

HRS 2016 Sampling Weights

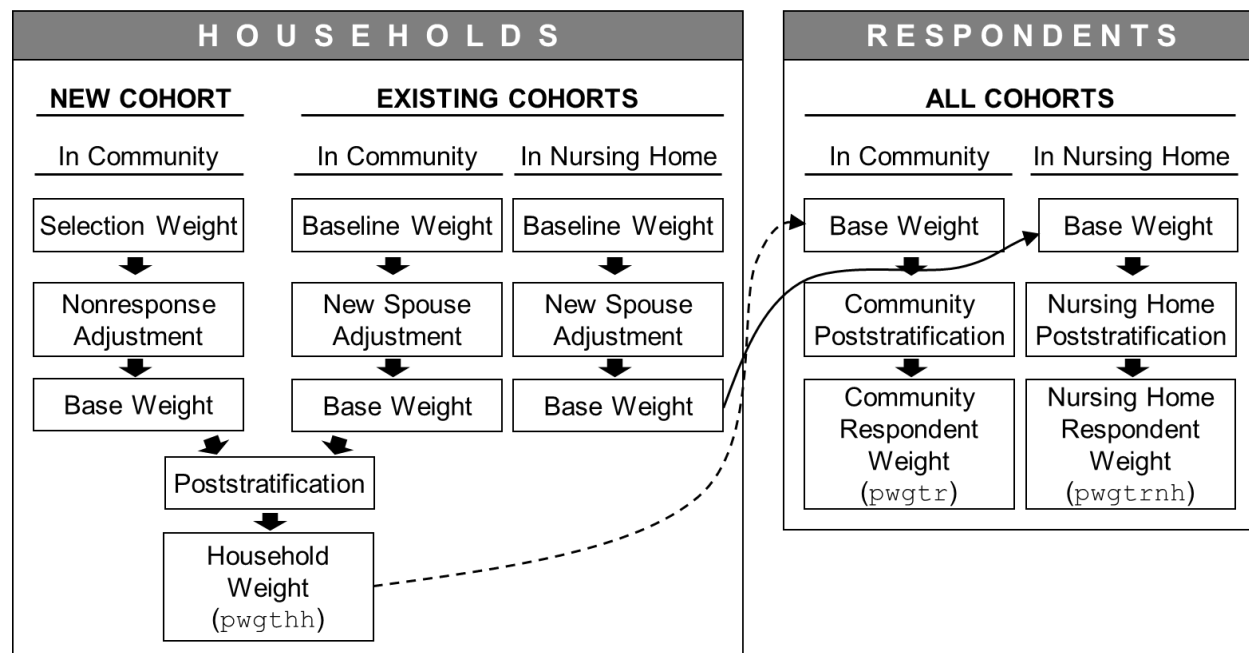
Sunghee Lee, Raphael Nishimura, Paul Burton, and Ryan McCammon

Introduction

The Health and Retirement Study (HRS) is a population-based longitudinal survey of the U.S. population aged 51 or older. The current design introduces a new age cohort every six years. In 2016, the Late Baby Boomer (LBB) cohort, born between 1960 and 1965, was newly recruited. This document describes the sampling procedure and creation of weights, contrasting in some places the 2016 recruitment (and implications for weighting) with that of previous cohorts.

To ensure population-level representation, HRS provides sampling weights to account for differential selection probabilities by race/ethnicity and birth cohort, and to correct for differential non-response. These weights are developed for households (or financial units) as well as for respondents. Although HRS does not sample from institutions at the time of recruitment, it does follow respondents who enter nursing homes or other institutions. The weighting process accounts for whether the residence is in the community or in a nursing home. The weight construction process is described in Figure 1.

Figure 1. Health and Retirement Study 2016 Weighting Steps



For a newly recruited cohort, HRS establishes a sample of households and interviews eligible individuals within it. The sampling frame includes addresses equating to housing units or households. It should be noted that the terms ‘households’ and ‘financial units’ have been used interchangeably in HRS data descriptions and documentation. However, precisely speaking, they are not the same. A household may have multiple financial units when the household contains more than one unrelated age-eligible person (i.e., financial unit). For most HRS-eligible households, this is not the case, and the terms financial units and households are, in practice, interchangeable. For households with more than one unrelated age-eligible person (i.e., more than one financial unit) one of these individuals is randomly selected. If the selected individual has a spouse, the spouse is automatically selected for HRS, even if he or she is not age-eligible. This report will use the term “households” for units used in sampling operations and the term “financial units” for units ascertained through screening work, as the number of potential financial units is unknown before screening interviews.

Weighting follows the fact that the sample is designed at the household level. Weights are first calculated at the financial unit level. For newly recruited cohorts, weights are computed for each financial unit, accounting for selection probabilities derived from the sample design, unit-level nonresponse as well as poststratification. The resulting weights are the baseline financial unit weights on which weighting in subsequent years will be based. By design, a newly introduced cohort is sampled only from the community. Thus, LBB households in HRS 2016 cannot enter the study from nursing homes. For prior cohorts, baseline financial unit weights that were established in the year corresponding to their introduction to the sample are adjusted for the new spouse status. That is, if there is a new spouse in a household and if the new spouse is in the same age cohort as the household, this household would have had more than one chance to be included in the sample retrospectively in the cohort introduction year. This is adjusted for in the process. If the existing cohort financial units are in the community, the new spouse adjusted weights will be poststratified and the resulting weights are made available in the data set (e.g., `pwgthh`). If the financial unit is in a nursing home, a financial unit weight of 0 is assigned under `pwgthh`. Regardless of cohort, the starting point of the respondent weights are poststratified financial unit weights for those in the community and base financial unit weights for those in nursing homes. At the respondent level, separate poststratification processes are applied based on the nursing home status, resulting in separate weight variables (e.g., `pwgtr` for community respondents; `pwgtrnh` for nursing home respondents).

