HEALTH AND RETIREMENT STUDY

Tracker 2006
Version 1.0
October 2007

TABLE OF CONTENTS

Data Description and Usage

1. Introduction

The HRS tracker file is created to facilitate the use of HRS data within and across waves. The file contains one record for every person who was ever eligible to be interviewed in any wave. Each record contains basic demographic information, interview status, and if, when and how an interview was conducted in each wave. Also included are cross-sectional weights and information on inter-respondent relationships, which are vital to almost all substantive analyses of the HRS data.

The current version of the tracker file (Tracker 2006, Version 1.0) covers all types of interviews (core, exit, and post-exit) through HRS 2006. It contains 30,889 records and 284 variables. Structurally, Tracker 2006 Version 1.0 retains all the information contained in Tracker 2004, Version 2.0.

This document, complementary to the tracker file codebook, consists of six sections. Section 2 describes the basic structure of Tracker 2006. Sections 3, 4, 5 and 6 discuss various issues related to the variables in the tracker. Section 7 provides some general instructions on merging the tracker file with other HRS data.

To report any problems regarding the tracker file or this document, please contact us via email at hrsquest@isr.umich.edu.

2. The Structure of Tracker 2006

2A. Variable Listing and Description

The 284 variables in Tracker 2006 are listed in Table 1 according to the order in which they appear in the tracker data and the codebook. First on the list are two primary identifiers, household identifier (HHID) and person number (PN), which together uniquely identify each HRS panel member. Following the two identifiers is a set of "constants" or permanent characteristics that do not vary across waves. These variables either provide some basic demographic information for a person, or describe his or her study membership.

However, it is important to note that in the HRS 2006 core interview the question wording and format for the self assessed race questions were changed. In short, respondents were able to identify themselves as part of more than one racial category and every respondent was asked. After the multiple mention questions about race, those who gave more than one answer were asked, "Do you consider yourself primarily…(first mention, second mention, etc.)? In the 2006 Tracker file (Version 1.0), the 2006 core data were only used to update race if it was missing data in the Tracker file for a given respondent. In cases where more than one race was mentioned, the follow-up question about the respondent's primary race was used.

After the primary identifiers and the constants are several supplemental variables that are related to the National Death Index (NDI) and the HRS Ancillary studies, including Off-Year Mail Surveys and Internet surveys.

The final group includes ten sets of wave-specific variables, with one set representing each of the ten HRS survey waves to date. The waves are differentiated by the leading letter in variable names, from A to K (see Table 2). These wave-specific variables document whether a panel member was in a particular sample/study, whether he or she provided an interview in a particular wave, and, if yes, when and how the interview was conducted. When there are two panel members in a sub-household, the wave-specific variables also describe the relationship between

them in the sub-household. Moreover, the variables include cross-sectional weights
that are often needed for both household-level and respondent-level analyses.

Table 1. Variables in the 2006 Tracker (Version 1.0): Listing and Description
```
----------------------------------------------------------------------
Name            Type  Length   Label
--------        ----  ------   -------------------------------------------
```

Group A. Primary Identifiers and Permanent Characteristics

```
HHID            Char    6       Household Identifier
PN              Char    3       Person Number
BIRTHMO         Num     2       Birthdate: Month
BIRTHYR         Num     4       Birthdate: Year
DEGREE          Num     1       Highest Degree of Education
FIRSTIW         Num     1       First Interview: Study Year
GENDER          Num     1       Gender
HISPANIC        Num     1       Hispanic Type
IMMGYEAR        Num     4       Year Immigrated to the U.S.
OVHHID          Char    6       Overlap Case: Old HHID
OVPN            Char    3       Overlap Case: Old PN
OVRESULT        Num     1       Overlap Case: Result Code
RACE            Num     1       Race/Ethnicity
SCHLYRS         Num     2       Number of Years in School
SECU            Num     1       Sampling Error Computation Unit
STRATUM         Num     2       Stratum ID
STUDY           Num     2       Study Membership
USBORN          Num     1       Born in the U.S.
WTCOHORT        Num     2       Birth Cohort Used for Creating Weights
VERSION         Num     1       Tracker File Version Number
```

Group B. Supplemental Variables

```
Mail99          Num     2       1999 Mailout Pilot Survey Status
CAMS01          Num     2       2001 CAMS Sample Status
HUMS01          Num     2       2001 Human Capital Mail Survey Status
CAMS03          Num     2       2003 CAMS Sample Status
DIAB03          Num     2       2003 Diabetes Sample Status
CAMS05          Num     2       2005 CAMS Sample Status
PDS05           Num     2       2005 PDS Sample Status
INTRNT03        NUM     1       2003 Internet Survey Sample Status
PENFLAG         Num     1       HRS Pension Data Flag
ADAMS1          Num     1       ADAMS1 Sample Status
NSCORE          Num     3       NDI Match Score
NMONTH          Num     2       NDI Month of Death
NYEAR           Num     4       NDI Year of Death
IMMGYEAR        Num     4       Year Immigrated to the U.S.
```

