by Rs, but too sensitive to release publicly. Other data come from linkages with: Social Security earnings and benefits, pensions, Medicare, and National Death Index information. These can be merged with the publicly available HRS data sets.

The restricted data sets from the Health and Retirement Study (HRS) hold a rich potential of research opportunities. The HRS reflects many analytical and policy interests of researchers in various disciplines. Restricted data files multiply opportunities for data analysis for these researchers.

These data files are restricted to researchers who meet eligibility requirements and who apply for use. They may not be distributed to the general public due to participant confidentiality concerns. Due to these concerns also, there are some restrictions on merging data.

Method

1. From the HRS Home Page click on Data on the top navigation bar. On the Data page click on HRS Restricted Data on the secondary navigation bar.

For questions contact either:

Procedural Questions
Cathy Liebowitz
Health and Retirement Study
3044 ISR
426 Thompson St.
Ann Arbor, Michigan 48104
Phone: 734-763-4180
Fax: 734-647-1186
E-mail: catlieb@umich.edu

Technical Questions
Michael A. Nolte
Health and Retirement Study
3100 ISR
426 Thompson St.
Ann Arbor, Michigan 48104
Phone: 734-936-1903
E-mail: manolte@isr.umich.edu

Restricted Data Files

Social Security Administrative Data
Income and benefit information for HRS participants derived from SSA earnings and benefits records

Geographical Data
Participant location information that includes state, county, zip code, and census tract

Industry/Occupation Data
Participant occupation and industry details

Pension Estimation Data
Comprehensive pension projection information derived from employers' records

Medicare Claims and Summary Data
Information on health care costs, use of services, and diagnosis

Cross-Year NDI Cause of Death
Death information on deceased HRS participants including month/year and cause

Health and Retirement Study Date of Interview Data
Exact day, month, and year of participant interview (variables not included in public data)

Human Capital Mail Survey (HUMS) 2001 Tuition Imputation
(variables not included in public data)

Figure 58. What's Available - Restricted Data Links http://hrsonline.isr.umich.edu/rda/reslis2.php

Overview of Eligibility Requirements and Application Process

A researcher must meet the following requirements to obtain access to HRS restricted data:

1. Hold a tenure or tenure track position, or equivalent at an institution with a DHHS Federal-Wide Assurance of Protection for Human Subjects

2. Current Receipt of Federal Research Funds as a PI or CoPI

3. Written Research Proposal

4. Restricted Data Protection Plan

5. Human Subjects Review

6. Agreement for Use of Restricted Data from the Health and Retirement Study

Unannounced Inspections

Restricted Data Users are subject to outside unannounced inspections to verify compliance with these agreements.

Recertification and Renewal

There is a yearly recertification procedure. Then, once the date specified in the contractual agreement for access to HRS restricted data has passed, there is renewal process. For detailed information see our website: http://hrsonline.isr.umich.edu/rda.

Frequently Asked Questions about Restricted Data Files

Q: I am a researcher who can't meet the eligibility requirements. Is there any way to get HRS Restricted Data for my use?
A: Try to work with a researcher at your institution who does meet the requirements. Also, the Michigan Center on the Demography of Aging has opened a restricted data facility that houses HRS Restricted Data. Researchers who cannot meet the usual restricted data requirements may be able to use the MiCDA Data Enclave. For more information about the enclave please review the website: http://micda.psc.isr.umich.edu/enclave.

Q: I am a tenured researcher with federal funding who works for a research organization that is not part of an educational institution. Is there any way I could qualify for access to HRS Restricted Data?
A: In many cases, the answer is yes. Contact the Health and Retirement Study for details.

Q: I am a researcher who works for a United States Government Agency. Is there any way that I (or my agency) can qualify for access to HRS Restricted Data?
A: Yes, contact HRS for details.

Q: I am a faculty member or student at an educational institution located outside the United States who wants to use HRS Restricted Data. Can I gain access to the restricted files listed on this site?
A: In all likelihood, probably not, since you must meet the basic requirements i.e., United States Government funding at an educational or research institution holding a DHHS Federal-Wide Assurance of Protection for Human Subjects (formerly Multiple Project Assurance).

Q: What computing environments and operating systems do you support?
A: We distribute our data in formats that are compatible with most computing environments and operating systems.
Environments: Standalone workstation (recommended), networked workstation, client-server.
Operating Systems: Windows XP/2000, MacOS X, all flavors of Unix. If other, contact HRS.

Q: How will the restricted data files be delivered?
A: The data are stored in a PGP self-decrypting archive (Windows) or PGP static key file (Unix, MacOS X) and are delivered on CD-ROM or diskette by courier service.

Pension Analysis

Researcher Contributions

For those interested in researching pensions and are curious about what other's have done with links to pension data see the What's Available page at the very bottom.

Figure 59. What's Available

At the very bottom of the page find the Researcher's Contributions for Pension data analysis.

Methods

Using the What's Available table (Figure 36.)

- URL: http://hrsonline.isr.umich.edu/data/avail.html
- From the HRS home page:
  - Click on Intro/Guide on the primary navigation bar.
  - Click on What's Available on the secondary navigation bar

Researcher Contributions[3]

Dataset Name	*Datasets Used*	*Authors*	*Latest Release*
401k File (Ver 1.0)	HRS 1992	Bob Peticolas	Jul 1999
HRS/AHEAD Second Home Ownership and Equity Corrections (Ver 1.0)	AHEAD 1993, AHEAD 1995, HRS 1996	Honggao Cao; F. Thomas Juster	Aug 2004
Imputed Medical Expenditures (Ver 1.0)	HRS 1994	Dan Hill	Apr 1999
Pension Present Value Database (Ver 1.0)	HRS 1992	Bob Peticolas; Tom Steinmeier	Sep 1999
RAND HRS Data (E)	HRS 1992 through HRS 2002	RAND	Apr 2004
Self-Reported Pension Wealth (Ver 1.0)	HRS 1992	Anonymous	Oct 1999
Tax Calculations for HRS 2000 & 2002 (Ver 1.0)	HRS 2000, HRS 2002 (All Cohorts)	S. Rohwedder, M. Hurd, P. Pantoja, P. StClair	Jun 2005

Figure 60. Researcher Contributions

Researchers with access to restricted data may find these helpful.

New Pension Estimation Program

by Bob Peticolas and Helena Stolyarova

Overview

Pension entitlements estimated by the new Pension Estimation Program can be expressed as cash flows, present values or both. The pension data and related respondent information are contained in Microsoft Access database files entitled HRS92.mdb and HRS98.mdb , corresponding to the data collected in 1992 and 1998, respectively. The Pension Estimation Program is free to registered users; the data required by the program are available to researchers who have negotiated a Restricted Data Agreement with the HRS. The user interface of the Pension Estimation Program is organized as a tabbed dialog. The application allows the researcher to modify many of the assumptions used in calculating pension benefits. The user can also specify the outputs required and the filenames to be used for output and log files.

