HEALTH AND RETIREMENT STUDY

Restricted Data

Respondent Cross-Year Detail Earnings

Data Description

Updated through Consents/Matches as of the HRS 2020 Wave
Released Date: September 2022

Table of Contents

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 the National Institutes of Health (U01AG009740), 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. 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. Prospective users of HRS restricted data have two access options:
- [Virtual Desktop Infrastructure](#) (recommended)
- Traditional Licensing Agreement (deprecated)

For instructions on how to proceed, visit the [HRS Restricted Data](#) page or contact the HRS Restricted Data Applications Processing Team (*[hrsrdaapplication@umich.edu](#)*) by email.

2b. Restricted Data Agreement

This restricted data set is intended for exclusive use by you and the persons specified in the *Confidentiality Agreement for Use of Restricted Data from the Health and Retirement Study.* Traditional licensing agreements that include multiple users are also bound by the *Supplemental Agreement with Research Staff for Use of Restricted Data from the Health and Retirement Study*.

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 *[hrsquestions@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 3450
> P.O. Box 1248
> Ann Arbor, Michigan 48106-1248

3. Data Set Content

The *Respondent Cross-Year Detail Earnings* data set contains Master Earnings File (MEF) earnings data for 23,785 respondents derived from the permission sets as summarized below:

Primary Cohort	Permission Set												Total
	1992	1998	2000	2004	2006	2008	2010	2012	2014	2016	2018	2020	
HRS	3,313	14	5	1,840	30	2,197	1,739	398	280	166	344	148	10,474
AHEAD[1]	11	12	7	93	10	426	252	0	0	22	25	4	862
CODA	0	478	57	158	379	263	116	2	0	26	80	7	1,566
War Baby	0	343	35	154	205	414	272	158	149	49	216	112	2,107
EBB	0	0	0	468	66	316	961	354	312	150	281	483	3,391
MBB	0	0	0	0	0	0	1,341	256	344	253	303	640	3,137
LBB	0	0	0	0	0	0	0	0	0	1,251	445	552	2,248
Total	3,324	847	104	2,713	690	3,616	4,681	1,168	1,085	1,917	1,694	1,946	23,785

Table 2: Respondent Counts

The data file contains one record per employer[2] per year for each consenting respondent who received wages during years 1978-2020,[3] and whose records were available in the MEF. See Appendix B for more information on cohort permission timing.

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**. Amounts greater or equal to $250,000 or less than or equal to -$250,000 were top/bottom coded. See Appendix A for further details.

---

[1] 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.

[2] Employer identifiers are scrambled and are unique to each respondent.

[3] **Detail earnings records for 1978 and 1979 are incomplete and should be used with caution, if at all**.

4. If You Need to Know More

This document is intended to serve as a brief overview and to provide guidelines for using this data set. If you have questions or concerns that are not adequately covered here or on our Web site, or if you have any comments, please contact us. We will do our best to provide answers.

4a. HRS Internet Site

Health and Retirement Study public release data and additional information about the study are available on the Internet. To access public data or to find out more about restricted data products and procedures, visit the HRS Web site.

4b. Contact Information

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

```
Internet: Help Desk at the HRS Web site (https://hrs.isr.umich.edu/help)

E-mail: hrsquestions@umich.edu

Postal Service:
     Health and Retirement Study
     The Institute for Social Research
     426 Thompson Street, 3450 ISR
     Ann Arbor, Michigan  48104
```

Appendix A: Data Management Notes

Table 3: Missing Data Codes			
Bottom/Top Code Category	**SAS/Stata Special Missing**	**Numeric Missing**	**Numeric Recode**
<= -$500,000	.a	-999998	
-$300,000 to 499,999	.b	-999997	
-$250,000 to 299,999	.c	-999996	
-1 to -49	.m	-2	-25
1 to 49	.n	-1	25
$250,000 to 299,999	.x	999996	272000
$300,000 to 499,999	.y	999997	370000
>= $500,000	.z	999998	1000000

To preserve respondent confidentiality, all data elements are rounded and top/bottom-coded. A code of zero (0) represents a true zero dollar amount. A code of .N represents an amount between 1 and 49 dollars, inclusive. A code of .M represents an amount between -1 and -49 dollars, inclusive. Amounts with an absolute value of 50 dollars or more were rounded to the nearest 100 dollars. An absolute value ending in 01 to 49 was rounded down, and a value ending in 50 to 99 was rounded up.

