

The Wealth of Wealthholders

John Ameriks, The Vanguard Group, Inc.

Andrew Caplin, New York University and NBER

Minjoon Lee, Carleton University

Matthew D. Shapiro, University of Michigan and NBER

Christopher Tonetti, Stanford GSB and NBER

Revised September 2023

This research is supported by a program project grant from the National Institute of Aging P01-AG026571. The Vanguard Group, Inc. supported the data collection through the Vanguard Research Initiative (VRI). Vanguard's Client Insight Group and IPSOS SA were responsible for implementing the VRI survey and provided substantial input into its design. The design of the VRI benefited from the collaboration and assistance of Annette Bonner, Joseph Briggs, Wandu Bruine de Bruin, Alycia Chin, Brooke Helppie McFall, Wendy O'Connell, and Ann Rodgers.

The Wealth of Wealthholders

ABSTRACT

Wealth, though crucial for modeling economic behavior and understanding well-being, is difficult to measure in surveys. This paper introduces a new, comprehensive account-by-account approach for eliciting asset holding. This approach is implemented in the Vanguard Research Initiative, a panel of wealthholders designed to yield high-quality measurements for a large sample of older Americans with significant financial assets. Because survey responses are linked to administrative account balances, this paper can show that the approach yields precise, unbiased estimates. Having accurate and dense data on the wealth of wealthholders provides sharper inferences on wealth management behavior as well as on relationships between wealth and economic behavior than is possible in leading datasets.

John Ameriks
The Vanguard Group, Inc.
P.O. 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
Carleton University
Ottawa, ON K1Y 2G1, Canada
minjoon.lee@carleton.ca

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
and NBER
tonetti@stanford.edu

1. Introduction

Wealth is a fundamental determinant of the well-being of households. This motivates a rising concern about whether older U.S. households enter retirement with enough financial savings (Poterba, 2014). Furthermore, understanding and predicting households' economic behaviors require information not only on the amount of wealth they own but also on its composition (e.g., Kaplan and Violante, 2014, show the importance of liquid wealth in predicting the marginal propensity to consume). High-quality measurement of wealth and portfolio allocations is essential in properly answering many economic questions.

Wealth, however, is hard to measure. Measurement issues of existing wealth data raise concerns about findings drawn from them (see Henriques and Hsu, 2014, and Bricker, Henriques, Krimmel, and Sabelhaus, 2016a, 2016b, for example, for gaps between the trends in wealth inequality from different data sources and efforts to reconcile them). Surveys have been the most conventional approach to measuring wealth. Unlike income, which respondents typically see periodically, measuring wealth has its own challenges. Respondents might not monitor their wealth at high frequency; wealth may be held in disparate accounts; and more generally, respondents might not have a good grasp of wealth concepts. Existing survey approaches have had limited success in overcoming these challenges (see Bucks and Pence, 2015, for a good summary of issues in existing survey measures and Section 2 of this paper and papers cited therein for discussions on specific issues). Relying on administrative data alone may not be an alternative because they often capture only a part of households' portfolios. The current status of wealth measurement calls for progress in measurement methodologies, and recently there has been an increasing amount of discussion on this issue (see Moffitt, Schoeni, Brown,

Chase-Lansdale, Couper, Diez-Roux, Hurst, and Seltzer, 2015 and Samphantharak, Schuh, and Townsend, 2018, among others).

In this paper, we suggest a way to advance the literature by developing improved measurement protocols that can be generally applicable in household surveys. There are two main innovations made in our approach. The first innovation is to use a new, comprehensive account-by-account approach for eliciting asset holding on surveys. Under this approach, respondents are asked to report their financial assets account-by-account. The aim of this approach 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.

The second innovation is to use a novel correction mechanism that helps respondents detect and correct mistakes in their responses. After collecting the number of accounts of each asset type, the survey presents a summary account inventory, which helps respondents understand the structure of their own financial portfolio and also detect and correct any mistakes in their responses. A similar snapshot of the financial portfolio is presented after collecting the balances of the provided accounts.

We demonstrate achievements from our approach by using the Vanguard Research Initiative (VRI), a linked survey-administrative dataset on account holders at a mutual fund company (Vanguard).¹ The VRI provides a unique opportunity to evaluate the quality of survey wealth measures by comparing them to corresponding administrative record data. The comparison reveals that our comprehensive account-by-account approach yields unbiased and

¹ The full VRI survey questionnaire can be found at https://ebp-projects.isr.umich.edu/VRI/surveys/VRI_survey1.pdf.

precise measures of wealth. In addition, the correction mechanism based on snapshots of portfolios turns out to be highly effective in reducing response errors. For those who used the correction mechanism, the gap between the post-correction survey and administrative measures tends to be significantly smaller than that from the pre-correction measures. Based on this evidence, we argue that the new approach we suggest in this paper provides a way to improve wealth measurement in household surveys.

The sampling framework used in the VRI also addresses another issue in existing household surveys, in that they often do not have ample observations of wealthholders, which we define in this paper as those with significant financial assets. Suppose we want to understand how older Americans, facing increasing responsibility for financing retirement due to the aging of the population, manage their financial assets and how these assets affect their behaviors including retirement. Though the transition from a defined benefit to a defined contribution retirement system has been underway for decades, about half of households approaching or in retirement have relatively low financial assets. Datasets designed to represent the population, therefore, have surprisingly little information on older Americans with wealth sufficient to finance a non-trivial fraction of their retirement consumption (see Gustman, Steinmeier, and Tabatabai, 2010, Poterba, Venti, and Wise, 2011, and Poterba, 2014).

The VRI fills this gap by producing an innovative new dataset containing a large number of households with significant financial assets to potentially use in retirement. To highlight the value of these data, this paper shows that having ample observations of wealthholders and accurate wealth measurement allows a better understanding of behavioral relationships that include wealth and portfolios. By way of examples, this paper shows how improved data can provide sharper inferences than in leading datasets such as the Health and Retirement Studies

(HRS) and Survey of Consumer Finances (SCF) on issues including liquidity constraints, asset location, and the relationship between financial wealth and retirement plans.

The remainder of the paper is structured as follows. In Section 2, we discuss issues in measuring wealth. Section 3 introduces the innovations in wealth measurement we implemented in the VRI. Section 4 establishes the high quality of the survey responses by comparing them to the corresponding administrative data. In Section 5, we demonstrate that the VRI has dense observations of wealthholders. Section 6 presents three examples where accurate and dense measures of the wealth of wealthholders help improve our understandings of their behaviors. Section 7 concludes.

2. Issues in Measuring the Wealth of Wealthholders

What makes wealth so hard to measure? The fundamental issue is that it is an elusive object, even for those who own it. It often has a complex structure, and the value of some items may change frequently. Hence, measuring it using a survey, which has been the most conventional approach, might result in poor measurement unless the survey is well designed to help respondents recognize what they own and map them to categories used in the survey. In existing surveys, however, questions are often asked in categories that are meaningful to economists but not as much so to respondents.² (See Juster and Smith, 1997 and Bucks and Pence, 2015, for a

² The HRS and SCF, for example, 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, 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

good summary of measurement issues in existing surveys.) Then the respondents need to perform a difficult task of summing and allocating balances over different categories. Reporting the composition of wealth (e.g., stock shares) may be even more difficult. Difficult questions to answer result in a higher item non-response rate or a high usage of unfolding brackets. (See Fries, Starr-McCluer, and Sundén, 1998 and Kennickell, 2015 for the item non-response rate and the share of bracket responses in wealth measures in the SCF and Juster and Smith, 1997, for those in the HRS.) Though with imputations one can replace missing values with the most likely values given other observable characteristics and bracket responses provide meaningful information without much cognitive burden, it goes without saying that the first-best solution is to elicit responses that are precise and do not require imputations.