New features

- User Interface: The parameter file is accessed by means of a visually intuitive graphical user interface which allows users to modify and save assumptions or restore original default settings
- Cash Flows option: This is a major enhancement allowing users to output a vector of actual cash flows paid to a respondent
- Intermediate Results options: This option allows users to store intermediate results (values of the variables and formulas) in a separate file
- Output options: Users are given the ability to specify a list of items to be written to the output file. For example, users interested in obtaining normal retirement benefit estimates will be able to select and save calculation output for this benefit type alone.
- Multiple Runs: Users now have the ability to increment certain parameters by user-defined values and perform multiple runs without restarting the application
- Application Data: All data files (coded plan descriptions, respondent information, parameter file, mortality tables) have been moved to Microsoft Access for easy database access

For Detailed Information Refer to:

Pension Estimation Program Users Guide by Bob Peticolas and Helena Stolyarova

at http://hrsonline.isr.umich.edu/papers/pep/PenCalcUser.pdf

Final Note — HRS Data Files

The HRS data files and the documentation to support them are vast. If you have any problems or questions regarding our study, our data, data analysis, or any of our tools, search the Web site, or contact us. We are here to help you.

Thank you for taking the time to investigate our study and for considering our resources.

5

HRS Glossary & Acronyms

Glossary

A

Age-Eligible Household: A household with at least one age-eligible respondent (each HRS household is supposed to have at least one age-eligible respondent, but there are a few exceptions).

B

Middle Boomers: born between 1954-1959 future HRS Respondents (plan to be added to cohort in 2010)

C

Case ID: "case identifier". (Refers to HRS 1992)

Case Identifier: [also "Case ID".] A unique five-digit number assigned to each individual respondent in Wave 1 of the original HRS. Case IDs from 1-9999 were given to respondents who are single. Primary respondents (R1s) received Case IDs ranging from 10,000 to 19,999, while Secondary respondents (R2s) received Case IDs ranging from 20,000-29,999. Variable one is always the Case ID, and is present in all individual-level files. (This only applies to HRS 1992)

Core file: an HRS data file excluding end-of-life interviews

Cross-wave: file/s of data that includes variables across multiple or all waves of HRS (ex. Tracker file)

Cross-sectional analysis: Data analysis within one wave of a data file

D

Data Alert: notices of errors, corrections, or problems in HRS early and final public data releases and associated documentation.

Data File: an HRS product containing the data from a study wave or mailout questionnaire.

E

Early Baby Boomer (EBB): Refers to the HRS respondents born from 1948-53. Their cohort was added to the study in 2004.

Early Release: first data set released following a survey, HRS early releases are provided within a very short time of the end of a data collection period. They are provided for the convenience of users who wish to begin analyzing data prior to a final release being available. Early release files are at the respondent level.

Exit files: based on an exit interview with a proxy for the respondent after the R has died.

F

Family Composition: Refers to persons related to either the R1 or the R2. In the HRS, there are separate files for children, siblings, and parents, whether they live in the household or not. Other relatives are present in the household listing if they live in the household.

G

Guides: Refers to the contributions of researchers on data use

H

HH: household

Household Identifier: [also "HHID"] Each HRS household (up to two respondents and persons living in the household with them) has a unique six digit Household ID. The Household ID is present in all files.

Household level: A household sampled by the HRS, including all persons who live there. Upon being interviewed, each sample household was assigned a Household Identifier (HHID). The HHID is stable, and uniquely identifies the original household across time.

Household-Level Files: Data files in which each record/line contains not information of an individual nature, but concerning the household.

Household Listing: One of the data files; a list of every person living in the household, one person to a record/line.

HRS: Health and Retirement Study

I

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

Imputation: created data based on similar responses by similar individuals

Individual-Level Files: Data files in which there is one respondent per record/line.

L

Line Number: A family member identifier. This is primarily for internal sample management purposes.

M

Mail-out file: based on questionnaires mailed out to a set of HRS respondents in off-years from core HRS biennial surveys

Married Pairs: See "pairs".

P

Pairs: [also "Married Pairs"] Households in which there are respondents (who are married or partnered).

Person Number: A three-digit identifier for each family member. The number is unique within each household (no other family members will have the same person number, but the numbers are reused in other households).

Proxy: Helper to the respondent who can answer survey questions accurately. Answers exit interview questions after death of respondent.

R

R: See "respondent" defined below.

R1: The first peron interviewed in a household; usually answers cover screen questions

R2: The second person interviewed in a household

R (Financial): The R that answers the financial questions

R (Family): The R that answers family questions (in single or remarried HH, may be same R as financial R)

Relationship Code: A code that represents the relationship of each family or household member to the respondents.

Respondent: [also "R"] A person surveyed by the HRS.

S

S/P: Spouse/partner (of a respondent, normally).

Subhousehold: All households are assigned a SUBHH of 0 in the first year of collection. A SUBHH of 0 indicates that the original household remains intact. A SUBHH of 1 or 2 recognizes households that have broken off from the original household due to the severance of partnership. A SUBHH of 3 indicates a deceased respondent, is considered to

now be in a household of their own.

Common Health and Retirement Study (HRS) Acronyms or Abbreviations

ADAMS	Aging, Demographics and Memory Study
ADLs	Activities of Daily Living
AHEAD	Study of Asset and Health Dynamics Among the Oldest Old (1890- 1923 cohort)
MBB	Middle Baby Boomers (1954-59 cohort)
CS	cover sheet or cover screen
CAI	computer assisted interviewing (computer program)
CAMS	Consumption and Activities Mail Survey, longitudinal off-year mail-out; includes questions about R activities and consumption
CODA	Children of the Depression (1924-1930 cohort)
DK	Don't know (the answer)
EBB	Early Baby Boomers (1948-1953 cohort)
FC	Transfer-from-child-level
H	Household-level
HIPPS	Health Insurance and Pension Plan survey
HHID	Household identifier
HP	Helper-level
HUMS	Human Capital and Educational Expenses Mail Survey
IADL	Instrumental Activities of Daily Living
INAP	Inapplicable (usually from preload information)
IW	Interview
IWR/IWER	Interviewer
ISR	Institute for Social Research
MC	Household-member-and-child-level
NDI	National Death Index
NIH	National Institutes of Health
NIA	National Institute on Aging
OYMO	Off-year mail out
PC	Proxy cognition
PN	Person number
PR	Preload (Information obtained from during a previous interview
R	Respondent or Respondent-level
RF	Refused to answer
SRC	Survey Research Center
SSA	Social Security Administration
SB	Sibling-level
TC	Transfer-to-child-level
VOL	The Respondent volunteers an acceptable response that was not among the choices presented to the R by the interviewer
WB	War Baby (1942-1947 cohort)

Notes:

PART TWO
TUTORIAL
USING THE HRS DATA AND HRS RAND DATA

6

Using the HRS RAND Data with STATA

Preparing to Load the RAND Data and Run Analyses

Your first steps are to begin setting up your system for saving the documentation, the data files, and your work.

