The Wealth of Wealthholders:
Evidence and Implications from the Vanguard Research Initiative

John Ameriks, Vanguard Group

Andrew Caplin, New York University and NBER

Minjoon Lee, University of Michigan

Matthew D. Shapiro, University of Michigan and NBER

Christopher Tonetti, Stanford University

December 18, 2014

The Wealth of Wealthholders:
Evidence and Implications from the Vanguard Research Initiative

ABSTRACT

This paper introduces the Vanguard Research Initiative (VRI)—a new panel study of clients of the Vanguard Group combining survey and administrative data—that is designed to yield high-quality measurements of a large sample of older Americans who face meaningful financial tradeoffs approaching and during retirement. The VRI features an account-by-account approach to asset measurement. The paper shows that this approach is highly successful in eliciting measures of wealth. It also shows that, relative to the Health and Retirement Study (HRS) and the Survey of Consumer Finances (SCF), the VRI provides a larger sample of older Americans with significant assets. Once eligibility for the VRI, based on having nontrivial assets and Internet access, is imposed on the HRS and SCF, the VRI sample has similar characteristics as the population, so inferences from it are not specific to Vanguard clients. The paper shows that the relationship between wealth and expected retirement date is very different in the VRI from in the HRS and SCF—mainly because the HRS and SCF have so few observations where wealth levels are high enough to finance substantial consumption during retirement.

John Ameriks
The Vanguard Group, Inc.
PO Box 2600
Valley Forge PA 19482-2600
john_ameriks@vanguard.com

Andrew Caplin
Department of Economics
New York University
New York, NY 10012
 and NBER
andrew.caplin@nyu.edu

Minjoon Lee
Department of Economics
 and Survey Research Center
University of Michigan
Ann Arbor, MI 48109-1248
minjoon@umich.edu

Matthew D. Shapiro
Department of Economics
 and Survey Research Center
University of Michigan
Ann Arbor, MI 48109-1248
 and NBER
shapiro@umich.edu

Christopher Tonetti
Graduate School of Business
Stanford University
Stanford CA, 94305-7298
tonetti@stanford.edu

1. Introduction

As defined benefit pension plans become rare and as the generosity of a pay-as-you-go Social Security system becomes increasingly limited by aging of the population, households are increasingly responsible for financing their own retirement. Hence, understanding how individuals' financial assets affect their retirement decisions and well-being in retirement is of utmost importance for understanding behavior and the welfare of the retired population, as well as policy changes that may affect them. Though the transition from a defined benefit to a defined contribution retirement system has been underway for decades, the majority of households approaching or in retirement have relatively low financial assets. Datasets designed to represent the population, therefore, will have surprisingly little information on older Americans with wealth sufficient to finance a non-trivial fraction of their retirement consumption. Our research fills this gap by producing an innovative new dataset containing a large number of households with meaningful financial assets that they can potentially use for retirement income. To highlight the value of these new data, the paper shows that the relationship between wealth and retirement plans differs dramatically over the range of wealth that is sufficient to sustain consumption in retirement compared to that in the population reflected in standard datasets such as the Health and Retirement Study and the Survey of Consumer Finance.

This paper presents results from a collaboration between the Vanguard Group Inc. and academic researchers to provide measurements and analysis surrounding the behavior, preferences, expectations, and information of older Americans managing spending in retirement. Specifically, the paper presents findings from the Vanguard Research Initiative (VRI), which provides high-quality, linked administrative and survey data on a large sample of households that face or will soon face the problem of managing assets in retirement. The VRI attempts to

improve on measurements of surveys that are justifiably called the gold standard for wealth

measurement—namely, the Health and Retirement Study (HRS) and the Survey of Consumer

Finances (SCF)—along multiple dimensions:

- First, we target the population of interest—i.e., older Americans with nontrivial financial assets. Even though the overall sample sizes are similar, the HRS and SCF actually have relatively small samples of this population of interest. The HRS—since it is representative of the entire age-eligible population—has many respondents with trivial wealth. The SCF—since it is representative of the overall population—has many respondents who are younger.

- Second, we make use of administrative account data to obtain a large sample of individuals with non-negligible financial wealth. Using a combination of administrative and survey data, we can address the question of whether—apart from having non-negligible wealth—the Vanguard population is different from the overall population. We draw respondents from two lines of business—those with individual and those with employer-sponsored accounts. We find that, especially for the employer-sponsored sample, the Vanguard population is broadly representative of the U.S. older population.

- Third, we take a comprehensive account-based approach to measuring assets. Under this approach, respondents are asked to report their financial assets account-by-account. The aim of this approach—which is used selectively in the HRS and SCF—is to get information from respondents in the form that they have it or think of it rather than by requesting responses using accounting or economic categories that may not be meaningful to them.

- Fourth, we employ a set of survey techniques designed to elicit more accurate survey measures of financial assets. Respondents give meaningful nicknames to their accounts. The survey provides a summary of accounts and balances at various stages, so respondents can check whether they missed or double-counted accounts or misreported balances. Respondents can then make corrections without having to reenter correct items.

- Fifth, we use the administrative data to validate the survey responses. We are able to show that our novel survey approach provides unbiased measures of the level of assets as opposed to the understatement typically observed in survey responses. Additionally, we can show that our correction mechanism does reduce the variance of response errors.

Given the cost and difficulty of collecting asset data from respondents, our use of account data

and survey data in tandem provides a roadmap for augmenting or replacing survey-based

measures of assets in large-scale surveys. Therefore, in addition to its specific findings, this

paper documents and analyzes an approach that could be applied very broadly for improving

measurement of wealth.[1]

There is a puzzling finding in the literature on wealth and retirement: even following very

large stock market declines—such as in 2000 to 2002 and 2007 to 2009—changes in wealth have

either a small or no effect on retirement or on retirement plans of older Americans. Comparing

changes in retirement rates between defined-contribution (DC) and defined-benefit (DB) pension

holders for the period 1992–1998 using the HRS, Sevak (2002) finds that DC pension holders

tended to reduce retirement age more during the stock market boom in the 1990s. Coronado and

Perozek (2003) and Kthitatrakun (2004), by comparing expected and actual retirement age, show

that a wealth gain caused by a stock market boom reduces retirement age compared to

households' expectations. Coile and Levine (2004) focus on aggregate labor supply measures

from the HRS and Current Population Survey and also find no evidence. Goda, Shoven and

Slavov (2012), Hurd, Reti and Rohwedder (2012) and Kezdi and Sevak (2004), using risky asset

holdings in the HRS data as a measure of exposure to the stock market, estimate the wealth effect

on the retirement decision and find no evidence of such an effect. Using pre- and post-crash

interviews from the CogEcon survey conducted in 2007 and 2009, McFall (2011) finds a

---

[1] This research is therefore related to an emerging program to augment or replace survey data
with administrative records, including private account data. See Gelman et al. (2014) for high-
frequency spending and income data; Kapteyn and Ypma (2007) for earnings data; Agarwal, Liu,
and Souleles (2007) for credit card data to measure the response of spending to income; Aguiar
and Hurst (2007) for linking administrative data on price paid to survey data on demographics
and time use. For the difficulties of measuring wealth and earnings in surveys, see Juster and
Smith (1997). For systematic bias in economic measurement in surveys, see Gorodnichenko and
Peter (2007) and Hurst, Li, and Pugsley (2014). See Krimmel, Moore, Sabelhaus, and Smith
(2013) for problems with the timeliness of asset data, which are addressed by the VRI approach.

relationship between wealth change and retirement age in the expected direction, though the estimated size of the effect is modest.[2]

The VRI is designed to have greater power to detect these effects by collecting a large amount of high-quality asset data for households where such changes in wealth might be relevant. Therefore, it addresses the problem, identified by Poterba (2014), Poterba, Venti, and Wise (2011) and Gustman, Steinmeier and Tabatabai (2010), that in the survey datasets commonly used in this literature, most households do not have meaningful retirement wealth or stock market exposure. In this paper, we estimate the relationship between wealth and retirement plans in the VRI, HRS, and SCF. We demonstrate that this relationship is highly non-linear and that we can estimate the effect of wealth at the relevant range of wealth levels to be significant only when we have dense observations in that range. We then show, for households with enough wealth to typically have significant stock market exposure, that the expected retirement horizon varies significantly with wealth.

Admittedly, such estimates from the VRI are "out of sample" for the population of older Americans in the US—most of whom have little wealth and little exposure to the stock market. Making such out-of-sample inferences is precisely the aim of the VRI. As noted at the outset of the paper, policy changes and changes in employer offering of pensions are pushing older Americans to save and invest for their retirement through 401(k) and similar accounts. To understand the ultimate effects of this transformation of the retirement landscape, data such as those from the VRI are essential. There is the concern, however, that the VRI respondents are

---

[2] Some studies use other sources of variations in wealth changes. Imbens, Rubin and Sacerdote (2001) use lottery windfall gains, while Holtz-Eakin, Joulfaian and Rosen (1993) use inheritance information in IRS data and Joulfaian and Wilhelm (1994) use inheritance data in PSID. Estimated effects are mostly modest, with the exception of Holtz-Eakin, Joulfaian and Rosen (1993), who find a sizeable effect.

different—not just because they have significant retirement savings, but because they are different from the population in terms of their demographics, socioeconomic status, or other characteristics. Put bluntly, are Vanguard clients so special that they are not a valid population for drawing inferences more generally? We address this question head-on in the paper. We show that Vanguard clients are different from the HRS and SCF respondents mainly because they have more financial wealth. For HRS and SCF respondents of similar wealth, education and other attributes are not that different. This is particularly true in the subsets of HRS and SCF respondents with 401(k) plans, compared to the Vanguard employer-sponsored sample. Indeed, our approach to sampling Vanguard respondents—drawing separately from individual clients and those in employer-sponsored plans—substantially obviates concerns about selection. The VRI employer-sponsored sample has a retirement/wealth relationship that looks quite similar to the overall VRI sample. This finding suggests that the findings from the VRI are driven by having dense observations of households with meaningful levels of wealth and stock market exposure, not by differences in households that select Vanguard as a financial institution.

2. Innovations in Wealth Measurement: Vanguard Research Initiative (VRI) Approach

What makes the VRI innovative? First, it surveys financial wealth by accounts, not by asset classes. Its aim is to ask households to report numbers that closely correspond to how they receive statements and to how they might classify assets. The approach avoids asking households to map their balances into accounting or economic constructs or, specifically, to do addition or distribution of amounts. Second, after each step of wealth section, the survey instrument shows the respondents the summary of their responses in tabular form and allows them to modify their answers. Third, the survey is integrated with administrative data.

Administrative data create the sample frame, allow validation of survey responses, and create a high-frequency panel of asset data. In this section of the paper, we describe the design of the VRI sample and how the wealth measurements are implemented in the VRI survey.

Table 1 shows in tabular form the main survey design elements and how they compare with those of the HRS and SCF. Section 4 provides a detailed comparison of these surveys.[3]

2.1. <u>The VRI Sample Design</u>

The administrative data and, more generally, the collaboration with Vanguard are critical in achieving the VRI objective of creating a large sample of older wealthholders. By construction, Vanguard clients have some wealth. Additionally, information in the Vanguard administrative data on customer type, account balances, age, geography, and use of the Internet are all essential for creating the sample. This information allows us to reach a large population of relevant households.

The population for the VRI is Vanguard Group account holders aged 55 and older who are Web-survey eligible (must be registered for use of the Vanguard website, have a valid email address, and have logged on in the past six months). We stratified the sample based on the following characteristics from the administrative data: individual versus employer-sponsored accounts; age; and administratively-single status. We sampled evenly from five-year age intervals from 55 to 74 and from 75 and above. For those under 65, we divided the sample evenly between the two client types. After age 65, those in the employer-sponsored line tend to exit this group as they roll over their employer-sponsored accounts into IRAs accounts (either at Vanguard or elsewhere). For this age group, we sample the types in the proportion they appear in the population.

---

[3] Ameriks et al. (2014) describes the VRI in greater detail. Readers interested in the specifics of sampling, testing, and design, as well as in more-detailed tabulations, are referred to that paper.

A variety of research questions are more difficult to answer in the context of multi-person households. There are relatively few single households in the Vanguard population. Thus, we felt it useful to oversample singles to secure an adequate sample size of singles. The administrative data contain an imperfect indicator of single status. In particular, Vanguard constructs a household indicator by using common address and joint registration. Being in a single-member household using this indicator is strongly, but not perfectly, correlated with the survey measure of single status. Using information on the relationship between the survey and administrative measures of single status in a pilot survey, we increased the sampling rate of administratively-single accounts in the production survey. See Ameriks et al. (2014, XX).

These sampling criteria are all imposed *ex ante* based on the administrative data. To draw the sample that we invited to complete the survey, we randomly selected from the specified populations of account holders. We monitored our success at hitting the desired sample proportions, but made no adjustments after drawing the sample. We did not impose quotas of any kind on responses.

2.2. Survey Measurement of Wealth in the VRI

A key innovation of the VRI approach is to elicit assets on a comprehensive, account-by-account basis. This section describes this approach. The next section will show that it yields highly accurate measurements of assets.[4] Appendix A shows screen shots of the wealth section for a hypothetical respondent. The steps in the wealth section are as follows.

---

[4] The VRI approach is unique in taking a comprehensive, account-based approach to wealth measurement. The HRS and SCF take approaches that mix the account- and asset-class approaches. For non-retirement assets, the HRS asks respondents to aggregate the balances across accounts into the following asset classes: stocks and stock mutual funds; bonds and bond mutual funds; checking, savings, and money market accounts; and CDs, government bonds, and Treasury bills. The SCF takes a mixed approach. For checking, savings/money market, and mutual funds, it asks for the number of accounts and the balance for each account. For CDs,

Step 1: Account Type. The respondents are shown a list of 15 account types divided into groups. The rows in Table 2 after Total Financial Assets show the types. The major groups are "Tax deferred-retirement accounts" (IRA, employer sponsored plans, pension with account balance, and other retirement assets); "Savings/Investment accounts that are not in a tax-deferred retirement plan or account" (checking, savings, money market mutual funds, CDs, brokerage, and directly held securities); "Insurance-related accounts" (annuities with cash value and life insurance with cash balance); "Educational accounts"; and "Other." The survey displays a table with these account types and a checkbox for having each type.

Step 2. Number of accounts. The survey shows a list of account types that the respondent has checked in step 1. The respondent is asked to indicate number of each type of account using a drop-down menu.

Step 3.   Nicknames of accounts; verification. The survey then shows a list of accounts. The respondent is asked to give a nickname for each account. After the respondent enters all the nicknames, the survey displays a summary table (see Appendix A, Figure A-A4). Respondents are then asked whether all the information is correct. If not, they are asked whether they want to correct the list of accounts (either add or delete an account type or change the number of

---

savings bonds, individual stocks, and brokerages, it asks for asset-class totals as in the HRS. For IRAs, it asks for an inventory of types of IRA (regular, Roth, rollover) and then asks for total by type.

For pensions, the HRS and SCF take a pension-by-pension approach. The SCF household head reports up to three separate pension accounts for each household member; the HRS respondent and spouse report up to three separate pension accounts. The HRS 2012 has taken a step toward creating a longitudinal record of pensions. The HRS asks about IRAs (up to three accounts per respondent and spouse) as part of the pension module. The bifurcated structure of the HRS wealth measures (household basis for non-retirement assets and individual basis for pensions and retirement accounts) results from a strategic design decision made at the outset of the HRS to collect pension data as part of the labor section rather than the wealth section.

accounts for any type). Depending on their answers, they are brought back to either step 1 or step 2.

Step 4. Balances. The survey then loops over accounts. Respondents are asked to input the balance of each account by its nickname.

Step 5. Summary table of balances; verification. The survey displays a summary table of accounts as well as a total (see Figure A-A6). For each account, there are checkboxes for "referred to records." There is also a checkbox at the bottom of the table that asks whether everything is correct. If the respondent checks "No, I need to go back and make an update," the screen updates with two checkboxes asking whether the respondent needs to add/delete accounts or correct the dollar amount. (Both can be checked. See Figure A-A7.) If the respondent indicates a need to correct amounts, the account summary table updates with a new column of checkboxes asking which need to be corrected. (See Figure A-A8.) The survey asks only for the required corrections. Specifically, if the respondent clicks on the "add/delete account" box, they are taken back to step 1 with all previous responses pre-filled. On the other hand, if the respondent needs to correct only the amounts, the survey returns to step 4. Once the respondent returns to step 5, the respondent is again asked if the answers are correct and again allowed to make corrections. There is no limit on the number of times respondents can go through the correction sequence.

After the respondent indicates that the summary table of balances needs no correction, the survey presents follow-up questions about the composition of the accounts. First, for accounts other than saving/checking/MMMF, the respondents are shown the table with balances and asked to enter the share of stock held in each account. The table updates and translates the share into dollars of stock for each account.

Finally, the respondent again sees the table with balances. The table presents a checkbox for indicating whether or not each account is held at Vanguard. This table excludes account categories not offered at Vanguard (e.g., life insurance). This step enables comparison of responses with the administrative data.

At the end of the wealth section, the survey displays a summary table of financial wealth combined with two pie charts showing the stock share in the overall portfolio and the share of wealth at Vanguard (see Figure A-A13 for an example). The survey prompts respondents to print out this page, if desired.

2.3.   Summary of VRI Wealth Measurements

Table 2 summarizes the distribution of financial assets from the survey. The mean of total financial assets (sum of accounts surveyed as described above) is over a million dollars. The median is about $660,000. Other than checking accounts, IRAs are the most common asset class and account for about one third of total assets. Employer-sponsored plans are by construction held by almost all employer-sponsored plan respondents, but are also common among individual Vanguard account holders. Similarly, mutual funds and brokerage accounts are significant non-retirement assets in the population of Vanguard account holders.

