RAND Center for the Study of Aging
A RAND LABOR AND POPULATION CENTER

RAND Enhanced Fat Files

The RAND Enhanced Fat Files contain most of the raw HRS variables from each questionnaire section, and collapses them into a single respondent-level dataset for each wave. The files are easily merged with one another and with RAND HRS dataset, which can expedite the preparation of analytic files from HRS.

There is one file per wave with one observation for each responding individual. There are 1992, 1993, 1994, 1995, 1996, 1998, 2000, 2002, 2004, 2006 and 2008 files. The data can be downloaded as 3 formats: SAS, Stata/SE or SPSS. The files correspond to those used in developing RAND HRS Version **K**.

The 1992 and 1994 fat files include all HRS raw variables. For all other years, all the raw variables at the respondent or household level are included. For the other-person data, such as those about children, siblings, or helpers, are not included.

The files pertaining to each year are stored in the corresponding sub-directory, e.g., HRS 2000 final release V1 files are in H00F1. Please see the README file of each year for details.

More detail documentations are available on RAND Center of Aging web site:
http://www.rand.org/labor/aging/dataprod.

1. What makes these files different from the Public Release HRS Data?
2. Additions and Changes to the Public Release Data
3. Lining up the years
4. Overlap cases and merging with RAND HRS Data
5. Public Release Versions
6. Downloading the Files

1. What makes these files different from the Public Release HRS Data?

The HRS surveys collect a considerable amount of information about elderly households, and utilize a complicated questionnaire structure to minimize the burden on any single respondent within the household. For example, the respondent who answered the household finance module may not be the same person who answered the module on children and other household members. This format complicates the way the data is structured upon its public release. For instance, since only one person in a couple household provides financial and family information, the raw HRS data provides these data at the household level, but many analyses use respondent level observations.

We have created public use files which simplify this data structure in a number of ways. We reorganized the data so that each observation represents one individual, and we merged the appropriate information from the various modules to each observation. Thus, household level information is present for each individual respondent.

Similarly, certain individual-level information was collected by asking one household member to answer about another. For example, the financial respondent provides earnings information about himself or herself, and about his or her spouse. We reassigned this data to "self" and "spouse/partner" variables for both respondents to clarify to whom the information applies.

Finally, we make the files easy to merge across waves and with the RAND HRS Data, and applied consistent variable naming conventions across waves that reflect those already used by the HRS data.

2. Additions and Changes to the Public Release HRS Data

The RAND Enhanced Fat Files include most of the public release data. "Other Person" files containing data on children, siblings, household members, helpers, and transfers between HRS respondents and their children, with observations for each other person, are distributed by ISR but have not yet been included in the RAND Enhanced Fat Files.

However, RAND Center of Aging has produced the RAND HRS Family data including the "Other Person" files. The RAND HRS Family data contains a cleaned, processed, and streamlined collection of variables related to the respondent's family. The files described here include a subset of available characteristics of all children of HRS respondents and spouses, data on children-in law, and data on grandchildren of the respondent.

Variables are added:

- Tracker file variables (version of tracker varies across years)
- Clarify the household level data for applying to respondent level data (for financial and family variables)
- Generally helpful variables or flag discrepancies
- Provide categorical bracket variables, if not provided in the raw data, that summarize unfolding bracket questions
- Correct raw data to reflect Data Alerts

2

3. Lining up the years

The table below summarizes the fat files, the first character of variable names, and RAND HRS waves associated with each interview year. References to "raw" data mean the data as publicly released by HRS.

Note that for years affected by overlap cases, there are two versions of the Fat Files, one with the original respondent IDs from the raw public release data, and the other with the most recent IDs as found on the Tracker and RAND HRS Data file. For most cases the IDs are identical on both files, and if you drop all overlap cases they are essentially the same.

In the raw public release data for 1992 and 1993 most variable names begin with "V". In order to make unique variable names when merging 1992 and 1993 data together, the variables in the 1993 Fat File that have names like Vnnnn, e.g., V1234, have been renamed to begin with "B", e.g., B1234.