1. Research Binder

Assemble a 3-inch binder to contain your project definition and variable list; printed data documentation; statistical worksheets; notes on your project; and other materials.

2. Computer Folders and Files

Set up the computer folders and files in advance and they are ready when you begin downloading.

Create new folder C: RANDHRS_ProjectName, save new folders into this folder including Data Documentation, Data Library, Text, and Graphics.

C: RANDHRS_ProjectName Folder

- Data Documentation Folder will contain:
  - RAND Data Documentation and Codebook
  - RAND Data Description

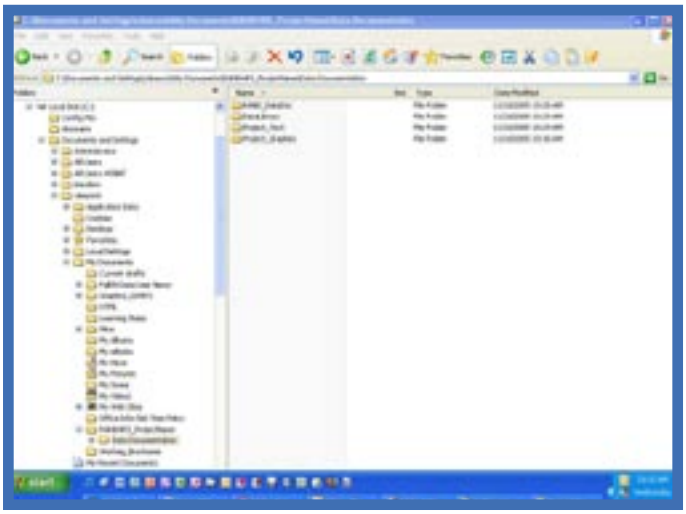


Figure 61. Computer File Folders It is easier in the long run to prepare file folders before downloading the data and data documentation.

EASY TIPS

Print the Data Description PDF and add it to your binder. Save time by reading the background of HRS and the RAND Data File construction. *Don't print RAND Data Documentation and Codebook without considering that it is over 1000 pages.* Minimize it instead, and keep it open for reference while using STATA.

The HRS RAND Data

The RAND data files are a processed set of the same variables across waves using the original HRS data. A sample of respondents from the original number of Rs is used. Variables are constructed by using information from sometimes more than one original HRS variable.

The RAND file is constructed at the respondent level, whereas the original HRS data have many levels of data, including: household, sibling, children, and parents. The RAND Data while easier to use initially will limit the researcher's possibilities for analysis in comparison to the HRS data files.

Data Library File Folder will contain:

- HRS RAND data file
- Data files you create
- STATA log.smcl files

Project Name Folder will contain:

- Project plan
- Plan text
- Feedback and ideas
- Abstract
- Working Paper
- Graphs

3. Obtaining RAND Data and Documentation Files

As you proceed through the following steps you will use the computer folders you have just set up to save the data and documentation files.

To obtain the RAND data you enter the Health and Retirement Study Home Page. This Web site has a wealth of information that will be helpful throughout your project.

The Health and Retirement Study Web Site

The HRS Home Page

The home page contains the latest news in product announcements, news items, general HRS information, data alerts (corrections to data already released), the latest HRS Data User Newsletter, and a link to the HRS Participant Web site.

Method

1. To view the home page of the HRS Web site (Figure 2) insert: http://hrsonline.isr.umich.edu into your internet browser address box.

   Bookmark this page.

2. Notice that on our Web site each page has two navigation bars that indicate where you are in the site page hierarchy. (Figure 3)

   – The top (primary) navigation bar has links to the main sections: Home, Intro/Guide, Documentation, Data, and Publications. For convenience, a hot link is also provided to the Search function (powered by Google).

   – The name of the section you're currently visiting is highlighted in burgundy on a blue background.

   – The secondary navigation bar below the top bar is specific to each section, providing links to the main subdivisions of that section.

HRS Products include all waves of data, documentation, and data links, including RAND data files and documentation.

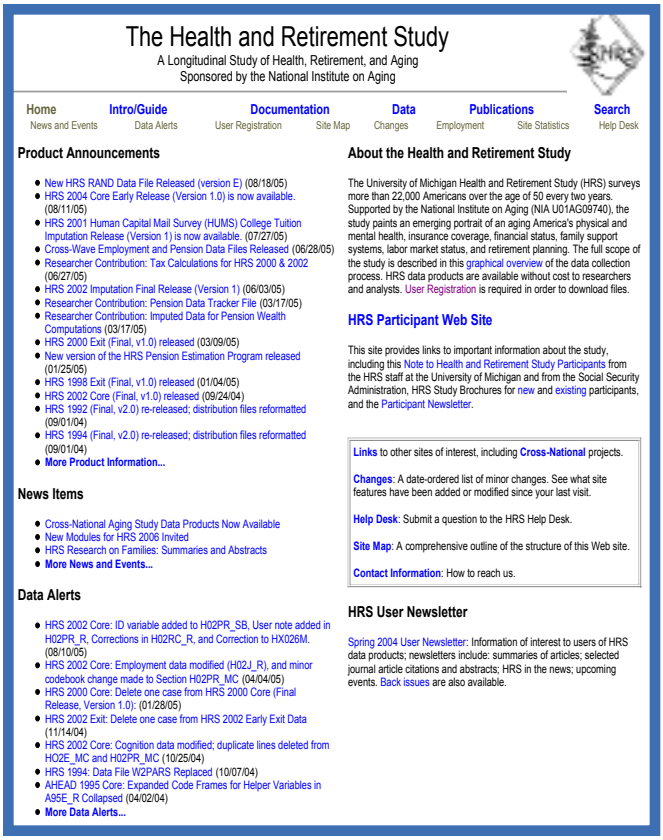


Figure 62. HRS Home Page (Enlarged)
This Web page is continually updated with our latest releases. *Bookmark this page.*

Figure 63. HRS Home Page (Regular)
Use the *Top Navigation Bar to find Web page categories and the *Secondary navigation bar for specific pages in that section.

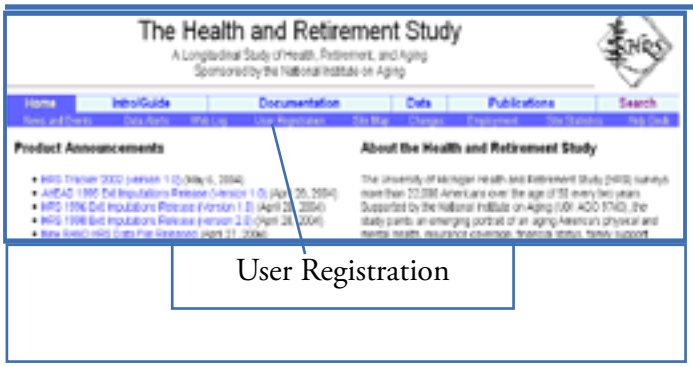


Figure 64. HRS Home Page — User Registration
Click here to go to registration page.