Group C. Wave-Specific Variables

```
xALIVE          Num     1       Wave X Vital Status
xCORES          Num     1       Wave X Co-residence Status
xCOUPID         Char    6       Wave X Wave Household
xCOUPLE         Num     1       Wave X Whether Coupled (Married or Partnered)
xFAMR           Num     1       Wave X Family Respondent
xFINR           Num     1       Wave X Financial Respondent
xINSAMP         Num     1       Wave X Sample Status
xIWWAVE         Num     1       Wave X Whether Interview Status
xIWLANG         Num     1       Wave X Interview Language
xIWMODE         Num     1       Wave X Interview Mode
xIWMONTH        Num     2       Wave X Interview Month
```

```
xIWTYPE       Num    2        Wave X Interview Type
xIWYEAR       Num    4        Wave X Interview Year
xMARST        Num    1        Wave X Marital Status
xNEWSP        Num    1        Wave X New Spouse-Partner Flag
xNURSHM       Num    1        Wave X Nursing Home Status
xPPN          Char   3        Wave X Spouse-Partner Person Number
xPROXY        Num    2        Wave X Proxy Status
xRESCODE      Num    2,4      Wave X Result Code
xSUBHH        Char   1        Wave X Sub-Household Identifier
xSUBHHIW      Num    1        Wave X Whether SUBHH Interviewed
xPENFLAG      Num    1        Wave X Whether Pension Data is Available
xWGTHH        Num    8        Wave X Weight: Household-Level
xWGTR         Num    8        Wave X Weight: Respondent-Level
xWGTRNH       Num    8        Wave X Weight: Nursing Home Resident
xWHY0WGT      Num    1        Wave X Why Zero Weight
xWHY0HWT      Num    1        Wave X Why Zero HH Weight
xWHY0RWT      Num    1        Wave X Why Zero R Weight
xAGE          Num    3        Age at Wave X Interview
xNDIFLAG      Num    1        NDI Wave X Alive or Deceased Flag
-----------------------------------------------------------------------
```
Note: xCORES and xMARST are available only for HRS 2004 and 2006. xWGTRNH is available only for HRS 2000 and 2002 (e.g., x = G or H). xWHY0HWT and xWHY0RWT are available only for HRS 2006.

2B. Naming Convention for the Wave-Specific Variables

The letters, A, B, C, D, E, F, G, H, J, and K are used as indicators of the HRS waves. Each wave-specific variable is named with a leading wave indicator according to the rule listed in Table 2.

All of the wave-specific variables share the same naming structures across waves (i.e., ANEWSP, BNEWSP, CNEWSP, DNEWSP, ENEWSP, FNEWSP, GNEWSP, HNEWSP, JNEWSP and KNEWSP are, respectively, new spouse or partner flags for HRS 1992, AHEAD 1993, HRS 1994, AHEAD 1995, HRS 1996, HRS 1998, HRS 2000, HRS 2002, HRS 2004, and HRS 2006). This feature has two advantages. First, it allows users to identify and link easily all the wave-specific variables. Second, it gives us the ability to name the variables specific to a future HRS data wave in a predictable manner (e.g., the variable for new spouse or partner flag for HRS 2008 will be named as LNEWSP.)

Table 2. HRS Data Waves and Wave Indicators

```
----------------------------------------------------
HRS
Data Wave                       Wave Indicator
----------                      --------------------------
HRS   1992                          A
AHEAD 1993                          B
HRS   1994                          C
AHEAD 1995                          D
HRS   1996                          E
HRS   1998                          F
HRS   2000                          G
HRS   2002                          H
HRS   2004                          J
HRS   2006                          K
----------------------------------------------------
```

2C. Character versus Numeric

Following a protocol used in all of the recent HRS data releases, the ID variables are stored in character format. In Tracker 2006, the ID variables include HHID, PN, OVHHID (Inter-Study or Inter-Respondent Overlap HHID), OVPN (Inter-study or Inter-Respondent Overlap PN), xPPN (Spouse or Partner Person Number in Wave x), xSUBHH (Sub-Household ID in Wave x), and xCOUPID (Wave-Household ID in Wave X).

Non-ID variables are numeric.

3. Permanent Characteristics

3A. Birth Dates, Race, Hispanicity, Gender, Years of School, Educational Degree, and Immigration Status

Basic demographic variables for a respondent may come from a variety of sources. Most often the information was obtained initially through a respondent's answers when he or she was first interviewed. In other cases, the information may come from a spouse or partner or some other knowledgeable person if the sample person was not interviewed. In some cases, it may have been revised, either by a spouse or partner in an exit interview, or by HRS staff based on investigations into conflicting information for the respondent. Efforts have been made to resolve all such conflicts in a reasonable way, but it is possible that users may find some differences between the variables in the tracker and other HRS data. The tracker data should be considered definitive.

3B. First Interview

Variable FIRSTIW contains the information on the first time a sample member actually provided an interview (either core or exit, self or proxy). It differs from a sample member's entry cohort or study membership in that a sample member entering a given entry cohort or study might not have been interviewed the first time he or she was eligible for an interview.

As of HRS 2006, there are 412 sample members who have never provided any interviews (self or proxy, core or exit); 28 of the 412 are new spouses or partners.

3C. Overlaps

Overlaps refer to cases that have multiple IDs and require special handling in constructing longitudinal files and in merging tracker to wave-specific files. The current HHID and PN reflect the current status of the case. Overlap cases also have a former HHID and PN in some previous wave, and these are given in the variables OVHHD and OVPN.

There are two basic types of overlaps in the HRS data. First, there were a number of original HRS 1992 (Wave 1) households eligible to be either an HRS or AHEAD household. An interview was attempted for all of them in HRS 1992. Afterwards, a random sub-sampling was performed, with 60% of the cases staying in HRS and the remaining going to AHEAD. We refer to the 134 cases transferred to AHEAD as HRS inter-study overlap. Among those 134 cases assigned to AHEAD, 106 were actually interviewed in AHEAD, and 25 were not.

Second, there are two cases (as of this version of the tracker) belonging to what we call household merge overlaps, and which result from intermarriage among respondents who entered the study in separate households. In one case, an AHEAD married sample member (OVHHHID='205906' and OVPN='010') became widowed from his original spouse (HHID='205906' and PN='020'), and married another AHEAD sample member (HHID='205864' and PN='010') in HRS 1998, gaining a new identity as HHID='205864' and PN='011'. In the other, an unmarried AHEAD sample member (OVHHID='205399' and OVPN='010') married another AHEAD sample member (HHID='208545' and PN='020') in HRS 2000, gaining a new identity as HHID='208545' and PN='010'.

Instructions on how to deal with the various types of overlaps when merging the tracker file with the HRS core data are provided in Section 7.

4. Ancillary Studies and Supplemental Variables

4A. 1999 Pilot Mail Survey

In 1999, 2001, 2003, and 2005, questionnaires were mailed to subsamples of the HRS. The 1999 mail survey was a pilot study and included questions contained in the core survey. The variable MAIL99 indicates whether an HRS respondent was selected for the 1999 mail survey, and if selected, whether the questionnaire was returned.