Amounts greater or equal to $250,000 were recoded as follows:

    $250,000 to 299,999 = .X;
    $300,000 to 499,999 = .Y;
    >= $500,000 = .Z.

Amounts less than or equal to -$250,000 were recoded as follows:

    -$250,000 to -299,999 = .C;
    -$300,000 to -499,999 = .B;
    <= -$500,000 = .A.

SPSS users may wish to convert the SAS special missing codes to SPSS equivalents. The following code fragment serves as an example:

```
* input WTOTCOMP as a string variable named xWTOTCOMP .
compute WTOTCOMP=0.
compute WTOTCOMP=number(xWTOTCOMP, F10).
recode xWTOTCOMP (convert) ('.n'=25) ('.x'=999996)
('.y'=999997) ('.z'=999998) into WTOTCOMP.
missing values WTOTCOMP (999996, 999997, 999998).
desc WTOTCOMP.
execute.
```

Appendix B: Installation (Traditional License Only)

A1. Distribution Set

This data set is packaged for distribution in a .ZIP file, XDetErn.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.

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.[4] 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 A1.

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 A1.

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 A1.

> **Note: MiCDA Enclave Virtual Desktop Environment users are given access to pre-built SAS, Stata and SPSS versions of this dataset; therefore the information in this appendix does not apply to such users.**

Table 1: Contents of Distribution Package

Directory[5]	File	Type
c:\xyrdetern	XDetErn.zip	zip file
c:\xyrdetern\docs\	XyrDetErn2020scr.txt	Codebook file (ASCII text)
c:\xyrdetern\data\	XyrDetErn2020scr.da (n=706261)	Data file (ASCII text)
c:\xyrdetern\sas\	XyrDetErn2020scr.sas	SAS program statements
c:\xyrdetern\spss\	XyrDetErn2020scr.sps	SPSS program statements
c:\xyrdetern\stata\	XyrDetErn2020scr.dct/.do	Stata dictionary and "do" files

---

[4] The built-in Windows decompression utility will not process AES-256bit encrypted zip files; it halts with "an unexpected error is keeping you from copying the file".

[5] 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.

A2. 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.

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:\xydetern\sas" and the data file is located in "c:\xydetern\data", you can run the file as is. A SAS system file will be saved to directory "c:\xydetern\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.

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:\xydetern\spss" and the data file is located in "c:\xydetern\data", you can run the file as is. An SPSS system file will be saved to directory "c:\xydetern\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.

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:\xydetern\stata" and the data file is located in "c:\xydetern\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.

A3. Non-Windows Environments

Non-Microsoft users should modify the default Windows file structure syntax to match that of their own operating system. The following examples should work for both Macintosh OS X and any Unix/Linux distribution. Open the SAS program file(s), SPSS syntax file(s) or the Stata do/dct files in an ASCII editor and make the changes indicated below.

SPSS in an OSX environment

In this example, we assume that:

- The username is "user1"
- The encrypted zip file containing the data set has been copied to the user's desktop.
- The user has decrypted/decompressed the zip file (use Stuffit for OS X) into a **Desktop** folder named **xydetern**
- The ASCII data file is stored in sub-folder **data** and the syntax file in sub-folder **spss**.

The commands in the syntax file should be modified to look like this:

```
FILE HANDLE xydetern /name='Desktop/xydetern/data/xydetern2020scr.da'
RECL=221.
DATA LIST FILE= xydetern/
HHID 1-6(A)
[rest of syntax file goes here]
.
execute.
SAVE /outfile
'Desktop/xydetern/spss/xydetern2020scr.sav'.
Execute.
```

STATA in an OS X Environment

In the following example we assume that:

- The username is "user1"
- The encrypted zip file containing the data set has been copied to the user's desktop.
- The user has decrypted /decompressed the zip file (use Stuffit for OS X) into a desktop folder named **xydetern**
- The ASCII data file is stored in sub-folder **data** and the .do/.dct files in sub-folder **stata**.

