

HEALTH AND RETIREMENT STUDY

Respondent Cross-Year Summary Earnings

Data Description and Usage

Version 4.0, December 2014

To the Restricted Data Investigator: This restricted data set is intended for exclusive use by you and the persons specified in the *Agreement for Use of Restricted Data from the Health and Retirement Study* and/or the *Supplemental Agreement with Research Staff for Use of Restricted Data from the Health and Retirement Study*. If there are any questions about this data set and its use, refer to the restricted data pages on the HRS Web site (<http://hrsonline.isr.umich.edu/rda>). You can also contact the HRS Help Desk via electronic mail (<mailto:hrsquestions@umich.edu>) or on the Web (<http://hrsonline.isr.umich.edu/help>).

This document may not be reproduced without the written consent of the staff of the Health and Retirement Study, the Institute for Social Research, the University of Michigan.

Table of Contents

1. Overview and Acknowledgments	2
2. Obtaining the Data	2
2a. Restricted Data Agreement	2
2b. Access to Restricted Data	2
2c. Publications Based on Restricted Data	3
3. Setup	3
3a. Distribution Set	3
3a-1. Windows Environment	3
3a-2. UNIX/Linux Environment.....	3
3a-3. Macintosh OS X Environment (10.4.x and above).....	3
3d. Program Statements	4
3d-1. Using the Files with SAS.....	4
3d-2. Using the Files with SPSS.....	4
3d-3. Using the Files with Stata.....	4
4. Data Set Content	5
5. If You Need to Know More	5
5a. HRS Internet Site	6
5b. Contact Information	6
Appendix: Cohort Permission Status.....	7

1. Overview and Acknowledgments

The Health and Retirement Study (HRS) is a national longitudinal study of the economic, health, marital, and family status, as well as public and private support systems, of older Americans. The HRS is a rich source of longitudinal, cross-sectional data for researchers and policymakers who study aging. Funding for the Health and Retirement Study is provided by the National Institute on Aging at NIH (U01 AG009740), with supplemental support from the Social Security Administration. The study is conducted by the Institute for Social Research (ISR) at the University of Michigan.

The HRS gratefully acknowledges the special assistance of the Social Security Administration's Office of Research and Statistics for their assistance in retrieving the administrative records of HRS respondents who gave consent for those records to be used for research purposes.

2. Obtaining the Data

2a. Restricted Data Agreement

This restricted data set is intended for exclusive use by you and the persons specified in the *Agreement for Use of Restricted Data from the Health and Retirement Study* and/or the *Supplemental Agreement with Research Staff for Use of Restricted Data from the Health and Retirement Study*. If you have questions about the content and use of this data set, refer to the HRS Restricted Data Web site or contact the [HRS Help Desk](#) by email.

2b. Access to Restricted Data

Although most HRS data sets are available to the public without restriction, certain HRS data sets contain sensitive respondent information and are only available under terms of a formal agreement negotiated between the researcher and HRS. If you are a researcher who has found a restricted data file that is of interest, you should begin the application process by reviewing the application materials and procedures on the HRS Web site (<http://hrsonline.isr.umich.edu>). If you decide to go forward, you should submit a preliminary application package containing a research proposal and data protection plan to HRS. If problems are found with your application, the HRS staff members will contact you to provide assistance in resolving them. Otherwise, once you have been notified by HRS that your application is complete, you may submit your application to your local Institutional Review Board and/or Contracting Authority. After IRB approval is obtained, you should submit the signed originals of the institutional approval documents to the HRS. At this point the application is complete and it is forwarded to the HRS Data Confidentiality Committee (DCC) for review and approval. If approval is granted by the DCC, the Principal Investigator of the HRS will sign the restricted data agreement, and you will be sent the restricted data product(s) by the HRS. Note: access codes for your restricted data product(s) are sent separately.

If your institutional or physical circumstances change, (e.g., new persons added to the project, hardware/software changes, or office moves) you should contact the HRS in order to modify the underlying agreement. A yearly re-certification of your restricted data agreement is required; if you wish to continue with analysis of the data after the expiration of the initial agreement, you will need to submit a renewal request. Note: An HRS agent will periodically audit your site in order to ensure that all conditions of the Restricted Data Agreement are being met.

2c. Publications Based on Restricted Data

Your restricted data agreement specifies that you will inform HRS of any papers, publications, or presentations based on this restricted data set. Please send a copy of such publications in PDF format via e-mail to hqsquestions@umich.edu with "Attn: Papers and Publications" in the subject line. If you wish, you may include a bibliographical reference.