4B. 2001 Mail Surveys (CAMS01 and HUMS01)

The 2001 mail survey consisted of two different questionnaires mailed to separate sub samples of HRS respondents. One version of the 2001 mail survey asked respondents about their consumption (household expenditures), time spent on activities, and prescription drugs. The survey is referred to as the 2001 Consumption and Activities Mail Survey (CAMS). The HRS data file [CAMS01_R] and documentation can be found on the HRS web site. The variable CAMS01 (in previous releases of the tracker this variable was called MBSTATUS) indicates whether a respondent was selected to receive a 2001 CAMS, and if selected, whether the questionnaire was returned. The second version of the 2001 mail survey, the Human Capital Mail Survey (HUMS) asked respondents about college expenses (human capital investments) associated with children who had attended college. The HRS data file [HUMS01_C] and documentation can be found on the HRS web site. The variable HUMS01 (in previous releases of the tracker this variable was called MCSTATUS) indicates whether a respondent was selected to participate in the 2001 Human Capital Mail Survey, and if selected, whether the questionnaire was returned.

4C. 2003 Mail Surveys (CAMS03 and Diabetes03)

The 2003 CAMS asked respondents about their consumption (household expenditures), and time spent on activities. The pool for the sub sample consisted of respondents who were alive, who had been selected for the 2001 CAMS, and not part of a household where their participation in another ongoing HRS mail survey was not complete. From this pool, if the respondent was eligible for CAMS but not eligible for the other 2003 mail survey, all eligible respondents were selected. If the respondent was eligible for CAMS and also eligible for the other 2003 mail sub sample, and they were part of a coupled household, all eligible respondents were selected. If the respondent was eligible for CAMS and also eligible for the other 2003 mail sub sample, and they were an uncoupled household, half of the eligible respondents were selected randomly for CAMS, and half for the other 2003 mail survey. The variable CAMS03 (in previous releases of the tracker this variable was called MDSTATUS) indicates whether a respondent was selected to participate in the 2003 Consumption and Activities Mail Survey, and if selected, whether the questionnaire was returned. There was no Human Capital Mail Survey (HUMS) in 2003.

The HRS 2003 Diabetes Study asked respondents about aspects of treatment and self-management of diabetes. The HRS 2003 Diabetes Study also collected a clinical biomarker of glucose control: glycosylated hemoglobin, or HbA1c. There were 3,194 interviewed respondents in the 2002 HRS who reported a diagnosis of diabetes (including respondents whose interviews were given by proxies). Of these, 680 were excluded from the 2003 Diabetes Study because of their participation in the Consumption and Activities Mail Survey (CAMS). This exclusion is random, with the slight exception that proxy cases from 2000 were ineligible for CAMS but eligible for the Diabetes Study, so they are represented at slightly higher rates prior to weighting adjustments. Of the 2,514 eligible 2002 participants, 133 were subsequently determined to have died prior to the October 2003 start of the Diabetes Study, and so are ineligible for inclusion in the sample. Of the 2,381 remaining eligible cases, 1,901 returned questionnaires, for a response rate of 79.8%. A total of 1,233 valid blood spots were returned for HbA1c assays. That is 64.9% of those who returned questionnaires, and 51.8% of all eligible cases. The variable DIAB03 in the tracker data indicates whether a respondent was selected to participate in the 2003 Diabetes Study (n=2,514), and if selected, whether the questionnaire was returned (n=4,415) or had died prior to the start of the study

(n=129). The variable is further broken down into respondents who returned a questionnaire with the HbA1c blood kit (n=1,233) and those who did not return the blood kit (n=668).

4D. 2005 Mail Surveys (CAMS05 and PDS05)

The 2005 CAMS mail survey asked respondents about their consumption (household expenditures), and time spent on activities. The pool for the sub sample consisted of respondents who were alive, who had been selected for the 2003 CAMS, and not part of a household where their participation in another ongoing HRS survey was not complete. From this pool, if the respondent was eligible for CAMS but not eligible for the other 2005 mail sub sample, they were selected for the 2005 CAMS. If the respondent was eligible for CAMS and also eligible for the other 2005 mail sub sample, and they were part of a coupled household, both halves of the couple were selected for the 2005 CAMS. If the respondent was eligible for CAMS and also eligible for the other 2005 mail sub sample, and they were an uncoupled household, half of the eligible respondents were selected randomly for CAMS 2005, and half for the other 2005 mail sub sample. In short, the 2005 CAMS sample consists of everyone who had participated in the 2003 CAMS survey, their spouses and/or partners, and a random sub sample of HRS respondents who were not in some other mail survey conducted in 2005. In "coupled" households, two versions of the questionnaire were sent to the household. One contained questions about household consumption and time spent on various activities; the other asked only about activities. The respondent who had been designated for the 2003 CAMS was given the full set of questions (consumption and activities), whereas their spouse/partner was given the booklet that asked only about activities. The variable CAMS05 (in previous releases of the tracker this variable was called MESTATUS) indicates whether a respondent was selected to participate in the 2005 Consumption and Activities Mail Survey (n=8,171), and if selected, whether the questionnaire was returned (n=5,815). This variable also breaks down the difference between respondent and spouse eligibility, and whether or not the questionnaire was returned in both of those groups.

