Active versus Passive Sample Attrition: The Health Retirement Study

Honggao Cao and Daniel H. Hill
Institute for Social Research
University of Michigan
426 Thompson Street
Ann Arbor, MI 48104
E-mail: honggao@isr.umich.edu

March 2005

Honggao Cao, Institute for Social Research, University of Michigan, 426 Thompson Street, Ann Arbor, MI 48104;
E-mail: honggao@isr.umich.edu

Abstract

This paper investigates sample attrition in the Health and Retirement Study (HRS). We compare attrition behavior in two of the HRS cohorts: original HRS cohort and AHEAD cohort. We distinguish attrition due to death (passive attrition) from attrition due to other causes (active attrition), examining potential effects of different attrition modes on the representativeness of the remaining samples. This distinction is justified based on a specification test on a multinomial logistic regression model. Among other results from the study are differences between passive and active attritors in a set of demographic, economic, and health measures, and a finding that active attrition occurring in the HRS is perhaps not selective and, thus, is statistically ignorable.

Keywords: active attrition, passive attrition, sample attrition, HRS

Introduction

Sample attrition is a significant problem facing many longitudinal surveys. By 1989, for example, the Michigan Panel Study of Income Dynamics (PSID) had lost about 50 percent of its initial 1968 membership (Fitzgerald, Gottschalk, and Moffitt, 1998). In studies conducted in developing countries, the annual attrition rate could be as high as 23 percent (Alderman, Behrman, Kohler, Maluccio, and Watkins, 2001). One of the concerns shared by researchers about sample attrition is that it may lead to selective samples, rendering the interpretation of research results problematic.

Several alternative strategies have been used to address this concern. One such strategy seeks to understand who drops out of a survey and why, hoping that by knowing the answer to this question one could find ways to reduce attrition. A second strategy, which is often referred to as sample "refreshment" or "replenishment" in the literature, tries to protect the integrity of a sample by replacing units who have dropped out with new units randomly sampled from the original population (Heeringa, 1997; Hirano, Imbens, Ridder, and Rubin, 2001; Ridder, 1992). A third and most common strategy examines whether the attrition occurring in a longitudinal study is statistically ignorable in a sense that the representativeness of the original sample is still intact even with the attrition (Fitzgerald, Gottschalk, and Moffitt, 1998; Alderman, Behrman, Kohler, Maluccio, and Watkins, 2001; Hill, 2001). When the attrition is found to be non-ignorable, this strategy also seeks alternative ways to correct attrition bias.

While differing in their focuses, most of the studies based on these strategies take as given how sample attrition takes place in a longitudinal setting, treating attrition as a single category unit nonresponse. They seldom distinguish, for example, attrition due to uncontrollable causes such as death (called "passive attrition" in this paper) from attrition due to choice (or

"active attrition"), ignoring the potentially different effects on their analyses of different causes or modes of attrition.

But the distinction between passive and active attrition is perhaps very important in survey research, for two reasons. First, to the extent that attrition due to death (or other uncontrollable causes) is mostly unavoidable, the efforts to reduce sample attrition should be directed only to those who are likely to drop out of a survey willingly or "actively". If someone drops out of a survey because he or she cannot be reached, for example, finding an effective way to contact respondents may be important to reduce attrition. On the other hand, if someone drops out because of lack of motivation, improving incentive schemes may help him or her stay. Second, when "attrition due to uncontrollable causes" is intertwined with the objectives of a survey—as is the case death-generated attrition in a survey of health or aging—differentiating active from passive attrition may offer an important way to monitor the validity of the survey.

In this paper we investigate active and passive sample attrition in the Health and Retirement Study (HRS), a national longitudinal survey conducted at the University of Michigan. We define passive attrition in the HRS as attrition induced by the death of a respondent, and active attrition as induced by any other causes. We compare active and passive attrition rates in the original HRS sample with those in its sister study, the Asset and Health Dynamics Among the Oldest Old (AHEAD), which is now an integral component of the HRS. We contrast some important health and economic measures for both active and passive attrition groups, speculating potential effects of different types of attrition on the representativeness of the study. We also examine whether attrition due to death and attrition due to choice are two distinct, independent outcomes, and whether people who drop out of the survey because of death and those who drop out of the survey because of choice are from two different groups.

The remainder of the paper is organized as follows. Section II reviews prior research on sample attrition. This review demonstrates that previous studies have rarely done anything distinguishing active from passive attrition. It also helps us formulate our empirical strategy. Section III describes the HRS and AHEAD data, presenting an overall picture of sample attrition in both the HRS and AHEAD. In Section IV we examine the impact of different types of attrition on the representativeness of some important health and wealth measures, and speculate whether attrition to the HRS and AHEAD samples are statistically ignorable. In Section V we estimate a multinomial logistic regression model to contrast active with passive attrition, testing whether active and passive attrition should be regarded as independent attrition outcomes, and whether people dropping out of the survey actively and passively are different in some of important individual characteristics. The paper concludes with a discussion of the results found in the study.

II. Prior Research on Sample Attrition

With recent developments in many longitudinal surveys, sample attrition in panel data has attracted increasing interest of statisticians and econometricians.[1] The existing research on sample attrition has focused on several interrelated areas. In the first area, researchers concentrate on identifying who drops out of a survey and why. Concerned with potentially negative impact of attrition on the representativeness of the remaining sample, a typical study in this area (e.g., Alderman, Behrman, Kohler, Maluccio, and Watkins, 2001; Jacomb, Jorm, Korten, Christensen, and Henderson, 2002) looks for reasons why attrition occur, explores the differences between attritors and non-attritors, and identifies who is more likely to attrite. Among the findings in this area are that attritors are different from non-attritors in some

[1] The Journal of Human Resources has devoted a complete issue (Volume 33, No. 2, 1998) to cover sample attrition.

important characteristics; attrition over time tends to be highly selective of people who are young, mobile, with lower cognitive scores, and among lower socioeconomic status.

A second area of interest of attrition research deals with so-called sample "refreshment" or "replenishment". In a hope of mitigating the effects of attrition, panel data are sometimes augmented or "refreshed" by replacing sample members who have dropped out of a survey with new units randomly drawn from the original population. Sample refreshment or replenishment generally serves two purposes: maintaining the sample size (Herringa, 1997), and using the replacement units to test models of attrition (Hirano, Imbens, Ridder, and Rubin, 2001; Ridder, 1992).

The majority of the studies on sample attrition concentrate on statistical features of the attrition problem. In this area, researchers often distinguish missing (or attrition) on observables from missing on unobservables (Fitzgerald, Gottschalk, and Moffitt, 1998; Heckman and Robb, 1985; Heckman and Hotz, 1989). According to this distinction, missing on observables occurs if the variable of interest is independent of the attrition process conditional on the explanatory variables, whereas missing on unobservables occurs if this conditional independence does not hold. When missing on observables happens, sample attrition is statistically ignorable, as an appropriate weighting or multiple imputation procedure can be used to generate unbiased inference. When missing on unobservables happens, however, the effect of the attrition on the variable of interest can only be assessed in a sample selection model that relies on a specification of the joint distribution of the error terms for both the response equation and the equation of the interest (Hausman and Wise, 1979; Ridder, 1990). Although many longitudinal surveys have experienced significant loss of original sample members, a consensus in the literature is that

selection biases introduced by attrition are very mild (Fitzgerald, Gottschalk, and Moffitt, 1998; Lillard and Panis, 1998; Alderman, Behrman, Kohler, Maluccio, and Watkins, 2001).