Part 1 below is focused on the newly recruited LBB cohort in 2016. Section 1.1 describes computation of selection weights and Section 1.2 describes the nonresponse adjustment for the LBB cohort. Part 2 describes the poststratification adjustments for differential nonresponse as applied to all cohorts.

Part 1. Baseline Weights for the LBB cohort

1.1 Selection Weights

The sample for the HRS 2016 screening work came from two sources: 1) a stratified multi-stage sample specifically drawn for HRS 2016 using multi-stage area-probability sampling similar to previous cohorts; and 2) a list of households screened out as LBB in HRS 2010 when the target cohort was MBB. The selection weights differed by source as described below.

1) HRS 2016 Sample

The conterminous U.S. was divided into primary sampling units (PSUs), which equated to metropolitan statistical areas, counties or groups of counties. PSUs were stratified into self-representing (SR) or non-self-representing (NSR) PSUs based on the population size of the newly recruited age cohort. SR PSUs were selected with certainty. NSR PSUs were selected using a probability proportional to size method.

Selected PSUs were further divided into census blocks. Using information from the decennial census, the number of new cohort households and the racial/ethnic composition were estimated within each block and used as the measure of size in the sample selection. Using the percentage of Hispanic and non-Hispanic Black households from the race/ethnicity information, blocks were classified into one of the following four strata: 1) non-minority ($<10\%$ Black & $<10\%$ Hispanic); 2) high Black ($\geq 10\%$ Black & $<10\%$ Hispanic); 3) high Hispanic ($<10\%$ Black & $\geq 10\%$ Hispanic); and 4) high Black & Hispanic ($\geq 10\%$ Black & $\geq 10\%$ Hispanic). To ensure that there were enough households to support the interviewer workload, blocks with small population sizes were combined within stratum within the PSU. The secondary sampling units (SSUs) were blocks or groups of blocks. SSUs were selected with probability proportional to size, while considering SSU stratum as well as the stratum of the associated PSU.

Housing units (HUs) were selected in the last stage of selection. Unlike previous cohorts, household sampling was informed by commercial data, which included age and race/ethnicity information at the address level. All addresses in the sampled SSUs were processed for a match with addresses in the commercial data. From this process, each address was classified into one of the six HU strata: 1) LBB-Black; 2) LBB-Hispanic; 3) LBB-Other race or No race info; 4) Not LBB; 5) No age info; and 6) No match. HU selection within sampled SSUs was based on selection probabilities assigned to HU strata. Therefore, there were a total of 24 strata combining stratification at the SSU-level and the HU-level.

The overall selection probability was a multiplication of PSU selection probability, SSU selection probability and HU selection probability. The selection weight was an inverse of the overall selection probability.

2) HRS 2010 Pre-screened LBB Sample

Pre-screened LBB households had two chances of being selected into the HRS LBB cohort: first in 2010 and second in 2016. While the selection probabilities of these households for 2010 were known, those for 2016 were not known and, hence, needed to be estimated. The estimation

mimicked the SSU- and HU-level stratification used in HRS 2016 by locating the pre-screened LBB households into their 24 combinations (= 4 SSU strata x 6 HU strata) and applied corresponding sampling rates to each of the combinations. The resulting selection probabilities for the pre-screened LBB sample subtracted the multiplication of the known selection probabilities in 2010 and the estimated selection probabilities in 2016 from their sum.

1.2 Nonresponse Adjustment

With the selection weights computed for the HRS 2016 sample and the HRS 2010 pre-screened LBB sample described above, we examined three nonresponse adjustment approaches for the HRS 2016 LBB screener and main interviews as follows.

1) Approach 1: SSU Response Rates

To simplify the process, this approach computed the nonresponse adjustment factor using the response rates to the screener interviews and to the main interviews. The response rates were first computed for each SSU. SSU-level response rates were combined into SSU stratum within PSU. The response rates were computed separately for the screener interview and the main interview and then were multiplied together. The inverse of the product of the screener response rate and the main interview response rate was used as a nonresponse adjustment factor.

2) Approach 2: Response Propensity on External Data

Similar to the HRS 2010 nonresponse adjustment, a binary response status of the screener interview was modelled on the HRS sample design information (SSU strata, HU strata and selection weights) and neighborhood characteristics at the census block group level (obtained from 2010 decennial census as well as American Community Survey 2012-2016 listed in Appendix 1) in a logistic regression model.

While this is similar to the approach used in the HRS 2010 nonresponse adjustment, the difference is that, instead of hand-selecting predictors, we used Least Absolute Shrinkage and Selection Operator (LASSO) to handle variable selection for the goal of increasing prediction accuracy. The model was separately fitted for the screener nonresponse (AUC= 63.3%) and the main interview nonresponse (AUC= 62.7%). From these, we obtained predicted response propensities. In order to reduce variation in the nonresponse adjustment factor, we created 10 classes based on deciles of the predicted response propensities and assigned the inverse of the unweighted mean predicted response propensities in each class as a nonresponse adjustment factor.

3) Approach 3: Response Propensity on HRS Key Outcome Variables Predicted with External Data

The nonresponse mechanism is ultimately understood with outcome variables. Under this frame, we modelled a binary response status on a set of key HRS outcome variables listed in Appendix 2. Because the key outcome variables can be obtained only for households who completed the main interview, we first predicted these outcome variables for both those who

completed and did not complete the main interview through a LASSO-based regression model where each outcome variable was regressed on the covariates in Appendix 1.

With the predicted outcomes, SSU strata, HU strata and selection weights, the binary response status was modelled in a logistic regression model separately for the screener interview (AUC= 61.6%) and the main interview (AUC= 60.8%). From these models, we obtained predicted response propensities. As done with Approach 2, a nonresponse adjustment factor was computed by using deciles of the predicted response propensities.

