

Health and Retirement Study

Labor Section Carry Forward Variables
Final, Version 2.0
September 2006

Data Description and Usage

1. Overview and Background

The *Labor Section Carry Forward Variables* (Final, Version 2.0) data release consists of information derived from the Health and Retirement Study (HRS), a national longitudinal study of the economic, health, marital, family status, and public and private support systems of older Americans. The National Institute on Aging provided funding (NIH U01 AGO9740), with supplemental support from the Social Security Administration. The Institute for Social Research (ISR) Survey Research Center (SRC) at the University of Michigan conducted the survey.

By receiving the data, which have been freely provided, you agree to use them for research and statistical purposes only and make no effort to identify the respondents. In addition, you agree to send HRS a copy of any publications you produce based on the data. See [Obtaining the Data](#) for additional details.

When an HRS respondent reports no change in employment status since the previous interview, certain questions relating to current employment are skipped. Additional questions are skipped if the respondent holds the same position or title in addition to the same employer. As a result, questions about certain job characteristics were asked only in HRS 1992 (Wave 1), HRS 2000 (Wave 5), HRS 2002 (Wave 6), or when a respondent changed employment/position. That is, a respondent in HRS 1998 (Wave 4) having the same employment characteristics as in 1992 would not have been asked employment questions in HRS 1994 (Wave 2) through HRS 1998 (Wave 4). As a service to users, the files in this data release carry forward the information from each prior wave whenever certain employment variables are "skipped" in the current wave.

2. Background

The *Labor Section Carry Forward* data set consists of respondent-level, cross-sectional files constructed from the employment sections of HRS 1994 (Wave 2), HRS 1996 (Wave 3), HRS 1998 (Wave 4), HRS 2000 (Wave 5), HRS 2002 (Wave 6), and HRS 2004 (Wave 7). The files include only HRS cohort members for HRS 1992 to HRS 1996, and both HRS and War Baby cohort members for HRS 1998 and HRS 2000. The HRS 2002 and 2004 files contain information for all employed respondents. In HRS 2000 (Wave 5)¹, some of those questions have been restored while others are not asked. In HRS 2002 (Wave 6), all employed respondents were asked the industry and occupation sequence, but were not asked certain questions relating to sick time and disability if their employment status was unchanged². For those respondents who reported working at the same employer or business, or who, for some questions, were at the same position, values are carried forward from the previous interview where the job was first reported.

3. Procedures

3a. Code Frame Reconciliation

As the skipped values were carried forward, study staff attempted to reconcile and unify code frames, since alternative responses were allowed for similar questions in different waves. For some questions, non-matching codes may stand for the same response, and in other cases, matching code

¹ The restored questions in HRS 2000 ask about the number of employees at the location and all locations, number of paid sick days, disability coverage, start date of employment, and union membership.

² The Industry, Occupation, and Job Requirements variables parallel those constructed for the [RAND HRS Data File](#).

values may have a different meaning in different waves. An example of different codes representing the same response is the *firm size at all locations* question. In this question the response indicating "only one location" is represented by different codes depending on the survey year. It is 999995 in HRS 1992, 9999997 in HRS 1994, 9999995 in HRS 1996, 99995 in HRS 1998, a combination of 0, 9995, 99995, and 999995 in HRS 2000, and a combination of -2, 0, 9999995 in HRS 2002, and -2 in HRS 2004. The unification of this code involves converting those codes to 9999995 designating "only one location" for all waves.

As an example of a situation where the same code has a different meaning in different waves consider the question *number of years of education needed for the job*. For this question the code 96 indicates "no education needed" in HRS 1994, but "Missing/Inap" in HRS 1992. The unification of this code involves converting the code 96 in HRS 1992 to blank (Missing/Inap) and in HRS 1994 to 0 (no education needed). It was not possible to apply the unification process to all codes with different interpretations. Therefore, it is critical for the user to continue to pay attention to the source of each of the observed values in the constructed variables and interpret them accordingly.