One key feature of most of the attrition analyses is their definition of attrition as a single category nonresponse. Seldom is a distinction made between attrition due to different causes, as demonstrated by an excellent example in Inkman (2001):

> "The existing approaches to the correction of the attrition bias have in common that they focus on a single nonresponse category: a sample unit either responds or not. This binary indicator prohibits any distinction between different nonresponse reasons. As an illustration consider the work by Harhooff, Stahl and Woywode (1998) who estimate firm survival and employment growth equations using firm panel data. The authors have information on two different causes of firm exit: voluntary liquidation and bankruptcy. While they this distinction in estimating competing risk models of the instantaneous firm exit probability (hazard), they merge these two nonresponse reasons into a single exit category in estimating the employment growth equation. This was done although the competing risk model revealed substantial differences between the determinants of voluntary liquidation and bankruptcy." (P.2)

Inkman suggested that "an obvious explanation for neglecting the different nonresponse mechanisms is the lack of appropriate panel data estimation techniques which account for nonresponse heterogeneity." (P.2) He continued to propose a GMM estimator to account for the heterogeneity.

While it is important to find a feasible way to directly account for nonresponse heterogeneity in the process of correcting attrition bias, in this paper we use a different strategy to approach the nonresponse heterogeneity issue. We explicitly explore the distinctions between two alternative modes of attrition in the HRS and AHEAD, testing whether attrition due to death differs from attrition due to other causes. We also assess whether different modes of attrition have different effects—if any—on the representativeness of some important health and economic measures.

III. Sample Attrition in the HRS and AHEAD

The Health and Retirement Study (HRS) is a biennial national longitudinal survey on older Americans (50+ years). It intends to "provide data for researchers, policy analysts, and program planners who are making major policy decisions that affect retirement, health insurance, saving and economic well-being" (HRS, 2004). The first wave of HRS data collection was conducted in 1992, when 12,544 original HRS cohort members—individuals born between 1931 and 1941 and their spouses—were interviewed. Since then, six more waves of HRS interviews (HRS 1994, 1996, 1998, 2000, 2002, and 2004) have been conducted, with the sample now consisting of three more cohorts including Asset and Health Dynamics Among the Oldest Old (AHEAD), Children of the Depression Age (CODA), and War Babies (WB). Originally designed as an independent study parallel to the HRS but targeting individuals born before 1923, the AHEAD cohort was first interviewed in 1993 and then in 1995, before it became an integral part of the HRS in 1998. Also joining the HRS in 1998 were two virgin cohorts—CODA and WB. The former targeted individuals born between 1924 and 1930, whereas the latter individuals born between 1942 and 1947.

This study concentrates only on the attrition behavior of the original HRS and AHEAD cohort members. We define attrition as a one-time event when a sample member drops out of the survey. Obviously, according to this definition, sample members re-entering the survey after having dropped out in a previous wave will not be treated as staying in the sample. We may regard those re-entry cases as a special type of sample refreshment and argue that they should be dealt with separately.

As shown in Table 1, in about ten years, the HRS cohort has lost more than one third of its original members, while the AHEAD has lost more than half. The marginal rate of attrition

was about 9% for the HRS cohort, while was more than 15% for the AHEAD cohort. The primary reason for this differential attrition rate was due to the fact that the AHEAD cohort was about 15 years older. Higher mortality in the AHEAD cohort has induced a much higher rate of passive attrition than in the HRS cohort. In fact, as shown in Table 3, attrition in AHEAD was predominantly passive (due to death)—especially in later waves, while attrition in HRS was predominantly active—especially in early waves.

Black sample members tended to attrite at a higher rate than non-Black members (Table 2). This was the case for both the HRS and AHEAD cohorts. By HRS 2002, for example, the HRS cohort has lost about 45% of its black members, 11 percentage points higher than the attrition rate for the HRS cohort as a whole. Similarly, the AHEAD has lost more than 60% of its black members, which was 8 percentage points higher than the attrition rate for the AHEAD cohort as a whole.

Hispanic sample members showed a higher attrition rate in the HRS cohort, but a lower attrition rate in the AHEAD. By HRS 2002, more than 45% of Hispanic members had attrited in the HRS cohort, about 12 percentage points higher than the attrition rate for the HRS cohort as a whole. In AHEAD cohort, about 53% of Hispanic sample members had attrited, compared to about 55% of attrition rate for the AHEAD cohort as a whole.

In terms of the mode of attrition, the percentage of passive attrition increased over time in both the HRS and AHEAD cohorts (Table 3). In the first follow-up wave for the HRS cohort (HRS 1994), less than 20% of total attrition was due to death of sample members. This number increased to 47% in HRS 2002. Similarly, in the first follow-up wave for the AHEAD cohort (AHEAD 1995), about 64% of total attrition was due to death, compared to 84% for the AHEAD cohort in HRS 2002. In addition, among all the attrition in the HRS cohort, the ratio of passive to

active attrition was uniformly higher for Black sample members than for non-Blacks. By contrast, the ratio of passive to active attrition was uniformly lower for Hispanic sample members than for non-Hispanics. The same results generally held for the attrition in the AHEAD cohort, but the difference in the ratio of passive to active attrition among different race and ethnicity groups was less salient.

IV. The Effects of Sample Attrition on Selected Demographic, Economic, and Health Measures

One inherent problem with a panel survey suffering from sample attrition is that information at a particular wave *t* is not available for those who attrite at wave *t*. To examine the impact of attrition on the quality of panel data, therefore, researchers often have to rely on lagged information (information at wave *t-1* or at baseline) for attrited sample members.

Table 5 displays mean values of a set of demographic, economic, and health measures in different waves of the HRS by attrition status and by the mode of attrition. For each measure examined at wave *t*, the following three comparisons can be made in the table: (a) wave (*t+1*) non-attritors versus the entire sample at wave *t* (Column (4) vs. Column (3)) ; (b) wave (*t+1*) attritors versus wave (*t+1*) non-attritors (Column (5) vs. Column (4)); and (c) wave (*t+1*) passive attritors versus wave (*t+1*) active attritors (Column (6) vs. Column (7)). Comparison (a) is intended to assess the impact of attrition on the representativeness of the remaining sample. The focus is on one statistical feature of the mean values between the overall sample at wave *t* and the overall sample at wave (*t+1*):[2] whether the mean difference is statistically significant. If the mean difference is not statistically different in a variable, the attrition effect on the variable may be statistically ignorable, and vice versa. Similarly, comparison (b) is to assess the mean

[2] Notice that the overall sample at wave (*t+1*) is the same as the non-attritor sample at wave *t*.

difference at wave t between the wave (t+1) attritiors and the wave (t+1) non-attritors. Statistically significant differences between the two groups of people are expected, and the differences may be used to validate the results found in many previous attrition studies: attritors and non-attritors are different in many important aspects. Finally, comparison (c) is to assess the effect of the mode of attrition, whether the death-induced attrition is different from the attrition induced by any other factors.