The 2005 Prescription Drug Study (PDS05) is the first wave of a two-wave mail survey designed to track changes in prescription drug utilization as Medicare Part D, the prescription drug benefit, 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, sources of information on Part D, and expectations of the impact of Part D on prescription drug cost, coverage, and health. The second wave, scheduled for 2007, will capture similar information post-implementation. The study sample is comprised of 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. The sample was drawn from respondents to the 2004 HRS Core Survey, including respondents for whom interviews were obtained by proxy. The new drug benefit under Medicare Part D is expected to have its greatest impact on persons without drug coverage, and persons of low income and wealth who may be eligible for subsidies ("extra help") through Social Security. Because such persons are a minority of the eligible cases, and have shown lower response rates to past self-administered mail questionnaires, the sample included over samples of these groups. There were 5,654 respondents in the 2004 HRS Core Survey who were eligible for participation in the Prescription Drug Survey. Of the eligible participants, 340 died prior to the October 2005 start of the first wave of the Prescription Drug Study and are ineligible for inclusion in the sample. The variable PDS05 in the tracker data indicates whether a respondent was selected to participate in the 2005 Prescription Drug Study (n=5,654), and if selected, whether the questionnaire was returned or a phone interview completed (n=4,684). The variable is further broken down into respondents who returned a questionnaire with the medication list (n=4,376) and those who did not return the medication list (n=308).

4E. ADAMS WAVE 1

The Aging, Demographics, and Memory Study (ADAMS) is a supplement to the Health and Retirement Study (HRS) funded by the National Institute on Aging (U01-AG009740) with the specific aim of conducting a population-based study of dementia. Data in ADAMS1 were collected between 2001 and 2005. The purpose was to gather additional information on respondents' cognitive status and assign a diagnosis of (1) dementia, (2) cognitive impairment but not demented (CIND), or (3) non-case. The data collected allow researchers to estimate the prevalence, predictors, and outcomes of dementia in the U. S. elderly population. A group of 1,770 HRS respondents, age 70 or older, was selected from the 2000 and 2002 waves based on the score on the self- or proxy-cognitive assessment measure. Among this group, we completed assessments with 856 respondents. Additional information about this study is available on the HRS website: http://hrsonline.isr.umich.edu/adams/

4F. NDI

In addition to the information on vital status obtained by HRS through tracking of respondents (see xAlive), the HRS seeks matches to the National Death Index for persons who are reported as deceased or who are not known to be alive through contact during tracking. The tracker file contains information derived from finder files submitted to the National Center for Health Statistics (NCHS), in 1995, 2000, 2002, and 2004. For all submitted cases that were flagged as valid by NCHS and verified by HRS staff, the Tracker file contains year and month of death, match score, and an alive/deceased flag (NYEAR, NMONTH, NSCORE, xNDIFLAG).

5. Wave-Specific Variables

Roughly speaking, the wave-specific variables in Tracker 2006 may be divided into three groups: those indicating a sample member's interview, sample, or study status in a given wave; those indicating the relationships between the sample members in a household in a given wave; and those describing other wave-specific information.

5A. Interview, Sample or Study Status

5A1. xIWTYPE, xIWWAVE, and xINSAMP

In most of the HRS data waves, a sample member could fall into one of the seven categories: (1) provided core interview; (2) eligible but did not provide core interview; (3) provided exit interview; (4) eligible but did not provide exit interview; (5) provided post-exit interview; (6) eligible but did not provide post-exit interview; and (7) not eligible in the wave. This information is captured in variable xIWTYPE, the first variable a user might want to turn to when checking a sample member's status in any given wave. By design, sample members in HRS 1992 or AHEAD 1993 can appear only in categories (1) and (2), and sample members in HRS 1994 or AHEAD 1995 or HRS 1996 will not be in categories (5) and (6).

Closely related to xIWTYPE is xIWWAVE, a flag variable which collapses the interview type information into a single YES (1) or NO (0) dichotomy: xIWWAVE is equal to 1 if a sample member provided any type of interview; otherwise it is equal to 0 if the sample member was eligible but did not provide an interview. (It is equal to 99 if he or she was not eligible at wave x.)

While both xIWTYPE and xIWWAVE indicate whether an interview was obtained for a sample member in a given wave, xINSAMP indicates whether a sample member was supposed to be interviewed in the wave, and, if not, why. By design, some sample members should not be in certain data waves. Examples include (1) that the (non-overlap) HRS members should not be in an AHEAD data wave, and vice versa, (2) that the CODA or WB respondents should not be in any data wave prior to HRS 1998, and (3) the EBB respondents should not be in any wave of data prior to HRS 2004.

Tracker 2006 defines eight types of sample status: (1) in the sample; (2) belonging to an entry cohort not interviewed this wave; (3) not yet entered, but belonging to

the cohort interviewed in this wave; (4) members of HRS-AHEAD overlap households in which no one was interviewed at AHEAD wave 1, and who were dropped from the sample (i.e., were not contacted) at subsequent waves; (5) no longer in the sample because a complete exit (and, if necessary, a post-exit interview) has been obtained; (6) permanently dropped from the sample per request of the sample person, his/her spouse/partner, or other gatekeeper; (7) a deceased sample member for whom the field staff was unable to find an informant eligible to do an exit or post-exit interview at a previous wave; and (8) permanently dropped from the sample for any other reason. Which of these categories to include in non-response analyses will depend on the specific purpose, but in general, those in categories (4), (6), (7), and (8) should be regarded as eligible and therefore included in the denominator when calculating non-response and attrition rates.

The relationship among xIWTYPE, xIWWAVE, and xINSAMP may be better seen in Table 3, where the concepts "in the wave" and "in the sample" are mathematically delineated.

5A2. xRESCODE

Part of the HRS tracking procedure is to note reasons for non-interviews. For each sample member eligible to be interviewed in a wave, the variable xRESCODE indicates the final disposition of the case. Of particular importance is the distinction between a non-interview for the current wave only and a permanent removal from the study. Each can generate one of several possible result codes indicating a specific reason.

Each sample member eligible to be interviewed in a wave will have a non-missing value on xRESCODE. The sample members of wave x with missing values on xRESCODE, therefore, will be those with a value of 2, 3, 4, 5, 6, 7, or 8 on variable xINSAMP.