Ameriks et al. (2014) describes how we collected data on non-account-based assets (housing, businesses, etc.). That paper also describes in greater detail the findings from the account-by-account approach. Notably, respondents were perhaps surprisingly willing to provide details on many accounts. The median respondent provided information on seven accounts. One quarter provided information on 12 or more accounts. The respondents were also willing to refer to records, with the strong majority referring to records for all accounts. Hence,

it appears that our approach gives us a comprehensive and accurate measure of assets. We

provide evidence for that contention in the next section.

3. Comparing Administrative and Survey Measures of Assets

A key feature of the VRI is its combination of administrative account data and survey

measurements of assets. As discussed above, the administrative data are a powerful tool for

obtaining a sample frame for a wealth survey. Additionally, administrative data can supplement

survey data by providing alternative measures of wealth, potentially at very high frequency. The

administrative data also can be used to verify the survey measures. This section of the paper

investigates the joint measurement properties of the survey and account data both to evaluate the

quality of the VRI and to guide future use of administrative account data in surveys.

3.1.   Quantifying Response Errors

The VRI contains administrative data on the account holders' total wealth and information about

its composition. The administrative data, though exact, are not perfect. The linking of accounts

to clients might not be perfect, especially for married clients. Additionally, the administrative

data are end-of-month, so intra-month transactions and changes in value can give rise to

discrepancies between survey and administrative data. Nevertheless, the administrative wealth

data give an unusually good reference point for evaluating the quality of the survey data and vice

versa.

The administrative data are, of course, limited to accounts at Vanguard. The survey was

designed to capture all assets. To facilitate comparison of survey and administrative data, at the

end of the account section of the survey the respondent is shown a table listing each account and

the survey report of its balance. Using the same format as shown in Figure A-A6 (used records),

the respondent is asked to check a box indicating whether or not the account is at Vanguard. In this section, the survey measure of Vanguard wealth relies on these survey responses. Figure 1 shows the distribution of the survey reports of Vanguard assets relative to the administrative data. For each decile of administrative assets, the figure shows a box and whiskers diagram of the distribution of the survey report of Vanguard assets. The responses are tightly bunched along the 45-degree line, though there are also substantial outliers. There is a slight over-reporting of assets in the survey relative to the administrative data. The fraction over-reported declining as assets increase.

To shed some light on the difference between the administrative and survey measures, Table 3 splits the sample by line of business and single status. The first line of each panel shows the survey data, the second line the administrative, the third line the survey minus the administrative data, and the last line the percent difference.[5] For the employer-sponsored sample, the median difference is $890, or 0.6%; for the individual client sample, the median difference is $2,623, or 1.4%. Yet, for both samples, the interquartile ranges of the differences are substantial.

A long-standing concern in wealth measurement is that assets are under-reported— because individuals forget about accounts and because they are reluctant to share account amounts (see Juster, Smith and Stafford (1999)). The VRI, with its account-by-account approach, builds on the insights of Juster and the designers of the HRS and SCF by presenting the

---

[5] The administrative data are the weighted average of the end of month before the survey and after the survey with the weight equal the fraction of the month elapsed on the survey date. Percentage difference is calculated in the following way. Let SW and AW denote the survey wealth and the administrative wealth. Following Davis and Haltiwanger's (1992) formulation from the gross flow literature, we define the percentage difference as $2 \times (SW - AW)/(SW + AW)$. The main advantage of this formula is that it can be applied even when either SW or AW is 0.

respondents with a detailed list of asset types, so that they do not neglect to report certain items. Remarkably, the VRI data show no evidence of such under-reporting on average, so this approach appears to be effective.

A potential reason for survey over-reports is that some accounts might not be linked to the survey respondent in the administrative data. Since the administrative records are at the account-holder level, they will not include a spouse's account if it is registered solely under the spouse's name. To address this issue, we conduct the same comparison only for singles, that is, respondents who report in the survey that they are not married or partnered. The results are reported in Table 3, Panels C and D. For singles, the tendency to over-report is essentially gone. For the singles in the individual account holder sample, median deviation is almost zero (-0.03%) and the interquartile range of the deviation is -2.9% to 2.2%. The difference is most acute for the individual client sample because employer-sponsored respondents are less likely to have a family-level relationship with Vanguard. In particular, note that the large upper tail of difference in the individual sample is dramatically reduced for singles relative to the overall sample in Panel B.[6]

3.2   Corrections and Wealth Measurement

In this section, we examine how the VRI's correction mechanism works to enhance the accuracy of the account data. The survey instrument not only captures the final responses, but also saves the initial answers. Therefore, for respondents who modified their answers after seeing the summary tables, we can check whether or not their answers got closer to the

---

[6] We are also able to examine whether checking records matters for accuracy of survey responses. Interestingly, checking records shrinks the deviation of administrative and survey reports, but being logged on to the Vanguard website during the survey does not play a significant role in this result. See Ameriks et al. (2014).

administrative data. Figure 2 summarizes the paths respondents took through the wealth section

given that they have multiple opportunities to correct their account inventories and balances:

>Path 1. No corrections. About two thirds of the sample (62.49%) completed the wealth
>section without making any corrections.
>
>Path 2. Inventory corrected before balance entered; balance not corrected. About 15% of
>respondents corrected their inventory (the first checkpoint in step 3 described in Section
>2.2), but did not correct balances.
>
>Path 3. Only balance corrected. About 11% of respondents corrected their balances (the
>second checkpoint in step 5) without either previously correcting their inventory or going
>back to correct after entering balances.
>
>Path 4. Inventory corrected, then balance corrected. About 5% of respondents corrected
>their inventory, entered their balances and then corrected their balances, but did not go
>back to revise inventory subsequent to entering balances.
>
>Path 5. Non-sequential corrections. About 6% of respondents made complex
>corrections.

Specifically, these respondents typically went back to the start of the wealth section to correct the

inventory of their accounts after having entered balances. Hence, about one third used the

correction mechanism in some way.

In Table 4, we again show the percentage difference between the survey and the

administrative Vanguard wealth, but for the initial and the final survey answers separately.

Respondents are grouped according to the correcting paths they took. Again, the comparisons

are done only for singles.

When respondents did not make any corrections, their initial responses were already very

close to the administrative information. The interquartile range is -3.3% to 2.6% for those who

made no corrections; for those who corrected account inventory only, it is very similar, -3.5% to

2.5%. For respondents who corrected their balances, their initial responses seem to be noisier.

Though the median percentage difference is close to that of those who do not correct balances,

the pre-correction interquartile range for those who correct balances is much larger. After the

corrections, however, the width of the interquartile range shrinks dramatically toward that with

no corrections. Indeed, the corrected range is a bit smaller than for those who made no corrections at all. Therefore, the correction mechanism did prove to be effective.

4. Representing Wealthholders versus Representing Households: VRI, HRS, and SCF

This paper studies households with non-negligible financial wealth—in particular with sufficient wealth to face meaningful choices about financing retirement from assets. The previous sections document that the VRI provides accurate and comprehensive data on this group. This section addresses two, interrelated questions. First, why is the VRI needed? The answer is that leading surveys aimed at measuring wealth contain remarkably few respondents in the relevant age range with significant levels of wealth. Second, is the VRI—having achieved the aim of representing such wealthholders in significant numbers—unrepresentative of the population apart from having targeted individuals with non-negligible wealth? We answer these questions through a detailed comparison of the VRI with the HRS and SCF.

4.1.   Comparing VRI, HRS, and SCF Design. Table 1 summarizes and compares the overall features of the VRI, HRS, and SCF. The VRI is composed of Vanguard clients at least 55 years old with non-negligible assets. The HRS is a representative sample of those at least 50 years old and their spouses. The SCF aims to be representative of wealth across all age groups. Because high-wealth individuals are hard to survey, its frame includes a list sample of high-income households. The VRI oversamples singles and, as discussed above, screens for Web-survey eligibility and stratifies the samples by Vanguard line of business. The HRS and SCF do not impose these screens, but we use relevant variables on the HRS and SCF to construct subsets that match VRI sampling criteria.

The last panel of Table 1 shows summary statistics for the three surveys for observations that meet the VRI age-eligibility (age 55 years or older). For HRS, we use the age of the financial respondent. The VRI is comparable in size to the HRS in this age range—about 9,000 households in the VRI and about 11,500 in the HRS. The SCF has less than a third the number of respondents in this age range compared to the VRI.

The VRI sample is much more affluent than the HRS or SCF samples. Of course, by design the VRI targets wealth holders while the HRS and SCF are representative, that is, they include the very large group of older Americans with very low assets. The next set of results explores these differences and shows the extent to which they derive from VRI sampling restrictions.

4.2.   Comparing VRI, HRS, and SCF Respondents. Table 5 shows the distribution by wealth and age of raw household counts in the VRI age-eligible range of 55 years and older for the VRI, HRS, and SCF.[7] It reminds us how little financial wealth the population of older households has. The total number of observations in the VRI and HRS are comparable, but their distributions of wealth are very different. Ninety percent of the VRI respondents have financial wealth more than $100,000, and one third of them have more than a million dollars. In contrast, the HRS distribution has a very fat left tail. One third of the HRS sample has a negligible amount of financial wealth (less than $10,000) and only about a third has more than $100,000.

The SCF, which is age-representative overall, has less than a third of the number of observations in the age-eligible range compared to the VRI and HRS. With the list sample of high-income households, the SCF has disproportionately high-wealth respondents. Even so,

---

[7] The wealth measure used in the comparisons is total net financial wealth. Values of houses and mortgages are excluded. See Appendix B for the definition of the total financial wealth for each survey and how we impose similar sampling screens in the VRI, HRS, and SCF.

given that the SCF is not aiming at the population near or after retirement, for most of the wealth-age bins with non-negligible wealth, the number of households in the SCF is much smaller than in the VRI.

The age distributions are also quite different across surveys. The VRI, by construction, has a similar number of observations for age bins 55-64 and 65-74, and about half the size for age 75+. The HRS has relatively more observations in the oldest age bin, while the SCF has about half in the youngest.

These tabulations illustrate vividly how the VRI is targeted for studying the financial decisions of those approaching or in retirement with non-trivial financial wealth. Given the stark differences in the VRI wealth distribution relative to the population, we need to understand the main determinants of these differences. In particular, *does the high affluence of the VRI sample derive mainly from our sampling screens or, even taking into account these screens, is a sample based on Vanguard clients very different from the U.S. population*? In the following, we try to disentangle these effects by examining the effect of *VRI eligible* screens in the HRS and SCF. The screen requires Internet eligibility and that households have at least $10,000 in a non-transactional financial account.

These screens are restrictive in the HRS and SCF samples in this age group. Table 6 shows how the screens affect the number of eligible households by age. For the HRS and SCF, the first columns of counts impose just age-eligibility. The second columns impose "VRI eligibility" (Internet eligibility and the $10,000 minimum balance in non-transactional financial accounts). The third column imposes "401(k) subset" (at least $10,000 in a DC pension account). Note that these screens are imposed *ipso facto* in the VRI for both employer-sponsored and

17

individual client groups.[8] For HRS and SCF, the screen yields relatively small subsets of age-eligible respondents. For HRS, only about a third satisfy VRI eligibility. In the SCF, a relatively larger fraction of households satisfy these conditions owing to oversampling of high-income households. The size of the 401(k) subset group is much smaller in both the HRS and the SCF. In VRI, the age distribution is flat by design. (Everywhere, there are few of the oldest groups represented in the employer-sponsored samples because most retirees roll over their 401(k) to an IRA and therefore are represented in the individual client sample.) In the HRS and SCF, the screen has more of a bite for older groups. See Appendix C for implications for wealth by age.

In Table 7, we show that the effects of the VRI screens are similar in the HRS and SCF in terms of weighted sample.[9] Imposing Internet eligibility alone reduces the weighted sample by about half in both HRS and SCF. The asset cut-off has a similar effect. Because these two conditions are highly correlated, there is an only incremental additional effect when taken together. Within the VRI-eligible samples in both the HRS and the SCF, only half of the weighted sample has at least $10,000 in DC pension accounts.

A key question is, after imposing comparable sampling screens, how similar are the characteristics of VRI compared to those of the subsamples of the HRS and SCF? The answer is that they are not so different under VRI-equivalent sampling screens. Table 8 shows the wealth distributions from the VRI, HRS and SCF. From this point forward, HRS and SCF tabulations

---

[8] The two screens in VRI are constructed to be mutually exclusive to avoid inviting respondents twice. Therefore, the second and third columns of VRI counts sum to the first column.

[9] Up to now, we have focused on raw counts of observations in order to give a concrete sense of the size of the samples across the surveys. Since the SCF oversamples high-income individuals, these households are assigned smaller sampling weights. Similarly, the HRS oversamples blacks and Hispanics (in order to make statistically significant inferences by groups) and residents of Florida (because of the cost saving in reaching older respondents there). In the following analysis, all the comparisons are made after weighting observations from the HRS and SCF with the corresponding sampling weights.

use sampling weights. With only age eligibility, median values from the HRS and SCF are an order of magnitude smaller than the corresponding numbers from the VRI. When we impose the VRI eligibility screen, the gaps are dramatically reduced, though there are still important differences. The remaining gap is smaller if the HRS and SCF subsamples are compared with employer-sponsored sample in the VRI. The 90th percentile from the VRI-eligible subsample of the SCF is actually larger than the one from the VRI employer-sponsored group. Recall that for the employer-sponsored group the potential self-selection issue is mitigated, since the availability of Vanguard funds in their retirement plan results from their employers' decision making. To more closely mimic the asset cut-off imposed on the employer-sponsored group in the VRI, we also made tabulations on the HRS and SCF subgroup composed of households with at least $10,000 in their 401(k) or similar pension accounts. The results are reported in the third row of the HRS and SCF panels. On average, the 401(k) subset of the HRS is wealthier than the overall HRS VRI-eligible sample, while the 401(k) subset in the SCF is less wealthy. The means of the 401(k) subsets in the SCF and HRS are closer to those of the VRI employer-sponsored sample, though the VRI is less right-skewed. Nonetheless, it is reassuring that there is broad similarity between the 401(k) subsets of the SCF and HRS and the VRI employer sample.

Appendix C provides a more detailed comparison across the surveys. It compares across dimensions including income and demographics. Compared to the total population of the HRS and SCF in the same age range, the VRI sample has much more wealth, a much higher education level, better health, and a greater likelihood of being coupled. Most of these differences, however, can be explained by the effect of the sampling screens we imposed in the VRI panel. What is special about the VRI sample is that it is selected for non-trivial asset holding and use of the Internet. Once these criteria are imposed, the VRI looks quite similar to the analogous

populations in the HRS and SCF. There is a bit of residual higher education, better health, and high wealth-to-income ratio in the VRI compared to the relevant HRS and SCF populations. Yet the principal differences between the VRI and the general populations do not appear to be attributable to selection to Vanguard participation *per se*. For the employer-sponsored sample, the differences in the characteristics essentially disappear once VRI-eligible criteria are imposed on the HRS and SCF.

4.3.   Stock Share

The extent of stock ownership looms large in discussions of how individuals will manage under defined-contribution retirement plans. The VRI wealth survey asks for stock share on an account-by-account basis. Table 9 compares the stock share of the VRI with those of the HRS and SCF. Panel A reports stock shares while Panel B reports stock amount. Again, we see the importance of having a relevant sample. Compared to the VRI, if we impose only age eligibility, the HRS and SCF have much lower stock shares across almost all of the distribution. Compared to the median share of 55% in VRI, the median share is 0% in the HRS and close to 0% in the SCF. Conditioning, however, on the VRI sample screens, the median shares in HRS and SCF are still lower, but much closer to those of VRI. The left tail in the HRS still shows less stock ownership, but SCF and VRI are similar across the distribution.[10] The picture is similar with regard to the amounts of stock in panel B of Table 9. Hence, as with the level of wealth, the Vanguard respondents are less unrepresentative once the screen is imposed. But again, note that the VRI has a much larger sample of stock holders, so any analysis of portfolios should be much more precise.

---

[10] Note that the HRS 2012 stock shares in 401(k) or similar accounts are not yet cleaned and imputed, so they are excluded (numerator and denominator) from these HRS stock shares. This issue can be addressed in a subsequent draft.

5. Wealth and Retirement: Lessons from Data on Wealthholders

We have established that the VRI approach leads to substantially larger samples of older households with relevant levels of wealth for key decisions surrounding retirement and well-being in older age. Having dense observations across the relevant ranges is particularly important if the relationships between wealth and other behaviors are non-linear. Poterba, Venti and Wise (2011) show that for the majority of households surveyed in the HRS, the lack of demand for additional annuity income simply comes from having very low annuitizable wealth. Similarly, there is a substantial literature on how wealth and shocks to wealth affect retirement (e.g., Sevak (2002), Bosworth and Burtless (2010), Goda, Shoven and Slavo (2012), McFall (2011), and Coronado and Dynan (2012), among others). Again, for the majority of households that approach retirement with little financial wealth, how levels or changes in wealth affect decision-making is a very different question than for those who have significant savings for retirement.

In this section, we demonstrate that for the relationship between expected retirement date and wealth, having data that are dense in the VRI wealth ranges yields substantially clearer inferences than is possible with existing datasets. In particular, we investigate the relationship between current accumulated financial wealth and how long individuals plan to keep working.