Our wealth survey starts with concepts that are easy to understand for respondents (closest to how they perceive their financial wealth), processes those responses in real time, and provides a simple but comprehensive overview of the current financial situation. Such a snapshot, on the one hand, allows respondents to learn about their own financial situation, and as we discuss below, many respondents indeed appreciate an opportunity to learn from their own responses. On the other hand, it can help respondents detect and correct mistakes they made, and we show that a correction mechanism combined with such a snapshot significantly improves the accuracy of responses. But for the existing household surveys, including HRS and SCF, given that they are not internet-based and do not deploy a comprehensive account-by-account

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.

approach, it is hard to provide an efficient way for respondents to detect and correct errors in responses.

Another key drawback of the existing survey measures is that often there is no benchmark that enables researchers to evaluate the quality of survey measurement. Without corresponding administrative data, it is not possible to detect any bias or judge the accuracy of the survey measures. This also hinders research on effective survey approaches in eliciting high-quality responses. Often, there are nonnegligible discrepancies between wealth distributions measured in different surveys, but these do not tell us which survey measure is more accurate (see Juster, Smith, and Stafford, 1999, Sierminska, Michaud, and Rohwedder, 2008, Pfeffer, Schoeni, Kennickell, and Andreski, 2016, and Eggleston and Gideon, 2017).

An alternative approach to measure wealth is to use administrative data (e.g., TIAA-CREF data by Ameriks and Zeldes, 2004 and data from a financial mobile application by Gelman, Kariv, Shapiro, Silverman, and Tadelis, 2014). This approach does not depend on the accuracy of human-provided responses, so measurement error is likely to be small.³ Administrative data, however, often capture only a part of a financial portfolio that is held at the financial institution providing the data. Those data have limited demographic variables as well as other measures including beliefs, preferences, work history, etc. These pose a limit in understanding the relationships between financial wealth and household behavior.

Depending on the research question at hand, limited representation of wealthholders in representative datasets can be another issue. Suppose that we study how older Americans use their financial wealth (e.g., annuitization, long-term care insurance purchase, etc.) to prepare for

³ But even administrative data are not entirely free from measurement errors. These may occur during processing the account data (see Browning, Crossley and Winter, 2014, that discusses usage of administrative data in measuring consumption expenditure).

risks they face in retirement. As we illustrate in Section 5, about one-third of households that are near or in retirement have negligible assets, and another one-third have less than \$100,000.

Hence, representative data sets such as the HRS do not provide enough observations for these research questions that are more relevant to wealthholders. See also Gustman, Steinmeier, and Tabatabai (2010), Poterba, Venti, and Wise (2011), and Poterba (2014) on the same point.

The SCF does a better job of capturing wealthholders by oversampling high-income individuals using a list sample based on tax return data (Kennickell, 2007). On the other hand, the SCF is representative of all age groups, and when it comes to a certain age group, the number of observations of wealthholders is still small (see Section 5). Also, the SCF is a repeated cross-section, which makes it harder to understand dynamic interactions between wealth and other economic behaviors.

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

We introduce three innovations in the wealth measurement implemented in the VRI. First, it surveys financial wealth by accounts, not by asset classes. Its aim is to ask respondents to report numbers that closely correspond to how they receive statements and how they might classify assets. The approach avoids asking respondents to map their balances into accounting or economic constructs, and does not require them to do addition or distribution of amounts.

Second, after each step where the survey instrument elicits the composition or amount of assets, it shows a summary of responses in tabular form and allows respondents 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.⁴ In

⁴ 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. This research is therefore related to an emerging

this section of the paper, we describe how the wealth measurements are implemented in the VRI survey.⁵

Table 1 shows in tabular form the main survey design elements and how they compare with those of the HRS and SCF. Section 5 provides a detailed comparison of these surveys.

3.1. 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. Appendix B shows screenshots of the wealth section for a hypothetical respondent. The steps in the wealth section are as follows.

Step 1: Account Type. 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, C.D.s, brokerage, and directly held securities); "Insurance-related accounts" (annuities with cash value and life

program to augment or replace survey data with administrative records, including private account data. See Gelman, Kariv, Shapiro, Silverman, and Tadelis (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.

⁵ Note that the VRI stratifies the sample between individual versus employer-sponsored accounts. The self-selection to be a Vanguard client is less of an issue for the latter sample as it reflects their employers' choice to provide retirement plans at Vanguard. For those under age 65, the VRI sampled evenly between the two client types. After age 65, we sample the types in the proportion they appear in the population. See Appendix A for the VRI sample design. For description of the VRI in greater detail including the specifics of sampling, testing, and design, as well as more-detailed tabulations of economic and demographics variables, see Ameriks, Caplin, Lee, Shapiro, and Tonetti (2014).

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. Respondents are asked to indicate the 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. Respondents are asked to give a nickname for each account. After the respondent enters all the nicknames, the survey displays a summary table (see Appendix B, Figure A-B4). 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 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-B6). 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-B7.) 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-B8.) 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 the 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-B13 for an example). The survey prompts respondents to print out this page, if desired. This summary was provided in the hope that this potentially useful measurement for survey respondents would increase the likelihood of their continued participation in the survey.

3.2. 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 one 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, mutual funds, and brokerage accounts are also significant assets in the population of Vanguard account holders.

Ameriks, Caplin, Lee, Shapiro, and Tonetti (2014) describes how the VRI collects 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.

4. 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. The administrative data can be used to verify the survey measures. Additionally, administrative data can supplement survey data by providing alternative measures of wealth, potentially at very high frequency. 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.

4.1. 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 cause 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 the 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-B6 (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 declines 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.⁶ For the employer-sponsored sample, the median difference is \$890, or 0.6%; for the individual client sample, the median

⁶ The administrative data are the weighted average of the end of months before the survey and after the survey with the weight equal to 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.

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 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, the 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.⁷

⁷ We are also able to examine whether checking records matters for accuracy of survey responses. Interestingly, checking records shrinks the difference between administrative and

4.2 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 modify their answers after seeing the summary tables, we can check whether or not their answers get closer to the 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 3.1), 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. These respondents typically went back to the start of the wealth section to correct the inventory of their accounts after having entered balances.

Overall, 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 correction paths they took. Again, the comparisons are done only for singles.

survey reports, but being logged on to the Vanguard website during the survey does not play a significant role in this result. See Ameriks, Caplin, Lee, Shapiro, and Tonetti (2014).

When respondents do not make any corrections, their initial responses are already very close to the administrative information. The interquartile range is -3.3% to 2.6% for those who make no corrections; for those who correct account inventory only, it is very similar, -3.5% to 2.5%. For respondents who correct 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, 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 make no corrections at all. Therefore, the correction mechanism did prove to be effective.

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

In addition to implementing new approaches to wealth measurement as described above, the VRI sample also focuses on households with nonnegligible financial wealth approaching or in retirement. This section addresses two questions related to this sampling choice. First, why is it needed? The answer is that leading surveys aimed at measuring wealth contain remarkably few respondents in that age range with significant levels of wealth. Second, how similar are VRI survey respondents to the U.S. population, once we condition on those who have similar wealth levels? We answer these questions through a detailed comparison of the VRI with the HRS and SCF.

5.1. 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 nonnegligible 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 Appendix A, 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 in 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 the 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 of 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 wealthholders while the HRS and SCF are representative, that is, they include the older Americans with very low assets, who are about half the population. The next set of results explores these differences and shows the extent to which they derive from VRI sampling restrictions.