5A3. xSUBHH, and xSUBHHIW

In HRS, sub-households, which are split from original households and identified uniquely by the same household ID (HHID) AND sub-household ID (xSUBHH), are usually regarded as the unit of household-level analysis. An original household may split off primarily for one of the two reasons: dissolved marriage or partnership, or death in the household. The sub-household ID for a deceased sample member eligible in a given wave is normally "3" or "4," which values are never assigned to any sample member still alive.

xSUBHHIW describes whether a unique sub-household is interviewed in a given wave. A sub-household is considered interviewed in a wave (xSUBHHIW=1) if any of the members in the sub-household provided an interview in the wave: whether the interview was a core, an exit, or a post-exit; and whether the interview was complete or partial. It is possible, therefore, for a sample member himself or herself not interviewed in a wave to have a value of 1 on xSUBHHIW. On the other hand, when no members in a sub-household were part of the sample in a wave (xINSAMP>1), xSUBHHIW will be missing for all the respondents in the sub-household.

Table 3. The Relationship among xIWTYPE, xIWWAVE, and xINSAMP
----

xIWTYPE	xIWWAVE	xINSAMP
Core Interview Provided (1)		
Exit Interview Provided (11)	Yes (1)	
Post-Exit Interview Provided (21)		

```
------------------|-----------------------|       Yes (1)
 Core Interview    |                       |
  Not Provided (5) |                       |
                   |                       |
 Exit Interview    |                       |
  Not Provided (15)|       No (0)          |
                   |                       |
 Post-Exit         |                       |
  Interview        |                       |
  Not Provided (25)|                       |
                   |                       |
------------------|                       |--------------------------
                   |                       |
 Not in the        |                       |      No (2,3,4,5,6,7,8)
  Wave (99)        |                       |
 ---------------------------------------------------------------------
```

5B. Inter-Respondent Relationship

5B1. xCOUPLE, xCOUPID, xCORES, and xMARST

When an eligible sample member in a given wave is coupled (married or living with a partner as if married) the variable xCOUPLE is equal to 1. Otherwise, it is 5. If one member of a couple is reported dead in wave x, the deceased respondent is assigned a different xSUBHH, with xCOUPLE = 1, while the surviving spouse or partner will be assigned xCOUPLE = 5 (unless the surviving spouse or partner has remarried or repartnered).

Labeled as "Wave Household ID," xCOUPID is created to link the sample members who are married or partnered in a given wave, or the sample members who were married or partnered in a prior wave but are no longer spouses or partners due to the death in the household. Specifically, for sample members with xCOUPLE = 1, xCOUPID is a concatenation of the person numbers of the sub-household members in ascending order. XCOUPID is also useful when a death has occurred. When one member of a couple is reported dead at wave x, and the partnership or marriage had not ended prior to the death, the deceased respondent is assigned xCOUPLE=1 and xCOUPID is established in the same way as if the couple had survived to facilitate linkage to the prior wave spouse or partner. The surviving spouse/partner gets an xCOUPLE=5 and an xCOUPID formed of his/her own PN followed by 000, unless he or she has formed a new couple.

xCORES indicates whether a respondent lives with a spouse or partner. xMARST indicates whether a respondent is married, divorced/separated, widowed, or never married. In this version of the tracker, these two variables are available only for HRS 2004 and HRS 2006. It is important to note that xMARST may be different from the marital status assigned in the 2006 Core Final release. The marital status variable in the Tracker file (xMARST) was constructed by looking at several different sources of data, whereas the marital status variable in the 2006 Core was not changed, in order to preserve the flow through the instrument for respondents.

The basic procedure for generating xCORES and xMARST is as follows. We first use core interview data to determine the marital status for those who provided core interviews at wave x, xMARST. This variable is then modified to achieve cross-wave consistencies based on one of the following: a) spouse/partner relationship; b) examination of cross sectional and longitudinal data, and c) extrapolation based on the data and sometimes common sense.

The "common sense" method stipulates that if a respondent was "never married" at wave t, then s/he was not married at any wave prior to t. Conversely, if a respondent was married, separated, or divorced at wave t, then s/he could not be "never married" at any wave after t. The "spouse/ partner relationship" method assumes that in a coupled household at wave t, if one respondent was married, the

other should be married at the wave as well; and if one respondent was not married, the other should not be married either. The "extrapolation" method allows one to use a respondent's marital status at wave (t-1) to "approximate" his or her status at wave t when the latter information is not available.

For respondents in single-respondent households, we classify them as "living alone" (i.e., xCORES = 3) if they were married (i.e., xMARST = 1), and classify them as "living arrangements unknown" (i.e., xCORES = 4) if they were divorced, separated, widowed, never married, or their marital status was unknown (i.e., xMARST > 1.)

For respondents in two-respondent households, we classify them as "married spouses living together" (i.e., xCORES = 1) if both respondents were married, and classify them as "living with someone regardless of marital status" (i.e., xCORES = 2) if both respondents were not married.

5B2. xFAMR and xFINR

HRS respondents provide information at an individual level and, for some questions, at a household level. In particular, many questions about family or about income and wealth are asked of only one respondent in a household. The respondent providing family information in a given wave is called the family respondent (or family R) in that wave, whereas the respondent providing financial data in a wave is called the financial respondent (or financial R) in the wave. The family R and financial R designations are reflected in variables xFAMR and xFINR, respectively. For a single-person household, the respondent is designated to be both family R and financial R.

5B3. xNEWSP

When an original sample member takes a new spouse or partner following his or her baseline interview, an attempt is made to interview the new person. New sample members entering in this way are indicated with xNEWSP=1 for the wave in which an interview was first attempted, and XNEWSP=0 for that person in all other waves. There are no new spouses or partners in baseline waves. Thus, no sample members in HRS 1992 or AHEAD 1993 are considered new spouses or partners, and similarly, no sample members from the CODA, WB (1998) and EBB (2006) entry cohorts are considered new spouses or partners.

5B4. xPPN

xPPN identifies the person number of the partner or spouse of a sample member if the sample member is part of a couple. For all the sample members still alive in a given wave, a concatenation of their own person numbers (PN) with xPPN in an appropriate order would generate xCOUPID. For deceased sample members, xPPN gives the person numbers of their spouses or partners in the last wave during which they were alive.