The VRI is designed as a panel, though this paper analyzes the first survey. To study the wealth/retirement relationship, we use the relationship between retirement expectations and wealth in the cross-section.[11] Thus, we build on the tradition of using expectations rather than

---

[11] The VRI holds the promise to examine reaction to events as the panel builds over time. We do, however, have a panel aspect even with the cross-section of wealth from the survey from the administrative data. We have done some exploratory work using the administrative data panel to

realizations as the outcome variable. See McGarry (2004), Chan and Stevens (2004), and

Szinovacz, Davey, and Martin (2014). The use of subjective probability variables relies on

substantial experience showing the validity of these measures in the HRS and other surveys. See

Dominitz and Manski (1997) and Hurd and McGarry (2002).

5.1.   Specification

We examine whether low financial wealth leads to plans for longer work, and at what point

current wealth seems adequate for retirement. To be clear, we present an exploratory analysis

that is designed to reveal how data such as the VRI can shed light on variables that determine

retirement decision-making. The estimates should not be taken as a structural relationship

because of the obvious joint determination of retirement and saving.

To measure current financial wealth in a way that is meaningful for thinking about

expected retirement, we construct normalized financial wealth W_i^R as

$$W_i^R = (0.06 \times W_i \times (1.03)^{(65-age_i)}) / Y_i$$

where W_i is annuitizable financial wealth, Y_i is current income, and age_i is the current age of the

main breadwinner of the household.[12] Normalized wealth is a rough-and-ready measure of how

much *current* wealth could replace current income assuming no additional saving. See Brown

(2001) for a similar measure, but converting flows to a stock. The calculation assumes a 0.06

---

examine the effect of the financial crisis on VRI respondents. Note that the VRI was collected in
2013. By then, the stock market had recovered from the 2008/9 decline. By consulting the
administrative data, we find that most VRI respondents invested passively over the financial
crisis. That is, their stock share moved by roughly the amount consistent with little rebalancing.
As a consequence of this prudent investment strategy and the recovery of the market, there is, in
fact, little lasting effect of the crisis on VRI respondents' wealth overall.

[12] Annuitizable financial wealth is the sum of retirement and non-retirement financial assets. To
put these on the same tax basis, we use another rough-and-ready approximation. Specifically,
we presume a 25 percent average tax rate on withdrawals from qualified plans. Note that we do
not have good data separating Roth and non-Roth treatment, so all qualified plans are combined
in this calculation. The main findings are robust with respect to the assumed tax rate.

annuitization rate and a 3 percent real rate of return. The use of a fixed rate of return and a uniform annuity rate is a simple way to put current wealth of future retirees into common units. We compound returns until age 65 rather than the expected retirement date to avoid putting expected years of work on both sides of the equation. We estimate the relationship

$$H_i = \beta_1(W_i^R)W_i^R + \beta_2 Y_i^R + Z_i \gamma + \varepsilon_i \tag{1}$$

where H_i is the difference between the expected age of retirement and current age, W_i^R is normalized financial wealth, Y_i^R is expected DB pension plus Social Security divided by current income, and Z_i is a vector of covariates (age, dummies for education and health, and marital status).[13] The coefficient $\beta_1(W_i^R)$ is a potentially non-linear function of normalized wealth.

5.2. Variables and Sample for Wealth-Retirement Analysis

We focus on estimates of this relationship in the VRI and HRS. We also show the same analysis using the SCF data, but due to the small number of observations in the relevant age group and lack of some variables used—health of respondents and expected Social Security income—the results are not entirely consistent with the specification used for the VRI and HRS and the estimated relationship is much less precise. In the VRI, expected retirement is measured using the response to a question, "At what age do you expect to completely retire?"[14] Both VRI and HRS have questions about current and expected pension and Social Security income. For

---

[13] We assume that DB pension is taxed at the same 25 percent average rate as distributions from qualified plans. To account for the partial non-taxability of Social Security benefits, we apply a 15 percent average tax rate to them. The main findings are again robust with respect to different tax rates assumed.

[14] In HRS, the expected retirement age is the result of a complex sequence starting with whether an individual plans to retire and at what age or date.

singles, Y_i^R is simply the sum of expected pensions and Social Security divided by current

income. For couples, it is this sum across the couple.[15]

For simplicity, we limit the sample to households with just one main breadwinner who

has not yet retired and is aged 65 or younger. For singles, anyone not retired and is aged 65 or

younger is in the sample. For single worker couples, the household is included if the worker is

aged 65 or younger. These include single-worker households or dual-worker households in

which one is now retired. For both these households and singles, the retirement decision is for a

single worker. The assets and income used in the analysis reflect any retirement income or

assets of the already-retired spouse. For dual-worker households, the joint retirement is more

complex. We only include households that appear to have only one primary breadwinner, and

we base the retirement decision on that household member.[16] There are 2,026 households in the

VRI sample and 1,053 in the HRS sample. See Appendix D for details.

5.3.1   Estimates: Entire Sample

Figure 3 compares the distribution of normalized wealth across the VRI and HRS. The curves

shown are kernel densities where the solid lines are for the VRI while the dashed lines are for the

HRS. Panel A shows the entire sample analyzed in this section, while Panel B examines the

employer-sponsored subsets. Panel A shows the stark difference in the wealth distribution

between two surveys documented in Section 4. Recall that normalized wealth is roughly the

extent to which current wealth could replace current income at retirement if all assets were

---

[15] If one member of the couple is retired, we use the current retirement income for that person
plus the expectations for the non-retired person.
[16] To determine the primary breadwinner, we use expected Social Security income and defined
benefit pensions as a proxy for who has larger lifetime earnings. If one of the members has
expected Social Security and DB pension at least four times larger than the other earner, he or
she is classified as the main breadwinner and the household is included in the sample.
Otherwise, the household is dropped from this analysis.

devoted to retirement income. Values close to one mean that a household can live comfortably from its assets. Values substantially greater than one provide a cushion for high expenses and the likelihood of bequests. Values close to zero mean that retirement must be financed by Social Security, DB pensions, or more saving. In the VRI, observations are dense and fairly uniformly spread in the range from 0 to 0.5, and observations with normalized wealth between 0.5 and 1 are not rare. A non-negligible fraction of households have normalized financial wealth larger than 1. In contrast, in the HRS the vast majority of the households have a replacement rate lower than 0.5. A trivial fraction of observations has a replacement rate close to or higher than 1. This observation confirms the point made by Poterba, Venti, and Wise (2011): relatively few households in the broad population have meaningful levels of potentially annuitizable wealth.

Now consider the relationship between this measure of current assets and plans for continued work. To capture the non-linear relationship between retirement horizon and wealth holdings without imposing a restrictive functional form, we estimate LOESS regressions.[17] Figure 4 shows the results for the VRI and HRS. Again, Panel A shows the entire sample analyzed in this section. Panel B examines the employer-sponsored subsets.[18] In Figure 4, "x" denotes HRS (orange/dashed) and "o" denotes VRI (blue/solid). The LOESS curve is shown as a line with the shaded area indicating the 95% confidence interval. The y-axis of Figure 4 is measured in expected remaining years of work (mean zero because it is a residual). In the VRI for the entire sample in Figure 4A, we see the clear negative relationship between normalized wealth and retirement horizon up to the full replacement rate around 1. Moving from zero

---

[17] LOESS is a bivariate procedure. To deal the covariates, we first project the retirement horizon on the variables in equation (1) excluding normalized wealth. The LOESS estimate is the regression of this residual on normalized wealth. For the HRS sample, both stages used sampling weights.

[18] The ranges of the horizontal and vertical axes are truncated to exclude outliers. Appendix D, Figure A-D1, shows the data in Figure 4A for the entire sample including outliers.

annuitizable wealth to annuitizable wealth that could replace current income corresponds to a reduction in expected years of work by about 1.7 years. After that level, the estimated relationship flattens out. (For very high levels of annuitizable wealth, the bulk of wealth likely will not be used for routine consumption in retirement.) Over the entire range, the estimates are quite precise. In the HRS, the estimated relationship is very different. It shows a negative relationship up to the replacement rate 0.3, a slightly positive correlation in the range of 0.3 to 0.4, and then becomes flat after that. The change in years worked is about the same as in the VRI, but it occurs at much lower levels of annuitizable wealth. Given the low density of data in this range, the flattening of the LOESS line for higher levels of wealth occurs by construction. The HRS data simply cannot capture how the relationship changes over this range because there are so few observations.

Having ample data over the relevant ranges of wealth clearly affects the precision of the estimates. The VRI confidence interval is narrower due to the larger number of observations. The HRS confidence interval gets wider after the replacement rate of 0.25, as the number of observations gets smaller very quickly for individuals with annuitizable wealth sufficient to replace even a quarter of their income prior to retirement.

5.3.2. Estimates: Employer-Sponsored Sample

One concern about the VRI design is that the behavior of Vanguard clients might be very different from that of the general population. We can address this issue by considering whether or not the behavior of the VRI employer-sponsored sample differs from that of the individual client sample. Because the employer-sponsored clients come to Vanguard owing mainly to their employers' choices, they are much less self-selected than the individual account holders. This prior is borne out by the Section 4 results, which show that the characteristics of the VRI

employer-sponsored sample are quite similar to subsets of the HRS and SCF with DC pension accounts. Figure 3B confirms that after imposing similar screens, the distribution of normalized wealth looks much more similar across the VRI and HRS.

In Figure 4B, we show the relationship between wealth and retirement plans for the employer-sponsored samples of the VRI and HRS. The general inference drawn by comparing the VRI and HRS for the entire sample also holds for this subset, though the HRS curve is somewhat closer to the VRI curve. The HRS relationship in Panel B has a steep decline for lower levels of wealth, but then goes essentially flat as in Panel A. Likewise, the change in retirement plans show in Panel B for the VRI is larger than in the HRS over the relevant range, e.g., 0.25 to 0.75, and the HRS LOESS line is below the VRI confidence interval in this range. Hence, although the HRS estimates are quite imprecise for the 401(k) subset in Panel B owing to the paucity of data, the basic message of the entire VRI sample holds in the employer-sponsored samples. Therefore, the key results derived from the VRI appear to be driven by having dense data over relevant wealth ranges and not by self-selection by individuals into a relationship with Vanguard.

In Appendix D, we estimate the version where we include Y_i^R in normalized wealth instead of treating it as a control. Since HRS households have significant pension and Social Security wealth, the support of retirement resources is different—but less so—from the VRI than for financial resources alone. Nonetheless, a similar picture emerges in the analysis that includes Y_i^R because of the difference in financial wealth.

5.3.3. Estimates: SCF

Figure 5 reports the result from the SCF for the entire sample (Panel A) and the 401(k) subset (Panel B). Due to a small number of households in the relevant age interval, we have only 233

observations satisfying all the criteria to be included in the analysis (we use only one replicate from each household). The SCF does not have expected Social Security benefit information, so the estimates are not entirely parallel with those for the VRI and HRS, which is why we do not plot the VRI in Figure 5. The small sample size makes the estimates extremely imprecise. The LOESS curve moves substantially, but not statistically significantly. The SCF was not designed to study retirement saving, so it is not a criticism of that dataset that it has little power to address the relationship between wealth and expected retirement. Nonetheless, our finding points to the importance of collecting data that are relevant for the question.

5.4.   Implications: How Would a Stock Market Crash Affect Retirement Plans?

To get a more specific reading of the value of the VRI sample, we calculate the difference in the retirement horizon associated with the reduction in wealth associated with a 40% drop in the stock market. With the estimates using HRS data, we can replicate the small wealth effects that are found in the existing literature. The VRI estimates imply much bigger effects.

Consider the effects of a stock market contraction on the representative household in the HRS using the LOESS estimate from the HRS data. In the HRS sample used in the LOESS estimation, the mean normalized wealth is 0.163 and the mean stock share is 17%. Hence, a 40% drop in the stock market would imply a change in normalized wealth of only $0.4 \times 0.163 \times 0.17 = 0.011$. Figure 6A shows the LOESS lines from Figure 4 without the confidence intervals and observations. (Note that the vertical axis in Figure 5 has a reduced range.) If we plug this change into the LOESS estimate from the HRS (see "HRS-overall" in Figure 6A), the implied difference in the retirement horizon is small, about 0.13 years. This estimate is, for example, similar to what Goda, Shoven and Slavov (2012) find with the whole sample of the HRS (0.042 years).

The literature also considers estimates using only stock holders in the HRS. To mimic this approach, we take a representative stockholder to have mean normalized wealth of 0.51 and mean stock share of 55%, which are the means from the VRI sample used in the LOESS estimation. We apply the 40% drop in the stock market to this representative agent and examine the implications under estimates from both the HRS and VRI samples. Because the representative stockholder has more wealth and higher stock exposure, the shock arising from the stock market will be larger than for the average HRS respondents overall. This shock is represented by the same horizontal distance and is located at the same position on the horizontal axis of Figure 6A, but it is positioned along either the VRI or HRS LOESS line for comparison. Hence, the "HRS-average stockholder" and "VRI-average stockholder" have the same wealth (horizontal) changes, but will have different retirement horizon (vertical) changes because of the different position and slope of the LOESS relation in the two datasets.

The change in wealth from a 40% decline in the stock market of the representative stockholder is large, $0.4 \times 0.51 \times 0.55 = 0.112$. The relationship estimated in the HRS is, however, very flat (and even slightly positive) at this wealth level, so the implied difference in the retirement horizon is about 0.04 of a year reduction (see "HRS-average stockholder" in Figure 6A). Again, the sign is the same and the magnitude is similar to the corresponding number in Goda, Shoven and Slavov (2012) (-0.008). Now, if we examine the same shock to the same household but draw the inference from estimates in the VRI, the picture is quite different. Since large wealth change is combined with the steep slope estimated in the relevant wealth range, the implied change in the retirement horizon amounts to about 0.3 of a year (see "VRI-average stockholder" in Figure 6A). Also, narrow confidence intervals in this wealth range that we observed from the VRI curves in Figure 3 imply that the estimated effect would be

statistically significant. Hence, the representative stockholder is so poorly represented in the HRS that estimates of effects of stock market crashes on expected retirement will be very misleading using HRS data.

We consider different levels of wealth and stock exposure to further explore the relationship in the VRI and to illustrate the importance of nonlinearity in it. Figure 6B shows the VRI LOESS curve and considers the effect of the 40% decline in the stock market on different agents. The "VRI-average stockholder" is the same as just discussed in Figure 6A. The "VRI-high stock exposure" considers a household with a 70% stock share instead of 55%, but with the same level of wealth as before. Hence, the horizontal shift is larger (0.143 normalized wealth) and the difference in retirement horizon is also longer (0.40 year). We also consider a wealthier individual (1.0 normalized wealth) with an average stock share (55%), for whom the horizontal change is 0.22 in normalized wealth. Because the LOESS relationship flattens at higher wealth levels, the predicted effect on the retirement horizon is smaller (0.15 year).

The analysis in this section makes it clear, with the HRS and SCF it is hard to capture the relationship between wealth and retirement behavior of those with high levels of annuitizable wealth. Those individuals are more likely to be affected by stock market performance. The VRI, because it targets those with significant wealth, sharply shows the relationship between assets and retirement planning. In the future, as most workers rely on DC pension plans, they will need to have sufficient DC wealth in order to sustain retirement consumption. Therefore, the 0.5 to 0.75 replacement rate range that is common for VRI respondents is the relevant range for projecting the behavior of the population under an institutional and policy regime in which DC plans are the dominant source of retirement income.

6. Conclusions

This paper has introduced a new approach and new dataset—the Vanguard Research Initiative—for measuring the wealth of wealthholders. Based on a partnership between academic researchers and the Vanguard Group, we have developed a new survey-administrative dataset that provides a large, high-quality sample of households that have substantial wealth for financing retirement. Wealth measurement is based on a comprehensive account-by-account approach that is designed to elicit accurate information in the form that respondents think about it and have at their disposal. The data infrastructure makes use of high-quality administrative data at all stages of the analysis—establishing the sample frame, sending invitations, evaluating selective responses, evaluating quality of survey responses, and—ultimately—providing a distinct dataset. By collecting survey and administrative data in tandem, this project aims to demonstrate how large-scale surveys can make increasing and effective use of administrative data for wealth measurement. Given the challenges and costs of collecting surveys, these advances should inform measurement practice going forward. Based on the approach presented in this paper, it may be possible to replace expensive, infrequent, and error-ridden survey measures of wealth with administrative account data.

The research also informs practice for collecting wealth data within surveys. In particular, the account-based approach to survey measurement of wealth yields measurements that are unbiased relative to administrative measurements. In contrast, many surveys appear to undercount assets. Additionally, the paper demonstrates that the correction mechanism significantly reduces the variance of errors relative to the administrative account data.

Administrative data are by definition free from reporting error, so tend to have much less measurement error. Administrative data alone, however, might not provide enough information

for research. In many cases, they do not include a rich set of important demographic variables. Sometimes they capture only a part of the household balance sheet (examples include the administrative Vanguard wealth data used in VRI and TIAA-CREF data used in Ameriks and Zeldes (2004)). Measurement error can also occur while processing data. Browning, Crossley and Winter (2014) provide a valuable summary of these issues. Hence, to get a better picture of households' economic conditions, it is often necessary to link survey data to administrative data so that we can address the shortcomings of both types of data. As a linked dataset with, on the one hand, detailed survey measures of household finance and other economically important characteristics and, on the other hand, monthly-frequency observations on balances and compositions of their Vanguard assets, the VRI enables us not only to validate survey responses with the administrative data, but also to conduct research that requires high-frequency data on financial situations.