Figure 65. Registration Page — User Registration
Click here to register to obtain HRS data.

Figure 66. The WinZip Files of RAND Data

User Registration

To register to use the HRS data start from the Home Page, find User Registration on the secondary navigation bar (Figure 62) and click the hot link. These data are free for public use.

Obtaining the HRS RAND Data and Documentation

Method

1. Register for a password (if you have not already done so) by clicking on the HRS Home Page>Data>Access to Public Data Sets and Files>Click HERE to register for Public Release files (a password will be sent to your e-mail quickly) or Click HERE if you have already registered for Public Release Files.

2. Enter User Name and Password

3. Download the HRS RAND Data >Datasets and Files >RAND HRS Data File (v.F) from the box – RAND Contributed Files

4. Choose the data file appropriate to your statistical package SAS or STATA (SPSS users download SAS).

- For example, the file randestata8se.zip indicates a RAND data file, version e, for STATA 8 (or higher) Special Edition. Check your software for it's exact name if you are unsure which to use.

5. Use WinZip to open the file you chose, save in RAND_Data_Library folder.

6. Save this original data set under another file name, for example randestata8se_wrkg.dta, so one copy remains unaltered, this prevents having to reload the data set again. This second file is your working data file.

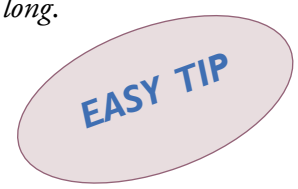
RAND HRS Data Documentation and Codebook and the Data Description

Get the RAND HRS Data Documentation and Codebook and the Data Description PDFs from the HRS Web site without a password. These are also available as zip files from Data Products with your password.

Method

1: Go to http://hrsonline.isr.umich.edu/meta/rand/index.html

(Optional) Download both into your Data Documentation folder and print the Data Description PDF for easy reference. ***Don't print the Data Documentation and Codebook without considering that*** *this document is over 1000 pages long.*




The RAND HRS Data (Version F)

1. Overview

1.1. Data Description

Figure 67. RAND HRS Data Documentation

Reading the Data Documentation and Codebook

Keep the Codebook minimized, but open while working in STATA. The Codebook is broken down into sections by discipline that makes finding variables easier. The Search function (clicking on the binoculars on the task bar) is very helpful in finding keywords or variables you may be searching for in the document.

Familiarize yourself with the history, goals, and cohorts, read the RAND Data Description and peruse the RAND Data Documentation and Codebook. Discover the variables available in your discipline, and/or any other variables that applies to your research.

In Addition —

Please refer to the guide *Using the RAND Versions of the HRS Data* for more information by the RAND Center for the Study of Aging.

Please refer to the guide Using the RAND Versions of the HRS Data by the RAND Center for the Study of Aging for more information. This documentation gives an overview of this product and another, referred to as the "RAND Fat Files," and how they can be used together. See pg. 73 for details.

http://www.rand.org/labor/aging/dataprod/randhrsprod2005.pdf

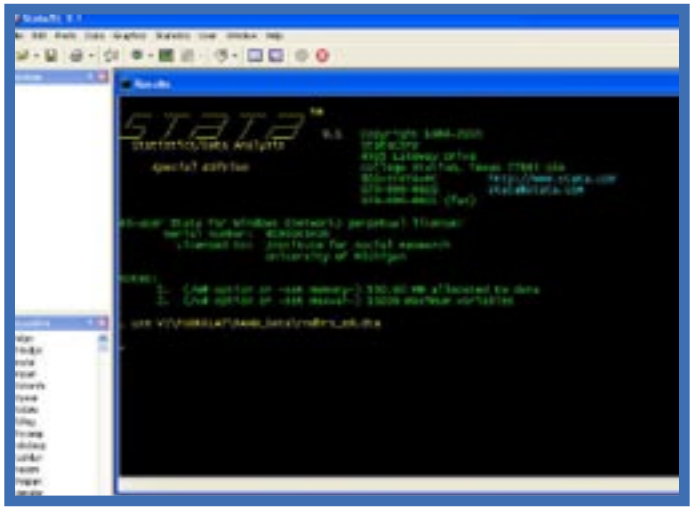


Figure 68. STATA Statistical Program Window
The RAND data variables are listed in the window in the lower left corner.

Increasing Memory for the Large RAND Data file

You may need to increase the memory in your program to load this file since it is about 190,000 KB. If you receive an error message that the file is too large to download into STATA: no room to add more observations r(901). Use these commands.

COMMANDS: set memory 190m, permanently

set maxvar 10,000, permanently

Invoking STATA with RAND Data using Windows

Method

1. Open the working STATA data file you saved in the Data Library file. (Remember to keep the original unaltered.)

>C:\randhrs_Data_Library\ randestata8se_wrkg.dta
The data take several minutes to load.
>C:\randhrs_Data_Library\ randestata8se.dta.
Reminder: STATA does not respond to any pathway with a space in it. Either runwordstogether or put_underscores_in_between for file names.

2. Open a log so you can review what you have done later, otherwise STATA doesn't save the work on the project. Use the following syntax to crate a log called "randproj." STATA automatically adds the suffix, .smcl (STATA markup control language).

COMMAND: log using randproj

3. After completing work in the file, if you need to save your work, save it to a new file name in STATA. On some platforms, STATA operates best by pasting in the Pathname of the folder from the Windows Explorer browser. Paste it into STATA, and then name the new file with a .dta suffix.

4. To copy and paste the folder name open the Explorer browser (by right-clicking on the mouse over the start Windows button). Navigate to the folder. Right click on the folder and Copy Pathname by left-clicking the mouse. Paste into the STATA Command window. Type COMMAND save into STATA, paste the destination folder, backslash, newfile.dta, like this:

COMMAND: save (paste name from Explorer window) C:\randhrs_Data_Library\newfile.dta

You have created and saved your working data file.

Finding RAND variables

Example variables Felt Lonely Waves 3-6

Method

1. Open the RAND Data Documentation and Codebook (RDDC) PDF.

2. Use the binoculars in the PDF to search for the keyword lonely. (Optional) Print the pages from the RDDC concerning these variables and put them in the binder under Data Documentation.

3. In Section B: Health find Felt Lonely variables 1-6 for the respondent. (This assumes you have read the introduction to the RDDC on how variables were constructed and the meaning of the code for each variable.)