Fat file names follow the following naming convention: [A/H][D]yy[E/F][vers]
-   A is for Ahead and H is for HRS
-   "D" if present, indicates the version of IDs will merge best with RAND HRS data
-   "yy" is the interview year
-   E is for early release and F is for final release
-   vers indicates the version of the HRS public release and of the RAND release

For example, h04f1a is the fat file for HRS 2004, HRS Final Release Version 1, RAND version A.

Interview Year	Fat File		Associated First Letter(s) of Variables (raw and Tracker)	RAND HRS Wave
	IDs as in public release	*IDs as on Tracker, RAND HRS*		
1992	h92f1b	hd92f1b	**A** (tracker) and **V** (raw, fat files)	Wave 1
1993	a93f2a	ad93f2a	**B** (tracker, fat files) **V** (raw)	Wave 2 (2A in codebook)
1994	h94f1a		**C** (tracker) and **W** (raw, fat files)	Wave 2 (2H in codebook)
1995	a95f2a	ad95f2a	**D**	Wave 3 (3A in codebook)
1996	H96f4a		**E**	Wave 3 (3H in codebook)
1998	h98f2b	hd98f2b	**F**	Wave 4
2000	h00f1c		**G**	Wave 5
2002	H02F2c		**H**	Wave 6
2004	h04f1a		**J**	Wave 7
2006	h06f2b		**K**	Wave 8
2008	h08h1b		**L**	Wave 9

4. Overlap cases and merging with RAND HRS Data

There are two versions of the 1992, 1993, 1995, and 1998 files, to accommodate the HRS-AHEAD (1992) and AHEAD-AHEAD (1993, 1995, and 1998) overlap cases.

For all these years, one version has the original HHIDPN, as found in the HRS public release data, and the other has HHIDPN as assigned on the RAND HRS Data:

1992 fat file variations
      h92f1b: with original HRS IDs
      hd92f1b: with IDs that merge best with RAND HRS

1993 fat file variations
      a93f2a: with original HRS IDs
      ad93f2a: with IDs that merge best with RAND HRS

1995 fat file variations
      a95f2a: with original HRS IDs
      ad95f2a: with IDs that merge best with RAND HRS

1998 fat file variations
      h98f2b: with original HRS IDs
      hd98f2b: with IDs that merge best with RAND HRS

The "hd" and "ad" versions of the 1992, 1993, 1995, and 1998 data will easily merge with the RAND HRS data.

Please see the RAND HRS Codebook, Overlap Identifiers, in Section A for background information on overlap cases. HRS/AHEAD overlap cases are identified by their AHEAD IDs in the RAND HRS, and in the Tracker files. AHEAD-AHEAD overlap cases are identified by the most recent AHEAD ID in the RAND HRS and in the Tracker file.

5. Public Release Versions

- HRS 1992 = V1.01
- AHEAD 1993 = V2.1
- HRS 1994 = V1.00
- AHEAD 1995 = V2.0
- HRS 1996 = V4.0
- HRS 1998 = V2.3
- HRS 2000 = V1.0 (with January and November 2005 HRS Data Alert applied)
- HRS 2002 = V2.0 (dropped 2 cases based on Tracker Documentation)
- HRS 2004 = V1.0
- HRS 2006 = V2.0
- HRS 2008 = V1.0

- Tracker 2.0 was used in developing the 1993 and 1994 fat files.
- Tracker 3.1 was used in developing the 1995, 1996, 1998, and 2000 fat files.
- Tracker 2004.v1 was used in developing the 2002 and 2004 fat files.
- Tracker 2006.v1 (released October 2007) was used in developing the 2006 fat files.
- Tracker 2008.v1 (released October 2010) was used in developing the 2008 fat files.

6. Downloading the Files

To access the data:

- You must first register with HRS, if you haven't already. Registration is free. You will receive a password from HRS within 24 hours, usually much sooner. When you receive your password, you can return here to continue.
- If you've already registered, login through the HRS Public File Download Area.
- Once you have logged in, follow the "Datasets and Files" link, then the "HRS Data Products" link.

  For example, the 2008 Fat File is called 2008 HRS Core Fat Files (Final) RAND (v.B)

  Available files include 1992, 1993, 1994, 1995, 1996, 1998, 2000, 2002, 2004, 2006 and 2008 data. These can easily be used with the current release of the RAND HRS Data file.

Please contact RANDHRSHelp (RANDHRShelp@rand.org) with any question.