The design of this VRI infrastructure was targeted at measuring the wealth of households with sufficient financial assets so they faced meaningful tradeoffs about whether to work longer, whether to annuitize, whether to buy long-term health insurance, how much to bequeath, and so on. In other papers that also leverage the VRI, we are investigating some of these questions in detail. In this paper, we make several substantive contributions beyond evaluating the quality of the VRI measurement. We show that the VRI is dense in data relative to the other excellent surveys with wealth data, namely the HRS and SCF, for the age and wealth levels where individuals face meaningful financial decisions related to planning for retirement and sustaining well-being during retirement. We also show that for one key variable—how much longer they expect to work—the VRI provides a very different picture than the HRS and SCF precisely

because it has sufficient observations with households with substantial financial wealth as they

approach retirement.

REFERENCES

Agarwal, Sumit, Chunlin Liu, and Nicholas S. Souleles (2007) "The Reaction of Consumer Spending and Debt to Tax Rebates: Evidence from Consumer Credit Data,*" Journal of Political Economy* 115, 986-1019.

Aguiar, Mark and Erik Hurst (2007) "Life-Cycle Prices and Production,*" American Economic Review* 97, 1533-1559.

Ameriks, John and Stephan P. Zeldes (2004) "How Do Household Portfolio Shares Vary with Age?" Unpublished manuscript.

Ameriks, John, Andrew Caplin, Minjoon Lee, Matthew D. Shapiro, and Christopher Tonetti (2014) "Combining Administrative Account Data and Surveys: Wealth Measurement Innovations from the Vanguard Research Initiative," In progress.

Bosworth, Barry P. and Gary Burtless (2010) "Recessions, Wealth Destruction, and the Timing of Retirement," *Center for Retirement Research Working Paper* 2010-22.

Brown, Jeffrey R. (2001) "Private Pensions, Mortality Risk, and the Decision to Annuitize," *Journal of Public Economics* 82, 29-62.

Browning, Martin, Thomas Crossley, and Joachim Winter (2014) "The Measurement of Household Consumption Expenditures," *IFS Working Paper*, W14/07.

Bucks, Brian and Karen Pence (2014) "Wealth, Pensions, Debt, and Savings: Considerations for a Panel Survey," Unpublished manuscript.

Chan, Sewin and Ann Huff Stevens (2004) "Do Changes in Pension Incentives Affect Retirement? A Longitudinal Study of Subjective Retirement Expectations," *Journal of Public Economics* 88, 1307 – 1333.

Coile, Courtney, and Phillip B. Levine (2004) "Bulls, Bears and Retirement Behavior," *Industrial and Labor Relations Review* 59, 408-429.

Coronado, Julia L. and Maria Perozek (2003) "Wealth effects and the consumption of leisure: Retirement decisions during the stock market boom of the 1990s," Federal Reserve Board Working Paper, 2003-20.

Davis, Steven J. and John Haltiwanger (1992) "Gross Job Creation, Gross Job Destruction, and Employment Reallocation," *Quarterly Journal of Economics* 107, 819-863.

Dominitz, Jeff and Charles F. Manski (1997) "Using Expectations Data to Study Subjective Income Expectations," *Journal of the American Statistical Association* 92, 855-867.

Gelman, Michael, Shachar Kariv, Matthew D. Shapiro, Dan Silverman, and Steven Tadelis (2014) "Harnessing Naturally Occurring Data to Measure the Response of Spending to Income," *Science* 345, 212-215.

Gustman, Alan L., Thomas L. Steinmeier, Nahid Tabatabai (2010) "What the Stock Market Decline Means for the Financial Security and Retirement Choices of the Near-Retirement Population," *Journal of Economics Perspective* 24, 161-182.

Goda, Gopi Shah, John B. Shoven, Sita Nataraj Slavov (2012) "Does stock market performance influence retirement intentions?" *Journal of Human Resources* 47, 1055-1081.

Gorodnichenko, Yuriy and Klara Sabirianova Peter (2007) "Public Sector Pay and Corruption: Measuring Bribery from Micro Data" *Journal of Public Economics* 91, 963-991.

Holtz-Eakin, Douglas, David Joulfaian, and Harvey S. Rosen (1993) "The Carnegie Conjecture: Some Empirical Evidence," *Quarterly Journal of Economics* 102, 413-435.

Hurd, Michael D. and Kathleen McGarry (2002) "The Predictive Validity of Subjective Probabilities of Survival," *Economic Journal* 112, 966 – 985.

Hurd, Michael, Monika Reti, and Susann Rohwedder (2009) "The Effect of Large Capital Gains or Losses on Retirement" in David A. Wise ed., *Developments in the economics of aging*, University of Chicago Press, 127-163.

Hurst, Erik, Geng Li, and Benjamin Pugsley (2014) "Are Household Surveys Like Tax Forms? Evidence from Income Underreporting of the Self-Employed," *Review of Economics and Statistics* 345, 212-215.

Imbens, Guido W., Donald B. Rubin, and Bruce I. Sacerdote (2001) "Estimating the Effect of Unearned Income on Labor Earnings, Savings, and Consumption: Evidence from a Survey of Lottery Players," *American Economic Review* 91, 778-794.

Joulfaian, David and Mark Wilhelm (1994) "Inheritance and Labor Supply," *Journal of Human Resources* 29, 1205 – 1234.

Juster, F. Thomas and James P. Smith (1997) "Improving the Quality of Economic Data: Lessons from the HRS and AHEAD," *Journal of the American Statistical Association* 92, 1268-1278.

Juster, F. Thomas, James P. Smith, and Frank Stafford (1999) "The Measurement and Structure of Household Wealth," *Labour Economics* 6, 253-275.

Kapteyn, Arie and Jelmer Y. Ypma (2007) "Measurement Error and Misclassification: A Comparison of Survey and Administrative Data," *Journal of Labor Economics* 25, 513-551.

Kennickell, Arthur B. (2000) "Wealth Measurement in the Survey of Consumer Finances: Methodology and Directions for Future Research," Working paper, Board of Governors of the Federal Reserve System.

Kezdi, Gabor and Purvi Sevak (2004) "Economic Adjustment of Recent Retirees to Adverse Wealth Shocks," *MRRC Working Paper* 2004-075.

Khitatrakun, Surachai (2004) "Wealth and the timing of retirement," Unpublished manuscript, University of Wisconsin-Madison.

Krimmel, Jacob, Kevin B. Moore, John Sabelhaus, and Paul Smith (2013) "The Current State of U.S. Household Balance Sheets," *Federal Reserve Bank of St. Louis Review* 95, 337-359.

McFall, Brooke Helppie (2011) "Crash and Wait? The Impact of the Great Recession on the Retirement Plans of Older Americans," *American Economic Review: Papers and Proceedings* 101, 40-44.

McGarry, Kathleen (2004) "Health and Retirement: Do Changes in Health Affect Retirement Expectations?" *Journal of Human Resources* 39, 624-648.

Poterba, James M. (2014) "Retirement Security in an Aging Population," *American Economic Review* 104, 1-30.

Poterba, James M., Steven Venti, and David Wise (2011) "The Composition and Drawdown of Wealth in Retirement," *Journal of Economic Perspective* 25, 95 – 118.

Sevak, Purvi (2002) "Wealth Shocks and Retirement Timing," Working paper, Michigan Retirement Research Center.

Szinovacz, Maximiliane E., Adam Davey, and Lauren Martin (2014) "Did the Great Recession Influence Retirement Plans?" *Research on Aging* 36, 1-31.

Table 1. Design of VRI, HRS, and SCF

	VRI	HRS	SCF
Sampling			
Population	Vanguard clients	U.S. Population	U.S. Population
Frequency	Multiple surveys per year; monthly admin. data	Biennial	Triennial
Panel/cross-section	Panel	Panel	Cross-section[1]
Main target	Age 55+ with non-negligible financial assets	Age 50+ and spouses	Representative of wealth
Oversampling	Singles	Blacks and Hispanics; Residents of Florida	High-income list sample
Additional screens	Internet eligible; Employer-sponsored and individual client samples		
Wealth measurement			
Account-based approach	Comprehensive	401(k)/IRA[2]	Transactional and pension accounts
Administrative data	Yes	No	No
Summary (age$\geq$55)			
Households	8,950	11,595	2,624
Median Financial Wealth	$663,100	$60,000	$33,200
Median Income	$121,481	$30,400	$42,610

Note: Table refers to most recent wave of each survey (VRI 2013, HRS 2012, and SCF 2013). Observations are restricted to respondents aged 55 and older. The VRI and SCF survey only one member of couples. The age of the household is determined by the age of respondent. The HRS surveys HRS respondents and their spouses. The age of the household is determined by the age of the financial respondent as defined by the HRS.

[1] The SCF occasionally (1983-89, 2007-09) has a panel structure.

[2] HRS implemented account-based approach for retirement accounts in 2012.

Table 2. Survey Financial Assets: All respondents

Account type	Mean	N	Conditional on having positive amount					
			Mean	Percentiles				
				10	25	50	75	90
Total financial assets	1,189,358	8,948	1,189,358	122,000	296,673	656,962	1,266,651	2,254,000
IRA	359,181	7,303	440,184	29,000	83,931	234,033	556,527	1,021,000
Employer sponsored	215,620	4,630	416,803	26,000	83,000	222,000	475,000	842,402
Pension	25,365	1,016	223,437	10,518	34,000	100,000	251,000	590,714
Other retirement asset	13,237	602	196,801	10,000	26,136	80,466	213,000	450,000
Checking	16,888	8,637	17,500	1,000	2,200	5,500	15,000	40,000
Saving	23,020	6,162	33,436	500	2,100	10,000	32,000	84,382
Money market	28,308	4,076	62,158	1,200	5,367	22,177	69,303	151,023
Mutual fund	231,577	3,942	525,777	8,500	30,000	114,000	309,000	690,000
Certificate of deposit	16,576	1,634	90,794	4,000	11,000	34,450	100,000	230,803
Brokerage	181,872	4,184	389,042	6,400	27,100	110,000	347,000	854,000
Directly held securities	22,634	1,801	112,477	2,000	10,000	30,000	100,000	235,664
Annuity	20,811	1,163	160,150	13,000	35,000	94,500	200,000	365,000
Life insurance	21,053	2,696	69,891	5,000	10,000	26,000	70,000	150,000
Educational related	3,022	613	44,119	3,400	8,300	20,000	48,000	100,000
Other accounts	9,930	429	207,165	1,500	10,000	46,000	195,000	478,000

Note: Pension, annuity, and life insurance are current cash values.

Table 3. Total Vanguard Assets: Survey versus Administrative Data

A. Employer-Sponsored (N=2,243)

	Mean	Percentiles				
		10	25	50	75	90
Survey	331,753	27,000	75,000	195,485	432,000	755,000
Administrative	299,540	29,519	69,668	181,375	400,707	656,832
Difference	32,213	-27,394	-4,093	890	12,999	95,978
% Difference	3.92%	-17.44%	-2.48%	0.63%	9.10%	47.83%

B. Individual client (N=6,705)

	Mean	Percentiles				
		10	25	50	75	90
Survey	517,724	29,000	87,017	260,000	615,081	1,178,158
Administrative	380,277	25,345	67,382	193,682	472,732	900,747
Difference	137,447	-23,315	-1,637	2,623	91,950	380,262
% Difference	18.53%	-14.42%	-1.20%	1.44%	32.89%	100.32%

C. Employer-Sponsored, Singles (N=585)

	Mean	Percentiles				
		10	25	50	75	90
Survey	240,488	22,000	49,000	125,000	300,000	574,000
Administrative	231,306	22,757	46,236	127,630	282,362	529,760
Difference	9,183	-24,297	-3,867	365	7,483	35,390
% Difference	2.05%	-22.06%	-3.04%	0.33%	6.21%	29.68%

D. Individual client, Singles (N=2,349)

	Mean	Percentiles				
		10	25	50	75	90
Survey	317,004	21,000	57,000	165,400	420,000	790,000
Administrative	305,997	22,501	58,759	160,638	406,609	744,563
Difference	11,008	-32,803	-4,180	-19	3,902	39,677
% Difference	-0.64%	-22.23%	-2.91%	-0.03%	2.18%	24.34%

Table 4. Comparison of Total Vanguard Wealth: Different Correction Paths (Singles only)

Correction paths	N	Measure	25 percentile	median	75 percentile
				Percent Difference	
None	1927	Final	-3.3	-0.0	2.6
Accounts only	426	Initial	-3.5	0.1	2.5
		Final	-3.5	0.1	2.5
Balances only	308	Initial	-12.2	-0.0	13.6
		Final	-2.6	-0.0	2.7
Accounts and balances (restarted)	121	Initial	-5.3	-0.1	12.1
		Final	-1.1	0.2	2.1
Accounts and balances (other paths)	153	Initial	-18.1	-0.1	2.7
		Final	-1.4	0.1	2.7

Table 5. Comparing VRI to Age-Eligible HRS and SCF Households (unweighted counts): Age and Financial Wealth

Age		<$0	$0-10K	$10K-100K	$100K-500K	$500K-1M	$1M-2.5M	>$2.5M	All
					Financial Wealth				
	VRI	48	36	292	1,147	871	762	181	3,337
55-64	HRS	1,459	586	933	897	287	160	41	4,363
	SCF	228	170	196	254	102	119	212	1,281
	VRI	16	19	258	1,117	985	1,066	377	3,838
65-74	HRS	746	487	727	817	290	162	35	3,264
	SCF	93	114	118	155	68	91	178	817
	VRI	2	4	95	549	461	472	192	1,775
> 74	HRS	800	712	1,030	927	284	172	43	3,968
	SCF	60	93	115	107	31	30	90	526
	VRI	66	59	645	2,813	2,317	2,300	750	8,950
Total	HRS	3,005	1,785	2,690	2,641	861	494	119	11,595
	SCF	381	377	429	516	201	240	480	2,624

Note: Numbers are raw counts (unweighted) of households. Note that only age-eligible households are included in the table. For SCF, only one replicate is included. For HRS, only those households surveyed in both the 2010 and 2012 waves are included. Age of HRS households based on financial respondent. Financial wealth is the sum of financial assets (both retirement and non-retirement assets) minus non-mortgage debt.

Table 6. Comparing Age-eligible VRI, HRS, and SCF Households (unweighted counts): VRI Sampling Screens

	VRI			HRS			SCF		
Age	All	Employer-Sponsored	Individual client	Age Eligible	VRI Eligible	VRI Eligible 401(k) subset	Age Eligible	VRI Eligible	VRI Eligible 401(k) subset
All	8,950	2,244	6,706	11,595	3,684	1,553	2,624	1,275	665
55-59	1,549	810	739	2,364	976	628	668	397	280
60-64	1,788	823	965	1,999	756	411	613	350	205
65-69	1,931	419	1,512	1,282	535	214	462	257	112
70-74	1,907	157	1,750	1,982	638	178	355	161	51
75-100	1,775	35	1,740	3,968	779	122	526	110	17

Note: Table shows total age-eligible number of households in total and after imposing the VRI-equivalent screen. VRI-eligible screen imposes Internet eligibility plus having at least $10,000 in any non-transactional financial accounts. The 401(k) subset imposes $10,000 wealth cut-off on DC type pensions. See text for details. See also the note to Table 5.

Table 7. Effect of Imposing VRI Sampling Screens: Fraction of weighted observations

Screens	HRS	SCF
Age-eligible	100%	100%
Internet eligibility	56%	58%
$10,000 asset cut-off	58%	45%
Internet eligible and $10,000 cut-off	41%	35%
401(k) subset	19%	18%

Note: Table shows the fraction of the sample in HRS and SCF (measured by the fraction of weighted observations) remaining after imposing VRI sampling screens. See text and note to Table 6 for descriptions of screens.

Table 8. Effect of Imposing VRI Sampling Screens: Wealth distribution

| | | Mean | Percentiles | | | | |
			10	25	50	75	90
VRI	All	1,206,594	115,337	292,000	663,100	1,286,000	2,291,235
	Employer-sponsored	847,349	65,050	185,600	496,350	1,029,700	1,856,005
	Individual client	1,326,807	140,100	330,636	715,790	1,383,209	2,421,840
HRS	Age eligible	293,596	-900	500	60,000	300,000	745,000
	VRI eligible	578,069	34,000	98,036	272,000	660,000	1,247,800
	VRI eligible, 401(k) subset	623,954	46,300	130,000	342,700	733,000	1,364,000
SCF	Age eligible	404,668	-6,300	320	33,200	220,550	794,700
	VRI eligible	970,294	28,860	96,350	262,100	792,400	2,109,000
	VRI eligible, 401(k) subset	871,897	18,000	76,870	219,500	674,000	1,953,500

Note: HRS and SCF tabulations use sampling weights.

Table 9. Stock Ownership

A. Share: VRI, HRS, and SCF (Percent)

| | Sample Screen | Percentiles | | | | | N |
		10	25	50	75	90	
	All	14.96	35.12	54.76	74.71	91.14	8905
VRI	Employer-sponsored	8.42	28.88	50.00	72.04	90.00	2233
	Individual client	18.55	37.37	56.06	75.33	91.52	6672
	Age eligible	0	0	0	40.32	81.48	11595
HRS	VRI eligible	0	0	29.20	70.75	90.54	3684
	VRI eligible, 401(k) subset	0	0	20.93	67.86	89.05	1553
	Age eligible	0	0	0.70	43.39	71.24	2624
SCF	VRI eligible	2.77	19.94	42.34	61.85	84.74	1275
	VRI eligible, 401(k) subset	6.98	21.51	40.66	61.04	83.33	665

Note: See text and note to Table 4 for sample screens. Respondents with less than $1000 in financial assets are coded as having a zero stock share.