File `XYDETERN.do` should be modified as follows:
Change…
```
      infile using c:\xydetern\stata\xydetern2020scr.dct
```
To…
```
      infile using /Users/user1/Desktop/xydetern/stata/xydetern2020scr.dct
```

Change…
```
      save c:\xydetern\stata\xydetern2020scr.dta
```
To…
```
      save /Users/user1/Desktop/xydetern/stata/xydetern2020scr.dta
```

File `XYDETERN.dct` should be modified as follows:

Change…
```
      dictionary using c:\xydetern\data\xydetern2020scr.da {…}
```
To…
```
      dictionary using /Users/user1/Desktop/xydetern/data/xydetern2020scr.da
{…}
```

Appendix C: Cohort Permission Status

Cohort	1992/94/96	1993/95	1998/2000	2004	2006: Asked If Face-to-face and…	2008: Asked if Enhanced Face-to-face and…	2010: Asked if Enhanced Face-to-face and…	2012: Asked if Enhanced Face-to-face and…	2014: Asked if Enhanced Face-to-face and…	2016: Asked if Enhanced Face-to-face and…	2018: Asked if Enhanced Face-to-face and…	2020: Asked if Enhanced Face-to-face and…
New spouses or never-interviewed R in any cohort	Asked all	Asked all	Asked all	Asked all	Asked all	Asked all	Asked all	Asked all	Asked all	Asked all	Asked all	Asked all
AHEAD (Born <=1923)	NO	Asked all	NO	NO	R was not asked in 2004	R refused, did not provide valid permission information, was not asked, or was non-response in 2004 or 2006	R refused, did not provide valid permission information, was not asked, or was non-response in 2006 or 2008	R refused, did not provide valid permission information, was not asked, or was non-response in 2006, 2008 or 2010	R refused, did not provide valid permission information, was not asked, or was non-response in 2006, 2008, 2010 or 2012	R refused, did not provide valid permission information, was not asked, or was non-response in 2006, 2008, 2010, 2012 or 2014	R refused, did not provide valid permission information, was not asked, or was non-response in 2006, 2008, 2010, 2012, 2014, or 2016	R refused, did not provide valid permission information, was not asked, or was non-response in 2006, 2008, 2010, 2012, 2014, 2016, or 2018
CODA (1924-1930)	NA	NA	Asked all	NO								
HRS (1931-1941)	Asked all	NO	NO	Asked all								
WAR BABY (1942-1947)	NA	NA	Asked all	Only asked refusers								
EARLY BOOMERS (1948-1953)	NA	NA	NA	Asked all								
MID BOOMERS (1954-1959)	NA	NA	NA	NA	NA	NA	Asked all					
LATE BOOMERS (1960-1965)	NA	NA	NA	NA	NA	NA	NA	NA	NA	Asked all		

Source Information

As seen in the previous table, 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 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, 2012, 2014, 2016, 2018, and 2020 for
  - New sample members (MBB, LBB)
  - Respondents (in all older cohorts) interviewed enhanced face-to-face who were not asked previously or refused or were non-response.

Information Received from SSA in Spring 2022

Consent Wave	Summary Earnings	Detail Earnings	Benefits (MBR and PHUS)	Form 831 SSI
2006	No	No	Yes, through 2018	Yes, through 2018
2008			Yes, through 2020	Yes, through 2020
2010				
2012				
2014	Yes, through 2020	Yes, through 2020		
2016				
2018				
2020				

Limits for Information Obtained Through Respondent Consent Forms

1. 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 and respondent cohort(s) covered by the agreement.
2. In 2006 the terms of the consent form were altered, allowing HRS to obtain updates of earnings and benefits data for all years up to and including the current year. The term of these consents was defined as year of signing through 2030.
3. In 2018 the respondent consent form[6] MOA terms were revised, limiting earnings permissions to year of signing through Survey Year + 6, and earnings permissions to year of signing through Survey Year + 12.
4. In future years, post-2004 respondents will need new earnings/benefits data consents once their Survey Year + 6/+12 anniversary has been reached.

---

[6] In 1998 through 2014, HRS-SSA agreements included provisions for obtaining benefits and earnings information for deceased spouses. The deceased spouse option was dropped in 2016.