5.2. 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.⁸ It reminds us how little financial wealth is owned by the lower half of the older household population. 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 of more than \$100,000, and one-third of them have more than one million dollars. In contrast, the HRS distribution has a very fat left tail. One-third of the HRS sample has

⁸ The wealth measure used in the comparisons is total net financial wealth. Values of houses and mortgages are excluded. See Appendix C for the definition of the total financial wealth for each survey and how we impose similar sampling screens in the VRI, HRS, and SCF.

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 VRI 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, given that the SCF is not aiming at the population near or after retirement, for most of the wealth-age bins with nonnegligible 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 relative 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 the *VRI-eligible* screens in the HRS and SCF. The screen requires Internet eligibility and that households have at least \$10,000 in non-transactional financial accounts.

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 the "VRI eligibility" (Internet eligibility and the \$10,000 minimum balance in non-transactional financial accounts). The third column imposes the "401(k) subset" (at least \$10,000 in a D.C. pension account), which yields a subsample that is comparable to the employer-sponsored sample in the VRI. Note that the conditions for the VRI eligibility are imposed *ipso facto* in the VRI for both employer-sponsored and individual client groups.⁹ For the HRS and SCF, the VRI eligibility screen yields relatively small subsets of age-eligible respondents. For the HRS, only about a third satisfy the VRI eligibility. In the SCF, a relatively larger fraction of households satisfy these conditions owing to the oversampling of high-income households. The size of the 401(k) subset group is much smaller in both the HRS and the SCF. In the VRI, the age distribution is flat by design. (For all three surveys, 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 D 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 samples.¹⁰ 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

⁹ The two subsamples in the 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.

¹⁰ 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.

conditions are highly correlated, there is only an 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 D.C. 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. 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 the 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' decisionmaking. To more closely mimic the asset cut-off imposed on the employer-sponsored group in the VRI, we also report results from the HRS and SCF 401(k) subset (the third row in each panel). 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 a broad similarity between the 401(k) subsets of the SCF and HRS and the VRI employer sample.

Appendix D 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 higher income, a much higher education level, better health, and a greater likelihood of being coupled. Most of these differences, again, can be explained by the effect of the sampling screens we imposed in the VRI panel. Once these screens are imposed, the VRI looks quite similar to the upper half of the wealth distribution in the HRS and SCF. There is a bit of residual higher education, better health, and a higher 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.

5.3. Stock Share

The extent of stock ownership looms large in discussions of how individuals will manage under D.C. pension plans. The VRI wealth survey asks for stock share on an account-by-account basis. Table 9 compares the stock ownership of the VRI with those of the HRS and SCF. Panel A reports stock shares, while Panel B reports stock amounts. Again, we see the importance of having a relevant sample. Note that the VRI has a much larger sample of stockholders, so any analysis of portfolios should be much more precise. Compared to the VRI, if we impose only age eligibility, about half of the HRS and SCF sample do not own stocks. Conditioning, however, on the VRI sample screens, almost all the observations in the SCF own stocks, while the left tail in the HRS still shows less stock ownership.¹¹ The median shares in the HRS and SCF are still lower but much closer to those of VRI with the VRI sample screens. The picture is similar with

¹¹ Note that the HRS 2012 stock shares in 401(k) or similar accounts are heavily imputed, so they are excluded (numerator and denominator) from these HRS stock shares.

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.

6. Wealth and Economic Behavior: The Importance of Understanding Wealth of Wealthholders

The VRI approach yields accurate estimates not only of the total wealth of households but also of its subcomponents. Recent theories on household behavior emphasize the importance of different roles played by various components of wealth (e.g., the importance of liquid wealth in predicting marginal propensity to consume; see Kaplan and Violante, 2014). Having an accurate measurement of wealth by category allows us to contribute to this literature. Furthermore, the good measurement of the composition of wealth allows us to contribute to the literature on portfolio management.

The VRI approach also leads to substantially larger samples of older households with relevant levels of wealth for many important decisions surrounding retirement and well-being in older age. Given the skewness in the wealth distribution, population-representative samples will have small samples of wealthholders. Therefore, it is very difficult to make inferences on the relationship between wealth and economic behavior. Barsky, Bound, Charles, and Lupton (2002) make this point compellingly for the study of racial differences in the income-wealth relationship (see also Charles and Guryan, 2011, for this point in a different context). Having dense observations across the relevant ranges of wealth is particularly important if the relationships between wealth and other behaviors are non-linear.

In this section, we use the VRI to provide new evidence surrounding issues of saving and portfolio choice that are difficult to address in even the current best surveys, which have few

observations of older households with nonnegligible financial assets. Specifically, we focus on the following three questions:

1. How does the share of households that are liquidity-constrained change by age and retirement status?
2. How do households locate stocks and bonds between tax-deferred accounts and taxable accounts?
3. What is the relationship between financial wealth and the retirement horizon of households?

Investigating these behaviors and relationships is crucial for understanding the dynamics of household wealth around retirement, which in turn have a fundamental impact on their financial well-being in retirement.

6.1. Wealthy-Hand-to-Mouth in Old Ages

Kaplan and Violante (2014) show the importance of measuring liquid wealth, as opposed to total wealth, in predicting the marginal propensity of consumption of households. Their study only includes households with a husband younger than 60 years old. Applying their definition of liquidity-constrained household, which is defined as liquid wealth being under half of the regular income stream, to older households is tricky for two reasons. First, wealth in tax-deferred accounts, which is classified as illiquid wealth in their analysis, becomes more liquid as households get older. Second, the frequency of income may change in retirement (e.g., receiving monthly Social Security checks instead of biweekly paychecks). Nonetheless, estimating changes in the liquid wealth to income ratio by age and retirement status can be the first step in understanding how their economic activities, including consumption, will react to changes in their economic environment.

Following Kaplan and Violante (2014), we define liquid wealth as the sum of money market, checking, savings, mutual funds, brokerage accounts, and directly held securities, net of credit card balances. Income is defined as the sum of wages, Social Security, pensions, and

disability benefits, from both members in the case of couples. We assume the frequency of income to be monthly based on that of Social Security. Households are considered to be liquidity-constrained if their liquid wealth is less than half of their monthly income. Liquidity-constrained households that own other forms of wealth (e.g., housing equity, tax-deferred accounts, etc.) are classified as wealthy-hand-to-mouth. Essentially all the liquidity-constrained households in the VRI are wealthy-hand-to-mouth, so we do not separately consider these two concepts. We also implement the same analysis using the SCF, which is used in Kaplan and Violante (2014), but focusing on the VRI age range (55 and above).

According to the above definition, 15 percent of the VRI sample is wealthy-hand-to-mouth. The corresponding number in the SCF sample is 26 percent. In addition, Table 10 shows clear patterns by age group and retirement status. In the VRI, the share of wealthy-hand-to-mouth households decreases from 30 percent for ages between 55 and 59 to 4 percent for ages above 75. It also decreases from 24 percent for those who are not completely retired to 7 percent for those who are completely retired. Even though somewhat less stark, we find the same qualitative pattern in the SCF sample.

This pattern is driven by changes in both income and liquid wealth. Not surprisingly, monthly income reduces over age and with retirement. The VRI results also show that the amount of liquid wealth substantially increases over age and with retirement. This change may be driven by withdrawal from tax-deferred accounts and receiving pensions in lump-sum, though with the cross-sectional data, we cannot establish it. These changes imply that households' marginal propensity to consume may vary significantly over age and with retirement. Similar patterns can be observed from the SCF, except for the increase in liquid wealth over age and with retirement. In the SCF, more than half of the sample in any age or retirement group has less than

\$10,000 in liquid wealth, making it hard to detect the changing patterns in liquid wealth found in the VRI. The results from the VRI-eligible SCF sample show an increase in liquid wealth over age and with retirement, but not as starkly as in the VRI, possibly due to the small size of the sample in particular at older ages.

