HEALTH AND RETIREMENT STUDY

2007 Prescription Drug Study
Final Release V1.0, March 2011
(Sensitive Health Data)

Data Description and Usage

1. Introduction

1a. Background

The HRS 2007 Prescription Drug Study (PDS) is the second wave of a multi-wave mail survey designed to track changes in prescription drug utilization as Medicare Part D, the prescription drug benefit administered by Medicare, is phased in. The baseline wave, administered in 2005, was intended to capture prescription drug use, coverage, and satisfaction prior to the implementation of Medicare Part D, as well as awareness of the new drug benefit and available subsidies. The second wave was designed to capture similar information post-implementation. An additional follow-up was administered in 2009 (the HRS 2009 Health and Well-Being Study).

The original sample for the PDS was drawn from respondents to HRS 2004. Of the 20,129 respondents to 2004, 14,242 met the criteria for selection into PDS. The study sample included HRS respondents born in 1942 or earlier (65th birthday in 2007), or already covered by Medicare or Medicaid at some time between 2002 and 2004. Respondents interviewed by proxy, and those in nursing homes were eligible for inclusion. Approximately 40% of those who were eligible were excluded because they were participants in the Consumption and Activities Mail Survey, which was fielded at the same time as PDS. A sample of 5,654 persons was drawn from the remaining eligible respondents, with oversamples of persons lacking prescription drug coverage or having low income and wealth.

The sample for the 2007 wave of the PDS consisted of everyone from the original 2005 sample who responded to either PDS 2005 or gave a core HRS interview in 2006 and was not known to be deceased at the beginning of the field period. The PDS 2007 sample originally consisted of 4,990 respondents.

The survey was mailed to selected respondents in October 2007. Non-respondents to the mail survey mailings were contacted by phone and asked to complete a telephone interview. The field period continued through September 2008.

After attempting contact with the sample, we determined that 211 persons died prior to the October 2007 start of the second wave of the PDS and were determined to be ineligible for inclusion in the sample. Of the 4,779 remaining eligible cases, 3,536 returned questionnaires or completed a telephone interview, for a response rate of 74%.

The HRS 2007 Prescription Drug Study was funded by the National Institute on Aging through a competing supplement to the Health and Retirement Study.

1b. File Structure

The 2007 Prescription Drug Study (Final Release V1.0) data consist of two respondent-level files (a sample data file and a questionnaire data file) and three multiple-record files (medication lists and related data). Each person is uniquely identified in the files by the combination of a household ID (HHID) and person number (PN). The primary identification variables used in the HRS (HHID, PN) are included in all files to facilitate merging with other HRS data sets.

The sample data file (PDS2007S_R) contains the 4,990 respondents who were selected for participation in the 2007 Prescription Drug Study, sample section weights, and relevant identifiers. This file is provided so that analysts can identify the non-respondents for analysis if desired.

The questionnaire data file (PDS07A_R) contains one record for each respondent or proxy informant who completed and returned a questionnaire (N=3,536).

The medication file (PDS2007E_M) contains detailed medication data for each respondent who completed section E of the questionnaire. This file includes drug names, information on dosage, cost, and adherence.

The medication files from Sections F and G of the questionnaire (PDS2007F_M & PDS2007G_M) contain the medications listed by respondents as additional prescription drugs that were not included in Section E, and over-the-counter medications, respectively.

Codebooks for each data file, the questionnaire, and this document may be viewed on the HRS public web site.

1c. Data Quality

1c-1. Recoded and constructed variables

The majority of variables in the questionnaire data file match the code frame structure of the questions in the Prescription Drug Study questionnaire. However, questions that allowed for multiple mentions have been recoded such that each possible response is represented as a dichotomous variable. In addition, new variables created to categorize responses to open ended questions have been made available in this data release.

Data from the medication sections of the questionnaire have been reorganized as a multiple-record file (multiple lines per observation).

1c-2. Extreme values and missing data

Extreme or unlikely values in the data were checked against the hard copies of the questionnaires to determine whether errors occurred during data entry. Confirmed cases in which the respondent reported an extreme value remain unchanged in the data.