5B5. xPROXY

The HRS makes every effort to obtain core interviews directly with respondents themselves. In cases where the respondent is unavailable, unable, or unwilling to do the interview, a proxy is sought to provide an interview on behalf of the respondent (unless the respondent refuses to allow a proxy). Exit and post-exit interviews, on the other hand, are always conducted with proxies. The variable xPROXY indicates self-interviews, and classifies proxies into two types: spouse/partner, and any others.

5C. Other Wave-Specific Variables

5C1. xALIVE

This variable contains information about vital status based on HRS tracking information. It provides five categories: (1) alive at the wave (xALIVE=1); (2) presumed alive as of the wave (xALIVE=2); (3) death reported in the wave

(xALIVE=5); (4) death reported in a prior wave (xALIVE=6), and (5) vital status unknown (xALIVE=9). Persons for whom no report of death has been made are classified on the following principles. If the person gave an interview in the wave or was contacted directly by an interviewer during the wave, or was reported to be alive by a spouse or partner, they are considered definitely alive (xALIVE=1). Similarly, anyone who met those criteria at a subsequent wave will be classified as xALIVE=1 at all prior waves. If contact of a less definitive type was made, and there was no report of death, we classify the respondent as presumed alive (xALIVE=2). In cases where no contact was made that could have shed light on vital status, we classify the person as unknown (xALIVE=9). This will always be the case for persons permanently removed from the study at a prior wave while still living.

No National Death Index information is used in constructing xALIVE. Both sources of information should be used to classify vital status according to the goals of any specific analysis.

5C2. xIWLANG, xIWMODE, xIWMONTH, xIWYEAR , and xAGE

These variables indicate the language (English or Spanish) in which the interview was conducted, the mode (in-person or telephone), the month and year in which the interview began and the age the respondent was at a specific interview. The variable xAge was not included in the tracker file prior to the 2004 wave of data.

5C3. xNURSHM

This variable indicates whether the respondent was in a nursing home at the time of the respondent's interview or the spouse/partner's interview. This can be reported either by the respondent or a proxy during the respondent's interview, or by a spouse or partner describing the respondent's whereabouts. There is no information available for an uncoupled respondent who was not interviewed by self or proxy, or for coupled respondents when neither of them gave an interview. The HRS cross-section sample weights (*xWGTR*) are constructed to match the non-institutionalized population. Nursing home residents in a wave therefore receive a zero weight at that wave. By design, HRS does not interview baseline respondents who are nursing home residents, unless they are part of a couple whose other member is part of the non-institutionalized population. The exception to this was for the 1993 baseline wave of AHEAD, where the decision was that interviews would not be attempted for nursing home residents, even if they were part of a couple. We did decide to include these AHEAD nursing home spouses or partners in subsequent waves, however, and attempted to interview them.

5C4. Changes to GNURSHM, HNURSHM, JNURSHM and KNURSHM

In the process of finalizing the 2004 sample weights, we discovered some incorrect classifications in nursing home status for that wave. The problem was first detected when we noticed that 15 Early Baby Boomer (EBB) respondents were recorded as residing in nursing homes. A review of those cases indicated that 14 of the 15 had been misclassified (either due to data entry error or misclassification of an assisted living facility as a nursing home). The remaining EBB respondent was in a hospice rather than a nursing home. This led to a broader review of the 2004 cases, and ultimately of 2002 and 2000 data.

Criteria used to identify 'suspicious' cases

We identified a set of key indicators that were highly predictive of misclassification in the 2004 wave. These include: 1) respondent type (self vs. proxy), 2) report of 'employee of institution' as a helper for ADLs and/or IADLs, and 3) prior wave nursing home status. We used this set of indicators to identify 'suspicious' cases in each of the 2000, 2002 and 2004 waves. Specifically, we flagged cases for review that were reported as residing in a nursing home in the current wave for review if the interview was conducted with a self-respondent, the R did not report receiving help from an employee of the institution, and the R was

not recorded as residing in a nursing home in the prior wave. In addition, for the 2002 and 2004 waves, cases were flagged for review if there was disagreement between the core report and the interviewer observation report on nursing home status. (No interviewer observations were available for 2000.)

Information used to reclassify cases

For all cases flagged for review, information was pulled from interview comments for nursing home-related questions in the core interview, as well as from interviewer call records. These comments were reviewed to make a determination as to whether the case should be reclassified. We also reviewed keystroke files for the flagged cases if the comments were inconclusive. We used the following decision rules for reclassifying cases from 'in nursing home' to 'not in nursing home':

> 1. The interviewer recorded a comment that indicated that the R was not residing in a nursing home or there was other evidence that the interviewer made a keying error (e.g., the interviewer tried to enter '0' in N115 and/or tried to back up and change the response in A023), or
>
> 2. The interviewer recorded a comment in an interview note or in the interviewer call records specifying that R lives in an assisted living facility, or
>
> 3. The call record included comments indicating that the R lives in a single family home, or
>
> 4. The call record included comments indicating that the R was currently working.

The following decision rules were used to reclassify cases from 'not in nursing home' to 'in nursing home':

> 1. R was recorded as residing in a nursing home in the Interviewer Observations, and
> 2. The interview comments or interviewer call record included comments indicating that R resides in a nursing home.

Respondents flagged as suspicious based on the above criteria for which neither the interview comments nor the interviewer call records contained evidence to contradict the Core assignment were left with their original assignment.

Change in 2006 questionnaire

Starting with the 2006 wave of HRS, we added an interviewer check at the nursing home question in order to reduce the number of entry errors and misclassifications. If the interviewer entered '1' at A028, indicating that the R is currently residing in a nursing home, a text box appeared informing the interviewer that s/he reported that R is currently residing in a nursing home. If the entry was correct, the interviewer could suppress the signal and go on with the interview. If the entry was incorrect, the interviewer was instructed to change the response in A028. This interviewer check substantially reduced the number of misclassifications.