Under each of Approaches 1 to 3, nonresponse adjusted weights were computed by multiplying the selection weights and a respective nonresponse adjustment factor. Table 1 compares nonresponse adjustment weights from these three approaches against the selection weights with respect to 1) the distributions of weights and 2) estimates and their corresponding standard errors of key outcome variables along with 3) design effects due to weighting based on 3,241 LBB financial units who completed the main interview.

Table 1. Distribution of Weights and Key Outcome Variables by Weight Type, LBB Financial Units Who Completed Main Interviews (n=3,241*)

		Base weights	Nonresponse adjusted weights		
			Approach 1. (SSU Response Rate)	Approach 2. (Propensity -- External Data)	Approach 3 (Propensity -- Key Outcomes)
Weight Distribution	Min	107	219	243	198
	Median	811	1981	1746	1765
	Max	6,747	67,630	31,181	28,257
	Mean	1,318	3,908	3,325	3,470
	Design effect due to weighting	2.08	2.80	2.36	2.37
HRS Outcome Variables: Estimate (deff)	curWorkPayHH	0.650 (2.04)	0.683 (2.42)	0.679 (2.21)	0.685 (2.23)
	depressionHH	0.329 (2.15)	0.318 (2.77)	0.314 (2.36)	0.311 (2.40)
	diabetesHH	0.189 (1.82)	0.183 (2.25)	0.181 (2.01)	0.177 (1.98)
	docVisitHH	9.105 (1.85)	9.312 (3.72)	9.024 (2.49)	8.932 (2.35)
	educHH_1	0.138 (1.84)	0.124 (2.07)	0.122 (1.84)	0.118 (1.82)
	educHH_2	0.518 (2.10)	0.497 (2.71)	0.489 (2.36)	0.489 (2.42)
	educHH_3	0.344 (2.14)	0.379 (2.85)	0.389 (2.52)	0.393 (2.55)
	employedHH	0.650 (2.04)	0.683 (2.41)	0.679 (2.21)	0.685 (2.23)
	impairLmtWrkHH	0.309 (2.08)	0.285 (2.42)	0.285 (2.25)	0.278 (2.29)
	medicaidCovHH	0.231 (1.91)	0.205 (2.04)	0.207 (2.03)	0.199 (2.00)
	ownHomeHH	0.530 (2.10)	0.552 (2.67)	0.563 (2.34)	0.571 (2.38)
	ownStockHH	0.223 (2.24)	0.259 (3.26)	0.263 (2.71)	0.264 (2.68)
	privateHlthInsHH	0.566 (2.07)	0.610 (2.54)	0.607 (2.24)	0.614 (2.29)
	regInternetUseHH	0.685 (1.99)	0.726 (2.26)	0.721 (2.06)	0.729 (2.08)

	selfRatedHealthHH	0.344 (2.06)	0.320 (2.56)	0.314 (2.25)	0.308 (2.29)
--	-------------------	--------------	--------------	--------------	--------------

* Base *n* for the analysis of each outcome variable differs slightly due to missingness on the corresponding outcome variable.

Overall, Approaches 2 and 3 yielded similar results in both weight distributions and outcome variable estimates. Regardless of the approach, adjusting for nonresponse resulted in the sample being somewhat more educated, affluent, and healthier. Contrasting Approaches 2 and 3, Approach 3 appeared to move estimates slightly further in those directions. For example, the rate of fair/poor health (selfRatedHealthHH) moved from 33.4% (base weight) to 32.0% (Approach 1), 31.4% (Approach 2) and 30.8% (Approach 3). Given the history of HRS and the similarities between Approaches 2 and 3 in the weight distribution, we decided that Approach 2 was the most appropriate and therefore was used for LBB nonresponse adjustment.

Part 2. Poststratification

HRS poststratification is done at the financial unit level as well as at the respondent level, as shown in Figure 1. For the financial unit poststratification and the community respondent poststratification, the American Community Survey (ACS) 1-year Public Use Micro Sample (PUMS) for 2016 was used for population control totals. For the nursing home respondent poststratification, population totals came from the combination of 2012-2017 ACS 5-year Summary File (SF) and decennial 2010 census data updated for intercensal years. It should be noted that for financial unit poststratification, we monitored the number of age-eligible unrelated persons within each household in the ACS PUMS and selected one as a financial unit and prepared the poststratification benchmark data accordingly.

In the HRS 2016 sampling effort, in order to increase the efficiency of targeting the LBB cohort and oversampling racial/ethnic minorities, the sample design for this newly recruited cohort was informed partially by commercial data that provided age and race/ethnicity information at the individual address level. The availability and accuracy of the commercial data are improving (Roth et al. 2018) but shown to be not random (Kalton et al. 2014). In fact, they are associated positively with higher socio-economic status (Pasek et al. 2014). Therefore, the HRS 2016 poststratification introduced additional dimensions related to socio-economic status focusing on the LBB cohort from the HRS 2016. The pre-screened LBB sample came from the sample designed in 2010 and were not part of the group subject to additional dimensions in the poststratification.

We describe the poststratification in HRS 2016 for financial units and for respondents separately below.

1) Financial Unit Poststratification

Prior to 2016, financial units were poststratified on four dimensions depending on the coupledness of the financial unit. Note that, for the coupled financial units, the financial unit cohort is determined by the older person and, for the non-coupled, by the individual.