When respondents did not provide an answer to a question, their answers were treated as missing. However, variables may contain missing values for other reasons. In an attempt to distinguish the various types of possible missing responses, meaningful missing codes are used in the SAS and Stata versions of the data. Generally, these codes distinguish data that are missing because a respondent chose not to respond to a question from those questions which the respondent was instructed to skip. The meaningful missing codes used in the data adhere to the classification in the table below.

Code	Reason for Missing Data
.D	Uncertain, can't say
.M	Missing
.R	Refused
.S	Skip (information not available due to skip patterns, typically from the preceding question
.X	Answer not applicable

1d. Sample Weights

PDS Questionnaire Weight

The PDS questionnaire weight is the product of the HRS core sampling weight, an adjustment factor for sample selection, and a non-response adjustment factor. The HRS sampling weight from the most recent preceding interview (HRS 2004 or 2006) was used as the base weight. The sample selection adjustment is the inverse of the sampling rate among eligibles in each of the sample strata defined by prescription drug coverage and household income and wealth levels in 2004. The non-response adjustment factor was obtained from a propensity model predicting the probability of completing the PDS questionnaire among those selected and eligible to participate. The propensity model was estimated by logistic regression and weighted by the base weight, adjusted for PDS sample selection. Predictor variables included age, sex, race/ethnicity, education, coupleness, self-rated health, number of physical limitations, type of drug coverage, and level of out-of-pocket spending on drugs. Predictor variables were taken from the most recent preceding interview, either 2004 or 2006. The inverse of the fitted probability of completion formed the non-response adjustment factor. Finally, the weights were post-stratified to closely match the HRS 2006 sample composition by age, gender, and race.

Section E Analysis Weight

There were 396 respondents who did not fill out Section E. Some of these had no drug use and so can be considered to have given a valid response, but others did have drug use and must be considered to have failed to respond to Section E. We seek to re-weight those who did respond to compensate for any differential non-response by level of drug use or other factors. To do this we need to determine how many medications each of these 396 should have reported on.

First, we determined those who had zero drug use by examining responses to earlier sections of the survey. There were 190 persons who did not respond to Section E, did not report any count of drugs at A3 or A4 and who checked "no" to question A1, indicating that they had not used prescription drugs in the past year. Another 39 said "yes" (or did not answer) at A1, but answered zero at A3 or A4, indicating they had not used drugs in the past month. An additional respondent gave no answer to A1 or A3, but indicated that they had zero expenditures on drugs. All of these 230 cases were assigned a count of zero drugs and considered to have validly responded to Section E.

Second, we determined the count of drugs used by 166 the persons who could not be definitely assigned to zero use but who did not respond to Section E. There were 133 persons in this group who

provided a non-zero count of drugs at A3 or A4. The remaining 33 were assigned an imputed count of drugs, of which 4 were imputed to zero. In the end, 162 respondents were considered to have had non-zero drug use but failed to respond to Section E, and 234 cases were considered to have zero use.

The medication analysis weight is the product of the PDS questionnaire weight and a non-response adjustment for non-response to Section E. The non-response adjustment factor is calculated for each level of self-reported drug count (from 1 to 10+). The inverse of the weighted probability of completing Section E forms the non-response adjustment factor.

2. Obtaining the Data

2a. How to Apply

The PDS 2007 data are considered a "Sensitive Health" data product and require an additional data use agreement in addition to the HRS user registration. However, if you have already received access to PDS 2005, you do not need to complete an additional request for the PDS 2007 data.

If you are not a registered HRS user or do not have PDS 2005 data access, follow these instructions:

1. Visit the HRS User Registration/File Download Web site. If you do not already have a username and password, you must register in order to enter the site.
2. Identify the desired data set (2007 Prescription Drug Study Data (v.1.0) under the *Requesting HRS Special Access Files* heading.
3. Download and complete the *Data Use Agreement* in hardcopy format (See Appendix).
4. Send a signed copy of the agreement via surface mail to:

   ```
   Health and Retirement Study
   DUA Review Committee
   426 Thompson Street, Room 3050 ISR
   Ann Arbor, Michigan 48104-2321
   ```

2b. Approval Process
The Health and Retirement Study will review your request, and based on the information that you provided when you registered at the HRS User Registration/File Download Web site, verify your identity and institutional affiliation. Once this authentication process has been completed to our satisfaction, we will authorize access to the Prescription Drug Study 2007 data. We will communicate with you at the email address that you provided when you registered at the download site.