B. Amount: VRI, HRS, and SCF (Dollars)

| | Sample Screen | Percentiles | | | | | N |
		10	25	50	75	90	
	All	30,000	113,800	326,162	712,200	1,397,710	8905
VRI	Employer-sponsored	13,500	65,428	221,443	551,365	1,047,212	2233
	Individual client	41,415	138,220	365,174	765,400	1,477,515	6672
	Age eligible	0	0	0	45,000	270,000	11595
HRS	VRI eligible	0	0	30,000	200,000	520,000	3684
	VRI eligible, 401(k) subset	0	0	15,000	150,000	453,700	1553
	Age eligible	0	0	0	78,000	360,000	2624
SCF	VRI eligible	3,000	22,750	105,000	357,000	1,227,600	1275
	VRI eligible, 401(k) subset	4,500	21,000	86,000	306,500	1,168,500	665

Figure 1. Administrative versus Survey Financial Assets at Vanguard
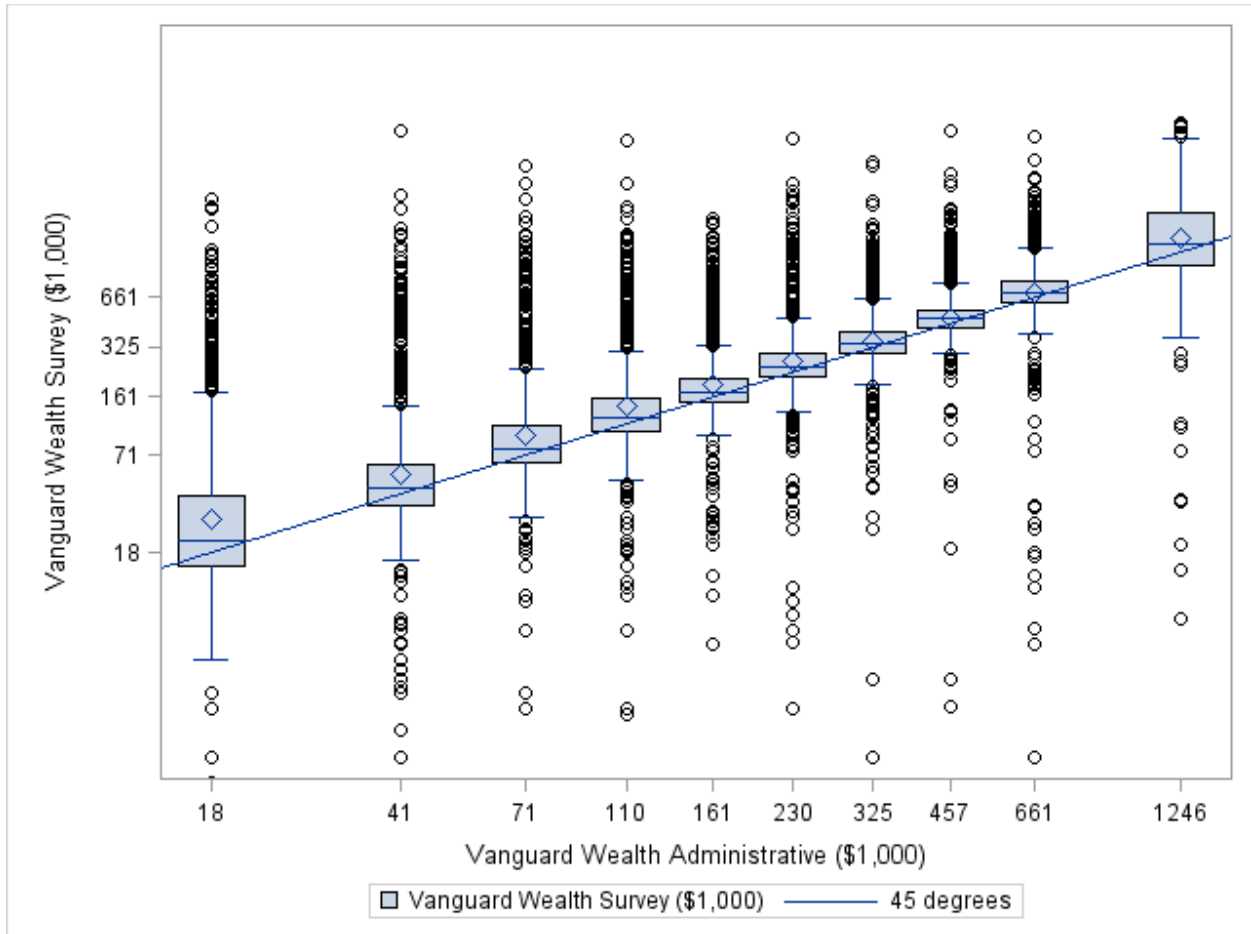
Note: The figure compares Vanguard administrative assets with survey report of Vanguard assets. See the text for how Vanguard assets are determined in survey. The chart shows box and whiskers figures for each decile of administrative assets (diamond is the mean; middle line is median; box is inter-quartile range [IQR]; outer lines upper and lower fences [1.5 times the IQR from the box]; and circles denote outliers). Amounts on the horizontal axis are medians of each decile ($1000). Log scale is used on both axes.

Figure 2. Correction Paths through Wealth Section.
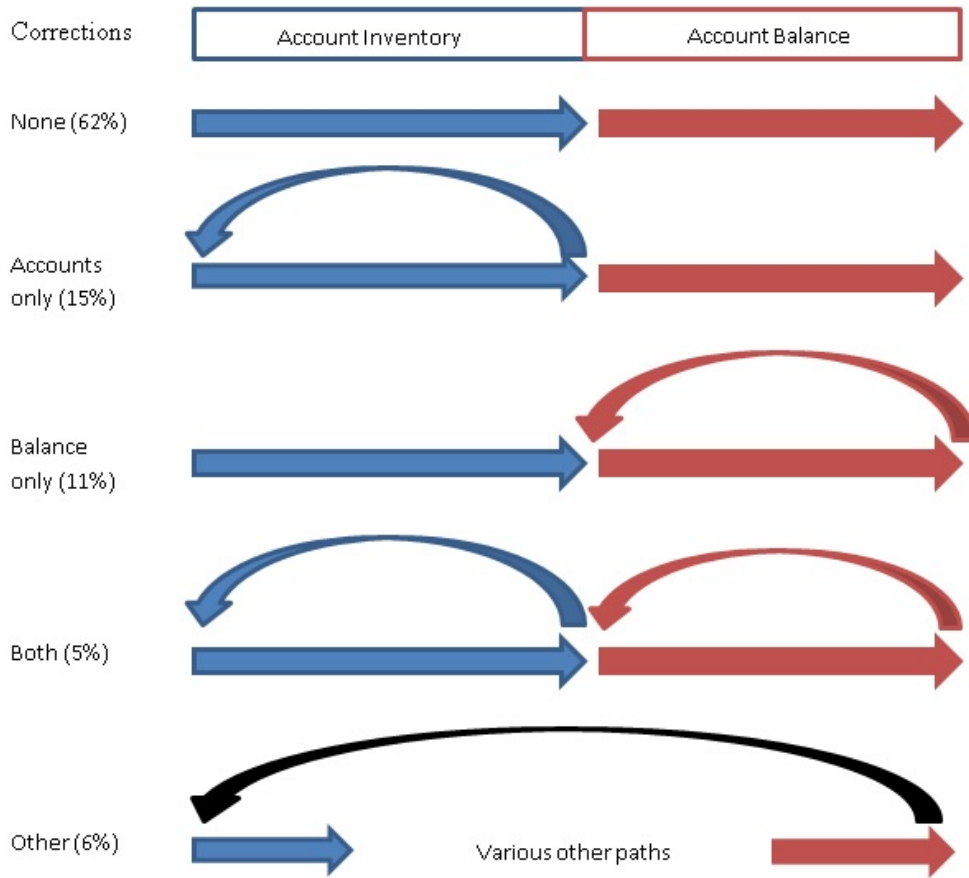
Note: The figure shows the fraction of respondents taking various paths through the account-based wealth section. Other includes those who started over and then took various paths to complete.

Figure 3. Distribution of normalized financial wealth (kernel estimation)
A. VRI vs HRS
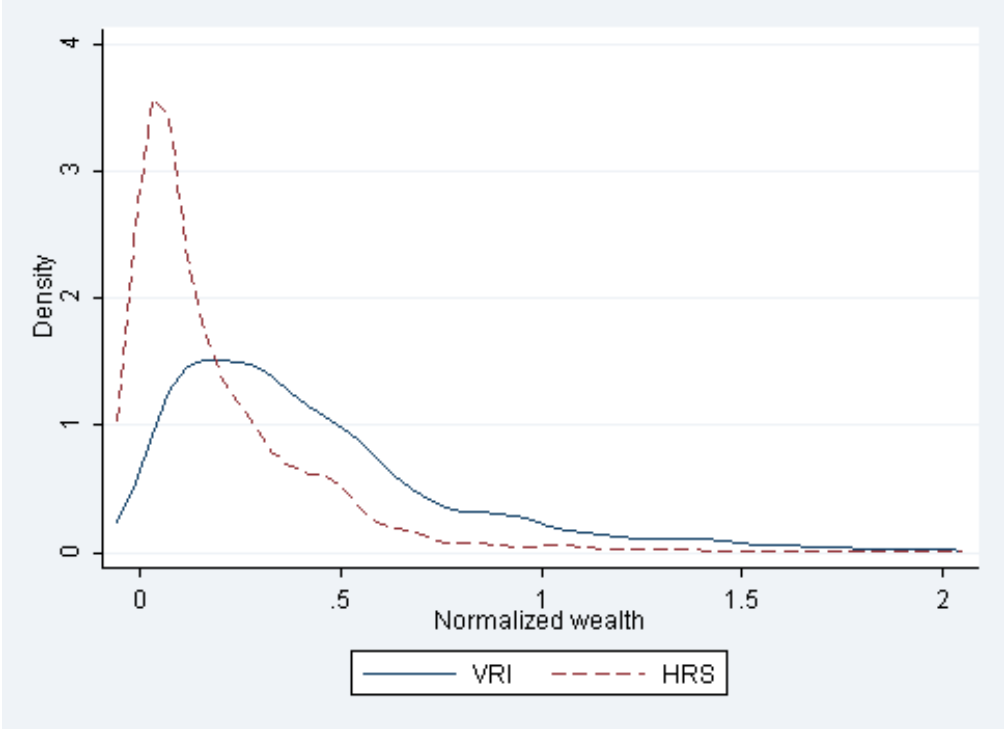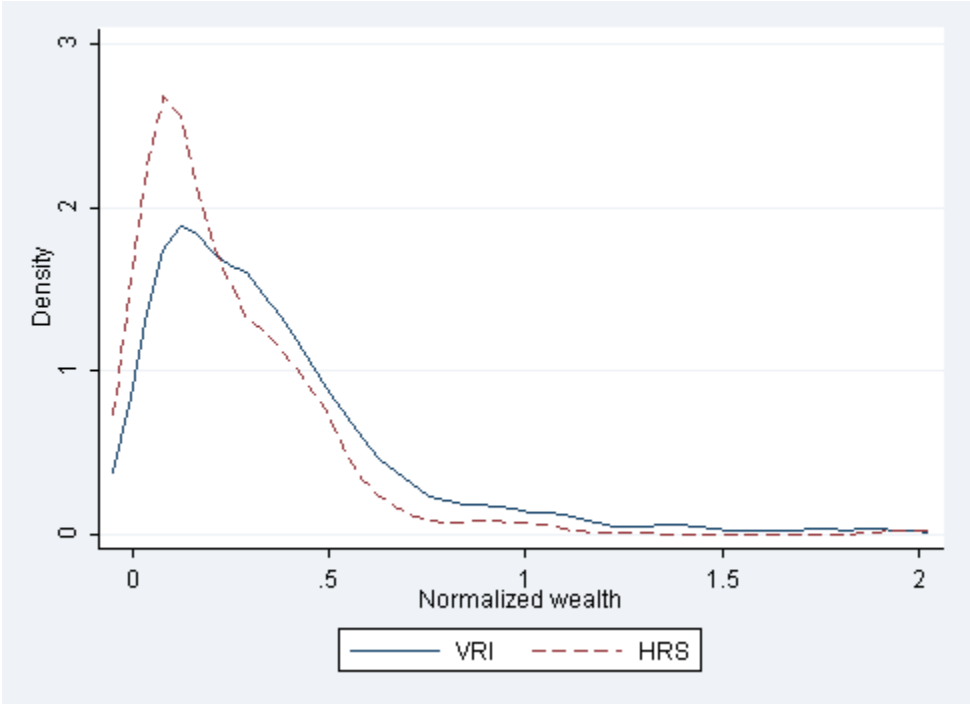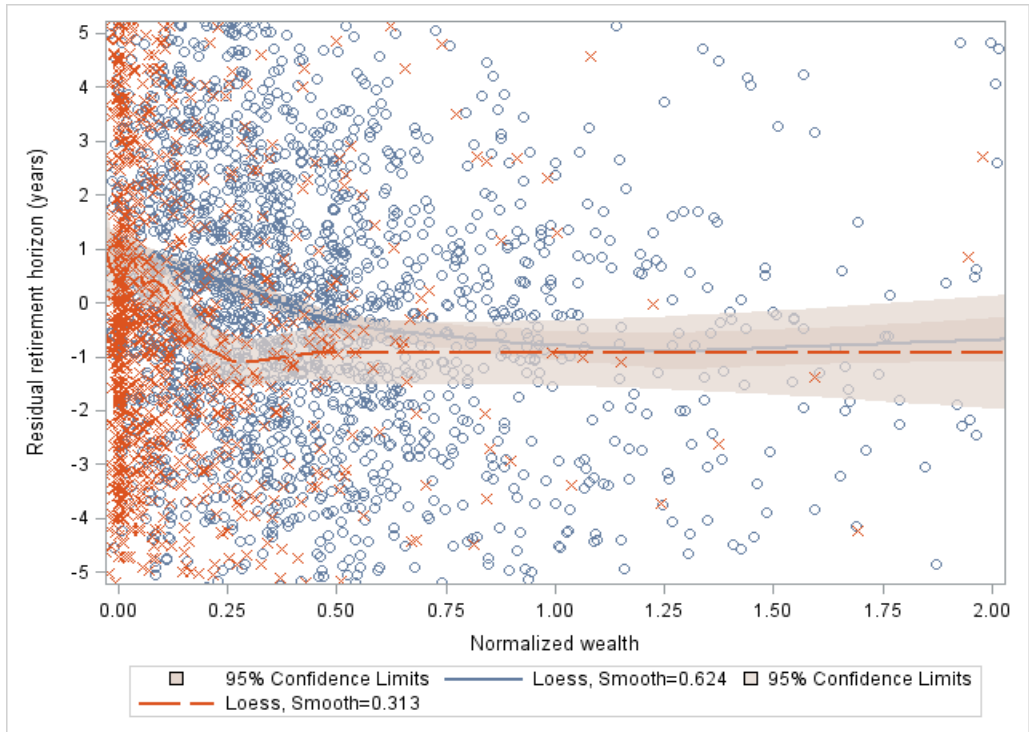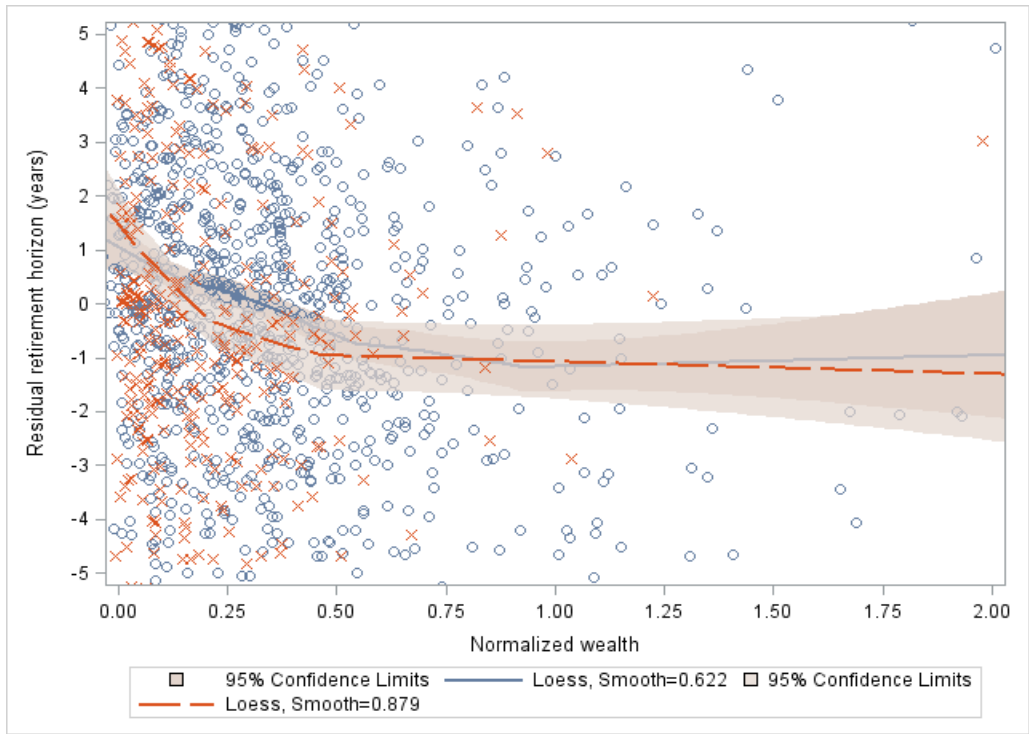

B. VRI employer-sponsored versus HRS 401(k) subset

Figure 4. Retirement horizon versus normalized financial wealth: LOESS
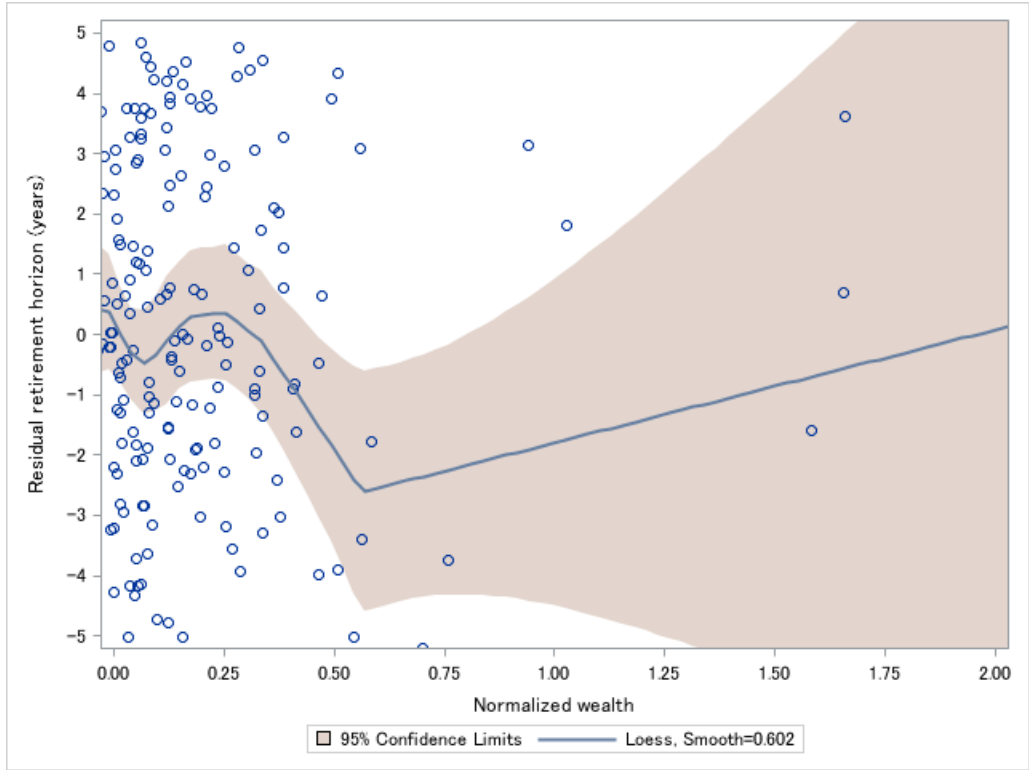A. VRI vs HRS

Note: x denotes HRS (orange/dashed line) and o denotes VRI (blue/solid line).

