HRS 2004 Core Early Imputations

Overview

HRS 2004 Core Early Imputations were created based on HRS 2004 Core Early Release (Version 1.0) and a set of imputation programs developed by the HRS. According to the *ad hoc* criteria described in Cao (2001a), imputations were performed for selected variables in the following 19 sections:

| HRS 2004 Core Early Imputations by Section ||
Section	Contents
CR	Physical Health
EH	Family Structure
ETC	Family Structure and Transfers
EFC	Family Structure and Transfers
FR	Parents, Siblings and Transfers
GHP	Functional Limitations and Helpers
HH	Housing
JR	Employment
KR	Last Job
LR	Job History
M1R	Disability for Reinterviews
M2R	Disability for Non-Reinterviews
NR	Health Services and Insurance
PR	Expectations
QH	Assets and Income
RH	Asset Changes
SR	Widowhood and Divorce
TR	Wills and Life Insurance
UH	Asset Reconciliation (HRS 2002 Assets)
UH	Asset Reconciliation (HRS 2004 Assets)

With a few exceptions, all the variables selected for imputation contain bracketed information. The standard imputation procedure starts with imputing "ownership" for each responding unit – whether it has a positive value on the variable being imputed. If a responding unit is determined to have a positive value on the variable, the procedure then goes on to impute the value, conditional on bracketed information available for the unit, and based on continuous reports provided by other units.

The actual missing value assignment scheme is essentially similar to what is often called in the literature "sequential random assignment," where data sequence is first sorted either completely randomly or based on some predicted scores from a simple regression model. A missing value is

then imputed based on an "immediate neighbor rule" as described in Cao (2001b). Wherever applicable, the covariates used in a regression model include age, gender, education, and marital status. For some housing and income variables in Sections H (Housing) and Q (Asset and Income), however, the covariates also include information more relevant to the variable being imputed. For instance, when imputing income from stocks/mutual funds, we include the value of stocks and mutual funds as a covariate. A "User Note" in the codebook indicates when covariates of this type were used..

Income and Asset Composites

HRS 2004 Core Early Imputations contain the following composites:

> JHHINC – household income in the last calendar year
> JASSETS – household assets at the time of the interview
> JHOME1 – primary home equity
> JHOME2 – second home equity
> JNETWRTH – sum of JASSETS, JHOME1, and JHOME2

For the definitions of these variables, please check the codebook.

Merging the Imputation Data with HRS 2004 Core

Before merging the Imputation data with HRS 2004 Core data, users will need to drop from the 2004 Core data eight respondents who do not fit the requirements of the HRS 2004 sample. (See HRS 2004 Early Core Release Data Description, http://hrsonline.isr.umich.edu/meta/2004/core/desc/h04dd.pdf (p. 12), for more detail.

The HHIDs and PNs of the eight respondents to be dropped are: 502333 010, 500075 010, 501383 010, 501679 010, 502587 010, 501851 010, 501353 010, and 501321 010.

Data Quality Control

Various measures have been taken to ensure the quality of HRS imputation data. These include verifying whether bracket responses are mathematically consistent (e.g., Min < Max), whether a valid response exceeds the default value for missing data, and whether a respondent provides—incorrectly—both an amount (or continuous report) and brackets. Imputed values were also checked against their underlying brackets, and the distribution of the imputed responses were compared with that of continuous responses. The consistency of the HRS asset variables across time was also tested.

As is the case with any large scale data generation efforts, the HRS 2004 Core Early Imputation Release data are not error free. If you find any problems with the data, please report to HRS via email at hrsquest@isr.umich.edu.

h2004idd.doc