2c. How to Download
Once you receive your approval notification, login to the HRS User Registration/File Download Web site. In the *HRS Special Access Files* box on the right hand side of the page you will now see links to the 2007 Prescription Drug Study (v.1.0) data set. Click on the appropriate link and follow instructions to download the PDS 2007 files.

2d. Publications Based on Sensitive Health Data

You must inform HRS of any papers, publications, or presentations based on this HRS sensitive health data set. Please send a copy of any such publications, with a bibliographical reference if appropriate, to this address:

```
Health and Retirement Study
Attn: Papers and Publications
The Institute for Social Research, Room 3050
426 Thompson Street
Ann Arbor, MI (USA) 48104
```

Papers and publications may also be submitted via electronic mail; send message and attachment(s) to hrsquest@isr.umich.edu. Please include "Attn: Papers and Publications" on the subject line.

The following citation should be included in any research reports, papers, or publications based on these data:

> In text: "The HRS 2007 Prescription Drug Study is sponsored by the National Institute on Aging (grant number NIA U01AG009740) and was conducted by the University of Michigan."

> In references: "Health and Retirement Study, 2007 Prescription Drug Study. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740). Ann Arbor, MI, (2011)."

3. Setup

3a. Distribution Set

The HRS 2007 Prescription Drug Study Final Release files are packaged for distribution both in a .zip file, pds2007.zip, and a self-decompressing .exe file, pds2007.exe. In order to keep the contents secure, the .zip and .exe files have been encrypted.

For users working in a Microsoft Windows environment, extract the files by running the self-decompressing archive file, pds2007.exe, from the command line or from Windows Explorer.

UNIX, Linux, and Macintosh users will extract the files from the standard .zip file, pds2007.zip.

Users who require special file formats should contact the HRS Help Desk (hrsquest@isr.umich.edu).

Extract the data file, the program statement file that matches your analysis environment, the data description (this file), and the codebook file.

3a-1. Windows Environment

Run the self-decompressing .exe file, pds2007.exe, from the command line or from Windows Explorer. When you are prompted for the pass-phrase, respond with the character string that you received via e-mail. The output from the self-extracting process will be the files listed below. If you have problems, please contact the HRS help desk.

3a-2. UNIX Environment

Use the file decompression software installed on your system (e.g. `gunzip`) to expand pds2007.zip. When you are prompted for the pass-phrase, respond with the character string that you received via e-mail to output the files listed below. If you have problems, please contact the HRS help desk.

3a-3. Mac Environment

For Mac users, we suggest that you download either StuffIt or The Unarchiver, free software that will handle password-protected .zip files on a Mac. You can first extract the files from the main .zip file, and then drag the separate files to the desktop, which you can then unzip. When you are prompted for the pass-phrase, respond with the character string that you received via e-mail to output the files listed below. If you have problems, please contact the HRS help desk.

3b. Filename Suffixes

For all data sets, filename suffixes mean the following:

```
File     Description
======   ==================================================
*.da     Data
*.dct    Stata dictionary
*.do     Stata "do" file
*.sas    SAS program statements
*.sps    SPSS program statements
*.txt    ASCII text files (codebook)
*.pdf    PDF format
```

3c. 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 files.

```
Directory            Contents
==================   ==================================
c:\pds2007\          Distribution file

c:\pds2007\docs\     pds07A_R.txt
                     pds07E_M.txt
                     pds07F_M.txt
                     pds07G_M.txt
                     pds07S_R.txt
                     pds2007dd.pdf (this file)
                     pds2007qnaire.pdf

c:\pds2007\data\     pds07A_R.da
                     pds07E_M.da
                     pds07F_M.da
                     pds07G_M.da
                     pds07S_R.da

c:\pds2007\sas\      pds07A_R.sas
                     pds07E_M.sas
                     pds07F_M.sas
                     pds07G_M.sas
```

```
                        pds07S_R.sas

  c:\pds2007\spss\      pds07A_R.sps
                        pds07E_M.sps
                        pds07F_M.sps
                        pds07G_M.sps
                        pds07S_R.sps

  c:\pds2007\stata\     pds07A_R.do
                        pds07A_R.dct
                        pds07E_M.do
                        pds07E_M.dct
                        pds07F_M.do
                        pds07F_M.dct
                        pds07G_M.do
                        pds07G_M.dct
                        pds07S_R.do
                        pds07S_R.dct
```