B. VRI employer-sponsored versus HRS 401(k) subset

Note: x denotes HRS (orange/dashed line) and o denotes VRI (blue/solid line).
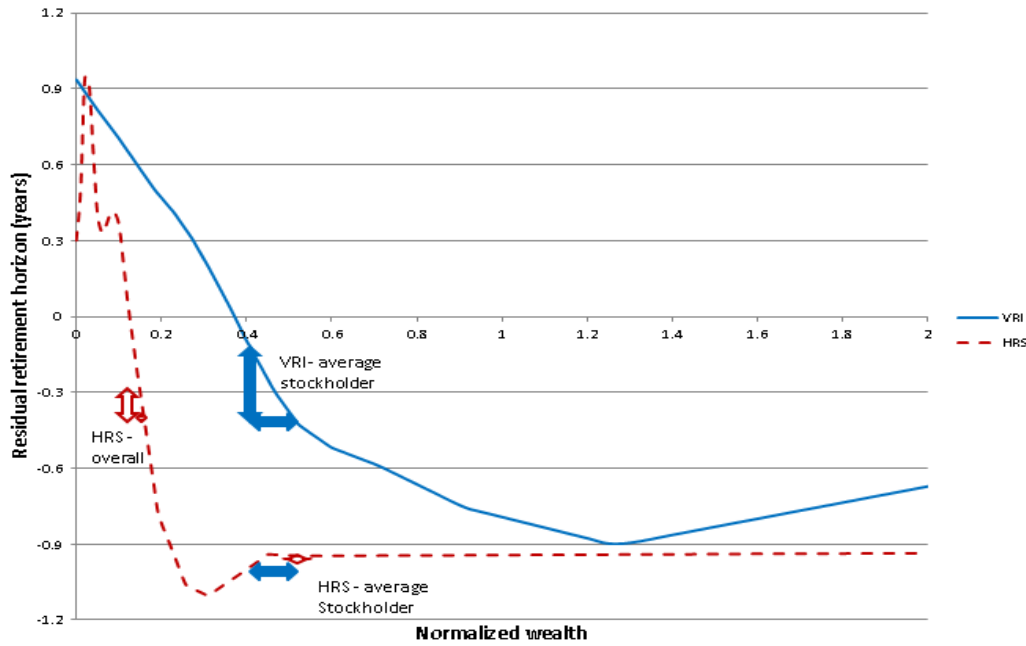
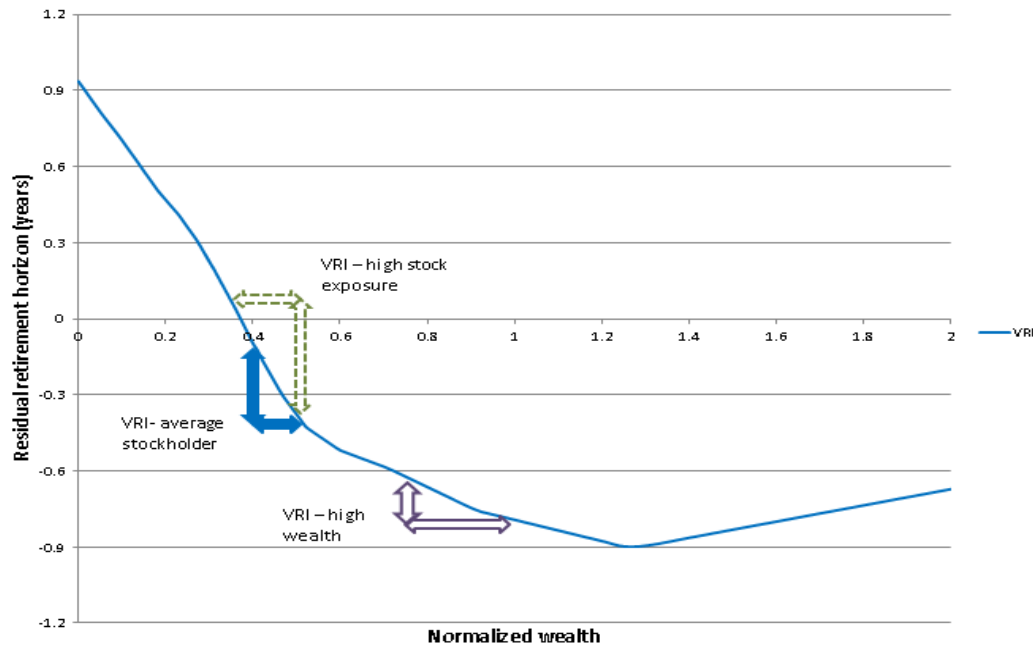Figure 5. Retirement horizon versus normalized financial wealth: LOESS
A. SCF

B. SCF 401(k) subset

Figure 6. Implied Changes in Retirement Horizon: 40% Decline in Stock Market

A. Comparison of VRI and HRS Estimates

B. Alternative Scenarios using VRI Estimates

Note: Lines are LOESS estimates from Figure 4 (confidence intervals and observation not shown). The figure shows the predicted change in the retirement horizon (years to retirement) resulting from a 40% decline in the stock market. In panel A, the HRS-overall applies the mean HRS wealth and stock share to the HRS estimates. The HRS-average stockholder applies the mean VRI wealth and stock share to the HRS estimates while the VRI-average stockholder applies the same mean VRI wealth and stock share to the VRI estimates. In panel B, the VRI-average stockholder is same as in panel A. The other two treatments show high stock exposure and high-wealth households. See text for details.

Appendix A. Account Sequence Example

Section 2.2 of the main text explained the structure of the wealth section of the survey in detail.

In this appendix, we show actual screen shots from the wealth section for a hypothetical

respondent who has two IRAs, one 401(k) pension, one checking account and one mutual fund

account.

The respondent starts the wealth section by entering all the types of accounts she has

(Figure A-A1). She answers how many accounts she has for each type using a drop-down menu

(Figure A-A2) and then gives each of the accounts a nickname (Figure A-A3). The survey

shows the summary of responses so far (Figure A-A4) and asks whether all the information

given is correct. If the respondent clicks no, then she can either add/delete the account type or

add/delete accounts within each type.

After this first check point, the survey then loops over the accounts and asks the balance

of each (Figure A-A5 is one example). After the loop, the survey displays a summary table of

account balances as well as a total (Figure A-A6). In this example, the respondent did not

provide a response to the balance question for the second IRA account ("Roth IRA"), so she sees

"No response provided" for Reported Value under that account. Let us say that the respondent

clicks "No" to "Is this correct?" under the summary table. Then the respondent is asked whether

she wants to add/delete accounts or correct balances (Figure A-A7). In this example, the

respondent chooses to correct balances, indicates that she wants to correct the balance for "Roth

IRA" (Figure A-A8), and then corrects the balance for that account (Figure A-A9). During the

corrections, the previously provided answers are shown above the question (in this case "Not

answered"). The respondent comes back to the summary screen again, indicates whether she referred to records to provide information on each account, and then confirms that all the responses are correct (Figure A-A10).

The survey then asks two follow-up questions for each account: stock share (Figure A-A11) and whether that account is held at Vanguard (Figure A-A12). Note that the survey does not ask these questions about the checking account that this respondent reported since it is a transactional account not offered at Vanguard. Based on these responses, the survey calculates the share of wealth held at Vanguard and the stock share of the total portfolio, and it shows these as charts along with the summary table of balances (Figure A-A13). The respondent can print this summary page as a record.

Figure A-A1. Types of Accounts

Which of the following types of investment, savings, and retirement accounts do you have? We are asking for a complete list of financial accounts that you own. Please answer for your own accounts only.

We have grouped the accounts listed below into categories to help make it easier for you to think about all of the various types that you may own. The purpose of this question is to obtain a comprehensive view of the various types of investment, savings and retirement accounts that households may own.

For each account that you own, please include it under only one type of account.

Tax-deferred retirement accounts
- ☑ IRA (including ROTH, traditional, an IRA rolled-over from an employer-sponsored plan)
- ☑ Employer-sponsored retirement plan account (401(k), 403(b), 457, etc.)
- ☐ Pension with an account balance which you can access as a lump sum
- ☐ Other type of tax-deferred retirement account (such as SEPs, Keoghs, etc.)

Savings/investment accounts that are not in a tax-deferred retirement plan or account
- ☑ Checking account
- ☐ Savings account
- ☐ Money market account
- ☑ Mutual fund account (other than money market mutual fund)
- ☐ Certificate of deposit (CD) portfolio (aggregate of all CD holdings)
- ☐ Brokerage account (including stocks, municipal, corporate, or other bonds, mutual funds, ETFs and other assets)
- ☐ Directly held securities or other financial assets (US Treasury Bonds or savings bonds at Treasury Direct, stocks, bonds or individual securities you own that are not at a brokerage, Dividend Reinvestment Programs.)

Insurance contracts/accounts with a cash value or balance
- ☐ Annuity accounts with a balance or cash value *(excluding immediate annuities reported in the previous section)*
- ☐ Life insurance with cash value *(excluding term life insurance)*

Educational Accounts
- ☐ Section 529 College Savings Plans or Coverdell Accounts

Other accounts
- ☐ Other accounts not specified above

Figure A-A2. Number of Accounts

You mentioned that your household has the following types of investment, savings, and retirement accounts. How many of each type does your household have?

For example, your household may have three checking accounts. In this case, you would enter '3' below for 'checking account'.

Or, for example, your household may have two IRAs (one owned by you, one owned by your spouse/partner) and one CD. In this case, you would enter a '2' below for 'IRA' and a '1' below for 'Certificate of Deposit (CD)'.

- Note, when you are counting, there is not a need to break out the subcomponents of an account. You can just count the overall account.

Please indicate the number after each.

		Number of accounts
Tax-deferred retirement accounts	IRA (including ROTH, traditional, rolled-over from an employer-sponsored plan)	2
	Employer-sponsored retirement plan (401(k), 403(b), 457, etc.)	1
Savings/investment accounts that are not in a tax-deferred retirement plan or account	Checking account	1
	Mutual fund account (other than money market)	1

Figure A-A3. Nickname Accounts

We will be asking you additional questions about each of the investment, savings, and retirement accounts you mentioned that your household has. To assist with this, it would be helpful if you give each account a "nickname." The nickname you assign could be any name, as long as it helps you keep track of which specific investment or savings account you are responding about in future questions.

Nicknames should be descriptive and are meant to help you remember the account types you have just selected - for example, if your household has two IRAs, one 401(k), and one checking account, you may elect to name your accounts as follows:

- *IRA 1:* My IRA
- *IRA 2:* Mary's IRA
- *Employer-sponsored retirement plan (401(k), 403(b), 457, etc.) 1:* Her 401(k)
- *Checking account 1:* Joint checking account at credit union

Please type in a nickname for each.

IRA 1:	Rollover IRA
IRA 2:	Roth IRA
Employer-sponsored retirement plan (401(k), 403(b), 457, etc.) 1:	Retirement
Checking account 1:	Chase
Mutual fund account (other than money market) 1:	Vanguard

Figure A-A4. Account Verification

Please scroll down to see a summary of your household's investment, savings, and retirement accounts. Please review this summary for accuracy - does this correctly reflect all of your household's investment, savings, and retirement accounts? What's most important is that nothing significant is forgotten or double-counted in the list.

If this information is not correct, you will be able to go back to the beginning of this section to update your information.

It is very important for the rest of the survey that your responses here be as complete and accurate as possible and we appreciate you taking the time to thoroughly review and update if necessary.

Please select one.

- ○ Yes - this is accurate and I am ready to continue
- ○ No – I need to go back to make an update

Summary of My Household's Investment, Savings and Retirement Accounts
Tax-deferred retirement accounts
IRA
1: Rollover IRA
2: Roth IRA
Employer-sponsored retirement plan (401(k), 403(b), 457, etc.)
1: Retirement
Pension with an account balance which you can access as a lump sum
None
Other type of tax-deferred retirement account (such as SEPs, Keoghs, etc.)
None
Savings/investment accounts not in a tax-deferred retirement plan or account
Checking account
1: Chase
Savings account
None
Money market account
None
Mutual fund account
1: Vanguard
Certificate of deposit (CD)
None
Brokerage account
None
Directly held securities or other financial assets
None
Insurance- and Education-related accounts
Annuity Accounts with a Balance or Cash Value
None
Life insurance with cash value
None
Educational-Related accounts
None
Other accounts
None

Figure A-A5. Account Balance

IRA 1:
Rollover IRA

Please enter your total balance in this account. You can reference any documents or records that may help you obtain this information. You may also give us your best estimate from memory. Please feel free to round, but try to be accurate at least to the nearest thousand dollars. For example, if the account balance was $24,823, you may enter '25000' below. We appreciate any effort you give to specify an amount as precisely as possible. The information you provide will be kept completely confidential.

$ 120,000

Figure A-A6. Balance Verification

Please refer to the below table and verify the balances you reported for each of your accounts, and indicate whether you referred to records or statements in supplying these figures.

	REPORTED VALUE	REFERRED TO RECORDS?	
		YES	NO
IRA 1: Rollover IRA	$120,000	○	○
IRA 2: Roth IRA	No response provided		
Employer-sponsored retirement plan (401(k), 403(b), 457, etc.) 1: Retirement	$400,000	○	○
Checking account 1: Chase	$15,000	○	○
Mutual fund account (other than money market) 1: Vanguard	$275,000	○	○
TOTAL	$810,000		

Is this correct?

○ Yes – this is accurate and I am ready to continue
○ No – I need to go back to make an update

Figure A-A7. Indicate What Type of Correction(s)

Please tell us which of these activities you need to do...

☐ I need to ADD and/or DELETE an account
☑ I need to fix the dollar amount of what I have already provided.

Figure A-A8. Indicate What Needs to Be Corrected

Please tell us which Account(s) you need to correct.

	REPORTED VALUE	For Which Account(s) Do You Need To Correct the Reported Value?
IRA 1: Rollover IRA	$120,000	☐
IRA 2: Roth IRA	No response provided	☑
Employer-sponsored retirement plan (401(k), 403(b), 457, etc.) 1: Retirement	$400,000	☐
Checking account 1: Chase	$15,000	☐
Mutual fund account (other than money market) 1: Vanguard	$275,000	☐

Figure A-A9. Correction of Previous Response(s)

Your Previous Response was... Not Answered

IRA 2:
Roth IRA

Please enter your total balance in this account. You can reference any documents or records that may help you obtain this information. You may also give us your best estimate from memory. Please feel free to round, but try to be accurate at least to the nearest thousand dollars. For example, if the account balance was $24,823, you may enter '25000' below. We appreciate any effort you give to specify an amount as precisely as possible. The information you provide will be kept completely confidential.

$ 150,000

Figure A-A10. Revised Balance Summary

Please refer to the below table and verify the balances you reported for each of your accounts, and indicate whether you referred to records or statements in supplying these figures.

	REPORTED VALUE	REFERRED TO RECORDS?	
		YES	NO
IRA 1: Rollover IRA	$120,000	●	○
IRA 2: Roth IRA	$150,000	●	○
Employer-sponsored retirement plan (401(k), 403(b), 457, etc.) 1: Retirement	$400,000	●	○
Checking account 1: Chase	$15,000	○	●
Mutual fund account (other than money market) 1: Vanguard	$275,000	●	○
TOTAL	$960,000		

Is this correct?