Reclassification outcomes

HRS 2006:
9 cases were changed from in a Nursing Home to not in Nursing Home
1 case was changed from not in a Nursing Home to in a Nursing Home

HRS 2004:

77 cases were changed from in a Nursing Home to not in Nursing Home
4 cases were changed from not in a Nursing Home to in a Nursing Home

HRS 2002:
97 cases were changed from in a Nursing Home to not in a Nursing Home
7 cases were changed from not in a Nursing Home to in a Nursing Home

HRS 2000:
23 cases were changed from in a Nursing Home to not in a Nursing Home

It is important to note that the Core data for HRS 2000, HRS 2002, and HRS 2004 were not changed. Therefore, there will be mismatches between the core waves of data and the Tracker file with regard to nursing home residency. The Tracker file should be considered as definitive.

5D. *HRS Sample Weights*

5D1. xWGTHH, xWGTR, xWGTRNH, xWHY0WGT, xWHY0HWT, and xWHY0RWT

xWGTHH is the sampling weight for analysis at the household level. xWGTR is the sampling weight for analysis at the respondent level; respondents included are those living in the community. xWGTRNH is the sampling weight for analysis at the respondent level; respondents included are those living in nursing homes. In this version of the tracker, xWGTRNH is available only for HRS 2000 and 2002.

Since the HRS sample is a multi-stage probability sample of the United States, with oversamples of Blacks and Hispanics, unbiased estimates of population parameters require the use of these sampling weights. The household level and respondent level weights (xWGTHH and xWGTR) are post stratified to the March Current Population Survey for the year of data collection. The nursing home sampling weight (xWGTRNH) is designed to adjust for nursing home residents interviewed in a given wave (2000 or 2002) to represent the national population of nursing home residents by race, sex, and age. Only nursing home residents interviewed in the wave have a non-zero value for xWGTRNH.

The (community dwelling) respondent-level weight xWGTR is zero when a respondent is not cohort eligible in wave x, is residing in a nursing home, or is deceased. There are also a very few cases with zero xWGTR because their birth cohort was not ascertained. The variable xWHY0WGT summarizes which of these conditions is responsible for a zero respondent level weight for cases with a core interview in wave x (available for waves 1992 through 2002).

xWHY0HWT explains why a household has a zero weight and xWHY0RWT explains why a respondent has a zero weight. For the 2006 wave of data the variables are: KWHY0HWT for households and KWHY0RWT for respondents.

It is important to note that HRS weights are generated based on WTCOHORT, not BIRTHYR. WTCOHORT is constructed using a "first-mention" or "entry-wave" birth year measure. This information may not be completely consistent with BIRTHYR, which takes into account corrections by the HRS staff to the birth year measure constructed in WTCOHORT.

5D2. Sample Selection Indicators and Sample Weights for 2004 Interview Components: Physical Measure and "Leave-Behind"

In the 2004 wave, HRS added two new components to the Core interview. These included a set of physical measures (lung function, grip strength, walking speed) and a "leave-behind" questionnaire (a self-administered questionnaire that was left with the respondent at the end of the interview) on one of two topics: psychosocial measures and disability vignettes. Each of these components was administered to a

subsample of the total Core sample. Sample weights were developed for each component to account for differential probabilities of selection into the subsample and non-response to the specific component. The sample weights for each component are provided on the 2006 Tracker file, along with sample selection indicators. These variables include:

Physical Measures:
PHYMSR04– sample selection indicator for physical measures
JPMSELWT – selection weight for the physical measures subsample
JPMWGTR – final respondent weight for the physical measures subsample

Psychosocial Leave-Behind:
PSYSOC04 – sample selection indicator for the psychosocial questionnaire
JWGTR_PS – final respondent weight for the psychosocial subsample

Disability Leave-Behind:
DISAB04 – sample selection indicator for the disability questionnaire
JWGTR_DB – final respondent weight for the disability subsample

A general description of the physical measures and leave-behind components can be found in the Data Description for the 2004 Core file, and the questionnaires for each are included under "Other" Content Area for the 2004 wave on the Questionnaire page on the HRS website. More detailed information on the selection indicators and sample weights is provided in the following documents (links below):

Physical Measures:
http://hrsonline.isr.umich.edu/meta/tracker/desc/PMWeight2004_Description_public.pdf

Leave-Behind Questionnaires:
http://hrsonline.isr.umich.edu/meta/tracker/desc/LBWeights2004_Description_public.pdf

6. Important Data Changes in Version 1 of Tracker 2006

6a. Changes to HIWTYPE
The following six cases have been recoded from '21. Post-exit interview obtained' to '25. Post-exit interview not obtained':

HHID	PN
030732	010
050365	020
201622	010
202259	020
204997	020
205307	010

6b. Changes to PN
The PN of the following two cases have been changed from '011' to '020':

HHID
501496
502153

6c. Changes to xALIVE
A correction has been made to xALIVE variables in all waves except 1992 - AALIVE. The changes all involve respondents who were not interviewed in a given wave and are based on result codes.

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 to

- not attempt to identify respondents
- not transfer data to third parties except as specified
- not share your username and password
- include specified citations in work based on HRS data
- provide information to us about publications based on HRS data
- report apparent errors in the HRS data or documentation files
- notify us 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 hrsquest@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 2000 HRS Core (Final, Version 1.0) 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. Our new, as of 9/1/2002, URL is:

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: 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

9. Merging the Tracker with other HRS Data

When merging the tracker with other HRS data, you should remember that the tracker
contains more records than other data releases for any given wave. The first step
is, therefore, to subset the records from the tracker that are present in the other
data file. xIWTYPE provides necessary information for performing this task. In
addition, since the tracker contains only one record per sample member, those with
multiple identities (i.e., multiple "hhid" and "pn" combinations) across waves may
not match with other HRS data without some manipulations before merging. In fact,
how to deal with the respondents with multiple identities across waves proves to be
the most complicated task when integrating the tracker with other HRS data.

In what follows, we provide some coding examples for merging the tracker with
various HRS data, covering most of the issues that users may encounter when
performing similar merges of their own. For simplicity, we assume that all the
data, including the tracker and other HRS data, are stored in the same directory
(or SAS data library).