3b. Flag Variables

In addition to the constructed carry forward variables, *Flag* variables have been created. Each flag variable indicates the source of an observed value, whether it is from HRS 1992 (Wave 1), HRS 1994 (Wave 2), HRS 1996 (Wave 3), HRS 1998 (Wave 4), HRS 2000 (Wave 5), HRS 2002 (Wave 6), or HRS2004 (Wave 7) for each of the variables. The flag variables have only two values for constructed variables from Wave 2 (1 or 2), three values for variables from Wave 3 (1, 2, or 3), four values for variables from Wave 4 (1, 2, 3, or 4), six values for variables from Wave 6 (1, 2, 3, 4, 5, or 6), and seven values for variables from Wave 7 (1, 2, 3, 4, 5, 6, or 7). For example, for the variable W3608 (the industry variable), if the respondent did not change her/his employment in Wave 2, the value in W3608s is taken from wave 1 and the flag FW3608 is set equal to 1, indicating that the source of the carry forward variable W3608s is Wave 1. If the person had changed her/his employment between Wave 1 and Wave 2, the value of the industry code W3608s is what is observed in Wave 2; FW3608 is set equal to 2, indicating that the data source is Wave 2.

3c. Variable Names

The constructed carry forward variables have the same name as the variables in their original data files, but those names are followed by an "s". For example, the constructed variable for W3608 (the industry question) in Wave 2 is W3608s. The source or flag variables have the same names as the variables in their original data files with an "F" preceding those names. For example, the flag variable for the industry code in Wave 2 is FW3608.

4. Distribution Information

4a. File Naming Conventions

Files are named beginning with "CF" (Carry Forward), followed by the wave year, a separator ("_"), and then one or two letters designating the original section. The following extensions are used for types of files that are distributed:

.DA data files

.SAS SAS program statements
 .SPS SPSS program statements
 .DO Stata DO statements
 .DCT Stata dictionary statements
 .TXT codebook

One of these file types is provided for each wave-level data set. For example:

CF2000_G.DA contains data derived from HRS 2000 Section G,
 CF2000_G.SAS contains corresponding SAS program statements,
 CF2000_G.SPS contains corresponding SPSS program statements,
 CF2000_G.DO contains corresponding Stata do statements,
 CF2000_G.DCT contains corresponding Stata dictionary statements, and
 CF2000_G.TXT contains the ASCII codebook.

4b. Data Files

Data files are provided in ASCII format, with fixed-length records. Use associated SAS, SPSS or Stata program statements to read the data into the analysis package of your choice. Each of the data files includes the HHID, PN, and xSUBHH identification variables, which should be used for merging with the original data files. Records are sorted in order by HHID and PN.

The files are packaged for download from our Web site in ZIP format. The distribution file, *LSCF.zip*, contains six sets of year-level files plus this document.

HRS Wave	SAS	SPSS	Stata	Data	N	Codebook
1994 (Wave 2)	CF1994_FA.sas CF1994_FB.sas	CF1994_FA.sps CF1994_FB.sps	CF1994_FA.do/.dct CF1994_FB.do/.dct	CF1994_FA.da CF1994_FB.da	11596	CF1994_FA.txt, CF1994_FB.txt
1996 (Wave 3)	CF1996_G.sas	CF1996_G.sps	CF1996_G.do CF1996_G.dct	CF1996_G.da	10964	CF1996_G.txt
1998 (Wave 4)	CF1998_G.sas	CF1998_G.sps	CF1998_G.do CF1998_G.dct	CF1998_G.da	13113	CF1998_G.txt
2000 (Wave 5)	CF2000_G.sas	CF2000_G.sps	CF2000_G.do CF2000_G.dct	CF2000_G.da	12455	CF2000_G.txt
2002 (Wave 6)	CF2002_J.sas	CF2002_J.sps	CF2002_J.do CF2002_G.dct	CF2002_J.da	18167	CF2002_J.txt
2004 (Wave 7)	CF2004_J.sas	CF2004_J.sps	CF2004_J.do CF2004_G.dct	CF2004_J.da	20147	CF2004_J.txt

4c. Directory Structure

While a particular setup is not required for using HRS files, we have traditionally suggested a directory structure. By using this directory structure, you will not have to change the path name in your data descriptor files. If you use a different structure, just change the directory references in the program statement files.