3d. Program Statements

The distribution package provides SPSS, SAS, or Stata program statements describing the 6 variables in the sample data set, 151 variables in the questionnaire data set, 24 variables from Section E, and 6 variables each from Sections F & G. These are designed to assist the user in creating system files from the ASCII data file. Files containing SPSS statements are named with a .sps extension, those containing SAS statements with a .sas extension, and those containing Stata statements with .do and .dct extensions. The ASCII file for the sample file contains 4,990 records. The ASCII file for the questionnaire file contains 3,536 records. System files will differ in size, but should contain the same number of records and variables.

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:\pds2007\sas" and the data file is located in "c:\pds2007\data", you can run the file as is. A SAS system file (*.sas7bdat, assuming that you are using SAS V7 or later) will be saved to directory "c:\pds2007\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:\pds2007\spss" and the data file is located in "c:\pds2007\data", you can run the file as is. An SPSS system file (*.sav) will be saved to directory "c:\pds2007\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:\pds2007\stata" and the data file is located in "c:\pds2007\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.

4. If You Need to Know More

This document is intended to serve as a brief overview and to provide guidelines for using the HRS 2007 Prescription Drug Study data. If you have questions or concerns that are not adequately covered here or on our restricted data Web site, or if you have any comments, please contact us. We will do our best to provide answers.

4a. HRS Internet Sites

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 at http://hrsonline.isr.umich.edu/.

Visit the HRS Restricted Data Web site at http://hrsonline.isr.umich.edu/index.php?p=resdat if you want to find out more about restricted data products.

4b. 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: hrsquest@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
```

Appendix: Data Use Agreement

Health and Retirement Study
Sensitive Health Data Use Agreement

Data Set Requested: _____

By signing this form and obtaining the requested data set from HRS, the User agrees:
1. To use the data set solely for statistical reporting and analysis.
2. Not to share these data with, or provide copies of these data to, any other person or organization. Note: Each research assistant/associate, graduate student or undergraduate student, or other individual working on the research project must sign and submit his/her own form.
3. To return or destroy the data set, and any derivative data files, upon request from HRS.
4. To make no attempt to link this data set with individually identifiable records from any source, or in any other way attempt to identify the persons in this or other HRS data sets.
5. That if the identity of any person or establishment in this data set is inadvertently discovered, then (a) no use will be made of this knowledge, (b) the Director of HRS will be advised of this incident immediately, (c) the information that would identify any individual or establishment will be safeguarded or destroyed, as requested by HRS, and (d) no one else will be informed of the discovered identity.
6. To employ the following guidelines when producing tabulations for distribution:
   - Magnitude Data: Ensure that no cells/strata with $n < 5$ are produced.
   - Frequency Data: Apply a marginal threshold of $n >= 5$ and cell threshold of $n >= 5$ to all tabulations.
   - Protecting against complementary disclosure: Additional cells may be suppressed, i.e., complementary disclosure, to make sure the primary suppressions cannot be derived by subtraction from published marginal totals.
7. Aggregate statistical summaries of the data and analyses (frequency tabulations, magnitude tabulations, means, variances, regression coefficients, and correlation coefficients): These are approved under this agreement and may be freely published by the User, subject to the provisions above.
8. To cite HRS as the data source in any publications or research based upon these data, and to provide a copy of any publications to the HRS. Please refer to the data description for the correct citation that should be included in any research reports, papers, or publications based on these data.
9. To maintain a secure computing environment for storage and use of this data set and any data sets derived from it.
10. To hold harmless and indemnify HRS and the University of Michigan, its agents and employees, for any claims of breaches of confidentiality arising out of his/her research, defined as failure to abide by any section of this agreement or any accidental or intentional violation of privacy of any contributor to any HRS data resource.

UserID (from HRS File Download site):

_____

_____   _____
Signature   Title

_____   _____
Name (printed or typed)   Employer/Institutional Affiliation

For Office Use Only	Date	Initials
Received		
Verified		
Approved		
Access Granted		