6.2 Asset Location between Tax-Deferred Accounts and Taxable Accounts

What types of financial assets to put in tax-deferred accounts (TDAs) versus taxable accounts (TAs) is an important financial choice for households engaging in asset accumulation for retirement. Stocks are typically considered tax-efficient as capital gains are not taxed until they are realized, while fixed-income assets such as bonds are considered tax-inefficient. Hence, to minimize the expected value of the tax, it has been shown that one should put stocks in TAs and fixed-income assets in TDAs as much as possible (Dammon, Spatt, and Zhang, 2004).¹²

Bergstresser and Poterba (2004) and Barber and Odean (2003), using the SCF and a brokerage account dataset, respectively, examine whether actual asset location patterns of households are close to this principle. They find that households tend to keep similar stock-bond allocations between TDAs and TAs, instead of putting the stock holdings in TAs as much as possible.¹³

Bergstresser and Poterba (2004), however, also show that the size of errors measured in terms of

¹² Shoven (1999), Poterba, Shoven, and Sialm (2001), and Shoven and Sialm (2003) point out that actively managed stock funds are not tax-efficient as they realize capital gains frequently. Dammon, Spatt, and Zhang (2004) show that holding actively managed funds in TDAs is still not an efficient option unless they substantially outperform other forms of tax-efficient stock investments such as passive index funds. Following the literature on asset location, we categorize assets into stocks and fixed-income assets, and study how these two types of assets are located. Also note that we do not distinguish tax-exempt bonds such as municipal bonds from general bond holdings, because in the VRI survey we did not ask about types of bond holdings. Whether we treat tax-exempt bonds as tax-efficient or tax-inefficient did not noticeably change the results from the SCF. On the other hand, according to the administrative data on the VRI sample's assets held at Vanguard, only about 2 percent of their TA wealth is held in such bonds.

¹³ Garlappi and Huang (2006) shows that if households care about the volatility of future tax as well as the expected level of it then having similar stock-bond allocations in TDAs and TAs can be closer to the optimal choice.

the amount of transactions households have to make to have the tax-efficient location is typically very small.

We revisit this question using the VRI. To be specific, we examine which of the following two hypotheses on the asset location patterns considered in the literature is closer to the actual behavior of households: having the tax-efficient location pattern (H1) and having the same stock-bond allocations in TDAs and TAs (H2). The large number of observations of households with nonnegligible amounts of financial wealth in both TDAs and TAs, as well as non-trivial portfolio allocation between stocks and bonds, allow us to tell which hypothesis is closer to reality more clearly. Accurate measures of wealth and stock share that are not based on imputation improve the comparison even further. We also compare the results from the VRI with those from the SCF to show that a clear comparison is obtained only with the VRI.¹⁴

Following the literature, in this analysis, we take the stock share in their overall financial portfolio (α) and the share of wealth held in TDAs (ψ) as parameters.¹⁵ We focus on households' decisionmaking on the share of stocks in TDAs (α_{ret}), which also determines the share of stocks in TAs given α and ψ . Under H1, the optimal value of α_{ret} is determined as

$$\begin{aligned}\alpha_{ret}^* &= 0 && \text{if } \Psi \leq \alpha \\ &= \frac{\alpha + \Psi - 1}{\Psi} && \text{if } \Psi > \alpha.\end{aligned}\tag{1}$$

Under H2, it is determined as

$$\alpha_{ret}^* = \alpha.\tag{2}$$

¹⁴ We do not use the HRS in this exercise since the stock share measures for the IRAs and DC pensions in the HRS are too noisy to make the exercise practical.

¹⁵ In the analyses we do not include assets held in checking and savings accounts as they are held to provide liquidity in short-run as opposed to for investment purposes.

We focus on how the values of α_{ret} observed in data are compared to α_{ret}^* in each hypothesis. We use all the households whose heads are aged between 55 and 64¹⁶ with a nonnegligible share of their financial wealth both in TDAs and TAs (i.e., ψ in between 0.05 and 0.95). We have 1,314 such households in the VRI and 481 in the SCF.

Figures 3 and 4 show the comparison between the actual α_{ret} in the data and α_{ret}^* from the two hypotheses for the VRI and SCF, respectively. In each figure, Panel A is the comparison under H1 and Panel B is under H2. The horizontal axis shows α_{ret}^* while the vertical axis shows the actual α_{ret} . In each of these figures, if the observations are closer to the 45-degree line, then the hypothesis is closer to what households actually do in the data.¹⁷

For the VRI, it is clear that H2 is a better description of actual behavior than H1. Under H1 (Figure 3A), there are many households whose share of the TDA wealth (ψ) is less than the share of stocks in the overall financial portfolio (α), hence α_{ret}^* being zero. But many of these households have positive stock holdings in the TDAs, contrary to the predictions under H1. On the other hand, observations cluster along the 45-degree line under H2, so H2 is a useful heuristic for describing behavior (Figure 3B).

Figure 4 shows the same calculations as Figure 3 for the SCF data. The relative sparseness of the SCF data is evident in comparing Figures 3 and 4. Qualitatively, the SCF

¹⁶ We focus on the age range that corresponds to the peak of the wealth accumulation and that is lower than the normal retirement age (65).

¹⁷ Note that all the observations are above the 45-degree line under H1. Since what H1 implies is that the households should put as few stocks as possible in the TDAs, if they are already doing it then they will be on the 45-degree line, otherwise they will be above that line. The observations can go either direction from the 45-degree line under H2, depending on the relative stock share between the TDAs and TAs.

findings are similar to the VRI, but pointing much less strongly to H2 because of the sparsity of the SCF data and the dispersion around the 45-degree line in Figure 4B.

These scatter plots, however, do not tell us the economic significance of the discrepancies between the actual and hypothesized behaviors. Following Bergstresser and Poterba (2004), we consider the amount of wealth that would have to be relocated to match the hypothesized behavior,

$$Dev_{ret} = |\alpha_{ret}^* - \alpha| W_{ret}. \quad (3)$$

Figures 5 and 6 show the kernel density estimation of the distributions of Dev_{ret} under the two hypotheses from the VRI and SCF, respectively. Based on this measure, we again confirm that H2 is a much better description of the actual behavior of the VRI sample than H1. Under H2, the density is much higher for the small values while much lower for the large values of Dev_{ret} , compared to under H1. The pattern from the employer-sponsored VRI sample (Figure 5B) is almost identical to that from the entire VRI sample (Figure 5A), showing that the finding is not specific to those self-selected into Vanguard. With the SCF, however, it is hard to judge which hypothesis is closer to the actual behavior. The density under H2 is slightly higher around zero, but for most of the ranges, the two curves overlap. In particular, the SCF does not give us much information about under which hypothesis it is more likely to observe sizeable deviations. The Kolmogorov-Smirnov test on the null that the two distributions are equal in the range above \$100,000 cannot be rejected at the 10 percent significance level. Using the VRI, on the other hand, the same null can be rejected at the 1 percent level. One reason that the two curves are hard to distinguish is that the overall size of Dev_{ret} are much smaller in the SCF under both hypotheses, reflecting the smaller amounts of financial assets the SCF sample has compared to the VRI. Having a large number of observations with non-trivial amounts of financial assets in

both TDAs and TAs allows us to make a clear judgment on which hypothesis is a better description of the actual behavior.

6.3. Wealth and Retirement Horizon

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 (see Sevak, 2002; Coronado and Perozek, 2003; Kthitatrakun, 2004; Kezdi and Sevak, 2004; Goda, Shoven and Slavov, 2012; Hurd, Reti and Rohwedder, 2012 for findings from the HRS; Coile and Levine, 2004 for analyses using aggregate labor supply measures; and McFall, 2011 for results from the CogEcon Study).¹⁸

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 exercise, we estimate the non-linear relationship between wealth and retirement plans in the VRI, HRS, and SCF to demonstrate that only in the range of wealth where we have dense observations can we detect a significant wealth effect on retirement.