The demographic measures considered included race and ethnicity dummies, age, gender, and year of education. The economic measures included household assets, household income, and two dummy variables representing a sample member's labor market status. The "household asset" measure was an aggregation of ten asset components, including real estate, businesses and farms, IRAs, stocks and mutual funds, bonds, checking and savings accounts, CDs, vehicles, other assets (e.g., jewels, coin collections, etc.), and debts. The "household income" measure was an aggregation of about forty components, including wage and salary, income from professional practice and trade, self-employment income, workers compensation, veterans' benefits, social security benefits, and income from various types of assets. Both the measures were based on the imputation data produced by the HRS (Cao, 2001).

The health measures examined in the table included self-reported health status (SRH), change in SRH, and cognitive score at the baseline.[3] Rated in five different categories, "Self-Reported Health Status" was defined with the following code frame: 1 = poor, 2 = fair, 3 = good, 4 = very good, and 5 = excellent. Notice that, with SRH defined this way, a larger value on the self-reported health implies better health, and vice versa. "Change in Self-Reported Health

[3] Preferably, we would like to examine wave-specific measures of individual cognitive functioning. However, because these measures were not consistent across waves (Ofstedal, McAuley, and Herzog, 2002), and there was no easy way to correct the inconsistencies without introducing other problems, we focus here on the cognitive measures at the baseline waves.

Status" was a variable about whether the current wave health status is better or worse than in the last wave. It could have one of the following three values: 1 = worse, 2 = about the same, and 3 = better. Obviously, with this code frame, a mean value of the variable greater than 2 implies improvement in health, and vice versa.

Cognitive score at the baseline was a composite generated from a series of cognitive tests administered at the baseline waves. For the HRS 1992, these included "immediate recall" of 20 short, concrete and high-frequency words (nouns) right after they were read to a respondent, "delayed recall" of the 20 words about 5 minutes after they were read, and a modified similarities test from the Wechsler Adult Intelligence Scale-revised (WAIS-R similarity test). Based on an approach similar to Knäuper, Belli, Hill, and Herzog (1997), the cognitive score in the HRS 1992 was created by adding the number of correct answers to each of the three tests, with a maximum value of 54.

For the AHEAD 1993, the cognitive tests included "immediate" and "delayed recalls" of 10 words, "series 7s test", and "tests of mental status." In the "series 7s test", respondents were asked to start at 100 and subtract by increments of seven for five trials. In the "tests of mental status", respondents were asked to count backwards from 20 for 10 continuous numbers, to name the day of the week and the date, to name the objects that "people usually use to cut paper" and the "kind of prickly plant that grows in the desert", and to name the president and vice president of the United States. Similar to that in the HRS 1992, the cognitive score in the AHEAD 1993 was created by adding the number of correct answers to each of the tests, with a maximum value of 35.

Sample attrition had very marginal effects on the mean values of the selected variables for the HRS cohort. Over the 13 variables examined in each of the 5 HRS waves, only self-

reported health showed—in both HRS 1992 and 1998—any mean differences between the overall sample and the sample comprised by the next wave non-attritors. This suggests that the representativeness of the HRS cohort has not been adversely affected by sample attrition, and any conceivable attrition biases to the measures for the HRS cohort are probably statistically ignorable.

The attrition effects for the AHEAD cohort were more visible, however. Among the 13 variables examined in each wave, age, education, cognitive score at the baseline, self-reported health, and change in the health status consistently showed mean differences between the overall sample and the sample comprised by the next wave non-attritors. Whether the differences in these variables are statistically ignorable remains to be seen.

Consistent with previous findings, attritors were different from non-attritors in both the HRS and AHEAD samples and in almost all the measures examined. On the average, Black and Hispanic respondents were more likely to attrite than White or Non-Hispanic respondents. Male respondents were more likely to attrite than female, and attritors tended to have less education. In the AHEAD cohort in particular, attritors tended to be older. In addition, attritors had lower cognitive score at the baseline interview, their self-reported health tended to be worse, and their health status tended to be worsening at the last wave of interview. They were also less likely to be still working. Further, in terms of household economic conditions, attritors had lower household income than non-attritors. Their household assets were also systematically lower.

The differences between attrition induced by different modes were almost as systematic as the differences between attritors and non-attritors. To a great degree, passive attritors were more like attritors, while active attritors were more like non-attritors. Passive attritors, for example, were generally older, had less education, had lower cognitive score at the baseline

interview, and were more likely to be male. Their health reported at the last wave of interview was worse than active attritors, and their health status was worsening. Passive attritors also had lower household income and assets, and were less likely to be still working.

V. Active versus Passive Attrition

The evidence documented above on the differences between active and passive attritors in the selected demographic, economic, and health variables prompts an obvious question: whether should active and passive attrition be treated as two distinct, independent outcomes in our attrition analysis? Alternatively, can we gain anything by differentiating attrition due to uncontrollable causes from attrition due to choice?

One straightforward procedure to address these questions is to model attrition in a multinomial logistic regression framework, and use a specification test (e.g., a X^2 test, or Hausman and McFadden (1984) IIA test) to examine whether a distinction between active and passive attrition makes sense. If the distinction is justified, the multinomial logistic framework is also useful in identifying the determinants of different modes of attrition.

To implement this procedure, we first classify the status of a HRS or AHEAD sample member i at wave t, Y_{it}, into one of the three categories: $Y_{it} = 1$ if i remains in the sample (i.e., interview provided at wave t), $Y_{it} = 2$ if i drops out of the sample because of death (passive attrition), and $Y_{it} = 3$ if i drops out of the sample because of any other reasons. We estimate a multinomial logistic regression model with this outcome variable on a set of explanatory variables—to be defined shortly—using pooled attrition data for the HRS or AHEAD cohort. The log likelihood from this regression model, called "the unconstrained model" for simplicity, is denoted as L_0. We then impose a restriction to the model that all the coefficients in the

equations for the two types of attrition are the same, and estimate the resulted "constrained model." The log likelihood from this constrained regression model is denoted as L_1. It follows that, with a null hypothesis that all the coefficients in the equations for the two types of attrition are the same—therefore, no distinction is needed between attrition due to death and attrition due to other causes—twice the difference between L_0 and L_1, or $2(L_0-L_1) \equiv S$, would be distributed as X^2 with a degree of freedom of k, where the k is the dimension of the vector of the explanatory variables.