For coupled financial units:

- a. Financial Unit Cohort x Race
- b. Person 1 Cohort x Person 2 Cohort¹
- c. Financial Unit Type x Race
- d. Multiple financial units within household

For non-coupled financial units:

- a. Financial Unit Cohort x Race
- b. Sex x Cohort
- c. Financial Unit Type x Race
- d. Multiple financial units within household

For LBB financial units sampled in 2016 (i.e., excluding LBB financial units from the pre-screened LBB sample in 2010), the following three dimensions were added in the poststratification:

- e. Region
- f. Coupledness x Education
- g. Coupledness x Labor force

Each of the financial unit poststratification variables is defined as follows:

- Cohort: 1) AHEAD; 2) CODA; 3) HRS; 4) WB; 5) EBB; 6) MBB; 7) LBB
- Race: 1) Hispanic; 2) Non-Hispanic Black; 3) Non-Hispanic Other
(Note: Race for the coupled financial units was determined considering both persons as follows. First, if any of the two persons is Hispanic, the household was considered as Hispanic. Then, if anyone is black, non-Hispanic black households. The rest was considered non-Hispanic other households. There were 39 coupled LBB financial units where the second person's race was missing. For those cases, we used imputation.)
- Sex: 1) Male; 2) Female
- Household type: 1) Married-Living together; 2) Partnered-Living together; 3) Other/Uncoupled-Male; 4) Other/Uncoupled-Female
- Multiple financial units within household: 1) Yes; 2) No
- Region: 1) Northeast; 2) Midwest; 3) South; 4) West; 9) Not in HRS 2016 LBB sample
- Coupledness: 1) Coupled; 2) Non-coupled; 9) Not in HRS 2016 LBB sample
- Education: 1) Less than High school or GED; 2) High school; 3) Some college; 4) Bachelor's degree; 5) Master's degree or higher; 9) Not in HRS 2016 LBB sample
(Note: For coupled LBB financial units, we aggregated two persons' education levels by taking the higher of the two as the financial unit education level. For those where the second person's education was missing (n=370), we used imputation.)
- Labor force: 1) 0 person in labor force; 2) 1 person in labor force; 3) 2 persons in labor force; 9) Not in HRS 2016 LBB sample

¹ Person cohort designations differ between married-couple financial units and other-couple financial units.

(Note: Person-level labor force status was based on questions that ask about this directly – “Are you doing any work for pay?”, for example. If both members of a coupled financial unit do not complete a core interview or if a person’s labor force status cannot be determined based on that interview, then information on the source of income in the last calendar year was used to determine person-level labor force status from the interview with the designated financial respondent. If there was no financial respondent or if income from employment activity, unemployment insurance, or worker's compensation cannot be determined, then the imputed information by RAND was used. For household-level labor force, person-level information was aggregated. There were 38 coupled LBB financial units where second the person’s labor force status was missing. For those, we imputed for the second person and created the financial unit level labor force.)

For the financial unit poststratification, we trimmed the weights that were below the 1st percentile or above the 99th percentile in order to decrease the effect of extreme weights. The resulting variable name is `pwgthh`.

We compared the distribution of poststratification control variables between the population (from ACS PUMS) and the HRS sample with no weights and with poststratification with and without LBB-specific dimensions (i.e., region, education and labor force). Note that the poststratification without LBB-specific dimensions was the same as the poststratification done prior to 2016. Table 2 includes this comparison for community financial units for all cohorts and for LBB financial units. The effect of poststratification was clear for LBB financial units. Without weighting, HRS LBB financial units under-represented those with married couples living together and over-represented other financial units with a female householder. They also over-represented racial/ethnic minority financial units, reflecting minority oversampling in the HRS 2016 sample. Even with the poststratification, for HRS LBB financial units, region, education and labor force were distributed quite differently than its population, when these characteristics were not included in the poststratification. In particular, the HRS LBB sample noticeably under-represented the Northeast and over-represented the South. In the population, 18.8% of the LBB financial units were in the Northeast and 37.2% in the South, whereas the figures were 13.1% Northeast and 44.0% South for the HRS LBB financial unit sample poststratified with dimensions used prior to 2016. When poststratification included LBB-specific dimensions, the HRS sample was distributed close to the population.

Table 2. Comparison of Population Distribution and HRS Sample Distribution of Control Characteristics in Poststratification, Financial Units in Community

Poststratification Control Characteristics	All Community Financial Units				LBB Financial Units			
	Pop	HRS (n=14,170)			Pop	HRS (n=3,235)		
	ACS	No Weight	Without LBB Dims	With LBB Dims	ACS	No Weight	Without LBB Dims	With LBB Dims
	%	%	%	%	%	%	%	%
Financial Unit Cohort								
AHEAD	1.6	1.5	1.6	1.6	-	-	-	-
CODA	4.0	5.6	4.1	4.0	-	-	-	-
HRS	15.6	23.9	15.8	15.6	-	-	-	-
WB	15.3	10.1	15.4	15.3	-	-	-	-
EBB	19.8	17.5	19.9	19.8	-	-	-	-
MBB	22.0	19.4	22.1	22.0	-	-	-	-
LBB	21.8	22.0	21.2	21.8	100.0	100.0	100.0	100.0
Financial Unit Type								
Married, living together	48.0	39.9	47.7	47.7	49.9	39.4	48.9	48.7
Partnered, living together	4.6	6.4	4.6	4.7	6.8	8.7	7.0	7.2
Other/Uncoupled, Male	16.8	15.5	16.7	16.7	18.8	18.4	18.6	18.7
Other/Uncoupled, Female	30.7	38.2	31.0	30.9	24.6	33.5	25.4	25.4
Race/Ethnicity								
Hispanic	11.0	17.2	11.1	11.2	15.6	23.0	16.1	16.4
Non-Hispanic Black	11.6	23.7	11.8	11.9	13.7	29.3	14.4	14.7
Non-Hispanic Other	77.5	59.2	77.1	76.9	70.7	47.7	69.5	68.9
Region								
Northeast	4.1	3.0	2.8	3.8	18.8	13.4	13.1	17.2
Midwest	4.8	3.9	4.2	4.7	21.9	17.8	19.6	21.5
South	8.1	10.3	9.3	8.4	37.2	46.6	44.0	38.7
West	4.8	4.9	4.9	4.9	22.1	22.2	23.2	22.6
Not in HRS 2016 sample	78.2	78.0	78.8	78.2	-	-	-	-
Education								
<High school or GED	2.5	3.7	2.8	2.6	11.5	16.7	13.0	12.1
High school	4.4	4.8	4.1	4.3	20.1	21.9	19.4	19.7
Some college	6.8	6.7	6.1	6.8	31.3	30.3	28.6	31.3
Bachelor's degree	4.7	3.8	4.1	4.6	21.5	17.3	19.4	20.9
Master's degree or higher	3.4	3.1	4.2	3.5	15.7	13.9	19.7	16.0
Not in HRS 2016 sample	78.2	78.0	78.8	78.2	-	-	-	-
Labor Force								
0 person in labor force	2.7	4.5	3.4	2.9	12.5	20.5	16.2	13.3
1 person in labor force	10.6	10.9	10.1	10.7	48.8	49.5	47.6	49.0
2 person in labor force	8.5	6.6	7.7	8.2	38.7	30.0	36.2	37.7
Not in HRS 2016 sample	78.2	78.0	78.8	78.2	-	-	-	-