Directory	Contents
=====	=====
c:\Lscfwd	File downloaded from Web site
c:\Lscfwd\codebook	codebook and documentation files
c:\Lscfwd\data	data files

c:\LSCFwd\sas	SAS command files
c:\LSCFwd\spss	SPSS command files
c:\LSCFwd\stata	Stata command files

Decompress the .zip file into the appropriate subdirectories. You will need approximately 7 MB of free space on your storage device to store the files.

5. Identification Variables

Identification variables for all *Labor Section Carry Forward* data files are stored in character format.

5a. HHID – Household Identifier

In the initial wave of data collection (in 1992 for the HRS sub-sample, in 1993 for the AHEAD sub-sample, and in 1998 for the WB and CODA sub-samples), each sample household was assigned a Household Identifier. HHID is stable across waves of data collection and uniquely identifies the original household and any households derived from that household in subsequent waves of data collection. HHID has six digits.

5b. PN – Person Number

In combination with HHID, PN uniquely identifies a respondent or respondent's spouse or partner. PNs are unique within an original household (HHID). The PN assigned to a particular respondent does not change across waves. PN has three digits.

5c. xSUBHH – Sub-household Identifier

In combination with HHID, xSUBHH uniquely identifies a household for a given wage. Sub-household identifiers can be different at each wave. xSUBHH has one digit..

6. 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 a .SPS extension, those with SAS statements with a .SAS extension, and those with Stata statements with .DO and .DCT extensions.

The statement files are named beginning with the same prefix as the corresponding data file. For example, SAS statements in the file CF2002_G.SAS go with the CF2002_G.DA data file.

6a. 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:\LSCFwd\sas" and the data file is located in "c:\LSCFwd\data", you can run the file as is. A SAS system file (*.SD2 or *.SAS7BDAT) will be saved to directory "c:\LSCFwd\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.

6b. 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, you must first open the *.SPS file in SPSS as an SPSS Syntax File.

If the *.SPS file is located in "c:\LSCFwd\spss" and the data file is located in "c:\LSCFwd\data", you can run the file as is. An SPSS system file (*.SAV) will be saved to directory "c:\LSCFwd\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.

6c. 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:\LSCFwd\stata" and the data file is located in "c:\LSCFwd\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.

Note that the variable names provided in the .DCT files are uppercase. If you prefer lowercase variable names, you may wish to convert the .DCT files to lowercase prior to use. You may do this by reading the .DCT file into a text or word processing program and changing the case. For instance in Microsoft Word, Edit, Select All, Format, Change Case, lowercase.

7. Obtaining the Data

7a. Registration and Downloading the Data

HRS data are available for free to researchers and analysts at the HRS Web site. In order to obtain public release data, you must first register at our Web site. Once you have completed the registration process, your username and password will be sent to you via e-mail. Your username and password are required to download any data files. 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. Registered users receive user support, information related to errors in

the data, future releases, workshops, and publication lists. The information you provide will not be used for any commercial use, and will not be redistributed to third parties.

7b. Conditions of Use

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

- not to attempt to identify respondents
- not to transfer data to third parties except as specified
- not to share your username and password
- to include specified citations in work based on HRS data
- to provide information to us about publications based on HRS data
- to report apparent errors in the HRS data or documentation files
- to notify us ([via our Web site](#)) of changes in your contact information

For more information concerning privacy issues and conditions of use, please read “Conditions of Use for Public Data Files” and “Privacy and Security Notice” at the Public File Download Area of the HRS Web site.

7c. Publications Based on Data

As part of the data registration process, you agree to include specified citations and to inform HRS of any papers, publications, or presentations based on HRS data. Please send a copy of any publications you produce based on HRS data, with a bibliographical reference, if appropriate, to the address below.