RAND Data Documentation and Codebook (RDDC) PDF

Section B: Health: Felt Lonely

Variables

R1FLONE **this is the respondent variable for the Wave of Interviews done in 1992**
R2FLONE**–1994**
R3FLONE–1996
R4FLONE –1998
R5FLONE – 2000
R6FLONE – 2002

Correcting Typographical Errors in STATA

When writing code for STATA it is easy to make a typographical error, click in the Command window and hit Page Up. The command that received the error message from STATA will reappear for you. Correct the COMMAND code and hit enter again.

```
. codebook r5flone

---------------------------------------------------------------------------------------------------
r5flone                                                    r5flone:w5 cesd: felt lonely
---------------------------------------------------------------------------------------------------

              type:  numeric (byte)
              label:  r5flone

              range:  [0,1]                units:  1
        unique values:  2                  missing .:  7149/26728
    unique mv codes:  5                    missing .*:  2076/26728

      tabulation:  Freq.   Numeric  Label
                   14281        0  0.no
                    3222        1  1.yes
                    7149        .
                       6        .d
                       6        .m
                       2        .r
                    2062        .s
```

This code tells you that r5flone is w5: wave 5 (2000 data collection) cesd (from the Center for Epidemiological Studies Depression scale): the abbreviation from the original question on the questionnaire is felt lonely. The type is a numeric variable (important for doing statistical analysis, as opposed to a string variable). The scale, the range of numeric values, is 0-1, the labels of the values are 0. no and 1. yes. The frequency: the number of people from the sample n answering 0. no is 14,281. The list of missing codes in the scale.

Missing Code	Code: Reason for Missing
.	Reference person did not respond to this wave
.D	Don't know
.R	Refused.
X	Does not apply (specifics depend on variable)
.Q	Data not available because of HRS and AHEAD survey instrument differences in Wave 2 or 3
.U	Reference person is not married (for spouse variables)
.V	Spouse did not respond this wave (for spousal variables)
.S	Information not available due to skip patterns, typically because the interview is by proxy respondent
.M	Other missing

The STATA Commands

Codebook

In addition to the information in the RDDC, you can obtain specific information about a variable using the codebook command within STATA.

Command: codebook r5flone

(The variables are in caps only in the documentation, not in the RAND STATA data file.)

The coding scheme varies across variables. Consult the codebook; look in the Appendix for details on individual variables. STATA introduced the ability to distinguish multiple types of missing values in its Version 8. The distinctions are therefore not available in Stata V7 versions of the RAND HRS.

Tab Command

Each STATA command gives you different, but valuable information, not only in a statistical analysis, but about a single variable. As seen in these examples. Tab gives you:

Tab Command

```
. tab r5flone

 r5flone:w5 |
 cesd: felt |
    lonely |      Freq.     Percent        Cum.
------------+-----------------------------------
       0.no |     14,281       81.59       81.59
      1.yes |      3,222       18.41      100.00
------------+-----------------------------------
     Total |     17,503      100.00
```

COMMAND ABBREVIATIONS

STATA uses the first letters of a command and understands. So instead of summarize, use sum. Or instead of regression, use reg.

Summarize Command

```
. summarize r5flone  (1 variable)

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
     r5flone |     17503    .1840827    .3875627         0          1


. sum r4flone r5flone  (2 variables)

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
     r4flone |     19322    .178967    .3833346         0          1
     r5flone |     17503    .1840827    .3875627         0          1
```

Describe Command

```
. describe r5flone

                storage  display     value
variable name   type     format      label      variable label
------------------------------------------------------------------------------
r5flone         byte     %8.0g       r5flone    r5flone:w5 cesd:
                                                felt lonely
```

Interrupting STATA

Hit Ctrl + Pause Break or STATA will continue to list the responses of all the respondents, the numerical value, and the response to this variable.

You can also hit the red button with an x in it that is on the STATA toolbar and STATA will interrupt the processing without exiting.

List Command

```
. list r5flone

     +---------+
     | r5flone |
     |---------  |
  1. |         . |
  2. |    1.yes |
  3. |         . |
  4. |    0.no |
  5. |    0.no |
     |---------- |
  6. |   1.yes |
  7. |   1.yes |
  8. |       .s |
  9. |   1.yes |
 10. |        . |
     |---------|
 11. |        . |
 12. |        . |
 13. |  0.no |
 14. |  0.no |
 15. |        . |
     |---------|
 16. |        . |
 17. |  0.no |
 18. |        . |
 19. |        . |
 20. |  0.no |
     |---------|
 21. |      .s |
 22. |        . |
--Break--
```

RAND Enhanced HRS/ AHEAD Fat Files

Another Data File from RAND

The Center has produced enhanced versions of "raw" HRS and AHEAD data files. Among the enhancements are:

* The HRS/AHEAD data are distributed in modules that are organized by questionnaire section. Some modules contain household-level variables, such as information on assets. We merged all modules into individual-level records. This eliminates the need to search through modules for particular variables and facilitates merging files across waves.

* The HRS/AHEAD collect certain individual-level information by asking one household member to answer about another. For example, the financial respondent provides earnings information for himself or herself and his or her spouse. We created new variables for "own" and "partner" information by re-assigning responses to the appropriate person.

There is a single Fat File for each year of HRS/ AHEAD which contains most of the "raw" or original variables, merged to the respondent level. The Fat Files are available on request with preliminary documentation. [Note: No documentation is available for 2000 and 2002 at this time.]

Method

To request these files please provide your name, address (no PO Box), and phone number to RANDHRShelp@rand.org

Available files include 1992, 1993, 1994, 1995, 1996, 1998, 2000, and 2002. These can easily be used with the current release of the RAND HRS. As part of the development of these data sets, the Center notes details of problems found in the data, and adds some generally useful variables. The RAND HRS Data file development begins with these enhanced "raw" files. For a more detailed view of how these files are enhanced, please see sample documentation for HRS

Figure 69. RAND Enhanced HRS/AHEAD Fat Files http://www.rand.org/labor/aging/dataprod/#fatfiles

1992/1994, AHEAD 1993/1995, HRS 1996, and HRS 1998. These provide information on what has been added to raw HRS data in these files, but they are not codebooks or questionnaires. Codebooks and questionnaires may be found on the HRS web site. Note these versions of the documentation have not been updated for the most recent release of the Fat Files.

Please refer to the guide Using the RAND Versions of the HRS Data by the RAND Center for the Study of Aging for more information. This documentation gives an overview of both products and how they can be used together.

Click on:
http://www.rand.org/labor/aging/dataprod/randhrsprod2005.pdf

7

Classical Ordinary Least Squares Regression

Using the HRS RAND Data with STATA

Classical Ordinary Least Squares Regression in STATA

This guide is not intended for a detailed explanation of statistics, their use, or interpretation. Examples are shown here to give an overview of STATA commands using the RAND data. A manual on Statistics should be consulted for setting up the models of any research statistical analysis.

This example is a simple regression without using the weights or strata of the sample (discussed in the following chapter (8), using the dependent variable — weight with independent variable — height in Wave 6.