2) Community Respondent Poststratification

Prior to 2016, respondents in the community were poststratified on the following two dimensions:

- a. Financial Units Cohort x Marital status
- b. Financial Unit Cohort x Sex x Race

For respondents in LBB households, following two dimensions were introduced additionally:

- c. Region x Education
- d. Region x Labor force

Each variable is defined as follows:

- Financial Unit Cohort: 1) AHEAD; 2) CODA; 3) HRS; 4) WB; 5) EBB; 6) MBB; 7) LBB
- Sex: 1) Male; 2) Female
- Race: 1) Hispanic; 2) Non-Hispanic Black; 3) Non-Hispanic Other;
- Region: 1) Northeast; 2) Midwest; 3) South; 4) West; 9) Not in HRS 2016 LBB sample
- Education: 1) Less than High school or GED; 2) High school; 3) Some college; 4) Bachelor's degree; 5) Master's degree or higher; 9) Not in HRS 2016 LBB sample;
- Labor force: 1) In labor force; 2) Not in labor force; 9) Not in HRS 2016 LBB sample

Because extreme weights were handled through financial unit weight trimming, respondent weights were not subject to trimming. The resulting variable name is `pwgtr`.

Table 3 compares the population and HRS sample on the control dimensions. Without the poststratification, HRS respondents, in particular LBB respondents in LBB financial units sampled in 2016, appeared to undercount those who were married, males, non-Hispanic other races, in the Northeast, with higher education or in the labor force than what the population data indicate. Given the importance of these characteristics in health and economic outcomes, they were accounted for through poststratification.

Table 3. Comparison of Population Distribution and HRS Sample Distribution of Control Characteristics in Poststratification, Respondents in Community Households

Poststratification Control Characteristics	All Community Respondents			LBB Financial Unit Respondents		
	Population	HRS (n=19,621)		Population	HRS (n=3,714)	
	ACS	No Weight	Poststrat	ACS	No Weight	Poststrat
	%	%	%	%	%	%
Financial Unit Cohort						
AHEAD	1.4	1.1	1.4	-	-	-
CODA	3.4	4.7	3.4	-	-	-
HRS	14.1	21.8	14.1	-	-	-
WB	14.4	11.3	14.4	-	-	-
EBB	19.5	17.6	19.5	-	-	-
MBB	23.0	20.6	23.0	-	-	-
LBB	24.1	22.9	24.1	100.0	100.0	100.0
Marital Status						
Married	60.1	53.9	60.1	57.0	47.2	56.1
Separated/Divorced	18.3	20.7	18.3	24.2	31.9	24.9
Widowed	12.8	17.4	12.8	3.1	4.5	3.2
Never Married	8.9	8.1	8.9	15.7	16.4	15.8
Sex						
Male	46.6	42.4	46.6	53.9	47.7	53.9
Female	53.4	57.6	53.5	46.1	52.3	46.1
Race/Ethnicity						
Hispanic	10.1	16.1	10.1	14.0	20.7	14.1
Non-Hispanic Black	10.5	21.3	10.5	12.9	28.3	13.3
Non-Hispanic Other	79.4	62.6	79.4	73.1	51.0	72.7
Region						
Northeast	3.7	2.5	3.7	18.9	13.2	18.9
Midwest	4.2	3.4	4.2	21.5	17.8	21.5
South	7.3	8.9	7.3	37.3	47.1	37.3
West	4.4	4.2	4.4	22.4	21.9	22.4
Not in HRS 2016 sample	80.4	81.1	80.6	-	-	-
Education						
<High school or GED	3.2	4.1	3.2	16.3	21.4	16.3
High school	4.8	4.6	4.7	24.4	24.5	24.4
Some college	5.8	5.4	5.7	29.5	28.6	29.5
Bachelor's degree	3.7	3.0	3.7	18.8	15.6	18.8
Master's degree or higher	2.2	1.9	2.2	11.1	10.0	11.1
Not in HRS 2016 sample	80.4	81.1	80.6	-	-	-
Labor Force						
Not in labor force	4.1	5.2	4.1	20.9	27.5	20.9
In labor force	15.5	13.7	15.4	79.1	72.5	79.1
Not in HRS 2016 sample	80.4	81.1	80.6	-	-	-

3) Nursing Home Respondent Poststratification

Nursing home residency status was determined based on a combination of respondent self-report, CMS Minimum Data Set (MDS) records, and nursing home address confirmation using data from Brown University’s LTC Focus project. Following from this approach, there were 306 age-eligible respondents from MBB or older cohort households residing in nursing homes. For these respondents, the HRS developed a poststratification strategy that uses ACS annual SF and intercensal estimates to the possible extent to obtain population nursing home resident counts and their distributions on the following dimensions:

- a. Sex x Age
- b. Sex x Race

Each variable in the control dimensions is defined as follows:

- Sex: 1) Male; 2) Female
- Age: 1) 55-64 years old; 2) 65-69 years old; 3) 70-74 years old; 4) 75-79 years old; 5) 80-84 years old; 6) 85-89 years old; 7) 90-94 years old; 8) 95 years old or older
- Race: 1) Black; 2) Non-Black

The resulting variable name is `pwgt_rnh`. Table 4 compares the population and the HRS sample of nursing home residents.

