

Master ID File Documentation

December 2009

The purpose of the HRS Master ID file is to simplify the process of merging HRS files. This document begins with a brief description of the Master ID file, includes a table of the variables contained in the Master ID file, and concludes with an example for merging HRS core data and Tracker file data by using the Master ID file.

The HRS uses two ID variables in combination to uniquely identify individuals: HHID and PN. Ideally, these would be fixed at entry into the study and never need to be changed. They would be identical for a single individual in every wave and in the Tracker file. In practice, a small but non-negligible number of individuals have had changes in their HHID or PN variables. The largest number of such changes happened almost immediately when a number of households from 1992 were transferred to the new AHEAD study in 1993 and given different IDs (referred to as the “overlap” cases in the Tracker file documentation). Other changes are documented in the Tracker file documentation or in the wave-level documentation (data descriptions).

Due to the changes in a small number of HHIDs and PNs, merging by HHID and PN across Core waves of data, or between the Tracker file and Core waves, typically produces a small number of cases that do not merge successfully. To simplify the process of merging, we have created a new Master ID file. This file mirrors the Tracker file in that it contains one record for every individual who was ever eligible for an HRS interview, and has the same number of observations with the same HHID and PN for each individual as in the tracker file. It then contains wave-specific HHID, PN, and SUBHH identification variables for every wave of HRS in Final Core release data, for every observation appearing in that wave. These wave-specific IDs match exactly the IDs as they appear in the current core release version of each wave.

Below is a list and description of the variables included in the Master ID file. All of the ID variables are in character format. For every wave specific variable, there will be a blank line in the data if a respondent did not complete a Core interview for that particular wave of Core public release data. Conversely, when a respondent did complete a core interview, the line for a given respondent will contain the value found in the Core public release data.

Variables in the Master ID File

Varname	Type	Length	Label	N
HHID	Char		6 HOUSEHOLD IDENTIFIER	31022
PN	Char		3 PERSON NUMBER	31022
AHHID	Char		6 1992 HOUSEHOLD IDENTIFIER	12652
APN	Char		3 1992 PERSON NUMBER	12652
ASUBHH	Char		1 1992 SUB-HOUSEHOLD IDENTIFIER	12652
BHHID	Char		6 1993 HOUSEHOLD IDENTIFIER	8222
BPN	Char		3 1993 PERSON NUMBER	8222
BSUBHH	Char		1 1993 SUB-HOUSEHOLD IDENTIFIER	8222
CHHID	Char		6 1994 HOUSEHOLD IDENTIFIER	11420
CPN	Char		3 1994 PERSON NUMBER	11420
CSUBHH	Char		1 1994 SUB-HOUSEHOLD IDENTIFIER	11420
DHHID	Char		6 1995 HOUSEHOLD IDENTIFIER	7027
DPN	Char		3 1995 PERSON NUMBER	7027
DSUBHH	Char		1 1995 SUB-HOUSEHOLD IDENTIFIER	10964
EHHID	Char		6 1996 HOUSEHOLD IDENTIFIER	10964
EPN	Char		3 1996 PERSON NUMBER	10964
ESUBHH	Char		1 1996 SUB-HOUSEHOLD IDENTIFIER	21384
FHHID	Char		6 1998 HOUSEHOLD IDENTIFIER	21384
FPN	Char		3 1998 PERSON NUMBER	21384
FSUBHH	Char		1 1998 SUB-HOUSEHOLD IDENTIFIER	21384
GHHID	Char		6 2000 HOUSEHOLD IDENTIFIER	19579
GPN	Char		3 2000 PERSON NUMBER	19579
GSUBHH	Char		1 2000 SUB-HOUSEHOLD IDENTIFIER	19579
HHHID	Char		6 2002 HOUSEHOLD IDENTIFIER	18167
HPN	Char		3 2002 PERSON NUMBER	18167
HSUBHH	Char		1 2002 SUB-HOUSEHOLD INDENTIFIER	18167
JHHID	Char		6 2004 HOUSEHOLD IDENTIFIER	20129
JPN	Char		3 2004 PERSON NUMBER	20129
JSUBHH	Char		1 2004 SUB-HOUSEHOLD INDENTIFIER	20129
KHHID	Char		6 2006 HOUSEHOLD IDENTIFIER	18469
KPN	Char		3 2006 PERSON NUMBER	18469
KSUBHH	Char		1 2006 SUB-HOUSEHOLD IDENTIFIER	18469
LHHID	Char		6 2008 HOUSEHOLD IDENTIFIER	17217
LPN	Char		3 2008 PERSON NUMBER	17217
LSUBHH	Char		1 2008 SUB-HOUSEHOLD IDENTIFIER	17217

Example: Merge using the Master ID file, HRS 1992 Core data, and the Tracker file

The Master ID file translates the HRS 1992 Core IDs to a common ID used in the tracker file; the following example picks up the "overlap" cases from the Tracker file.

*First, rename hhid and pn in HRS92 core data ;

```
data health;
set in.health;
rename hhid=ahhid;
rename pn=apn;
run;
```

*Next, sort both files (Master Id File and HRS92) by ahhid and apn ;

```
proc sort data=health;    /*the file health is HRS92 core data in this example ;
by ahhid apn;
run;
```

```
proc sort data=in.MasterIdFile out=MasterIdFile;
by ahhid apn;
run;
```

*Next, merge HRS92 Core and the Master ID file together;

```
data xwv_health;
merge MasterIdFile health(in=y);
by ahhid apn;
if y;
run;
```

*Next merge to the Tracker file and the HRS92 (with the new Master ID file IDs):

```
Sort both data sets by hhid pn;
proc sort data=xwv_health;
by hhid pn;
run;
```

```
Proc sort data=in.trk2004 out=trk2004;
by hhid pn;
run;
```

```
data health_trkr;
merge xwv_health(in=y) trk2004;
```

```
by hhid pn;  
if y;  
run;
```