

Data Management – unpacking the pieces

Everyone has their own way of organizing files, but here is one way to organize the RAND HRS data products.

Set up a folder to store the RAND HRS files, e.g., a RandHRS folder somewhere on your C-drive that makes sense for you, and make folders for Sas, Stata or SPSS data and for the programs distributed with the RAND HRS Data file. So you might have a RandHRS folder and the following sub-folders within it:

Sas = for Sas data files

Stata = for Stata data files

SPSS = for Spss data files

Programs = for programs distributed with the RAND HRS Data file

Assuming this is the way you set things up, here are some suggestions for how to “unpack” the data:

- *Unpacking the RAND HRS Data files*

- For SAS users
 - Unzip the contents of randHsas.zip to the RandHRS folder
 - Move formats.sas7bcart, sasfmts.sas7bdat, and rndhrs_h.sas7bdat to the Sas folder
- For Stata users
 - Unzip the contents of randHstataI.zip or randHstataSE.zip, as appropriate for your version of Stata, to the RandHRS folder.
 - Move the rndhrs1h.dta-rndhrs8h.dta or rndhrs_h.dta file(s) to the Stata folder
- For SPSS users
 - Unzip the contents of randHspss.zip to the RandHRS folder
 - Move rndhrs_h.sav to the SPSS folder
- For all users
 - randhrsh.pdf and rnd_hdd.pdf can be left where they are. Randhrsh.pdf is the codebook for the Rand HRS data, and rnd_hdd.pdf is the data description, which gives a brief description of the files and the different formats they come in.
 - Move rndpgmH.zip to the Programs folder and unzip it. It will create sub-folders Pgm, MrgH, Mac, and Fmt. The programs that were used to develop derived variables are under Pgm and those that pulled all sections together are under MrgH. Mac will hold Sas macros in text files called ____.mac and Fmt will have value labels in text files called ____.fmt. Under the Pgm folder, additional folders hold the programs by section, e.g., the Health folder holds the programs that developed health variables. In most of these there is a file called [folder-name].pgms, e.g., health.pgms, which is a text file that briefly describes what the programs do.
- **Note for SAS users:** the “Distribution and Technical Notes” section of the introduction in the codebook describes how you might use some of the macros available in the Mac folder. It also describes how you might use the setuphrs.inc file,

which can help set up libnames once you edit it to reflect where you've put things. If you think you'd like to use these, but have questions, send us email (RANDHRSHelp@rand.org).

- ***Unpacking the Rand Fat Files***

- Copy the full CD to your hard drive, under the RandHRS folder.
- In each folder on the CD are all the things that pertain to a particular wave, e.g., A93f2 for the 1993 data. The wave-specific folders may include documentation if there is any (it will be the PDF file), a "readme" file, ____.cont files which are text files with the SAS contents of the data files, and a folder called Fmts which has files that contain SAS value label definitions (*.fmt) and format statements that assign the value labels to variables (*.format). The files called *_sas.zip, *_stataSE.zip and *_spss.zip contain encrypted data.
- If you are using SAS, select the *_sas.zip file. If you are using Stata, select the *_stataSE.zip file. If you are using SPSS, select the *_spss.zip file. Unencrypt the selected file, using the pass-phrase gotten from the HRS web site. Please see the main Readme.txt file on the Fat File CD for directions on obtaining the pass-phrase. This will create a Sas, StataSE, or SPSS folder in the wave-specific folder, which will contain the Sas, Stata or SPSS data sets.
- For ease of accessing these, we suggest that you move the unencrypted Sas, Stata or SPSS data files from the Sas, StataSE, or SPSS folders within the wave-specific folders to the Sas, Stata or SPSS folder immediately under the RandHRS folder. That way they are all in one place, along with the rndhrs_h file.
- For each wave, there is at least one fat file with most of the raw data, which has a name such as h00f1c ("h" for HRS or "a" for Ahead, 00 for 2000, "f" for final, 1 for version 1, "c" for Rand version C). There are also at least two other files that contain income and wealth imputation data, e.g., incm00f1 and wlth00f1 for 2000. These include the income and wealth variables found in the RAND HRS Data file but also have more detailed information, such as the imputed data for each component.
- For some waves, there are two sets of each of these files, one with IDs as they were in the original data, and one with IDs as they appear in the RAND HRS Data file. These are the same except for the overlap cases. For example there are h98f2b and hd98f2b files for 1998, with the original and RAND HRS version H IDs, respectively. hd98f2b will merge better with the RAND HRS Data file (Versions D through H). There are also two sets of income and wealth files, incm98f2/wlth98f2 and incf98f2/wlth98f2, with original and RAND HRS Version H IDs, respectively. Please see the README.* files for a description of the specific files for each wave.