Health and Retirement Study
Attn: Papers and Publications
The Institute for Social Research, Room 3050
P.O. Box 1248
Ann Arbor, MI (USA) 48106-1248

Alternately, you may contact us by e-mail at hrequest@isr.umich.edu with “Attn: Papers and Publications” in the subject line.

8. If You Need to Know More

This document is intended to serve as a brief overview and to provide guidelines to using the *Labor Section Carry Forward* (Final, Version 2.0) data. 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.

8a. HRS Internet Site

Health and Retirement Study public release data and additional information about the study are available on the Internet. To access the data and other relevant information, point your Web browser to the HRS Web site: <http://hrsonline.isr.umich.edu/>

8b. 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

E-mail: hrequest@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

Table 1: List of Carried Forward Variables by Wave

Variables	Carried forward to...	Skipped if...
<i>Industry</i>	W2, W3, W4, W5	same employment
<i>Occupation</i>	W2, W3, W4, W5	same employment and same title
<i>Number of Paid Sick Days</i>	W2, W3, W4, W6	same employment and same title
<i>Disability Coverage</i>	W2, W3, W4, W6	same employment and same title
<i>Decision About Pay and Promotion</i>	W2, W3, W4, W5	same employment and same title
<i>Number of People Make Pay and Promotion Decisions for</i>	W2, W3, W4, W5	same employment and same title
<i>Starting Pay</i>	W2	same employment and same title or self-employed
<i>Start of This Kind of Work, Month</i>	W2, W3, W4, W5	same employment and same title or self-employed
<i>Start of This Kind of Work, Year</i>	W2, W3, W4, W5	same employment and same title or self-employed
<i>Start of This Employment, Month</i>	W2, W3, W4, W6	same employment
<i>Start of This Employment, Year</i>	W2, W3, W4, W6	same employment
<i>Number of Employees at Location (Continuous)</i>	W2, W3, W4, W6	same employment
<i>Number of Employees at Location (Bracketed)</i>	W2, W3, W4, W6	same employment
<i>Number of Employees at All Locations (Continuous)</i>	W2, W3, W4, W6	same employment
<i>Number of Employees at All Locations (Bracketed)</i>	W2, W3, W4, W6	same employment
<i>Union Membership</i>	W2, W3, W4, W6	same employment and same title
<i>Employer Contributions</i>	W3, W4, W5	same employment
<i>Intend to Participate</i>	W3, W4, W5	same employment
<i>Eligible to Participate</i>	W3, W4	same employment
<i>Number of Years to become Eligible</i>	W3, W4	same employment
<i>Number of Years of Education Needed for the Job</i>	W2, W3, W4, W5	same employment and same title or self-employed currently and self-employed in previous IW
<i>Amount of Experience Needed for the Job</i>	W2, W3, W4, W5	same employment and same title or self-employed currently and self-employed in previous IW
<i>Number of Years of Education Needed for the Job</i>	W2, W3, W4, W5	same employment and same title or self-employed currently and self-employed in previous IW
<i>Amount of Experience Needed for the Job</i>	W2, W3, W4, W5	same employment and same title or self-employed currently and self-employed in previous IW
<i>Unit of the Amount of Experience</i>	W2, W3, W4, W5	same employment and same title or self-employed currently and self-employed in previous IW
<i>Job Requires Physical Efforts</i>	W2, W3, W4, W5	same employment and same title or self-employed currently and self-employed in previous IW
<i>Job Requires Lifting Heavy Loads</i>	W2, W3, W4, W5	same employment and same title or self-employed currently and self-employed in previous IW
<i>Job Requires Stooping, Kneeling, or Crouching</i>	W2, W3, W4, W5	same employment and same title or self-employed currently and self-employed in previous IW
<i>Job Requires Good Eyesight</i>	W2, W3, W4, W5	same employment and same title or self-employed currently and self-employed in previous IW
<i>Job Requires Intense Concentration</i>	W2, W3, W4, W5	same employment and same title or self-employed currently and self-employed in previous IW
<i>Job Requires Skill in Dealing with Other People</i>	W2, W3, W4, W5	same employment and same title or self-employed currently and self-employed in previous IW