Begin with the /stata command regress

use 2 variables, a dependent variable and an independent variable

In this example weight is the dependent variable

and height is the independent variable.

```
. regress r6weight r6height

      Source |       SS           df       MS      Number of obs   =    17820
-------------+----------------------------------   F(  1, 17818)   = 6533.55
       Model |  1523080.11         1  1523080.11   Prob > F        =  0.0000
    Residual |  4153675.42     17818  233.116815   R-squared       = 0.2683
-------------+----------------------------------   Adj R-squared   =  0.2683
       Total |  5676755.52     17819  318.578794   Root MSE        =  15.268

------------------------------------------------------------------------------
    r6weight |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    r6height |    91.93559   1.137389    80.83   0.000     89.7062    94.16498
       _cons |   -77.50446   1.921973   -40.33   0.000    -81.27171   -73.7372
------------------------------------------------------------------------------
```

Tests for Linear Regression Diagnostics

The following tests assume that you use the first STATA command shown on This Page — predict double resid, residuals to derive the other tests.

1. Residual Normality

a.) Use command predict double resid, residuals; then summarize resid

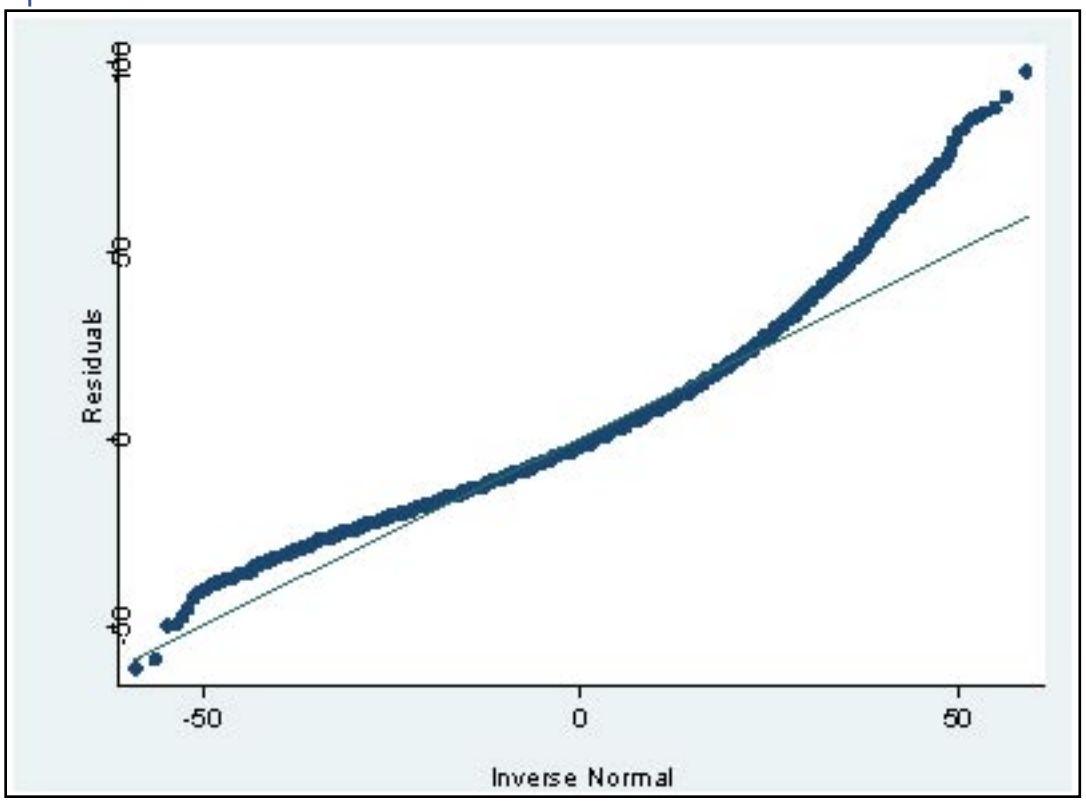
```
. predict double resid, residuals
(8908 missing values generated)

. summarize resid

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
       resid |     17820   -1.30e-14    15.26774  -61.27843   97.81415
```

b.) To plot the residuals on a graph use STATA command qnorm resid

```
. qnorm resid
```



Note: STATA graphs are saved by cutting from the open graphic STATA window and pasting it into a Word document.

Tests for Linear Regression Diagnostics

2. Heteroskedasticity of Residuals

Two commands that are useful are

a.) Use the STATA command hettest to obtain the Breusch-Pagan / Cook-Weisberg test

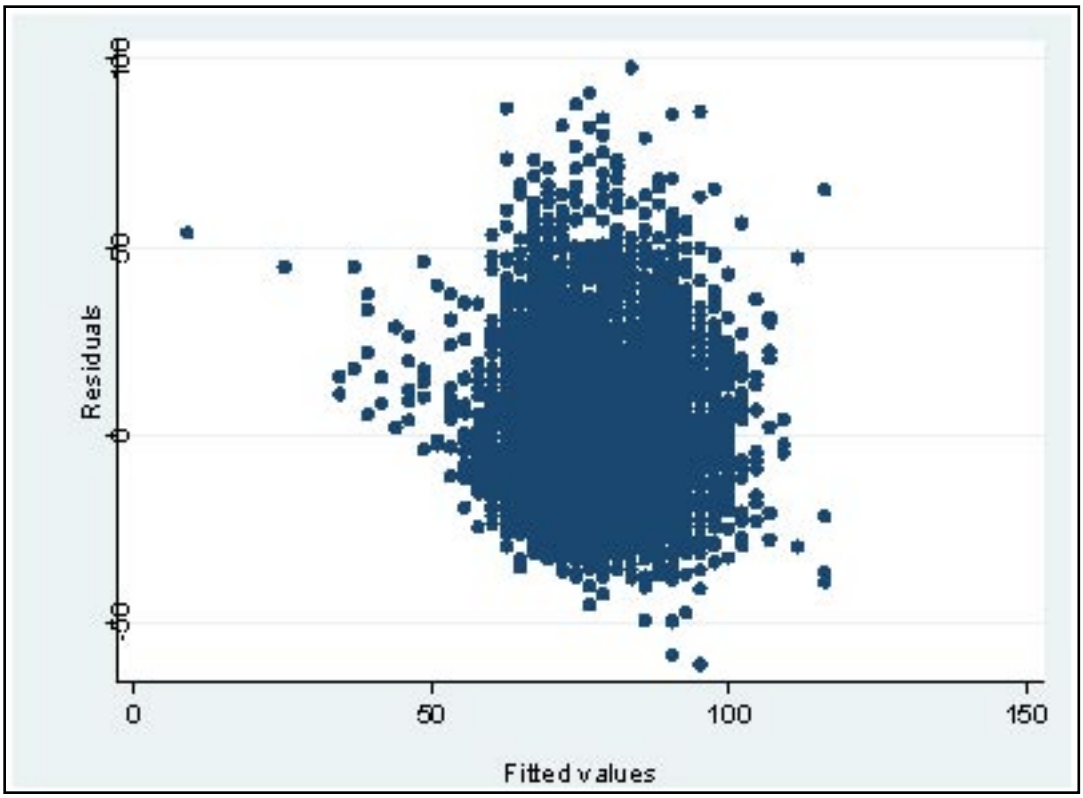
```
. hettest

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: fitted values of r6weight

         chi2(1)      =    14.51
         Prob > chi2  =   0.0001
```