The standard approach to studying how wealth shocks affect labor supply is to observe behavior in panel data.¹⁹ The VRI is designed as a panel, though this paper analyzes the cross-

¹⁸ 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 the PSID. Estimated effects are mostly modest, with the exception of Holtz-Eakin, Joulfaian and Rosen (1993), who find a sizeable effect.

¹⁹ 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 as the administrative

sectional data from the first survey. As a preliminary and suggestive exercise, we use the relationship between retirement expectations and wealth in the cross-section. Thus, we build on the tradition of using expectations rather than realizations as the outcome variable (see McGarry, 2004; Chan and Stevens, 2004; and Szinovacz, Davey, and Martin, 2014). The use of subjective expectation 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).

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

$$W_i^R = (0.06 \times W_i \times (1.03)^{(65 - age_i)}) / Y_i \quad (4)$$

where W_i is annuitizable financial wealth, Y_i is current income, and age_i is the current age of the main earner of the household.²⁰ 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 6 percent 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 the current wealth of future retirees into common

account data have a monthly panel structure. We have done some exploratory work using the administrative data panel to 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.

²⁰ 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.

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

$$H_i = \beta_1(W_i^R)W_i^R + \beta_2Y_i^R + Z_i\gamma + \varepsilon_i \quad (5)$$

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).²¹ The coefficient $\beta_1(W_i^R)$ is a potentially non-linear function of normalized wealth.

We focus on estimates of this relationship in the VRI and HRS. In Appendix E, 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?"²² Both VRI and HRS have questions about current and expected pension and Social Security income. For 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.²³

For simplicity, we limit the sample to households with just one main earner who has not yet retired and is aged 65 or younger. For singles, anyone not retired and aged 65 or younger is

²¹ 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.

²² 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.

²³ 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.

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 earner, and we base the retirement decision on that household member.²⁴ There are 2,026 households in the VRI sample and 1,053 in the HRS sample. See Appendix E for details.

Figure 7 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 the two surveys documented in Section 5. Recall that normalized wealth is roughly the extent to which current wealth could replace current income at retirement if all assets were 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 nonnegligible 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

²⁴ To determine the primary earner, 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 earner and the household is included in the sample. Otherwise, the household is dropped from this analysis.

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 a nonnegligible amount of potentially annuitizable wealth.

We then examine 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.²⁵ Figure 8 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.²⁶ In Figure 8, "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 is measured in the expected remaining years of work (mean zero because it is a residual). In the VRI, for the entire sample in Figure 8A, we see a clear negative relationship between normalized wealth and retirement horizon up to the full replacement rate of around 1. Moving from zero 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

²⁵ LOESS is a bivariate procedure. To deal with 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.

²⁶ The ranges of the horizontal and vertical axes are truncated to exclude outliers. Appendix E, Figure A-E1, shows the data in Figure 8A for the entire sample including outliers.

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.

Section 5 showed that the characteristics of the VRI employer-sponsored sample are quite similar to subsets of the HRS and SCF with DC pension accounts. Figure 7B confirms that after imposing similar screens, the distribution of normalized wealth looks much more similar across the VRI and HRS. In Figure 8B, 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 shown in Panel B for the VRI is larger than in the HRS over the range 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.²⁷

As the analysis in this section makes 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. The bottom line is that the VRI has far more potential in exploring the effect of wealth on the retirement behavior of the population under an institutional and policy regime in which DC plans are the dominant source of retirement income.

7. Conclusions

This paper shows a way forward for collecting wealth data within surveys. We propose a comprehensive account-by-account approach that is designed to elicit accurate information in the form that respondents think about wealth and have at their disposal. The account-based approach to survey measurement of wealth yields measurements that are precise and unbiased relative to administrative measurements. In contrast, many surveys appear to undercount assets. The paper also demonstrates that the correction mechanism that is integral to our account-by-account survey design significantly reduces errors, as measured relative to the administrative account data.

The design of the VRI infrastructure is targeted at measuring the wealth of households with sufficient financial assets so they face wealth allocation and accumulation decisions concerning whether to work longer, how to manage a financial portfolio, whether to annuitize, whether to buy long-term care insurance, how much to bequeath, and so on. In this paper, we

²⁷ In Appendix E, 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 less different 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.

show that the VRI is dense in data on older Americans in the upper half of the wealth distribution compared to other excellent surveys with wealth data, namely the HRS and SCF. We show that for issues including financial portfolio management and the relationship between financial wealth and retirement expectations, the VRI indeed provides a very different picture from the HRS and SCF precisely because it has high-quality asset data from sufficient observations of households who are approaching their retirement with substantial financial wealth.

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)** "Vanguard Research Initiative: Survey 1 Documentation and Tabulations," Vanguard Research Initiative Working Paper, <https://ebp-projects.isr.umich.edu/VRI/papers.html>.
- Barber, Brad, and Terrance Odean (2003)** "Are Individual Investors Tax Savvy? Evidence from Retail and Discount Brokerage Accounts," *Journal of Public Economics* 88, 419-442.
- Barsky, Robert, John Bound, Kerwin Kofi Charles, and Joseph P. Lupton (2002)** "Accounting for the Black-White Wealth Gap: A Nonparametric Approach," *Journal of the American Statistical Association* 97, 663-673.
- Bergstresser, Daniel and James Poterba (2004)** "Asset Allocation and Asset Location: Household Evidence from the Survey of Consumer Finances," *Journal of Public Economics* 88, 1893-1915.
- Bricker, Jesse, Alice Henriques, Jacob Krimmel, and John Sabelhaus (2016a)** "Estimating Top Income and Wealth Shares: Sensitivity to Data and Methods," *American Economic Review: Papers and Proceedings* 106, 641-645.
- Bricker, Jesse, Alice Henriques, Jacob Krimmel, and John Sabelhaus (2016b)** "Measuring Income and Wealth at the Top Using Administrative and Survey Data," *Brookings Papers on Economic Activity*, 261-312.

- 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," *Annual Review of Economics* 6, 475-501.
- Bucks, Brian and Karen Pence (2015)** "Wealth, Pensions, Debt, and Savings: Considerations for a Panel Study," *Journal of Economic and Social Measurement* 40, 151-175.
- 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.
- Charles, Kerwin Kofi and Jonathan Guryan (2011)** "Studying Discrimination: Fundamental Challenges and Recent Progress," *Annual Review of Economics*, 3, 479-511.
- 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.
- Dammon, Robert M., Chester S. Spatt and Harold D. Zhang (2004)** "Optimal Asset Location and Allocation with Taxable and Tax-Deferred Investing," *Journal of Finance* 59, 999-1038.
- 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.
- Eggleston, Jonathan S. and Michael Gideon (2017)** "Evaluating Wealth Data in the Redesigned 2014 Survey of Income and Program Participation," Social, Economic, and Housing Statistics Division Working Paper, 2017-35.
- Fries, Gerhard, Martha Starr-McCluer, and Annika E. Sundén (1998)** "The Measurement of Household Wealth using Survey Data: An Overview of the Survey of Consumer Finances," *Consumer Interests Annual* 44, 224-231.
- Garlappi, Lorenzo and Jennifer C. Huang (2006)** "Are Stocks Desirable in Tax-Deferred Accounts?," *Journal of Public Economics* 90, 2257-2283.
- 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.
- Henriques, Alice M. and Joanne W. Hsu (2014)** "Analysis of Wealth Using Micro- and Macrodata: A Comparison of the Survey of Consumer Finances and Flow of Funds Accounts," in Dale W. Jorgenson, J. Steven Landefeld, and Paul Schreyer eds., *Measuring Economic Sustainability and Progress*, National Bureau of Economic Research, Inc., 245-276.