Table 4. Comparison of Population Distribution and HRS Sample Distribution of Control Characteristics in Poststratification, Respondents in Nursing Homes

Poststratification Control Characteristics	Population	HRS (n=306)	
	ACS	No Weight	Poststrat
	%	%	%
Sex x Age			
Female 55-64 years old	5.2	1.6	5.2
Female 65-69 years old	4.1	3.3	4.1
Female 70-74 years old	5.8	3.9	5.8
Female 75-79 years old	7.5	6.2	7.5
Female 80-84 years old	10.1	13.1	10.1
Female 85-89 years old	15.0	15.4	15.0
Female 90-94 years old	11.4	13.4	11.4
Female ≥ 95 years old	5.8	9.2	5.8
Male 55-64 years old	6.4	4.3	6.4
Male 65-69 years old	4.3	2.9	4.3
Male 70-74 years old	4.5	2.6	4.6
Male 75-79 years old	4.2	3.9	4.2
Male 80-84 years old	5.5	6.5	5.5
Male 85-89 years old	5.9	5.2	5.9
Male 90-94 years old	3.2	5.9	3.2
Male 95 years old	1.1	2.6	1.1
Sex x Race			
Female Black	7.7	9.8	7.7

Female Non-Black	57.1	56.2	57.1
Male Black	6.8	7.5	6.8
Male Non-Black	28.4	26.5	28.4

Appendix 1. Predictors of Response Propensities

Variable name	Definition	Values	Unit Level	Source
LAND_AREA	Land area (sq.mi.)	Numeric	Census block group	Census geography
Tot_Population_ACS_12_16	Total Population	Numeric	Census block group	ACS 5 years estimate (2012-2016)
Tot_Housing_Units_ACS_12_16	Total Housing Units	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_URBANIZED_AREA_POP_CEN_2010	Percentage of the population that lives in a densely settled area containing 50,000 or more people	Numeric	Census block group	2010 Census
pct_URBAN_CLUSTER_POP_CEN_2010	Percentage of the population that lives in a densely settled area containing 2,500 to 49,999 people	Numeric	Census block group	2010 Census
pct_Females_ACS_12_16	Percentage of the population that is female	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_Pop_under_5_ACS_12_16	Percentage of the population that is under five years old	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_Pop_5_17_ACS_12_16	Percentage of the population that is between 5 and 17 years old	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_Pop_18_24_ACS_12_16	Percentage of the population that is between 18 and 24 years old	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_Pop_25_44_ACS_12_16	Percentage of the population that is between 25 and 44 years old	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_Pop_45_64_ACS_12_16	Percentage of the population that is between 45 and 64 years old	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_Tot_GQ_CEN_2010	Percentage of the population living or staying in a group living arrangement owned by an entity providing housing to residents, who are usually unrelated	Numeric	Census block group	2010 Census
pct_Non_Inst_GQ_CEN_2010	Percentage of the population who live in group quarters and are primarily eligible, able, or likely to participate in labor force while residents. Noninstitutional group quarters include college/university student housing, military quarters, and other noninstitutional facilities.	Numeric	Census block group	2010 Census

pct_Hispanic_ACS_12_16	Percentage of the population that identify as "Mexican", "Puerto Rican", "Cuban", or "another Hispanic, Latino, or Spanish origin"	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_NH_White_alone_ACS_12_16	Percentage of the population that indicate no Hispanic origin and their only race as "White" or report entries such as Irish, German, Italian, Lebanese, Arab, Moroccan, or Caucasian	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_NH_Black_alone_ACS_12_16	Percentage of the population that indicate no Hispanic origin and their only race as "Black, African Am., or Negro" or report entries such as African American, Kenyan, Nigerian, or Haitian	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_NH_AIAN_alone_ACS_12_16	Percentage of the population that indicate no Hispanic origin and their only race as "American Indian or Alaska Native" or report entries such as Navajo, Blackfeet, Inupiat, Yup'ik, or Central/South American Indian groups	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_NH_Asian_alone_ACS_12_16	Percentage of the population that indicate no Hispanic origin and their only race as "Asian Indian", "Chinese", "Filipino", "Korean", "Japanese", "Vietnamese", or "Other Asian"	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_NH_NHOPI_alone_ACS_12_16	Percentage of the population that indicate no Hispanic origin and their only race as "Native Hawaiian", "Guamanian or Chamorro", "Samoan", or "Other Pacific Islander"	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_Othr_Lang_ACS_12_16	Percentage of the population aged 5 years and over that speaks a language other than English at home	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_Not_HS_Grad_ACS_12_16	Percentage of the population aged 25 years and over that are not high school graduates and have not received a diploma or the equivalent	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_College_ACS_12_16	Percentage of the population aged 25 years and over that have a college degree or higher	Numeric	Census block group	ACS 5 years estimate (2012-2016)