9A. HRS 1992 Merges

Merging the tracker with the HRS 1992 data is difficult because of the HRS-AHEAD
overlaps explained earlier. Two different strategies are available for working
with HRS 1992 data. First, one may consider the entire, original HRS sample as the
data source of the analysis. In this strategy, one wants to retain all the data
records as released in the HRS 1992 core data, that is, N=12,652. Alternatively,
one may consider only the non-overlap cases as the HRS respondents, leaving the HRS
overlaps to AHEAD. In this strategy, the information in the HRS 1992 core data for
the overlap cases will be dropped, and N=12,521.

While we have no intention to force users to take one strategy over the other, it
is suggested that the second strategy be used whenever the HRS 1992 data are
involved in a longitudinal analysis. The overlap cases are no longer treated as
part of the HRS sample in the later waves.

The following examples illustrate these two strategies when merging with dataset
HEALTH

9A1. Retaining the Entire, Original HRS Sample

(i) SAS Code

```
   data tracker;
     set dat.trk2006;
     if ovresult=1 then do;    /* Convert AHEAD IDs into Original HRS IDs */
       hhid=ovhhid;
       pn=ovpn;
     end;
     run;

   proc sort data=tracker;
     by hhid pn;
     run;

   proc sort data=dat.health out=health(keep=hhid pn v1);
     by hhid pn;
     run;

   data healtht;
     merge tracker health(in=inh);
```

17

```
      by hhid pn;
      if inh;
      run;


(ii) Stata Code

    use trk2006
    replace HHID=OVHHID if OVRESULT==1
    replace PN=OVPN if OVRESULT==1
    sort HHID PN
    save tracker, replace

    use health
    sort HHID PN

    merge HHID PN using tracker
    keep if _m==3
    save healtht, replace


(iii) SPSS Code

    GET FILE "c:/temp/health.sav".
    SORT CASES BY hhid pn.
    execute.
    save outfile='C:/temp/health2.sav'.

    GET FILE "C:\temp\trk2006.sav".
    * Convert AHEAD IDs into Original HRS IDs.
    if (ovresult =1)hhid=ovhhid.
    if (ovresult =1)pn=ovpn.
    SORT CASES BY hhid pn.
    execute.
    save outfile='C:/temp/tracker.sav'.

    MATCH FILES /FILE="C:/temp/health2.sav"
        /IN=h
        /FILE='C:\temp\tracker.sav'
        /BY hhid pn
      select if h=1.
       EXECUTE.

    save outfile='C:/temp/healtht.sav'.
```

9A2. Keeping Non-Overlaps Only

(i) SAS Code

```
    proc sort data=dat.trk2006 out=tracker;
      by hhid pn;
      run;

    proc sort data=dat.health out=health(keep=hhid pn v1);
      by hhid pn;
      run;

    data healtht;
      merge tracker(in=int) health(in=inh);
```

```
      by hhid pn;
      if ovresult=0 and inh;                 /* Keep only Non-Overlap Cases */
      run;
```

(ii) Stata Code

```
   use trk2006
   sort HHID PN
   save tracker, replace

   use health
   sort HHID PN

   merge HHID PN using tracker
   keep if _m==3 & OVRESULT==0         /* Keep Only Non-Overlap Cases */
   save healtht, replace
```

(iii) SPSS Code

```
   GET FILE "c:/temp/health.sav".
   SORT CASES BY hhid pn.
   execute.
   save outfile='C:/temp/health3.sav'.

   GET FILE "C:\temp\trk2006.sav".
   SORT CASES BY hhid pn.
   execute.
   save outfile='C:/temp/tracker2.sav'.

   MATCH FILES /FILE="C:/temp/health3.sav"
       /IN=h
       /FILE='C:\temp\tracker2.sav'
       /BY hhid pn.
   select if h=1 and ovresult=0.
   EXECUTE.

   save outfile='c:/temp/healtht.sav'.
```

9B. HRS 1998 Merges
In this example, we merge the tracker file with dataset h98a_r.

(i) SAS Code

```
   data tracker;
     set dat.trk2006;
     if ovresult=8 and hhid not eq "205864" then hhid=ovhhid;
     if ovresult=8 and hhid not eq "205864" then pn=ovpn;
     /* Convert AHEAD Inter-Respondent Overlap ID into Original */
     run;

   proc sort data=tracker;
     by hhid pn;
     run;

   proc sort data=dat.h98a_r out=h98a_r (keep=hhid pn f968 f970);
     by hhid pn;
     run;
```

```
   data h98a_rt;
     merge tracker h98a_r(in=inh);
     by hhid pn;
     if inh;
     run;
```

(ii) Stata Code

```
   use trk2006
   replace HHID=OVHHID if OVRESULT==8 & HHID~="205864"
   replace PN=OVPN if OVRESULT==8 & HHID~="205864"

   sort HHID PN
   save track, replace

   use h98a_r
   keep HHID PN F968 F970
   sort HHID PN

   merge HHID PN using track
   keep if _m==3
   save h98a_rt, replace
```

(iii) SPSS Code

```
   GET FILE "C:\temp\ trk2006.sav".

   * Convert AHEAD Inter-Respondent Overlap ID to original.
         if (ovresult =8) and (HHID <> '205864') hhid=ovhhid.
   if (ovresult =8) and (HHID <> '205864') pn=ovpn.
   execute.
   SORT CASES BY hhid pn.
   execute.
   save outfile='C:/temp/tracker.sav'.

   GET FILE "C:\temp\h98a_r.sav"/keep hhid pn f968 f970.
   SORT CASES BY hhid.
   execute.
   save outfile='C:\temp\h98a_r2.sav'.

   MATCH FILES /FILE="C:/temp/h98a_r2.sav"
    /IN=h
    /FILE='C:\temp\tracker.sav'
    /BY hhid pn.
   select if h=1.
   EXECUTE.
   save outfile='C:/temp/h98a_rt.sav'.
```