- 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.
- 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.
- Kaplan, Greg and Giovanni L. Violante (2014)** "A Model of the Consumption Response to Fiscal Stimulus Payments," *Econometrica* 82, 1199-1239.
- 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 (2007)** "The Role of Over-Sampling of the Wealthy in the Survey of Consumer Finances," Federal Reserve Board Survey of Consumer Finances Working Papers.
- Kennickell, Arthur (2015)** "Dirty and Unknown: Statistical Editing and Imputation in the SCF," *Statistical Journal of the IAOS*, 31, 435-445.
- Kezdi, Gabor and Purvi Sevak (2004)** "Economic Adjustment of Recent Retirees to Adverse Wealth Shocks," Michigan Retirement Research Center Working Paper 2004-075.
- 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.
- Moffitt, Robert, Robert F. Schoeni, Charles Brown, P. Lindsay Chase-Lansdale, Mick P. Couper, Ana V. Diez-Roux, Eirk Hurst, and Judith A. Seltzer (2015)** "Assessing the Need for a New Nationally Representative Household Panel Survey in the United States" *Journal of Economics and Social Measurement* 40, 1-26.
- Pfeffer, Fabian T., Robert F. Schoeni, Arthur Kennickell, and Patricia Andreski (2016)** "Measuring Wealth and Wealth Inequality: Comparing Two U.S. Surveys," *Journal of Economic and Social Measurement* 41, 103-120.
- Poterba, James M. (2014)** "Retirement Security in an Aging Population," *American Economic Review* 104, 1-30.

- Poterba, James M., John Shoven, and Clemens Sialm (2001)** "Asset Location for Retirement Savers," in William Gale, John Shoven, and Mark Warshawsky eds., *Private Pensions and Public Policies*, Brookings Institution Press, 290-331.
- Poterba, James M., Steven Venti, and David Wise (2011)** "The Composition and Drawdown of Wealth in Retirement," *Journal of Economic Perspective* 25, 95 – 118.
- Samphantharak, Krislert, Scott Schuh, and Robert M. Townsend (2018)** "Integrated Household Surveys: An Assessment of U.S. Methods and an Innovation," *Economic Inquiry* 56, 50-80.
- Sevak, Purvi (2002)** "Wealth Shocks and Retirement Timing," Michigan Retirement Research Center Working Paper 2002-027.
- Shoven, John (1999)** "The Location and Allocation of Assets in Pension and Conventional Savings Accounts," National Bureau of Economic Research Working Paper, 7007.
- Shoven, John and Clemens Sialm (2003)** "Asset Location in Tax-Deferred and Conventional Savings Accounts," *Journal of Public Economics* 88, 23-38.
- Sierminska, Eva, Pierre-Carl Michaud, and Susann Rohwedder (2008)** "Measuring Wealth Holdings of Older Households in the US: A Comparison using the HRS, PSID, and SCF," mimeo.
- 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 ≥ 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 the 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 the 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 an account-based approach for retirement accounts in 2012.

Table 2. Survey Financial Assets: All respondents

Account type	Conditional on having a positive amount							
	Mean	N	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	Percent Difference		
			25 percentile	median	75 percentile
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		Financial Wealth							All
		<\$0	\$0-10K	\$10K-100K	\$100K-500K	\$500K-1M	\$1M-2.5M	>\$2.5M	
55-64	VRI	48	36	292	1,147	871	762	181	3,337
	HRS	1,459	586	933	897	287	160	41	4,363
	SCF	228	170	196	254	102	119	212	1,281
65-74	VRI	16	19	258	1,117	985	1,066	377	3,838
	HRS	746	487	727	817	290	162	35	3,264
	SCF	93	114	118	155	68	91	178	817
> 74	VRI	2	4	95	549	461	472	192	1,775
	HRS	800	712	1,030	927	284	172	43	3,968
	SCF	60	93	115	107	31	30	90	526
Total	VRI	66	59	645	2,813	2,317	2,300	750	8,950
	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

Age	VRI			HRS			SCF		
	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 the 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 a \$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

			Percentiles					
			Mean	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)

		Percentiles					N
	Sample Screen	10	25	50	75	90	
VRI	All	14.96	35.12	54.76	74.71	91.14	8905
	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
HRS	Age eligible	0	0	0	40.32	81.48	11595
	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
SCF	Age eligible	0	0	0.70	43.39	71.24	2624
	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)

		Percentiles					N
	Sample Screen	10	25	50	75	90	
VRI	All	30,000	113,800	326,162	712,200	1,397,710	8905
	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
HRS	Age eligible	0	0	0	45,000	270,000	11595
	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
SCF	Age eligible	0	0	0	78,000	360,000	2624
	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

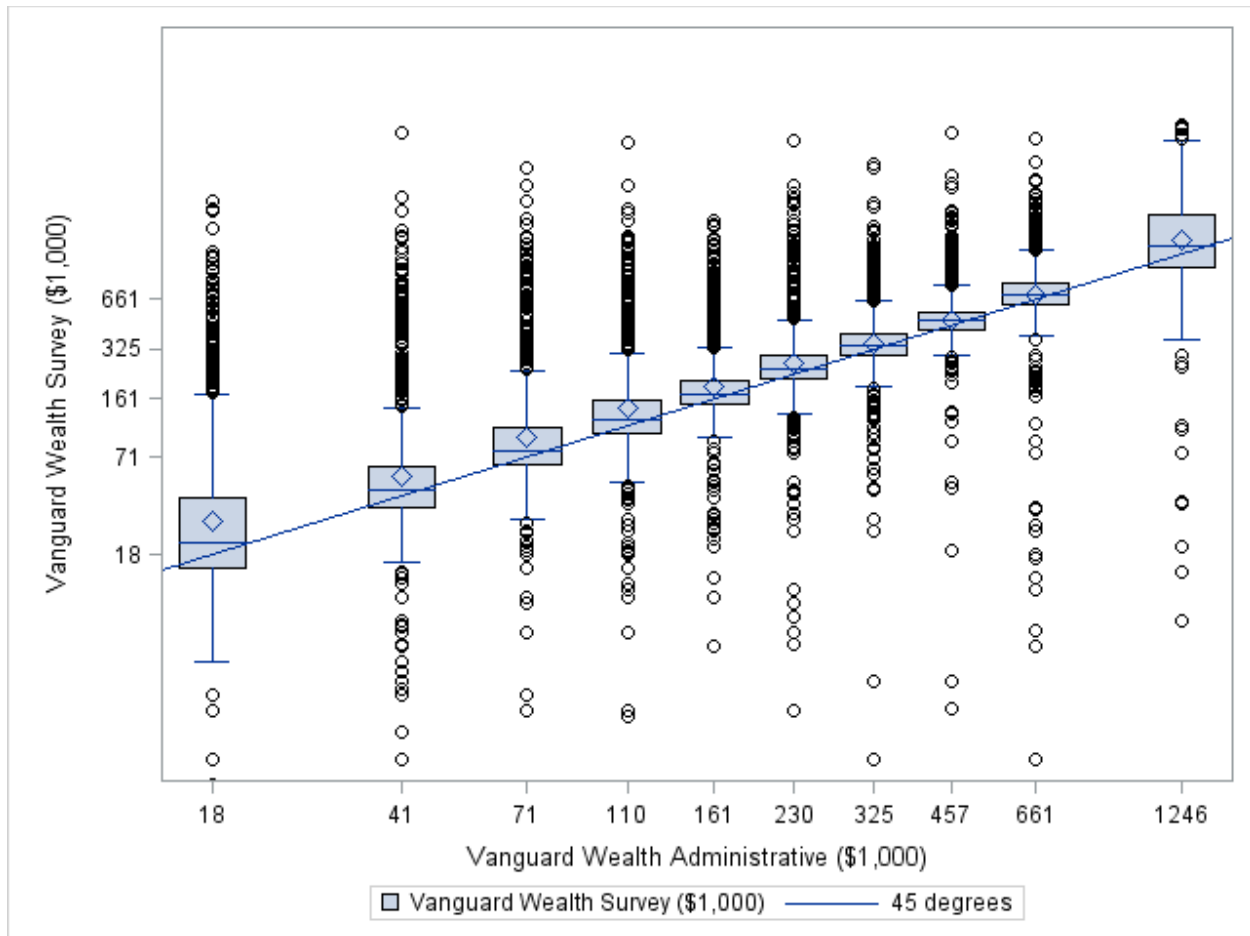
Table 10. Share of Wealthy-Hand-to-Mouth Households by Age Group and Retirement Status