For the explanatory variables, we include all the demographic, economic, and health measures examined earlier. In addition, we include two dummy variables indicating whether a sample member had spouse at the last wave, and whether the last interview was provided by a proxy. The idea to have the spouse dummy is to check whether presence of the spouse or partner increases the probability of staying in the sample. Presumably, this variable should have a stronger effect on active attrition than on passive attrition. The dummy variable for proxy interview is relevant in the model because a respondent may be less willing or able to stay in the sample in the future if he or she has to rely on a proxy to provide interview.

To control for the wave-specific effects on attrition, we also include in the model an index of data wave. This index is equal to "1" if the attrition outcome, Y_{it}, is for HRS 1992, "2" if Y_{it} is for HRS 1994, etc. Finally, to facilitate interpretation of the results, we convert "self-reported health" (SRH) and "change in SRH"—two multi-category health measures discussed earlier—into dummy variables, using "SRH = poor" and "Change in SRH = worse" as the reference groups, respectively. Other reference groups excluded from the model are "Other Races", "Non-Hispanic", and "Neither Working or Retired at Wave (t-1)."

The summary statistics of the outcome and explanatory variables are given in Table 6, while the results from the specification test described above are given in Table 7.

Strong evidence was found in favor of a distinction between active and passive attrition. In both the model using the HRS cohort data and the model using the AHEAD cohort data, the null hypothesis that there is no need for such a distinction was rejected resoundingly, based on the X^2 values calculated from the unconstrained and constrained regressions. This evidence paved a way for us to report in Table 8 the results from the unconstrained regression model. Had the null hypothesis not been rejected, attrition due to death and due to other causes would have to be combined. A logistic regression model with a dichotomous outcome would suffice for our analysis, and we would need an alternative explanation for the evidence documented earlier about the differences between active and passive attritors.

Active and passive attritors were indeed different, however. Compared to those still in the sample, passive attritors tended to be non-Hispanic, male, older when they were initially selected for the survey, and with lower cognitive score at the baseline wave. They were also more likely to be single the last time they were interviewed, and their last interviews were more likely to be provided by proxies. In their last interviews, they reported that they had poorer health, their health status getting worse, and that they were not working. All these results were intuitively reasonable.

Passive attritors seemed to have more education than those still in the sample. This result, however, was spurious due to strong correlations between education and health variables: the sign of the education variable would be negative once all the health variables were removed from the model. Therefore, the "net" effect of education on passive attrition was probably actually negative.

Active attritors differed from passive attritors in several important aspects. First, unlike in the passive attrition equations, which depicted a very consistent profile for passive attritors in both the HRS and AHEAD cohorts, the profile for active attritors was different in the HRS cohort from in the AHEAD cohort. Sample members who were older at their entry to the survey, for example, were less likely to attrite in the HRS cohort. But this age effect did not show up in the AHEAD cohort. The race/ethnicity effects were also different in the two cohorts: while non-White or Hispanic sample members were more likely to attrite actively in the HRS cohort, no such effect was found in the AHEAD.

Second, health status played no role in influencing active attrition. This was the case in both the HRS and AHEAD cohorts, suggesting that, in terms of their health status, active attritors were not different from those still in the sample.

Perhaps the most important result from comparing active and passive attritors was that active attritors were more like those still in the sample than different from passive attritors, as evidenced by the fact that most of the variables in the active attrition equations were not statistically significant. Obviously, this finding is a very good omen for any future efforts to further study the effects of sample attrition in the Health and Retirement Study, as active attrition occurred so far does not appear to be selective and thus is perhaps statistically ignorable.

VI. Conclusion

Over about ten years, the Health and Retirement Study has experienced significant sample attrition to two of its core cohorts. By HRS 2002, the original HRS cohort has lost more than one third of its original members, while the AHEAD cohort has lost more than half. This attrition, however, does not appear to have significantly influenced the representativeness of the

remaining samples, based on the mean comparisons between attritors and non-attritiors in a set of demographic, economic, and health measures. This result is generally in line with the findings from several major attrition studies—including Fitzgerald, Gottschalk, and Moffitt (1998), Lillard and Panis (1998), and Alderman, Behrman, Kohler, Maluccio, and Watkins (2001), which state that selection biases introduced by attrition are very mild.

Nevertheless, attritors and non-attritors are different in some important characteristics. In general, Black and Hispanic respondents are more likely to attrite than White or Non-Hispanic respondents. Male respondents are more likely to attrite than female, and attritors tend to have less education. In the AHEAD cohort in particular, attritors also tend to be older. In addition, attritors tend to have lower cognitive functioning score, and with poorer health.

Significant attrition heterogeneity has been found in our study, as suggested by the differences between active and passive attritors in our multinomial logistics regression model. Compared to those who are still in the sample, for example, passive attritors tend to be non-Hispanic, male, older, single, and with poorer health. By contrast, in many aspects, active attritors look just like those still in the sample. Their health statuses are about the same, so are their household income and assets, and their labor market status.

The finding that active attritors are similar to non-attritors has a significant implication for HRS data users: active attrition does not appear to be selective; attrition bias induced by active attrition—if any—is probably statistically ignorable.

References

Alderman, H., Behrman, J., Kohler, H., Maluccio, J., and Watkins, S. (2001). "Attrition in Longitudinal Survey Data." Demographic Research, 5, Germany: Max Planck Institute for Demographic Research.

Cao, H. (2001). "Impute: A SAS Application System for Missing Value Imputations, With Special Reference to HRS/AHEAD Income and Asset Imputations." HRS/AHEAD Documentation Report, Institute for Social Research, University of Michigan. Retrieved November 22, 2004, from http://hrsonline.isr.umich.edu/docs/userg/dr-007.pdf

Fitzgerald, J., Gottschalk, P., and Moffitt, R. (1998). An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics. Journal of Human Resources, 33(2):251-299.

Harhooff, D., Stahl, K, and Woywode, M. (1998). "Legal Form, Growth and Exit of West German Firms: Empirical Results for Manufacturing, Trade, and Service Industries", Journal of Industrial Economics, 46 (4), 453-488.

Hausman, J., and McFadden, D. (1984). "Specification Tests for the Multinomial Logit Model." Econometrica, 52 (5), 1219-1240.

Health and Retirement Study (2004). "An Overview of Health and Retirement Study Components." Retrieved December 24, 2004 from http://hrsonline.isr.umich.edu/intro/sho_uinfo.php?hfyle=overview&xtyp=2

Heeringa, S. (1997). Russia Longitudinal Monitoring Survey (RLMS): Sample Attrition, Replenishment, and Weighting in Rounds V-VII. Mimeo, Retrieved October 31, 2004, from http://www.cpc.unc.edu/projects/rlms/project/samprep.pdf

Hill, D. (2001). "Panel Attrition Bias in Health Event Models." Unpublished Manuscript, Institute for Social Research, University of Michigan.

Hirano, K., Imbens, G., Ridder, G., and Rubin, D. (2001). "Combining Panel Data Sets with Attrition and Refreshment Samples." Econometrica, 69, 1645-1659.