b.) Plot the fitted values versus the residuals on a graph

```
. rvfplot
```

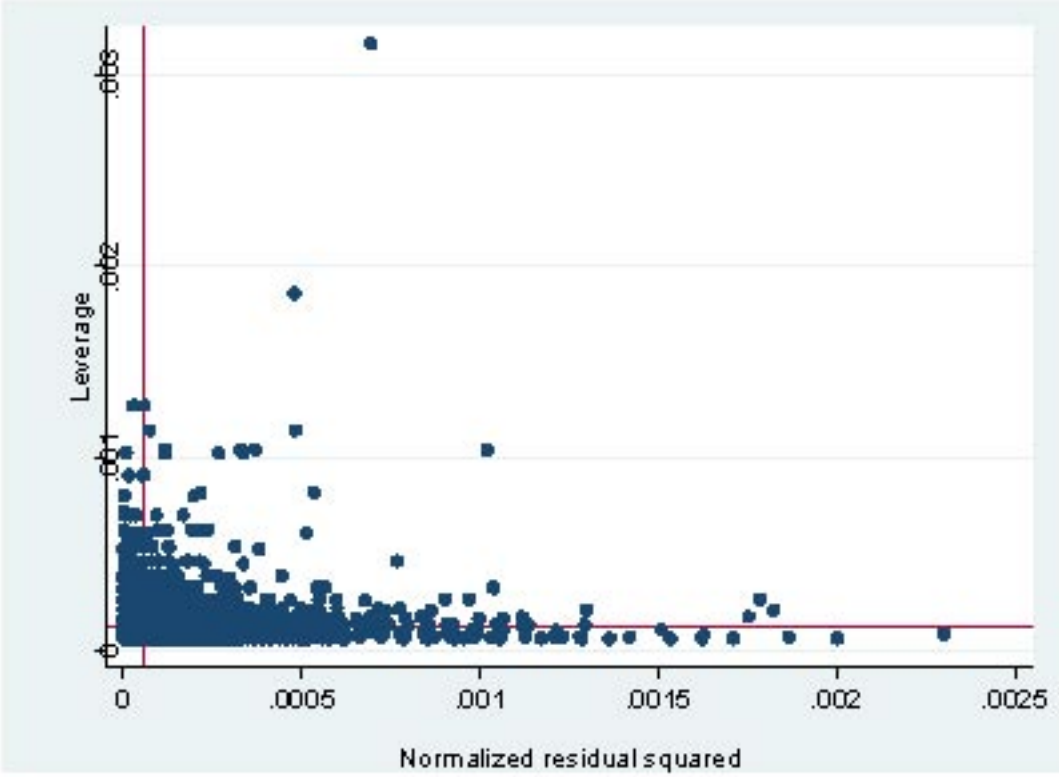


Tests for Linear Regression Diagnostics

3, Leverage versus The Residual Plot

To plot the leverage vs. the residuals on a graph use the command lvr2plot

```
. lvr2plot
```



4, Multicollinearity Tests

Use command pwcorr

```
. pwcorr r6weight r6height

             | r6weight r6height
-------------+------------------
   r6weight  |   1.0000
   r6height  |   0.5180   1.0000
```

Use command vif

```
. vif

    Variable |      VIF       1/VIF
-------------+----------------------
    r6height |       1.00    1.000000
-------------+----------------------
    Mean VIF |       1.00
```

Sort Command in STATA

When there is a variable to break down into groups for analysis commonly the sort command is used.

In this example, the variable ragender, the respondent's gender is used to sort the resgression of weight against height.

Command by ragender, sort: regress r5weight r5height

```
. by ragender, sort: regress r5weight r5height
-------------------------------------------------------------------------------------------------------------
-> ragender = 1.male

      Source |       SS           df       MS              Number of obs =    8082
-------------+----------------------------------           F(  1,  8080) = 1838.35
       Model |  375779.052         1  375779.052           Prob > F      =  0.0000
    Residual |  1651644.34      8080  204.411429           R-squared     =  0.1853
-------------+----------------------------------           Adj R-squared =  0.1852
       Total |  2027423.39      8081  250.887687           Root MSE      =  14.297

------------------------------------------------------------------------------
    r5weight |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    r5height |   92.63707   2.160584    42.88   0.000     88.40177    96.87237
       _cons |  -78.33022   3.829089   -20.46   0.000    -85.83622   -70.82422
------------------------------------------------------------------------------

-------------------------------------------------------------------------------------------------------------
-> ragender = 2.female

      Source |       SS           df       MS              Number of obs =   11208
-------------+----------------------------------           F(  1, 11206) = 1148.33
       Model |  272150.098         1  272150.098           Prob > F      =  0.0000
    Residual |  2655770.69     11206  236.995421           R-squared     =  0.0929
-------------+----------------------------------           Adj R-squared =  0.0929
       Total |  2927920.78     11207  261.258212           Root MSE      =  15.395

------------------------------------------------------------------------------
    r5weight |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    r5height |   72.18836   2.130261    33.89   0.000     68.01267    76.36405
       _cons |  -46.37204   3.464321   -13.39   0.000    -53.16272   -39.58136
```

User Registration

To register to use the HRS data start from the Home Page, find User Registration on the secondary navigation bar (Figure 62) and click the hot link. These data are free for public use.

Obtaining the HRS RAND Data and Documentation

Method

1. Register for a password (if you have not already done so) by clicking on the HRS Home Page>Data>Access to Public Data Sets and Files>Click HERE to register for Public Release files (a password will be sent to your e-mail quickly) or Click HERE if you have already registered for Public Release Files.

2. Enter User Name and Password

3. Download the HRS RAND Data
>Datasets and Files >RAND HRS Data File (v.E)
from the box – RAND Contributed Files

4. Choose the data file appropriate to your statistical package SAS or STATA (SPSS users download SAS).

- For example, the file randestata8se.zip indicates a RAND data file, version e, for STATA 8 (or higher) Special Edition. Check your software for it's exact name if you are unsure which to use.

5. Use WinZip to open the file you chose, save in RAND_Data_Library folder.

6. Save this original data set under another file name, for example randestata8se_wrkg.dta, so one copy remains unaltered, this prevents having to reload the data set again. This second file is your working data file.

Using Under the Condition, And, or If Statements in STATA

When creating a variable you will frequently use a statement using these symbols:

1. Under the condition ==
2. And &
3. If |
4. Values are:
< less
> greater
= equal
~= unequal

8

HRS RAND Data Weights and Strata

Weighting the HRS RAND Data

Weights, Clustered and Stratified Sampling

The HRS includes oversamples of African American and Hispanic respondents. Adjustments need to be made to match the HRS data at any given wave to estimated population totals. For this reason, the data must be weighted to yield unbiased estimates of population parameters. The HRS sample is also clustered and stratified, and unless this is taken explicitly into account standard errors based on the assumption of simple random sampling will be biased (generally downward).