	VRI				SCF				SCF (VRI-eligible)			
	WHtM share	Liquid wealth (median)	Monthly Income (median)	N	WHtM share	Liquid wealth (median)	Monthly Income (median)	N	WHtM share	Liquid wealth (median)	Monthly Income (median)	N
A. Age group												
55-59	0.30	34,963	8,333	1,550	0.32	2,350	4,396	578	0.27	20,820	7,524	345
60-64	0.22	68,250	7,383	1,790	0.30	2,750	4,058	561	0.29	17,700	6,594	318
65-69	0.13	114,338	5,838	1,933	0.22	9,400	3,466	488	0.08	55,070	6,256	244
70-74	0.08	154,152	4,615	1,908	0.26	5,800	2,824	352	0.10	71,000	6,002	158
75+	0.04	238,351	3,942	1,776	0.19	8,100	2,198	525	0.07	79,402	4,396	114
B. By retirement status												
Not completely retired	0.24	58,200	8,150	4,271	0.30	7,000	4,929	1,311	0.23	28,000	7,524	842
Completely retired	0.07	180,000	4,250	4,686	0.22	4,400	2,325	1,153	0.13	50,000	4,734	337

Notes: Households are classified as wealthy-hand-to-mouth if their liquid wealth is less than half of the monthly income. Retirement status is based on self-report.

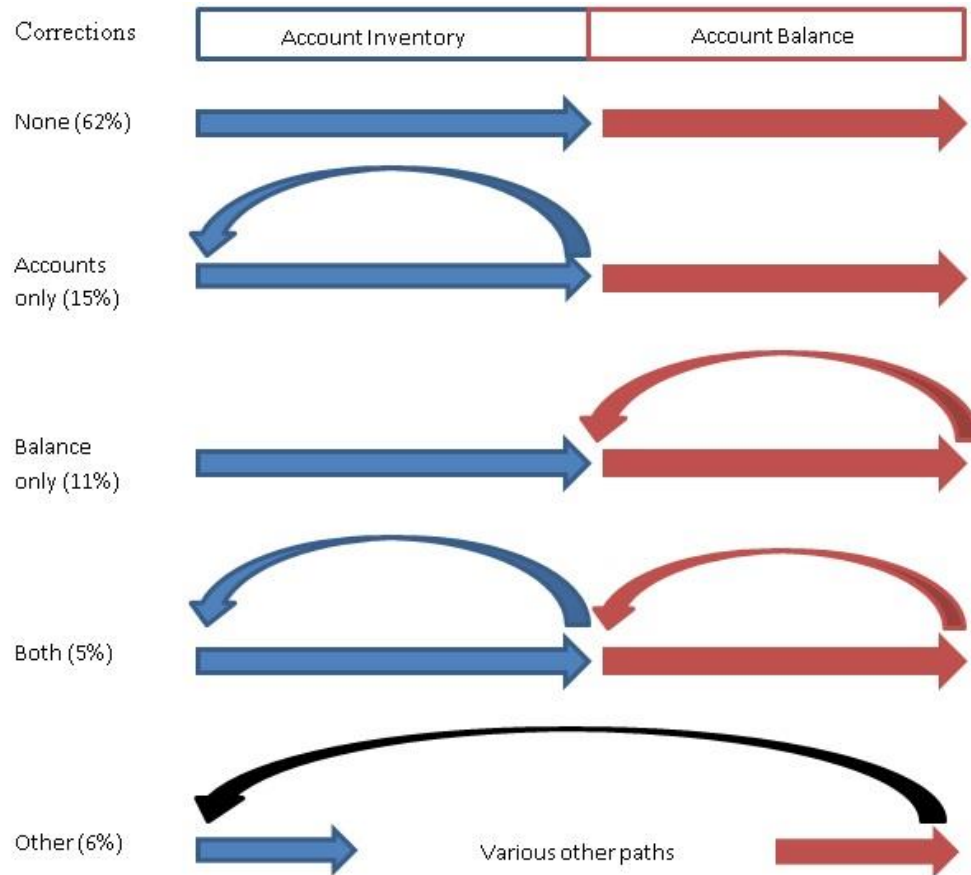
For SCF, we include households with respondents at least 55 years old from the 2013 wave. SCF results use sampling weights.

Figure 1. Administrative versus Survey Financial Assets at Vanguard



Note: The figure compares Vanguard administrative assets with the survey report of Vanguard assets. See the text for how Vanguard assets are determined in the 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.

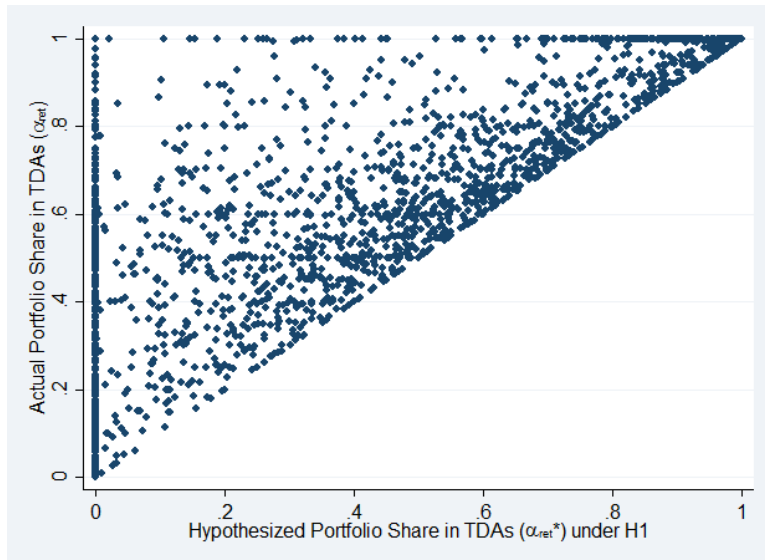
Figure 2. Correction Paths through Wealth Section.