pct_Pov_Univ_ACS_12_16	Percentage of the population for whom poverty level is determined, excluding institutionalized people, people in military group quarters, people in college dormitories, and unrelated individuals under 15 years old	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_Prs_Blw_Pov_Lev_ACS_12_16	Percentage of the population that are classified as below the poverty level given their total family or household income within the last year, family size, and family composition	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_One_Health_Ins_ACS_12_16	Percentage of the population that have one type of health insurance coverage, including public or private	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_No_Health_Ins_ACS_12_16	Percentage of the population that have no health insurance, public or private	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_Diff_HU_1yr_Ago_ACS_12_16	Percentage of the population aged 1 year and over that moved from another residence in the U.S. or Puerto Rico within the last year	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_ENG_VW_SPAN_ACS_12_16	Percentage of all occupied housing units where a Spanish or Spanish Creole language was assigned as the household language and no one ages 14 years and over speaks English only or speaks English "very well"	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_ENG_VW_INDOEURO_ACS_12_16	Percentage of all occupied housing units where an Indo-European language was assigned as the household language and no one ages 14 years and over speaks English only or speaks English "very well"	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_ENG_VW_API_ACS_12_16	Percentage of all occupied housing units where an Asian and Pacific Island language was assigned as the household language and no one ages 14 years and over speaks English only or speaks English "very well"	Numeric	Census block group	ACS 5 years estimate (2012-2016)

pct_ENG_VW_OTHER_ACS_12_16	Percentage of all occupied housing units where a language other than English, Spanish, Indo-Euro, or API was assigned as the household language and no one ages 14 years and over speaks English only or speaks English "very well"	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_Rel_Family_HHD_ACS_12_16	Percentage of all occupied housing units where at least 2 members are related by birth, marriage, or adoption; same-sex couple households with no relatives of the householder present are not included	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_MrdCple_HHD_ACS_12_16	Percentage of all occupied housing units where the householder and his or her spouse are listed as members of the same household; does not include same-sex married couples	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_Female_No_HB_ACS_12_16	Percentage of all occupied housing units with a female householder and no husband of householder present	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_Sngl_Prns_HHD_ACS_12_16	Percentage of all occupied housing units where a householder lives alone	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_HHD_PPL_Und_18_ACS_12_16	Percentage of all occupied housing units where one or more people are ages 18 years or under	Numeric	Census block group	ACS 5 years estimate (2012-2016)
avg_Tot_Prns_in_HHD_ACS_12_16	Average number of persons per occupied housing unit	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_Rel_Under_6_ACS_12_16	Percentage of family-occupied housing units with a related child under 6 years old	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_HHD_Moved_in_ACS_12_16	Percentage of all occupied housing units where the householder moved into the current unit in the year 2010 or later	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_PUB_ASST_INC_ACS_12_16	Percentage of all occupied housing units that receive public assistance income (general assistance and Temporary Assistance to Needy Families)	Numeric	Census block group	ACS 5 years estimate (2012-2016)
avg_Agg_HH_INC_ACS_12_16	Average aggregate household income	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_Tot_Occp_Units_ACS_12_16	Percentage of all housing units that are classified as the usual place of residence of the individual or group living in it	Numeric	Census block group	ACS 5 years estimate (2012-2016)

pct_Owner_Occp_HU_ACS_12_16	Percentage of occupied housing units with an owner or co-owner living in it	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_Single_Unit_ACS_12_16	Percentage of all housing units that are in a structure that contains only that single unit	Numeric	Census block group	ACS 5 years estimate (2012-2016)
ct_MLT_U10p_ACS_12_16	Percentage of all housing units that are in a structure that contains 10 or more housing units	Numeric	Census block group	ACS 5 years estimate (2012-2016)
ct_Mobile_Homes_ACS_12_16	Percentage of all housing units that are considered mobile homes	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_Crowd_Occp_U_ACS_12_16	Percentage of occupied housing units that have more than 1.01 persons per room	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_NO_PH_SRVC_ACS_12_16	Percentage of occupied housing units that do not have a working telephone and available service	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_No_Plumb_ACS_12_16	Percentage of all housing units that do not have complete plumbing facilities	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_Recent_Built_HU_ACS_12_16	Percentage of all housing units that are in a building that was constructed in 2010 or later	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_TEA_MailOutMailBack_CEN_2010	Percentage of addresses from which a Census form was expected to be delivered for mail return that were in a Mailout/Mailback type of enumeration area in the 2010 Census	Numeric	Census block group	2010 Census
pct_BILQ_Mailout_count_CEN_2010	Percentage of all addresses in a 2010 Census mailback area that received the 2010 Census English/Spanish bilingual Mailout/Mailback form	Numeric	Census block group	2010 Census
avg_Agg_House_Value_ACS_12_16	Average aggregate value of an owner-occupied housing unit (in dollars)	Numeric	Census block group	ACS 5 years estimate (2012-2016)
pct_Census_Mail_Returns_CEN_2010	Completed 2010 Census mail forms received from addresses in a mailback type of enumeration area (Mailout/Mailback and Update/Leave areas) out of all addresses from which a Census form was expected to be delivered for mail return	Numeric	Census block group	2010 Census
pct_Deletes_CEN_2010	Percentage of addresses in a 2010 Census mailback area deleted because it was determined to not correspond to a valid housing unit (Mailout/Mailback and Update/Leave areas)	Numeric	Census block group	2010 Census