As an alternative, you may transmit publications in paper format by postal mail:

Health and Retirement Study
Attn: Papers and Publications
The Institute for Social Research, Room 3410
P.O. Box 1248
Ann Arbor, Michigan 48106-1248

3. Setup

3a. Distribution Set

This data set is packaged for distribution in a .ZIP file, `XSumErn.zip`. In order to keep the contents secure, the .ZIP file has been encrypted using WinZIP 256 bit AES encryption. Extract the data file(s), the program statement file(s) matching your analysis environment, the data description (this file), and the codebook file. If you require a special file format or experience system problems, please contact the [HRS Help Desk](#). If all files are decompressed, they will require approximately 90MB of free space on your storage device.

3a-1. Windows Environment

Copy the ZIP file to the Windows folder where you plan to do your work. Do not use the built-in Windows decompression utility; instead you should use a third-party tool such as **7-ZIP** or **WinZIP** to extract the ZIP folder contents. When you are prompted for the pass-phrase, respond with the character string that you received via e-mail. The output will be the files listed in Table 1.

3a-2. UNIX/Linux Environment

Copy the ZIP file to the folder where you plan to do your work. Use the ZIP file decompression software installed on your system, (e.g. **7-zip**, **gunzip**) to decrypt and decompress the ZIP folder contents. When you are prompted for the pass-phrase, respond with the character string that you received via e-mail. The output will be the files listed in Table 1.

3a-3. Macintosh OS X Environment (10.4.x and above)

Copy the ZIP file to the folder where you plan to do your work and use **stuffit** to decrypt and decompress the ZIP folder contents. When you are prompted for the pass-phrase, respond with the character string that you received via e-mail. The output will be the files listed in Table 1.

Directory¹	File	Type
c:\xyrsumern	XSumErn.zip	zip file
c:\xyrsumern\docs\	XSumErn.txt	Codebook file (ASCII text)
c:\xyrsumern\data\	XSumErn.da (n=24939)	Data file (ASCII text)
c:\xyrsumern\sas\	XSumErn.sas	SAS program statements
c:\xyrsumern\spss\	XSumErn.sps	SPSS program statements
c:\xyrsumern\stata\	XSumErn.dct/.do	Stata dictionary and "do" files

3d. Program Statements

Each data file comes with associated SPSS, SAS or Stata program statements to read the data. Files containing SPSS statements are named with an .sps extension, those with SAS statements with an .sas extension, and those with Stata statements with .do and .dct extensions.

3d-1. Using the Files with SAS

To create a SAS system file for a particular data set, two file types must be present for that data set -- .sas program statement files and .DA data files. To create a SAS system file, load the *.sas file into the SAS Program Editor. If the *.sas file is located in "c:\xyrsumern\sas" and the data file is located in "c:\xyrsumern\data", you can run the file as is. A SAS system file will be saved to directory "c:\xyrsumern\sas". If the files are not located in the specified directories, you will need to edit the *.SAS file to reflect the proper path names prior to running the file.

3d-2. Using the Files with SPSS

To create an SPSS system file for a particular data set, two file types must be present for that data set -- .sps program statement files and .da data files. To create an SPSS system file, open the *.sps file in SPSS as an SPSS Syntax File. If the *.sps file is located in "c:\xyrsumern\spss" and the data file is located in "c:\xyrsumern\data", you can run the file as is. An SPSS system file will be saved to directory "c:\xyrsumern\spss". If the files are not located in the specified directories, you will need to edit the *.sps file to reflect the proper path names prior to running the file.

3d-3. Using the Files with Stata

To use Stata with a particular data set, the following three file types must be present for that data set -- .dct files, .do files, and .da data files. Files with the suffix ".da" contain the raw data for Stata to read. Files with the suffix ".dct" are Stata dictionaries used by Stata to describe the data. Files with the suffix ".do" are short Stata programs ("do files") which you may use to read in the data. Load the .do file into Stata and then submit it. If the *.do and .dct files are located in "c:\xyrsumern\stata" and the data file is located in "c:\xyrsumern\data", you can run the .do file as is. If the files are not located in these directories, you must edit the *.do and *.dct files to reflect the proper path names before you run the files.

¹ While a specific folder arrangement is not required for using HRS files, we have traditionally suggested a directory structure for the Windows environment. By using this directory structure, you will not have to change the path name references in your data descriptor files. If you do not use Windows, or if you use a different directory structure, just change the directory references in the program files.