Inkmann, J. (2001). "Accounting for Nonresponse Heterogeneity in Panel Data." Discussion Paper, Department of Economics and Center of Finance and Econometrics, University of Konstanz, Germany. Retrieved October 31, 2004, from http://econometrics.wiwi.uni-konstanz.de/CoFE/Papers/dp01_03.pdf

Jacomb, P., Jorm, A., Korten, A., Christensen, H, and Henderson, S. (2002). "Predictors of Refusal to Participate: A Longitudinal Health Survey of the Elderly in Australia." *BMC Public Health*. 2002; 2**:** 4.

Lillard, L., and Panis, C. (1998). "Panel Attrition from the Panel Study of Income Dynamics: Household Income, Marital Status, and Mortality." Journal of Human Resources, 33, 437-457.

Knäuper, B., Belli, R., Hill, D., and Herzog, A. (1997). “Question Difficulty and Respondents’ Cognitive Ability: The Effect of Data Quality.” Journal of Official Statistics, 13 (2), 181-199.

Ofstedal, M.B, McAuley, G., and Herzog, A. (2002). “Documentation of Cognitive Functioning Measures in the Health and Retirement Study.” HRS/AHEAD Documentation Report, Institute for Social Research, University of Michigan. Retrieved November 22, 2004, from http://hrsonline.isr.umich.edu/docs/userg/dr-006.pdf

Ridder, G., (1992). "An Empirical Evaluation of Some Models for Non-random Attrition in 35 Panel Data." Structural Change and Economic Dynamics 3:2, 337-355.

Table 1. Sample Attrition in the HRS:
HRS Cohort 1992-2002 and AHEAD Cohort 1993-2002

	Sample Size	Number of Attrition	Marginal Attrition Rate (%)	Cumulative Attrition Rate (%)	Cumulative Retention Rate (%)
HRS Cohort					
HRS 1992	12,544	--	100.0	--	100.0
HRS 1994	11,317	1,227	90.2	9.8	90.2
HRS 1996	10,315	1,002	91.1	17.8	82.2
HRS 1998	9,499	816	92.1	24.3	75.7
HRS 2000	8,741	758	92.0	30.3	69.7
HRS 2002	8,039	702	92.0	35.9	64.1
AHEAD Cohort					
AHEAD 1993	8,222	--	100.0	--	100.0
AHEAD 1995	6,952	1,270	84.6	15.4	84.6
AHEAD 1998	5,697	1,255	81.9	30.7	69.3
AHEAD 2000	4,660	1,037	81.8	43.3	56.7
AHEAD 2002	3,731	929	80.1	54.6	45.4

Table 2. Sample Attrition in the HRS By Race/Ethnicity:
HRS Cohort 1992-2002 and AHEAD Cohort 1993-2002

Panel 1. HRS Cohort

	Sample Size	Number of Attrition	Marginal Attrition Rate (%)	Cumulative Attrition Rate (%)	Cumulative Retention Rate (%)
White					
HRS 1992	9,965	--	--	--	100.0
HRS 1994	9,074	891	8.9	8.9	91.1
HRS 1996	8,347	727	8.0	16.2	83.8
HRS 1998	7,744	603	7.2	22.3	77.7
HRS 2000	7,158	586	7.6	28.2	71.8
HRS 2002	6,615	543	7.6	33.6	66.4
Black					
HRS 1992	2,079	--	--	--	100.0
HRS 1994	1,819	260	12.5	12.5	87.5
HRS 1996	1,603	216	11.9	22.9	77.1
HRS 1998	1,421	182	11.4	31.6	68.4
HRS 2000	1,281	140	9.9	38.4	61.6
HRS 2002	1,150	131	10.2	44.7	55.3
Hispanic					
HRS 1992	1,158	--	--	--	100.0
HRS 1994	973	185	16.0	16.0	84.0
HRS 1996	856	117	12.0	26.1	73.9
HRS 1998	776	80	9.3	33.0	67.0
HRS 2000	709	67	8.6	38.8	61.2
HRS 2002	635	74	10.4	45.2	54.8

Note: “Black” and “Hispanic” are not mutually exclusive.

Table 2. Sample Attrition in the HRS By Race/Ethnicity: HRS Cohort 1992-2002 and AHEAD Cohort 1993-2002 (Continued)

Panel 2. AHEAD Cohort

	Sample Size	Number of Attrition	Marginal Attrition Rate (%)	Cumulative Attrition Rate (%)	Cumulative Retention Rate (%)
White					
AHEAD 1993	6,937	--	--	--	100.0
AHEAD 1995	5,915	1,022	14.7	14.7	85.3
AHEAD 1998	4,875	1,040	17.6	29.7	70.3
AHEAD 2000	4,005	8,70	17.8	42.3	57.7
AHEAD 2002	3,233	772	19.3	53.4	46.6
Black					
AHEAD 1993	1,114	--	--	--	100.0
AHEAD 1995	898	216	19.4	19.4	80.6
AHEAD 1998	717	181	20.2	35.6	64.4
AHEAD 2000	562	155	21.6	49.6	50.4
AHEAD 2002	431	131	23.3	61.3	38.7
Hispanic					
AHEAD 1993	487	--	--	--	100.0
AHEAD 1995	402	85	17.5	17.5	82.5
AHEAD 1998	338	64	15.9	30.6	69.4
AHEAD 2000	292	46	13.7	40.0	60.0
AHEAD 2002	230	62	21.2	52.8	47.2

Note: “Black” and “Hispanic” are not mutually exclusive.

Table 3. Active versus Passive Sample Attrition in the HRS: HRS Cohort 1992-2002 and AHEAD Cohort 1993-2002

	Sample Size	Number of Attrition	Passive Attrition		Active Attrition	
			N	%	N	%
HRS Cohort						
HRS 1992	12,544	--	--	--	--	--
HRS 1994	11,317	1,227	226	18.4	1,001	81.6
HRS 1996	10,315	1,002	246	24.6	756	75.4
HRS 1998	9,499	816	246	30.1	570	69.9
HRS 2000	8,741	758	281	37.1	477	62.9
HRS 2002	8,039	702	332	47.3	370	52.7
AHEAD Cohort						
AHEAD 1993	8,222	--	--	--	--	--
AHEAD 1995	6,952	1,270	810	63.8	460	36.2
AHEAD 1998	5,697	1,255	945	75.3	310	24.7
AHEAD 2000	4,660	1,037	838	80.8	199	19.2
AHEAD 2002	3,731	929	780	84.0	149	16.0

Table 4. Active versus Passive Sample Attrition in HRS By Race/Ethnicity: HRS Cohort 1992-2002 and AHEAD Cohort 1993-2002

Panel 1. HRS Cohort

	Sample Size	Number of Attrition	Passive Attrition		Active Attrition	
			N	%	N	%
White						
HRS 1992	9,965	--	--	--	--	--
HRS 1994	9,074	891	158	17.7	733	82.3
HRS 1996	8,347	727	167	23.0	560	77.0
HRS 1998	7,744	603	181	30.0	422	70.0
HRS 2000	7,158	586	217	37.0	369	63.0
HRS 2002	6,615	543	243	44.8	300	55.2
Black						
HRS 1992	2,079	--	--	--	--	--
HRS 1994	1,819	260	64	24.6	196	75.4
HRS 1996	1,603	216	64	29.6	152	70.4
HRS 1998	1,421	182	59	32.4	123	67.6
HRS 2000	1,281	140	56	40.0	84	60.0
HRS 2002	1,150	131	74	56.5	57	43.5
Hispanic						
HRS 1992	1,158	--	--	--	--	--
HRS 1994	973	185	16	8.6	169	91.4
HRS 1996	856	117	13	11.1	104	88.9
HRS 1998	776	80	20	25.0	60	75.0
HRS 2000	709	67	17	25.4	50	74.6
HRS 2002	635	74	29	39.2	45	60.8

Note: “Black” and “Hispanic” are not mutually exclusive.