● Yes – this is accurate and I am ready to continue
○ No – I need to go back to make an update

Figure A-A11. Account-by-account Stock Share

Thinking about all of the investment, savings, and retirement accounts that your household currently has, what percentage, if any, of each account is held in stocks or stock market investments? In other words, what percentage of the underlying assets or funds in each account is held in stock investments?

Please note: Checking accounts, Savings accounts, Money Market accounts, CDs and Life insurance are not displayed in the table below since they have no stock/stock market investment value. The amounts on the far right of the table will compute after you click out of the box where you enter the approximate percentage.

	APPROXIMATE PERCENTAGE HELD IN STOCKS/STOCK MARKET	IMPLIED VALUE OF STOCK INVESTMENTS IN THIS ACCOUNT
IRA 1: Rollover IRA	50 %	$60,000
IRA 2: Roth IRA	100 %	$150,000
Employer-sponsored retirement plan (401(k), 403(b), 457, etc.) 1: Retirement	25 %	$100,000
Mutual fund account (other than money market) 1: Vanguard	100 %	$275,000

Figure A-A12. Which Accounts at Vanguard

Which accounts are currently held at Vanguard?

Please select one for each row.

| Accounts	Held at Vanguard	
	Yes	No
IRA 1: Rollover IRA	●	○
IRA 2: Roth IRA	●	○
Employer-sponsored retirement plan (401(k), 403(b), 457, etc.) 1: Retirement	○	●
Mutual fund account (other than money market) 1: Vanguard	●	○

Figure A-A13. Summary Table and Charts

Thank you very much for providing this detailed information about your financial assets. Before continuing with the final sections of the survey, we would like to present you with a summary of your assets. You may wish to print this for your own use:

FINANCIAL ASSETS	
Retirement	**Total: $670,000**
IRA	$270,000
Employer-sponsored retirement plans (401(k), 403(b), 457, etc.)	$400,000
Pension with an account balance which you can access as a lump sum	--
Other (such as SEPs, Keoghs, etc.)	--
Non-retirement	**Total: $290,000**
Checking	$15,000
Savings account	--
Money market account	--
Mutual fund account	$275,000
CDs	--
Brokerage account	--
Other stocks and bonds	--
Annuity accounts	--
Life insurance with cash value	--
Section 529 College Savings Plans or Coverdell Accounts	--
Other accounts	--
Total	**$960,000**

Percent of Financial Assets at Vanguard vs. Elsewhere
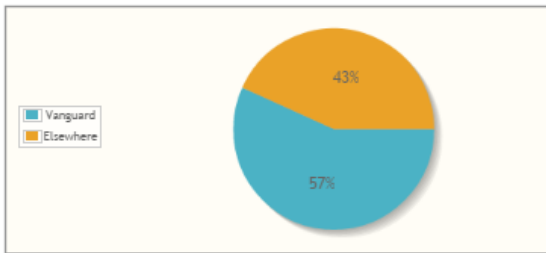
Percent of Financial Assets in Stocks vs. Not in Stocks
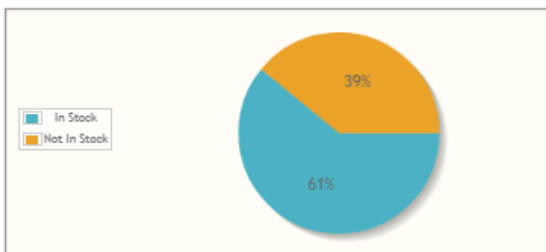
The Wealth of Wealthholders:
Evidence and Implications from the Vanguard Research Initiative
December 2014

Appendix B. Definition of concepts

This appendix defines concepts used for the VRI and how we measure them in the HRS and SCF.

<u>Total financial wealth</u>.   In the VRI, total financial wealth is the sum of all financial account balances (the items listed in Table 2) plus miscellaneous financial items (in non-account, cleanup questions) minus non-mortgage debt. For the SCF, financial wealth is total financial assets (FIN in the public version of data) minus non-mortgage debt (sum of CCBAL, INSTALL and ODEBT in the public version of data). For the HRS, financial wealth is the sum of total financial wealth (atof in RAND version), IRA wealth, and employer-sponsored plan and pension account balances. For the HRS 2012, we constructed these variables using RAND definitions. (We are grateful to Margaret Lay for sharing her construction of these variables.)

<u>Web-survey eligibility</u>. For the VRI, respondents are Web-survey eligible if the client is registered for Web access with Vanguard, if the registration has a valid email address, if the client logged in to the Vanguard Website at least once in the last six months, and if the client was not been recently included in another survey by Vanguard, and if the client had not requested exclusion from contacts for surveys. We need to simulate this set of screens in the HRS and SCF in order to select comparable respondents. We designate HRS respondents as Web-survey eligible if they use the Internet regularly. In the SCF, respondents are designated Web-survey eligible if they use the Internet to obtain information about borrowing/investing.

Asset cut-off. In the HRS, we impose a $10,000 cut-off on total financial assets net of checking, saving and money market balances. In the SCF, we impose a $10,000 cut-off on the sum of IRA, mutual funds and account type pensions.

The Wealth of Wealthholders:
Evidence and Implications from the Vanguard Research Initiative
December 2014

Appendix C. Detailed Comparisons: VRI, HRS and SCF

This appendix compares the VRI with the most recent waves of the HRS (2012) and SCF (2013) in more detail. It compares surveys along dimensions including wealth, income and demographics. For each dimension, we also provide comparisons conditional on age groups to control for the effect of different age compositions across surveys.

Recall that the age distribution differs across the samples. Table A-C1 compares median value of wealth by age group to see whether the difference in the overall wealth distribution is caused by differences in age. Even after imposing the similar sampling screens, the VRI sample has a higher median wealth for almost all the age groups. Again, the gap is much smaller when the HRS and SCF samples are compared with the employer-sponsored sample of the VRI. For the HRS, the gap shrinks further if we condition on respondents with at least $10,000 in 401(k)s or similar pension accounts. (Statistics for the age group 65+ under employer-sponsored conditions or 401(k) subset conditions are not very informative due to the small number of observations.)

Income. Tables A-C2 and A-C3 compare household annual income across samples. Compared to the overall population of the HRS and SCF, the VRI sample is not only wealthier, but also has higher income. The difference in income is, however, much smaller than the difference in wealth. If we impose the VRI screens, except for the oldest age group, income levels from the SCF are actually higher than the VRI; those from the HRS are quite comparable to those from the VRI. As a result, the wealth-to-income ratio is much higher for the VRI sample, as shown in Tables A-C4 and A-C5. This suggests that the high level of wealth of in the

VRI sample is not just due to the high level of lifetime income. They likely also save more, though other differences (e.g., inherited wealth) might be relevant.

Demographics. Table A-C6 compares education, health and marital status across samples. Tables A-C7, A-C8 and A-C9 compare the distributions of each of these variables by age bins. The VRI sample has a very high education level. Approximately 70% of the sample has a college degree with over half of those having an advanced degree. The education level is higher for the individual client sample. In contrast, only about 30% of that sample has a college degree in the HRS and the SCF. If we impose the VRI-equivalent screen, however, this gap almost disappears when compared to the employer-sponsored sample in the VRI. The college degree rates from the SCF and HRS are, under VRI-eligible conditions, similar to the VRI rate. For the HRS, the gap is further reduced for the 401(k) subset. Compared to the individual client sample, the HRS and SCF rates are still lower, though the gap is reduced considerably under the VRI-eligibility condition.

The VRI respondents are much healthier than the overall population with more than 70% reporting that their health is either excellent or very good. The corresponding percentage in the total HRS is about 40%. The SCF uses a different four-point scale, without the "very good" category. The fraction of respondents with excellent health is much higher in the VRI (31%) than in the SCF (18%). The gap is much smaller, though does not fully disappear, after imposing the VRI sampling screens on the HRS and the SCF.

The fraction of coupled households (defined as either married or partnered) in the VRI is 67%, which is roughly what was targeted by oversampling administrative singles. Even after this oversampling of singles, the fraction of coupled households is larger than that in the overall sample of the HRS and the SCF. Without imposing the VRI screens, the corresponding

percentages are about 51% in the HRS and 53% in the SCF. After imposing the VRI sampling criteria, coupled rates from the HRS and the SCF overshoot the VRI levels for most of the age groups owing to the VRI's oversampling of singles.

Table A-C10 compares retirement rates. Because the incidence of retirement changes so much with age, it makes sense to compare by age groups. Overall, once the VRI screens are imposed, the retirement rates are quite similar across the SCF and VRI. HRS respondents retire somewhat earlier relative to both the SCF and the VRI.

Table A-C1. Effect of Imposing VRI Sampling Screens: Median wealth by age

Age	VRI			HRS			SCF		
	Total	Employer-sponsored	Individual client	Age Eligible	VRI Eligible	VRI eligible, 401(k) subset	Age Eligible	VRI Eligible	VRI eligible, 401(k) subset
All	663,100	496,350	715,790	60,000	272,000	342,700	33,200	262,100	219,500
55-59	518,289	428,280	607,900	55,000	226,400	283,000	21,940	208,700	197,070
60-64	601,556	521,245	669,000	58,600	276,000	364,000	36,580	236,100	225,100
65-69	715,627	574,250	750,750	83,000	350,000	435,000	57,000	299,400	463,500
70-74	746,000	671,000	755,550	64,000	310,000	434,000	52,000	410,700	348,000
75-100	726,604	605,300	729,950	50,000	284,000	334,500	27,000	275,500	143,000

Note: HRS and SCF tabulations use sampling weights.

Table A-C2. Effect of Imposing VRI Sampling Screens: Income distribution

			Percentiles				
		Mean	10	25	50	75	90
VRI	All	121,481	27,004	50,000	82,017	125,000	191,616
	Employer-sponsored	122,800	42,370	65,000	100,000	146,000	218,201
	Individual client	121,040	24,000	45,000	76,655	119,133	180,000
HRS	Age eligible	65,856	8,476	15,384	30,400	70,300	145,604
	VRI eligible	110,274	17,532	31,600	63,000	123,240	230,000
	VRI eligible, 401(k) subset	134,119	25,927	48,001	87,030	153,010	262,000
SCF	Age eligible	90,848	13,189	22,320	42,601	85,221	160,296
	VRI eligible	177,786	36,219	54,785	91,308	160,296	295,229
	VRI eligible, 401(k) subset	197,214	43,625	66,959	101,453	173,484	320,592

Note: HRS and SCF tabulations use sampling weights.

Table A-C3. Effect of Imposing VRI Sampling Screens: Median income by age

	VRI			HRS			SCF		
Age	Total	Employer-sponsored	Individual client	Age Eligible	VRI Eligible	VRI eligible, 401(k) subset	Age Eligible	VRI Eligible	VRI eligible, 401(k) subset
55-64	92,100	100,000	84,943	50,500	84,003	97,000	57,785	94,351	96,380
65-74	79,704	100,698	75,130	29,756	46,659	62,051	45,654	91,308	115,657
75-	71,755	73,343	71,703	18,660	30,432	38,437	28,407	66,553	92,322

Note: HRS and SCF tabulations use sampling weights.

Table A-C4. Effect of Imposing VRI Sampling Screens: Wealth to income ratio

| | | Mean | Percentiles | | | | |
			10	25	50	75	90
VRI (SCF measure)	All	42.97	1.95	4.28	8.37	15.15	24.13
	Employer-sponsored	57.63	0.96	2.25	4.93	8.87	14.31
	Individual client	38.05	2.74	5.31	9.77	17.17	26.30
HRS	Age eligible	44.89	-0.04	0.04	1.46	5.95	16.39
	VRI eligible	95.97	0.59	1.50	3.80	10.39	24.49
	VRI eligible, 401(k) subset	25.30	0.64	1.54	3.35	8.04	17.38
SCF	Age eligible	3.13	-0.21	0.02	0.76	3.34	7.94
	VRI eligible	5.70	0.42	1.20	3.01	6.51	13.00
	VRI eligible, 401(k) subset	4.02	0.26	1.01	2.21	4.90	8.24

Note: HRS and SCF tabulations use sampling weights.

Table A-C5. Effect of Imposing VRI Sampling Screens: Median wealth to income ratio by age

| | VRI | | | HRS | | | SCF | | |
Age	Total	Employer-sponsored	Individual client	Age Eligible	VRI Eligible	VRI eligible, 401(k) subset	Age Eligible	VRI Eligible	VRI eligible, 401(k) subset
55-64	5.90	3.79	7.13	1.01	2.70	2.88	0.53	2.24	2.01
65-74	9.53	5.16	10.1	1.71	5.89	5.88	1.01	4.38	3.27
75-	11.36	9.36	11.11	2.55	9.08	9.85	0.92	4.87	1.41

Note: HRS and SCF tabulations use sampling weights.

Table A-C6. Effect of Imposing VRI Sampling Screens: Education, Health, and Marital Status.

		VRI			HRS			SCF		
		Total	Employer-Sponsored	Individual client	Age Eligible	VRI Eligible	VRI eligible, 401(k) subset	Age Eligible	VRI Eligible	VRI eligible, 401(k) subset
Education	College grad.	32.18%	33.69%	31.67%	14.25%	22.62%	23.26%	16.26%	27.43%	25.87%
	Post grad.	38.45%	26.24%	42.53%	14.64%	26.36%	30.54%	14.32%	28.39%	28.55%
Health	Poor	0.84%	0.53%	0.94%	7.60%	2.25%	1.71%	10.32%	2.50%	2.42%
	Fair	4.77%	3.48%	5.20%	19.10%	11.10%	9.01%	26.19%	15.67%	17.02%
	Good	21.77%	22.33%	21.58%	31.81%	29.39%	30.29%	45.34%	55.46%	53.51%
	Very good	41.84%	42.25%	41.71%	31.43%	41.30%	42.27%			
	Excellent	30.78%	31.42%	30.57%	10.06%	15.95%	16.71%	18.14%	26.37%	27.05%
Marital Status	Coupled	67.21%	73.88%	64.97%	52.46%	69.89%	77.82%	53.18%	71.04%	74.97%
	Single	32.79%	26.12%	35.03%	47.54%	30.11%	22.72%	46.82%	28.96%	25.03%

Note: HRS and SCF education is based on years of schooling (college grad is exactly 16 years and post-grad is more than 16 years). VRI education is based on degree attainment. SCF health has a four-point scale, while VRI and HRS health have five-point scales. HRS and SCF tabulations use sampling weights.

Table A-C7. Effect of Imposing VRI Sampling Screens: Fraction with College Degree by Age

	VRI			HRS			SCF		
Age	Total	Employer-sponsored	Individual client	Age Eligible	VRI Eligible	VRI eligible, 401(k) subset	Age Eligible	VRI Eligible	VRI eligible, 401(k) subset
55-64	68.38%	57.61%	78.69%	32.12%	48.92%	50.30%	40.83%	61.96%	60.04%
65-74	73.08%	66.83%	74.18%	26.67%	46.78%	55.18%	39.48%	66.64%	68.12%
75-	69.52%	54.27%	69.82%	21.28%	46.03%	64.19%	20.85%	52.82%	29.06%

Note: Education is based on attainment. HRS and SCF tabulations use sampling weights.

Table A-C8. Effect of Imposing VRI Sampling Screens: Fraction with Very Good or Excellent Health by Age

	VRI			HRS			SCF		
Age	Total	Employer-sponsored	Individual client	Age Eligible	VRI Eligible	VRI eligible, 401(k) subset	Age Eligible	VRI Eligible	VRI eligible, 401(k) subset
55-64	75.61%	73.43%	77.70%	43.82%	57.82%	59.73%	19.81%	25.92%	24.77%
65-74	75.35%	74.30%	75.54%	43.69%	58.74%	57.26%	23.67%	32.43%	38.77%
75-	61.13%	74.29%	60.87%	34.85%	51.38%	56.25%	10.96%	8.91%	0.28%

Note: SCF does not have 'Very Good' category, so the fraction captures respondents with Excellent health only. HRS and SCF tabulations use sampling weights.

Table A-C9. Effect of Imposing VRI Sampling Screens: Fraction Married or Partnered by Age

	VRI			HRS			SCF		
Age	Total	Employer-sponsored	Individual client	Age Eligible	VRI Eligible	VRI eligible, 401(k) subset	Age Eligible	VRI Eligible	VRI eligible, 401(k) subset
55-64	66.05%	73.72%	58.69%	58.88%	72.05%	77.28%	58.45%	71.78%	73.27%
65-74	68.65%	74.82%	67.57%	56.60%	69.95%	79.06%	56.26%	72.70%	78.88%
75-	66.26%	65.72%	66.26%	36.46%	60.74%	80.57%	40.23%	60.82%	97.12%

Note: HRS and SCF tabulations use sampling weights.

Table A-C10. Effect of Imposing VRI Sampling Screens: Retirement Rate by Age

	VRI			HRS			SCF		
Age	Total	Employer-sponsored	Individual Client	Age Eligible	VRI Eligible	VRI eligible, 401(k) subset	Age Eligible	VRI Eligible	VRI eligible, 401(k) subset
All	55.80%	17.78%	68.52%	63.99%	53.23%	36.70%	56.56%	33.92%	16.87%
55-59	9.43%	4.57%	14.75%	24.42%	19.61%	13.84%	19.88%	7.65%	5.34%
60-64	26.68%	12.39%	38.86%	50.25%	42.05%	34.10%	38.62%	24.56%	15.90%
65-69	62.14%	34.13%	69.91%	76.50%	73.16%	66.15%	59.72%	44.39%	34.15%
70-74	81.23%	57.96%	83.31%	87.18%	85.16%	80.70%	77.06%	67.07%	49.44%
75-100	91.38%	74.29%	91.72%	91.57%	92.95%	90.84%	92.16%	87.37%	69.44%

Note: HRS retirement rate includes respondents with partial retirement. For SCF retirement rate variable 'OCCAT1' in the public version of data is used. Households are defined to be retired if 'OCCAT1=3', which also includes disabled, age +65 and not working, etc. HRS and SCF tabulations use sampling weights.