4. Data Set Content

Version 4.0 of the *Respondent Cross-Year Summary Earnings* data set contains Master Earnings File (MEF) earnings data for 24,939 respondents derived from 8 permission sets as summarized below. See the Appendix for additional information on the timing of cohort permissions.

Primary Cohort	Permission Set								Total
	1992/94/96	1993/95	1998/2000	2004	2006 ²	2008	2010	2012	
HRS	3831		24	2305	80	2515	1865	354	10974
AHEAD ³	22	4276	23	116	14	499	284		5231
CODA			646	186	503	280	125	1	1741
War Baby			437	203	449	600	287	100	2076
EBB				607	150	811	1092	239	2899
MBB							1718	298	2016
n.a.							2		2
Total	3853	4276	1130	3414	1196	4705	5373	992	24939

This version of the *Respondent Cross-Year Summary Earnings* data set contains the same categories of variables as earlier versions. The file contains 1951-2013 annual earnings data based on the most recent information received for each respondent. When information is available for an individual respondent at two or more points in time, the most recent earnings value is used. The file includes:

- Earnings adjusted for taxable maximum
- Wage quarters of coverage pattern
- Total QC for year
- Self employment income quarters of coverage
- Agriculture income quarters of coverage (1955-1977)
- Agriculture income employment code (begins 1978)

To help preserve respondent confidentiality, earnings data have been rounded to the nearest \$100, except for fields that were originally between \$0.01 and \$49.99, which have been recoded to .N. Entries that were originally coded as 0 remain as 0. Observations that are missing due to absence of information for a given year are coded .M.

The file is keyed on Household Identifier and Person Number.

5. If You Need to Know More

This document is intended to serve as a brief overview and to provide guidelines for using the *Cross-Year Respondent Summary Earnings* data set. If you have questions or concerns that are

² Prospective permissions began in 2006

³ Note: This data set treats HRS Wave 1 “overlap” cases as members of the AHEAD cohort. There were a number of original HRS Wave 1 households eligible to be either an HRS or AHEAD household. An interview was attempted for each in HRS Wave 1. Afterwards, a random sub-sampling was performed, with 60% of the cases staying in HRS and the remaining going to AHEAD. We refer to the 134 cases given over to AHEAD as the HRS/AHEAD overlap.

not adequately covered here or on our Web site, or if you have any comments, please contact us. We will do our best to provide answers.

5a. HRS Internet Site

The Health and Retirement Study maintains a Web site (<http://hrsonline.isr.umich.edu>) that provides access to public release data, questionnaires, codebooks, sample and response rate documentation, a bibliography, user guides, and extensive background information. If you want to find out more about restricted data products and procedures, visit the restricted data section of the HRS Web site (<http://hrsonline.isr.umich.edu/rda>).

5b. Contact Information

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

Internet: Help Desk at our Web site (<http://hrsonline.isr.umich.edu/help>)

E-mail: hqsquest@isr.umich.edu

Postal service:

Health and Retirement Study
The Institute for Social Research, Room 3050
The University of Michigan
P.O. Box 1248
Ann Arbor, MI 48106-1248

FAX: (734) 647-1186

The *Respondent Cross-Year Summary Earnings* dataset is derived from Master Beneficiary Record (MBR) data for primary beneficiary and other (secondary) insured. The inputs to this dataset were received at different points in time over the past decade from the Social Security Administration:

- In 1992, 1994 or 1996 for HRS cohort members
- In 1993 or 1995 for AHEAD cohort members
- In 1998 or 2000 for CODA and War Baby cohort members as well as new spouses
- In 2004 for EBB and HRS cohort members, new spouses, and War Baby 1998/2000 refusers
- In 2006, 2008, 2010 or 2012 for respondents interviewed face-to-face who were not asked previously or refused or were non-response. These permissions are prospective and provide for biennial updates through 2030.

Prior to 2006, respondent permissions were based on separate agreements between HRS and SSA. This meant that information was obtained only for the time period covered by the agreement. In 2006 and thereafter, the terms of the permission agreement allow HRS to obtain biennial updates of earnings and benefits data through the current year. Since each new record set obtained from SSA reflects the most recent information available, the latest version of *Respondent Cross-Year Summary Earnings* will always be the most accurate. If current information for a given respondent is missing, the most recent data elements from past datasets are used.