Table 4. Active versus Passive Sample Attrition in HRS By Race/Ethnicity: HRS Cohort 1992-2002 and AHEAD Cohort 1993-2002 (Continued)

Panel 2. AHEAD Cohort

	Sample Size	Number of Attrition	Passive Attrition		Active Attrition	
			N	%	N	%
White						
AHEAD 1993	6,937	--	--	--	--	--
AHEAD 1995	5,915	1,022	655	64.1	367	35.9
AHEAD 1998	4,875	1,040	782	75.2	258	24.8
AHEAD 2000	4,005	8,70	708	81.4	162	18.6
AHEAD 2002	3,233	772	640	82.9	132	17.1
Black						
AHEAD 1993	1,114	--	--	--	--	--
AHEAD 1995	898	216	136	63.0	80	37.0
AHEAD 1998	717	181	142	78.5	39	21.5
AHEAD 2000	562	155	121	78.1	34	21.9
AHEAD 2002	431	131	118	90.1	13	9.9
Hispanic						
AHEAD 1993	487	--	--	--	--	--
AHEAD 1995	402	85	49	57.6	36	42.4
AHEAD 1998	338	64	41	64.1	23	35.9
AHEAD 2000	292	46	36	78.3	10	21.7
AHEAD 2002	230	62	53	85.5	9	14.5

Note: “Black” and “Hispanic” are not mutually exclusive.

Table 5. Effects of Sample Attrition on Selected Demographic, Health and Economic Variables in the HRS: Mean Values by Whether Attrited Next Wave and by the Type of Attrition

Panel 1. HRS Cohort

HRS Wave	Variable	All (3)	Non-Attritors (4)	Attritors		
				All Next Wave Attritors (5)	Passive Attritors Next Wave (6)	Active Attritors Next Wave (7)
HRS 1992	White	0.86	0.87	0.80 [b]	0.81	0.80
	Black	0.10	0.10	0.14 [b]	0.18	0.13
	Hispanic	0.06	0.06	0.11 [b]	0.05 [c]	0.12
	Age	55.56	55.55	55.58	56.69 [c]	55.36
	Male	0.48	0.47	0.54 [b]	0.59	0.53
	Year of Education	12.34	12.40	11.81 [b]	11.25 [c]	11.92
	Cognitive Score	19.68	19.82	18.27 [b]	16.74 [c]	18.57
	Household Income ($)	62,026	62,296	59,349	44,997 [c]	62,144
	Household Assets ($)	186,537	187,926	172,756	122,879	182,467
	Self-Reported Health (SRH)	3.49	3.53 [a]	3.15 [b]	2.16 [c]	3.35
	Change in SRH	2.03	2.04	1.95 [b]	1.73 [c]	1.99
	Still Working	0.67	0.67	0.63 [b]	0.37 [c]	0.68
	Already Retired	0.09	0.09	0.09	0.14 [c]	0.08
	N	12,544	11,317	1,227	226	1001
HRS 1994	White	0.87	0.87	0.81 [b]	0.74 [c]	0.83
	Black	0.10	0.10	0.14 [b]	0.19 [c]	0.12
	Hispanic	0.06	0.06	0.09 [b]	0.04 [c]	0.10
	Age	55.55	55.56	55.44	57.98 [c]	57.11
	Male	0.47	0.46	0.54 [b]	0.68 [c]	0.50
	Year of Education	12.40	12.45	11.86 [b]	11.68	11.92
	Cognitive Score	19.82	19.93	18.73 [b]	17.91	18.97
	Household Income ($)	62,296	62,226	63,059 [b]	39,104 [c]	56,721
	Household Assets ($)	187,926	187,504	192,520	180,672	225,241
	Self-Reported Health (SRH)	3.43	3.46	3.12 [b]	2.14 [c]	3.40
	Change in SRH	1.96	1.97	1.89 [b]	1.66 [c]	1.96
	Still Working	0.60	0.61	0.53 [b]	0.34 [c]	0.59
	Already Retired	0.16	0.16	0.15	0.20	0.14
	N	11,317	10,315	1,002	246	756
HRS 1996	White	0.87	0.88	0.82 [b]	0.80	0.83
	Black	0.10	0.09	0.15 [b]	0.17	0.14
	Hispanic	0.06	0.06	0.08 [b]	0.07	0.08
	Age	59.44	59.42	59.65	60.19 [c]	59.45
	Male	0.46	0.46	0.50 [b]	0.54	0.49

	Year of Education	12.45	12.48	12.03 [b]	11.66	12.16
	Cognitive Score	19.93	20.04	18.55 [b]	17.60 [c]	18.91
	Household Income ($)	57,402	58,128	48,431 [b]	31,920 [c]	54,779
	Household Assets ($)	236,129	243,871	140,456 [b]	81,041 [c]	163,297
	Self-Reported Health (SRH)	3.44	3.47	3.06 [b]	2.33 [c]	3.33
	Change in SRH	1.93	1.94	1.8 [b] 3	1.59 [c]	1.92
	Still Working	0.54	0.54	0.48 [b]	0.28 [c]	0.56
	Already Retired	0.26	0.25	0.27	0.31	0.25
	N	10,315	9,449	816	246	570
HRS 1998	White	0.88	0.88	0.86	0.84	0.86
	Black	0.09	0.09	0.11	0.12	0.10
	Hispanic	0.06	0.05	0.06	0.04	0.07
	Age	61.29	61.26	61.61 [b]	62.21 [c]	61.27
	Male	0.46	0.46	0.51 [b]	0.60 [c]	0.45
	Year of Education	12.48	12.53	11.95 [b]	11.40 [c]	12.26
	Cognitive Score	20.04	20.13	18.93 [b]	17.92 [c]	19.52
	Household Income ($)	63,026	63,315	59,550	38,077	71,822
	Household Assets ($)	274,030	278,967	214,814 [b]	117,851 [c]	270,232
	Self-Reported Health (SRH)	3.23	3.27 [a]	2.77 [b]	2.06 [c]	3.17
	Change in SRH	1.88	1.89	1.77 [b]	1.55 [c]	1.89
	Still Working	0.48	0.49	0.36 [b]	0.18 [c]	0.47
	Already Retired	0.35	0.34	0.40 [b]	0.51 [c]	0.33
	N	9,449	8,741	758	281	477
HRS 2000	White	0.88	0.88	0.84 [b]	0.80 [c]	0.87
	Black	0.09	0.09	0.13 [b]	0.17 [c]	0.10
	Hispanic	0.05	0.05	0.07	0.04	0.08
	Age	63.21	63.19	63.48 [b]	64.40 [c]	62.76
	Male	0.46	0.45	0.49	0.52	0.46
	Year of Education	12.53	12.56	12.06 [b]	11.66 [c]	12.37
	Cognitive Score	20.13	20.24	18.80 [b]	18.19	19.28
	Household Income ($)	62,156	63,009	51,946 [b]	37,349 [c]	63,444
	Household Assets ($)	326,491	332,923	249,539	138,837 [c]	336,744
	Self-Reported Health (SRH)	3.34	3.37	2.90 [b]	2.36 [c]	3.33
	Change in SRH	1.89	1.90	1.78 [b]	1.64 [c]	1.90
	Still Working	0.41	0.42	0.36 [b]	0.21 [c]	0.48
	Already Retired	0.41	0.41	0.40	0.45 [c]	0.36
	N	8,741	8,039	702	332	370