To allow these features of the HRS to be taken into account, the stratification is specified by the variable raestrat and the clustering by the variable raehsamp.

The Weight, Clustering and Stratum Variables

- The RAND data has respondent, spousal, and household weight variables (boxes to the right).
- The clustering variable is raehsamp.
- The stratum variable is raestrat.

These variables are specified using the STATA command: svy
(See STATA manual Survey for further explanation if needed.)

Command Example: svyset raehsamp [pweight=r1wtresp], strata (raestrat)
prepares the survey data for analysis. This command is run before a statistical analysis such as a regression or a chi2.

Weights — Respondent and Spousal All Six Waves

File	Variable	Type
1	R1WTRESP	Cont
2	R2WTRESP	Cont
3	R3WTRESP	Cont
4	R4WTRESP	Cont
5	R5WTRESP	Cont
6	R6WTRESP	Cont
1	S1WTRESP	Cont
2	S2WTRESP	Cont
3	S3WTRESP	Cont
4	S4WTRESP	Cont
5	S5WTRESP	Cont
6	S6WTRESP	Cont

Household Analysis Weights — All Six Waves

File	Variable	Type
1	R1WTHH	Cont
2	R2WTHH	Cont
3	R3WTHH	Cont
4	R4WTHH	Cont
5	R5WTHH	Cont
6	R6WTHH	Cont

Using the Household Weight and Strata Variables with a Multi-Variate Regression

Use the svyset command in STATA

1.) Prepare the data for analysis

Command svyset raehsamp [pweight=r6wthh], strata (raestrat)
This example uses household variables from Wave 6
and, therefore, the household weight from Wave 6.

```
. svyset raehsamp [pweight=r6wthh], strata (raestrat)

      pweight: r6wthh
        VCE: linearized
    Strata 1: raestrat
       SU 1: raehsamp
       FPC 1: <zero>
```

2.) Run the analysis command
Command: svy: regress h6icap h6ahous h6amort

```
. svy: regress h6icap h6ahous h6amort
(running regress on estimation sample)
Survey: Linear regression

Number of strata   =        52                  Number of obs      =     18167
Number of PSUs     =       104                  Population size    =  72060957
                                                Design df          =        52
                                                F(   2,     51)    =     22.00
                                                Prob > F           =    0.0000
                                                R-squared          =    0.0890

------------------------------------------------------------------------------
             |             Linearized
      h6icap |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    h6ahous |   .0940416   .0218934     4.30   0.000     .0501094    .1379739
    h6amort |   .0586711   .0378661     1.55   0.127    -.0173129     .134655
      _cons |  -664.0502   2482.031    -0.27   0.790    -5644.61     4316.509
------------------------------------------------------------------------------
```

ADDING HOUSEHOLD WEIGHTS BY WAVE

Use the svyset command in STATA

1.) PREPARE THE DATA FOR ANALYSIS

COMMAND SVYSET RAESAMP [PWEIGHT=R4WTHH], STRATA (RAESTRAT)
THIS EXAMPLE USES HOUSEHOLD VARIABLES FROM WAVE 4
AND, THEREFORE, THE HOUSEHOLD WEIGHT FROM WAVE 4.

```
. svyset raehsamp [pweight=r4wthh], strata (raestrat)

      pweight: r4wthh
          VCE: linearized
    Strata 1: raestrat
        SU 1: raehsamp
       FPC 1: <zero>
```

Command .svy: regress h4icap h4ahous h4amort

```
. svy: regress h4icap h4ahous h4amort
(running regress on estimation sample)

Survey: Linear regression

Number of strata   =        52                  Number of obs    =     21384
Number of PSUs     =       104                  Population size  =  69243916
                                                Design df        =        52
                                                F(   2,     51)  =     23.53
                                                Prob > F         =    0.0000
                                                R-squared        =    0.0179

------------------------------------------------------------------------------
             |             Linearized
      h4icap |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    h4ahous |   .0464027   .0101238     4.58   0.000     .0260879    .0667176
    h4amort |   .1544688   .0372891     4.14   0.000     .0796428    .2292949
      _cons |    7639.09   2101.669     3.63   0.001     3421.782     11856.4
------------------------------------------------------------------------------
```

Variables Used in Calculations

h4icap Household capital income — sums income from business, rental, dividend and interest, and trust fund or royalties plus rental expenses.

h4ahous Value of primary residence.

h4amort Value of all mortgages.

TWO-WAY TABLE USING WEIGHTS AND STRATUM

```
. svyset raehsamp [pweight=r3wtresp], strata (raestrat)

      pweight: r3wtresp
          VCE: linearized
    Strata 1: raestrat
        SU 1: raehsamp
       FPC 1: <zero>

. svy: tabulate r3flone ragender
(running tabulate on estimation sample)

Number of strata   =        52                  Number of obs      =       16342
Number of PSUs     =       104                  Population size    =  42189380
                                                Design df          =          52

-----------------------------------------
r3flone:w |
3 cesd:   |
felt      |      ragender: r gender
lonely    |   1,male  2,female     Total
----------+------------------------------
    0,no  |   .3709     .4686     .8395
   1,yes  |   .0527     .1078     .1605
          |
   Total  |   .4236     .5764         1
-----------------------------------------
  Key:  cell proportions

Pearson:
   Uncorrected   chi2(1)       =  115.8666

   Design-based  F(1, 52)      =   88.0084     P = 0.0000
```

To test for independence for a two-way contingency table, test the homogeneity of row (column) proportions. Command svy: tabulate, row

```
. svy: tabulate, row

Number of strata   =        52                  Number of obs      =       16342
Number of PSUs     =       104                  Population size    =  42189380
                                                Design df          =          52

-----------------------------------------
r3flone:w |
  3 cesd: |
    felt  |      ragender: r gender
  lonely  |   1,male  2,female     Total
----------+------------------------------
    0,no  |   .4418     .5582         1
   1,yes  |   .3284     .6716         1
          |
   Total  |   .4236     .5764         1
-----------------------------------------
  Key:  row proportions

Pearson:
   Uncorrected   chi2(1)       =  115.8666

   Design-based  F(1, 52)      =   88.0084     P = 0.0000
```

OUR SPONSORS

Primary support for the HRS comes from the National Institute on Aging at the National Institutes of Health, and from the Social Security Administration.

For more information, feel free to write us at:

University of Michigan

Survey Research Center

Health and Retirement Study

P.O. Box 1248

Ann Arbor, MI 48106-1248

Phone:
734-936-0314

Fax:
734-647-1186

E-mail:
hrsquest@isr.umich.edu

Web site:
http://hrsonline.isr.umich.edu

The insert for the side spine of a binder is to the right.

Getting Started with the Health and Retirement Study