pct_Census_UAA_CEN_2010	Percentage of addresses in a 2010 Census Mailout/Mailback area where the initial mail form was returned to the Census with the postal code "Undeliverable as Addressed"	Numeric	Census block group	2010 Census
pct_Mailback_Count_CEN_2010	Percentage of valid, occupied addresses in mailback areas in the 2010 Census	Numeric	Census block group	2010 Census
pct_FRST_FRMS_CEN_2010	Percentage of all addresses in a 2010 Census mailback area for which the first Mailout/Mailback form mailed was completed and returned	Numeric	Census block group	2010 Census
pct_RPLCMNT_FRMS_CEN_2010	Percentage of all addresses in a 2010 Census mailback area for which the replacement Mailout/Mailback form mailed was completed and returned	Numeric	Census block group	2010 Census
Med_House_Value_BG_ACS_12_16	Median House value for the block group	Numeric	Census block group	ACS 5 years estimate (2012-2016)
Mail_Return_Rate_CEN_2010	Number of mail returns received out of the total number of valid occupied housing units (HUs) in the Mailout/Mailback universe which excludes deleted, vacant, or UAA units.	Numeric	Census block group	2010 Census
Low_Response_Score	Score predicting that a block group will produce a low mail return rate	Numeric	Census block group	2010 Census
pct_Civ_unemp_16_24_ACS_12_16	Percentage of civilians between the ages of 16 and 24 in the labor force that are unemployed	Numeric	Census Tract	ACS 5 years estimate (2012-2016)
pct_Civ_unemp_25_44_ACS_12_16	Percentage of civilians between the ages of 25 and 44 in the labor force that are unemployed	Numeric	Census Tract	ACS 5 years estimate (2012-2016)
pct_Civ_unemp_45_64_ACS_12_16	Percentage of civilians between the ages of 45 and 64 in the labor force that are unemployed	Numeric	Census Tract	ACS 5 years estimate (2012-2016)
pct_Pop_Disabled_ACS_12_16	Percentage of the population who have one or more disabilities	Numeric	Census Tract	ACS 5 years estimate (2012-2016)
pct_Born_foreign_ACS_12_16	Percentage of the population who were not a citizen of the United States at birth	Numeric	Census Tract	ACS 5 years estimate (2012-2016)
NumberOfAdults	Number of adults in the household	Numeric	Address	MSG
NumberOfChildren	Number of children in the household	Numeric	Address	MSG
AsianSurname	Indicator that appended surname is an Asian surname	0 = No, 1 = Yes	Address	MSG

HispanicSurname	Indicator that appended surname is a Hispanic surname	0 = No, 1 = Yes	Address	MSG
Target18to24	Indicator of presence of 18 to 24 years old in household	0 = No, 1 = Yes	Address	MSG
Target25to34	Indicator of presence of 25 to 34 years old in household	0 = No, 1 = Yes	Address	MSG
Target35to64	Indicator of presence of 35 to 64 years old in household	0 = No, 1 = Yes	Address	MSG
Education_1	Indicator of high school diploma	0 = No, 1 = Yes	Address	MSG
Education_2	Indicator of some college	0 = No, 1 = Yes	Address	MSG
Education_3	Indicator of bachelor degree	0 = No, 1 = Yes	Address	MSG
Education_4	Indicator of graduate degree	0 = No, 1 = Yes	Address	MSG
Gender1_M	Indicator of male	0 = No, 1 = Yes	Address	MSG
Gender1_U	Indicator of unknown gender	0 = No, 1 = Yes	Address	MSG
Income_A	Indicator of income between \$1,000-\$14,999	0 = No, 1 = Yes	Address	MSG
Income_B	Indicator of income between \$15,000-\$24,999	0 = No, 1 = Yes	Address	MSG
Income_C	Indicator of income between \$25,000-\$34,999	0 = No, 1 = Yes	Address	MSG
Income_D	Indicator of income between \$35,000-\$49,999	0 = No, 1 = Yes	Address	MSG
Income_E	Indicator of income between \$50,000-\$74,999	0 = No, 1 = Yes	Address	MSG
Income_F	Indicator of income between \$75,000-\$99,999	0 = No, 1 = Yes	Address	MSG
MaritalStatus_M	Indicator of Married	0 = No, 1 = Yes	Address	MSG
ownrent_R	Indicator of Rental	0 = No, 1 = Yes	Address	MSG
race_B	Indicator of Black	0 = No, 1 = Yes	Address	MSG
race_H	Indicator of Hispanic	0 = No, 1 = Yes	Address	MSG
MSGmissingflag	Indicator of missing data in MSG variables	0 = No, 1 = Yes	Address	MSG

Appendix 2. Health and Retirement Study Key Outcome Variables for Head of Household

Variable name	Definition	Values	Unit Level	Source
employedHH	Head of household Employment status	0 = Not employed, 1 = Employed	Household	HRS2016 [J005M1-J005M5]
curWorkPayHH	Head of household Currently Working for Pay	0 = No, 1 = Yes	Household	HRS2016 [J020]
regInternetUseHH	Head of household Regular Internet Use	0 = No, 1 = Yes	Household	HRS2016 [W303]
diabetesHH	Head of household Chronic - Diabetes	0 = No, 1 = Yes	Household	HRS2016 [C010]
impairLmtWrkHH	Head of household Impairment Limit Work	0 = No, 1 = Yes	Household	HRS2016 [M502]
docVisitHH	Head of household Number of Doctor Visits in the past 2 years	Numeric	Household	HRS2016 [N147]
depressionHH	Head of household CIDI Depression	0 = No, 1 = Yes	Household	HRS2016 [C150/C167]
ownHomeHH	Head of household Own primary residence	0 = No, 1 = Yes	Household	HRS2016 [H004/H008/H014/X079]
medicaidCovHH	Head of household Medicaid Coverage	0 = No, 1 = Yes	Household	HRS2016 [N005]
privateHlthInsHH	Head of household Private Health Insurance	0 = No, 1 = Yes	Household	HRS2016 [N023]
ownStockHH	Head of household Own Stock	0 = No, 1 = Yes	Household	HRS2016 [Q316/Q513_1-Q513_3]
educHH_1	Head of household Education - Less than HS	0 = No, 1 = Yes	Household	HRS2016 [B014/B015/B016]
educHH_2	Head of household Education - HS/Some College	0 = No, 1 = Yes	Household	HRS2016 [B014/B015/B016]
educHH_3	Head of household Education – College	0 = No, 1 = Yes	Household	HRS2016 [B014/B015/B016]
selfRatedHealthHH	Head of household Self-rated Health	0 = Other, 1 = Fair/Poor	Household	HRS2016 [C001]