Note: Weighted results. All variables were dichotomous except for otherwise stated. “Black” and “Hispanic” are not mutually exclusive. “Household Income” and “Household Assets” were in the 2000 dollars. “Self-Reported Health” were a variable with the following values: 1 = poor, 2 = fair, 3 = good, 4 = very good, and 5 = excellent. “Change in Self-Reported Health” was a variable about whether the current wave health status was better or worse than in the last wave. It had the following three values: 1 = worse, 2 = about the same, and 3 = better. Superscript *a* indicates a statistically significant (p = 0.05) difference between Columns (3) and (4), *b* between Columns (4) and (5), and *c* between Columns (6) and (7).

Table 5. Effects of Sample Attrition on Selected Demographic, Health and Economic Variables in the HRS: Mean Values by Whether Attrited Next Wave and by the Type of Attrition (Continued)

Panel 2. AHEAD Cohort

AHEAD Wave	Variable	All (3)	Non-Attritors (4)	Attritors		
				All Next Wave Attritors (5)	Passive Attritors Next Wave (6)	Active Attritors Next Wave (7)
AHEAD 1993	White	0.90	0.91	0.87 [b]	0.88	0.86
	Black	0.08	0.08	0.11 [b]	0.10	0.12
	Hispanic	0.04	0.03	0.04	0.04	0.04
	Age	77.18	76.77 [a]	79.36 [b]	80.55 [c]	77.17
	Male	0.40	0.39	0.44 [b]	0.46 [c]	0.39
	Year of Education	11.06	11.18 [a]	10.40 [b]	10.28	10.62
	Cognitive Score	20.06	20.31 [a]	18.72 [b]	18.24 [c]	19.60
	Household Income ($)	31,109	31,525	28,883 [b]	26,760 [c]	32,808
	Household Assets ($)	140,149	146,994	103,522 [b]	97,042	115,497
	Self-Reported Health (SRH)	2.97	3.05 [a]	2.50 [b]	2.26 [c]	2.95
	Change in SRH	1.34	1.33	1.39 [b]	1.41	1.36
	Still Working	0.09	0.11 [a]	0.04 [b]	0.02 [c]	0.07
	Already Retired	0.87	0.85 [a]	0.93 [b]	0.95 [c]	0.90
	N	8,222	6,952	1,270	810	460
AHEAD 1995	White	0.91	0.91	0.90	0.90	0.90
	Black	0.08	0.07	0.08	0.09	0.07
	Hispanic	0.03	0.04	0.03	0.03	0.04
	Age	78.77	78.22 [a]	81.14 [b]	82.16 [c]	77.82
	Male	0.39	0.38	0.43 [b]	0.45 [c]	0.36
	Year of Education	11.18	11.34 [a]	10.52 [b]	10.54	10.45
	Cognitive Score	20.31	20.68 [a]	18.67 [b]	18.45 [c]	19.39
	Household Income ($)	33,247	34,449	27,972 [b]	28,313	26,864
	Household Assets ($)	260,475	264,285	243,756	223,815	308,389
	Self-Reported Health (SRH)	2.97	3.09 [a]	2.47 [b]	2.30 [c]	3.01
	Change in SRH	1.77	1.81 [a]	1.60 [b]	1.53 [c]	1.82
	Still Working	0.07	0.08	0.04 [b]	0.03 [c]	0.08
	Already Retired	0.76	0.75	0.77	0.77	0.77
	N	6,952	5,697	1,255	945	310
AHEAD 1998	White	0.91	0.91	0.90	0.91	0.87
	Black	0.07	0.07	0.09	0.08	0.12
	Hispanic	0.04	0.04	0.03	0.03	0.04

	Age	80.48	79.81 [a]	83.39 [b]	84.02 [c]	80.60
	Male	0.38	0.37	0.42 [b]	0.44 [c]	0.34
	Year of Education	11.34	11.48	10.72 [b]	10.67	10.92
	Cognitive Score	20.68	21.07 [a]	18.99 [b]	18.70 [c]	20.26
	Household Income ($)	33,377	34,687	27,616 [b]	26,313 [c]	33,337
	Household Assets ($)	212,350	212,910	209,884	218,468	172,192
	Self-Reported Health (SRH)	2.78	2.89 [a]	2.32 [b]	2.23 [c]	2.68
	Change in SRH	1.71	1.75 [a]	1.53 [b]	1.49 [c]	1.70
	Still Working	0.07	0.08 [a]	0.02 [b]	0.01 [c]	0.05
	Already Retired	0.75	0.75	0.75	0.76	0.72
	N	5,697	4,660	1,037	838	199
AHEAD 2000	White	0.91	0.92	0.88 [b]	0.88	0.92
	Black	0.07	0.07	0.09 [b]	0.10	0.05
	Hispanic	0.04	0.04	0.04	0.04	0.04
	Age	81.90	81.35 [a]	84.09 [b]	84.56 [c]	81.29
	Male	0.37	0.36	0.40 [b]	0.41	0.36
	Year of Education	11.48	11.65 [a]	10.80 [b]	10.83	10.60
	Cognitive Score	21.07	21.47 [a]	19.50 [b]	19.31 [c]	20.61
	Household Income ($)	32,228	33,791	26,061 [b]	25,654	28,460
	Household Assets ($)	229,411	244,208	171,009 [b]	177,108	135,050
	Self-Reported Health (SRH)	2.85	2.99 [a]	2.32 [b]	2.25 [c]	2.70
	Change in SRH	1.70	1.74 [a]	1.52 [b]	1.49 [c]	1.72
	Still Working	0.06	0.06	0.03 [b]	0.03	0.03
	Already Retired	0.71	0.72	0.70 [b]	0.70	0.69
	N	4,660	3,731	929	780	149

Note: Weighted results. All variables were dichotomous except for otherwise stated. "Black" and "Hispanic" are not mutually exclusive. "Household Income" and "Household Assets" were in the 2000 dollars. "Self-Reported Health" were a variable with the following values: 1 = poor, 2 = fair, 3 = good, 4 = very good, and 5 = excellent. "Change in Self-Reported Health" was a variable about whether the current wave health status was better or worse than in the last wave. It had the following three values: 1 = worse, 2 = about the same, and 3 = better. Superscript *a* indicates a statistically significant ($p = 0.05$) difference between Columns (3) and (4), *b* between Columns (4) and (5), and *c* between Columns (6) and (7).