The Wealth of Wealthholders:
Evidence and Implications from the Vanguard Research Initiative
December 2014

Appendix D. Estimating Retirement/Wealth Relationship

HRS sample. Table A-D1 shows how many observations we lose in the HRS by imposing each additional condition on samples used. As we have seen from Table 5, the majority of the HRS samples are older than 65. Among those households in which the main breadwinner satisfies the age condition, some are retired while some have dual main breadwinners. In addition, for many households that are not retired, responses for the expected retirement age are missing. [1] All of these conditions account for the small sample size used in the HRS.

LOESS curve and scatter plots including outliers. In Figure A-D1, we show the estimated relationship between retirement plan and wealth from the VRI (Panel A) and the HRS (Panel B) for the full range.

Estimation with future DB pension and Social Security income included in the normalized wealth. In the LOESS estimation in Section 5, expected DB pension and Social Security income are included as a control (Y_i^R). Here, we estimate another version of the model where we define the normalized wealth as the sum of the replacement rate from the annuitizable financial wealth and that from the expected annuity income (Y_i^R). Figure A-D2 shows the distribution of newly defined normalized wealth and Figure A-D3 shows the new LOESS estimates. For both figures, Panel A is for the entire sample used in Section 5. Panel B is for the employer-sponsored subsets.

---

[1] Some breadwinners who are not retired report that they are not currently working, leading to missing responses for expected retirement age. In addition, questions about retirement age are asked only when the respondents said that they plan to retire or stop working.

Figure A-D2A shows that the VRI sample still has higher replacement rates, though the gap is less stark than in Figure 3A. The VRI has many observations in the range between 1 and 2, while for the HRS, most of the observations have normalized wealth smaller than 1. The LOESS estimate (Figure A-D3A) shows basically the same relationship as the baseline model (Figure 4A). With the VRI sample, we can estimate a negative and statistically significant relationship for a wider range (between 0 and 2), while the HRS sample shows a steeper slope up to about 0.5 but then becomes flat and statistically insignificant. With the employer-sponsored subset, the distributions of normalized wealth are pretty similar across the VRI and HRS (Figure A-D2B). Figure A-D3B shows that conditioning on this subset does not affect the estimated relationship between wealth and retirement plan for the VRI, while for the HRS, the estimates get very noisy due to the small number of observations.

Table A-D1. HRS Sample Size for Retirement Horizon Analysis: Effect of Each Condition

Condition	Number of observations
(1) None	11,595
(2) Main breadwinner age ≤ 65	5,206
(3) (2) + Main breadwinner not retired, No dual breadwinner	2,442
(4) (3) + Have expected retirement age	1,053

Figure A-D1. Retirement horizon versus normalized financial wealth: LOESS
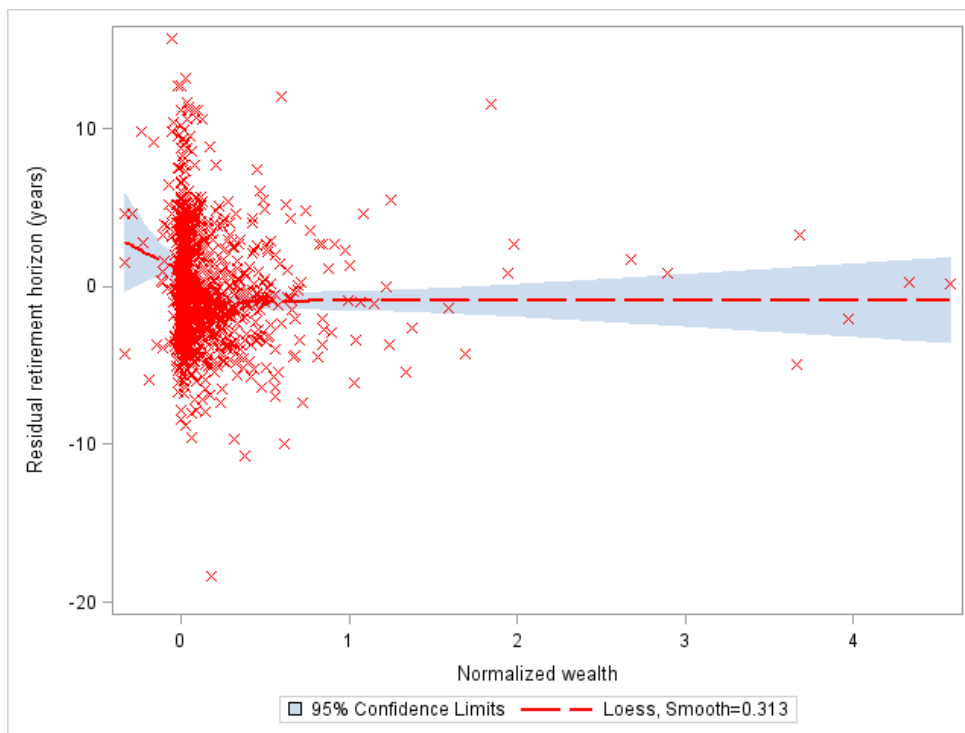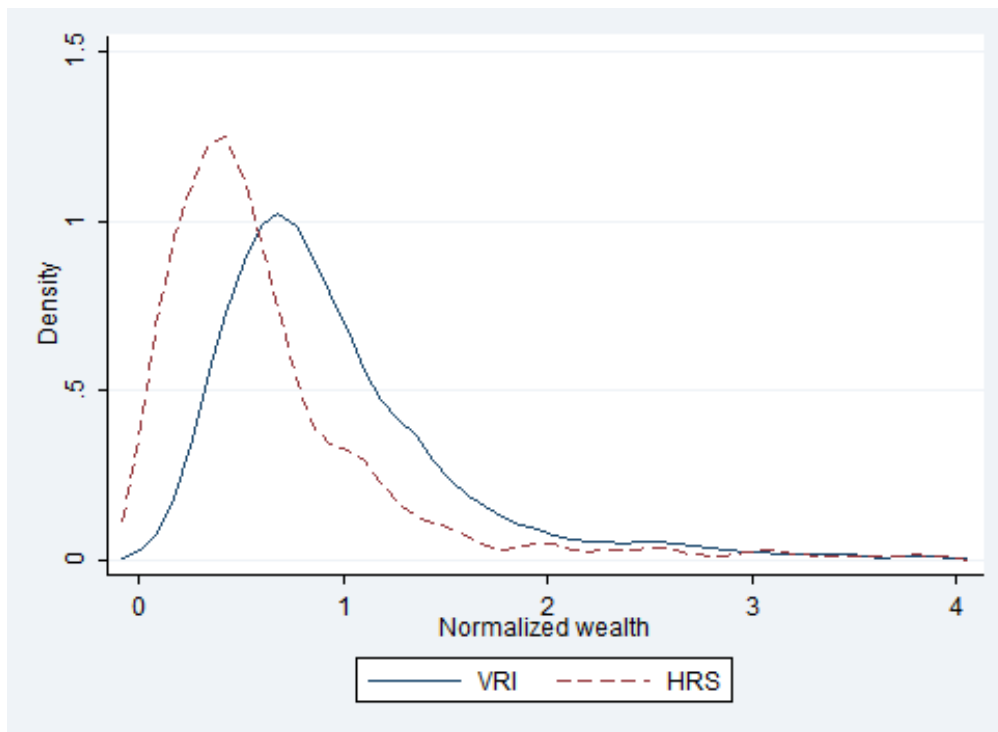(full range of data)
A. VRI
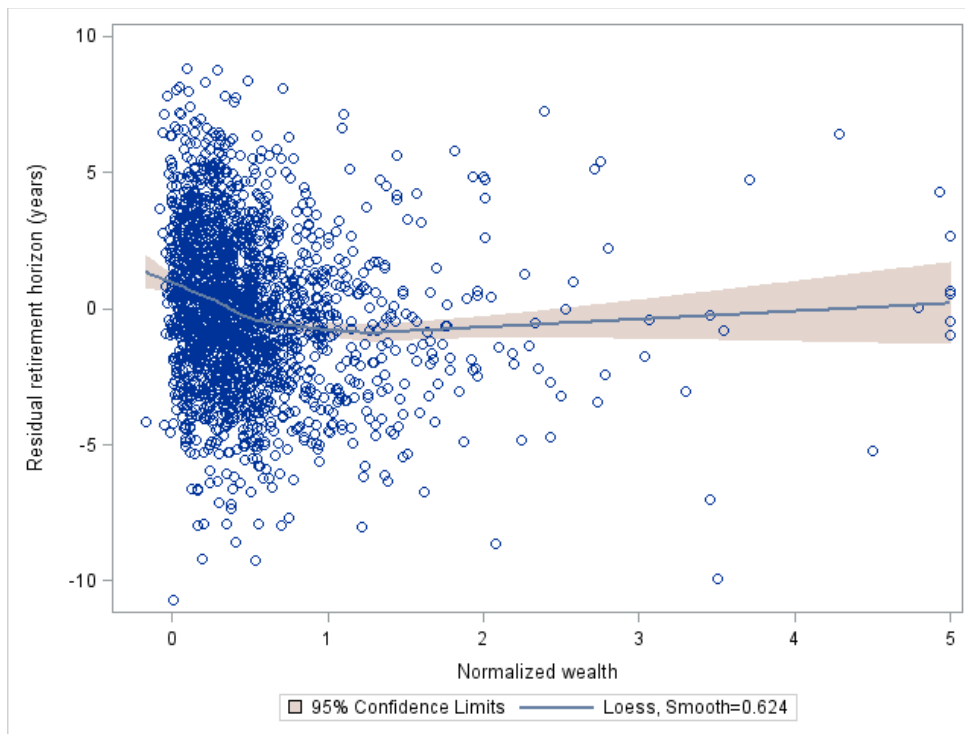
B. HRS

Figure A-D2. Distribution of normalized financial wealth (including future DB pension and SS income)

A. VRI vs HRS

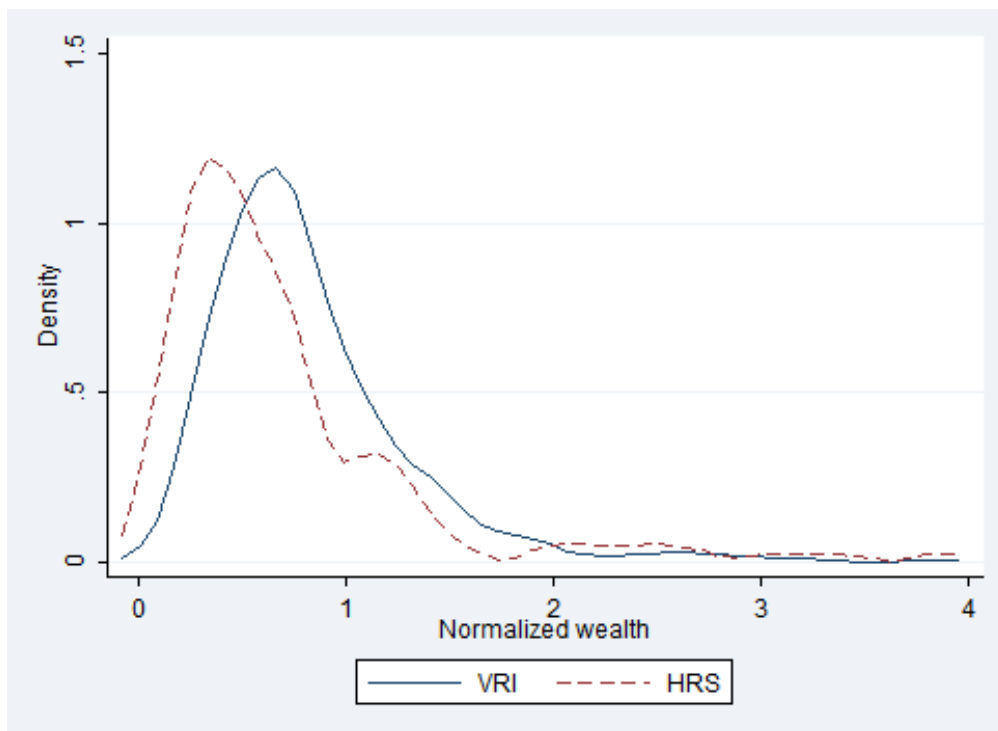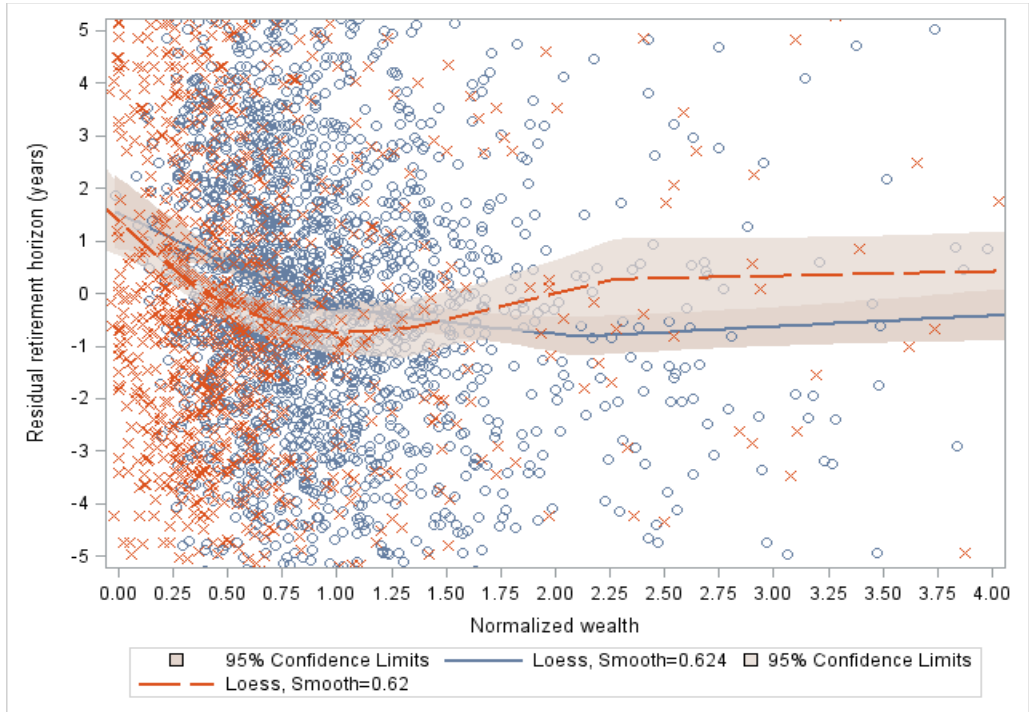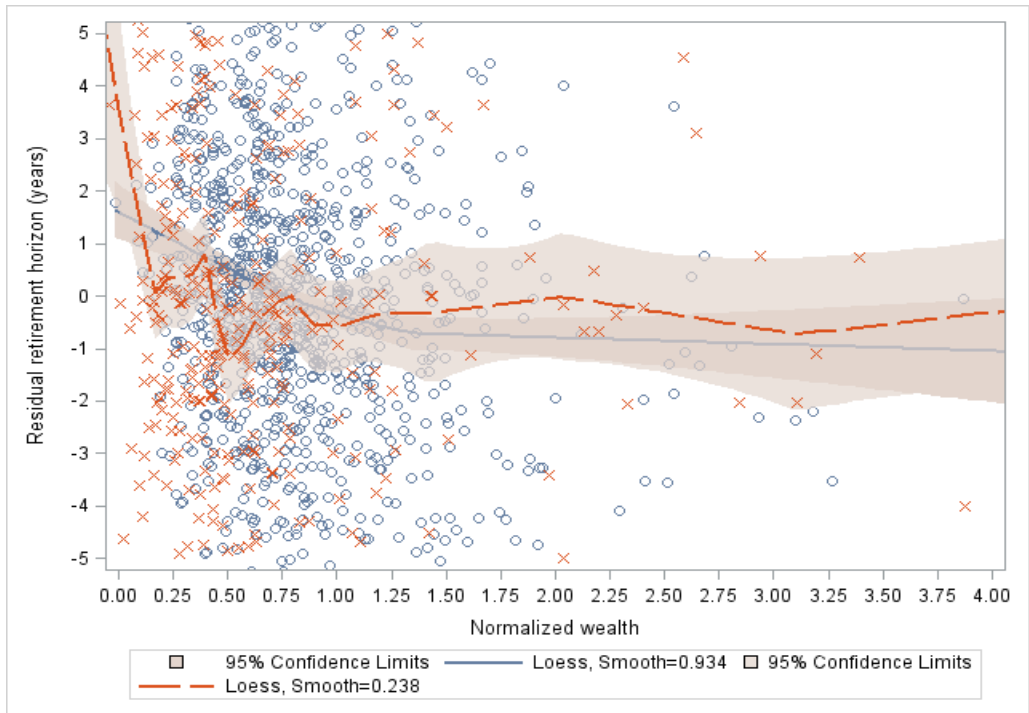B. VRI employer-sponsored versus HRS 401(k) subset

Figure A-D3. Retirement horizon versus normalized financial wealth: LOESS
(Normalized wealth including future DB pension and SS income)
A. VRI vs HRS

Note: x denotes HRS (orange) and o denotes VRI (blue).

B. VRI employer-sponsored versus HRS 401(k) subset

Note: x denotes HRS (orange) and o denotes VRI (blue).