Table 6. Summary Statistics of the Dependent and Explanatory Variables in the Pooled Data

Panel A. HRS Cohort

Variable Name		Mean	Std. Dev.	Minimum	Maximum
Active Attrition		0.06	0.24	0	1
Passive Attrition		0.02	0.15	0	1
Still in Sample		0.92	0.27	0	1
White		0.87	0.33	0	1
Black		0.10	0.30	0	1
Other Races		0.03	0.18	0	1
Hispanic		0.06	0.28	0	1
Non-Hispanic		0.94	0.23	0	1
Male		0.47	0.50	0	1
Age at Entry		55.55	3.24	39	74
Years of Education		12.43	3.01	0	17
Cognitive Score at the Baseline		19.90	6.65	0	53
Wave Index t		2.83	1.41	1	5
Coupled Respondent at Wave (t-1)		0.74	0.44	1	1
Proxy Interview at Wave (t-1)		0.05	0.22	0	1
Self-Rated Health (SRH) at Wave (t-1)	Poor	0.07	0.26	0	1
	Fair	0.15	0.35	0	1
	Good	0.29	0.45	0	1
	Very Good	0.31	0.46	0	1
	Excellent	0.18	0.39	0	1
Change in Health at Wave (t-1)	Worse than last Wave	0.18	0.38	0	1
	About the Same	0.70	0.46	0	1
	Better than Last Wave	0.12	0.33	0	1
Working at Wave (t-1)		0.55	0.50	0	1
Retired at Wave (t-1)		0.24	0.43	0	1
Neither Working or Retired at Wave (t-1)		0.23	0.42	0	1
Household Income at Wave (t-1)		61,438	94,433	0	5,339,961
Household Assets at Wave (t-1)		243,314	746,933	-911,656	50,200,000
N		52,416			

Table 6. Summary Statistics of the Dependent and Explanatory Variables (Continued) in the Pooled Data

Panel B. AHEAD Cohort

Variable Name		Mean	Std. Dev.	Minimum	Maximum
Active Attrition		0.04	0.20	0	1
Passive Attrition		0.14	0.34	0	1
Still in Sample		0.82	0.38	0	1
White		0.91	0.29	0	1
Black		0.08	0.30	0	1
Other Races		0.02	0.13	0	1
Hispanic		0.04	0.19	0	1
Non-Hispanic		0.96	0.19	0	1
Male		0.39	0.49	0	1
Age at Entry		76.56	6.04	47	103
Years of Education		11.23	3.54	0	17
Cognitive Score at the Baseline		20.45	5.60	1	35
Wave Index t		2.94	1.50	1	5
Coupled Respondent at Wave (t-1)		0.45	0.50	0	1
Proxy Interview at Wave (t-1)		0.13	0.33	0	1
Self-Rated Health (SRH) at Wave (t-1)	Poor	0.13	0.34	0	1
	Fair	0.24	0.43	0	1
	Good	0.31	0.46	0	1
	Very Good	0.23	0.42	0	1
	Excellent	0.09	0.29	0	1
Change in Health at Wave (t-1)	Worse than last Wave	0.51	0.50	0	1
	About the Same	0.39	0.49	0	1
	Better than Last Wave	0.11	0.31	0	1
Working at Wave (t-1)		0.08	0.26	0	1
Retired at Wave (t-1)		0.78	0.41	0	1
Neither Working nor Retired at Wave (t-1)		0.15	0.36	0	1
Household Income at Wave (t-1)		32,399	42,638	0	1,209,420
Household Assets at Wave (t-1)		205,184	745,479	-1,055,966	48,197,740
N		25,531			

Note: Weighted results. “Black” and “Hispanic” are not mutually exclusive. “Household Income” and “Household Assets” were in the 2000 dollars.

Table 7. Active versus Passive Attrition: Specification Test Based on Multinomial Logit Model

Model	L_0	L_1	$S \equiv 2(L_0-L_1)$	Degree of Freedom (k)	$Pr(X^2(k)>S)$
HRS	-12,716.3	-13,179.8	926.9	20	0.000
AHEAD	-11,658.2	-12,017.9	719.5	20	0.000

Note: The null hypothesis is that the distinction between active and passive attrition would not make any differences in the model. This hypothesis is rejected for both the model using pooled HRS data and the model using pooled AHEAD data.

Table 8. Active versus Passive Attrition in the HRS and AHEAD Cohorts: Multinomial Logistic Regression Coefficients

Explanatory Variable		HRS Cohort		AHEAD Cohort	
		Passive Attrition	Active Attrition	Passive Attrition	Active Attrition
White		-.121	-.314*	.135	-.270
Black		.030	-.096	.129	-.227
Hispanic		-.501*	.448*	-.260*	-.203
Male		.562*	.049	.444*	-.129
Age at Entry		.062*	-.022*	.090*	-.006
Years of Education		.034*	-.007	.024*	-.037*
Cognitive Score at the Baseline		-.014*	-.015*	-.024*	-.016
Wave Index t		.160*	-.151*	.200*	-.168*
Coupled Respondent at Wave (t-1)		-.197*	-.097	.085	.136
Proxy Interview at Wave (t-1)		.579*	.791*	.845*	.547*
Self-Rated Health (SRH) at Wave (t-1)	Fair	-.873*	-.112	-.666*	.117
	Good	-1.654*	-.006	-1.160*	.088
	Very Good	-2.185*	-.194	-1.500*	-.050
	Excellent	-2.426*	-.140	-1.543*	-.156
Change in Health at Wave (t-1)	About the Same	-.374*	.068	-.200*	.169
	Better than Last Wave	-.144	-.034	.051	.052
Working at Wave (t-1)		-.542*	.022	-.544*	-.187
Retired at Wave (t-1)		-.041	-.011	-.020	.064
Household Income at Wave (t-1)		-.001	.000	-.002	.000
Household Assets at Wave (t-1)		-.000	-.000	.000	.000
Constant		-5.954*	-.500	-8.645*	-1.233
Log Likelihood		-12,716.3		-11658.2	
N		52,416		25,531	

Note: Weighted results. The comparison outcome is “Still in sample.” Both “Household Income at Wave (t-1)” and “Household Assets at Wave (t-1)” are in the 2000 dollars.