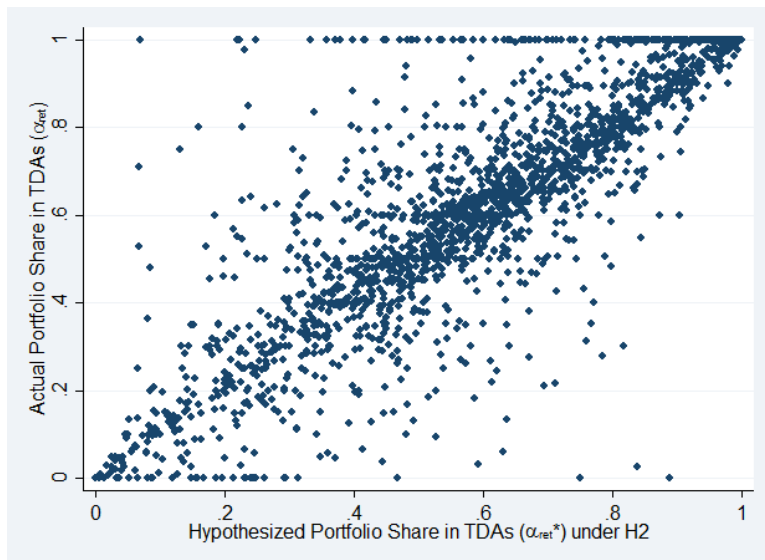
Note: The figure shows the fraction of respondents taking various paths through the account-based wealth section. Other includes those who started over and then took various paths to complete.

Figure 3. Stock shares in TDAs: comparison between actual behavior and hypotheses (VRI)

A. Under H1: Tax-Efficient Allocation



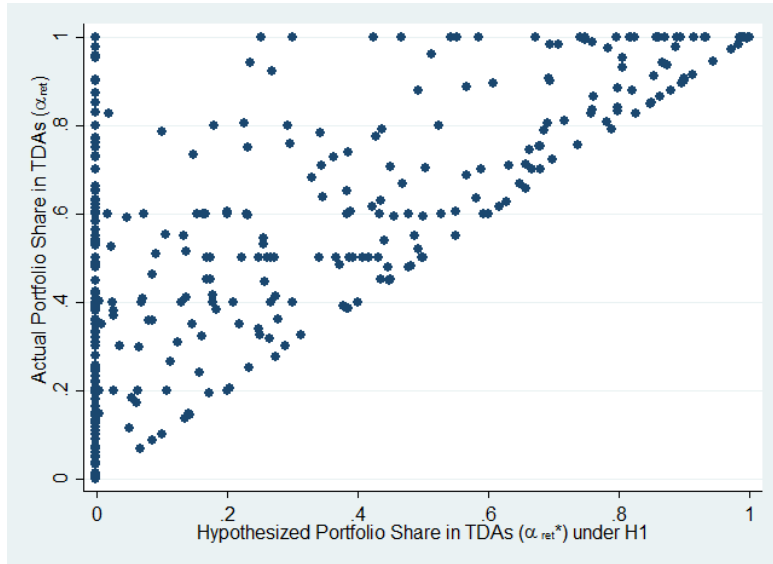
B. Under H2: Equal Allocation



Note: The horizontal axis shows the stock share in tax-deferred accounts (TDAs) implied by each hypothesis while the vertical axis shows the actual stock share in TDAs that the households have. $N = 1,314$.

Figure 4. Stock shares in TDAs: comparison between actual behavior and hypotheses (SCF)

A. Under H1: Tax-Efficient Allocation



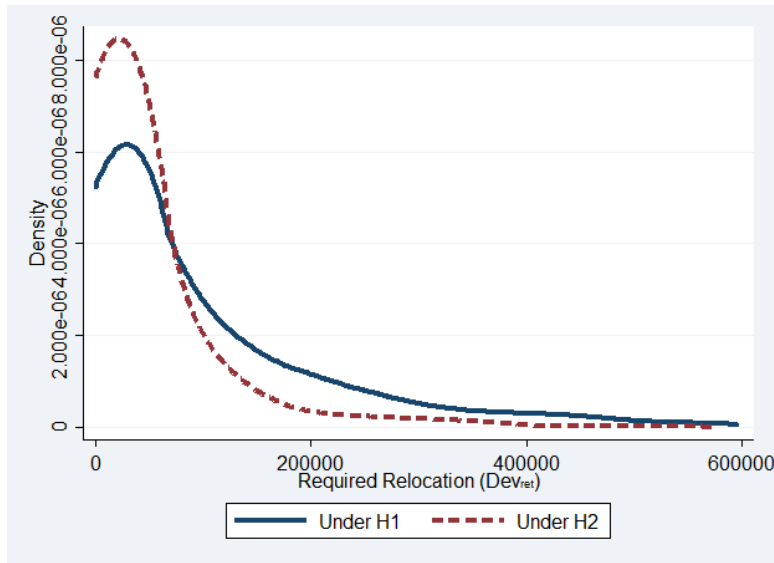
B. Under H2: Equal Allocation



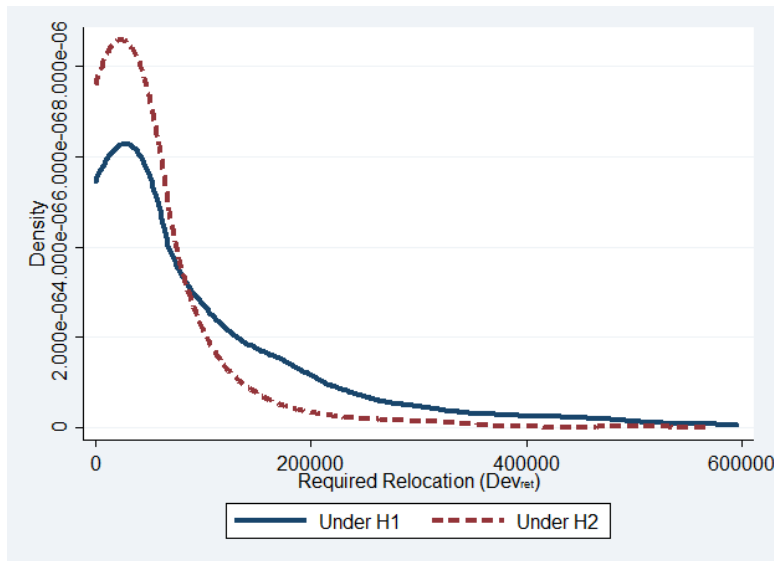
Note: The horizontal axis shows the stock share in the tax-deferred accounts (TDAs) implied by each hypothesis while the vertical axis shows the actual stock share in the TDAs that the households have. $N = 481$.

Figure 5. Size of Relocation Required to Have Asset Location Implied by Two Hypotheses (VRI)

A. Entire sample

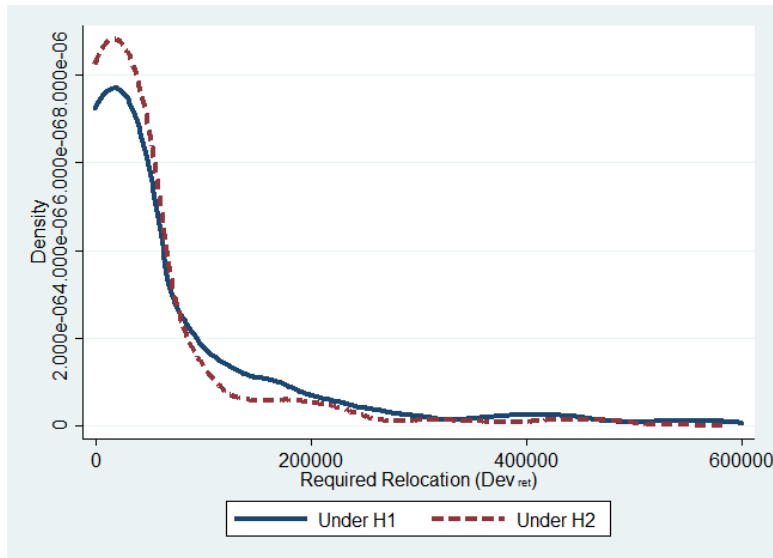


B. Employer-sponsored sample



Note: The curves show the kernel density estimations of the amount of relocation required from the actual asset location patterns to make them consistent with the patterns suggested by the two hypotheses. The blue curve corresponds to H1 and the red curve corresponds to H2.

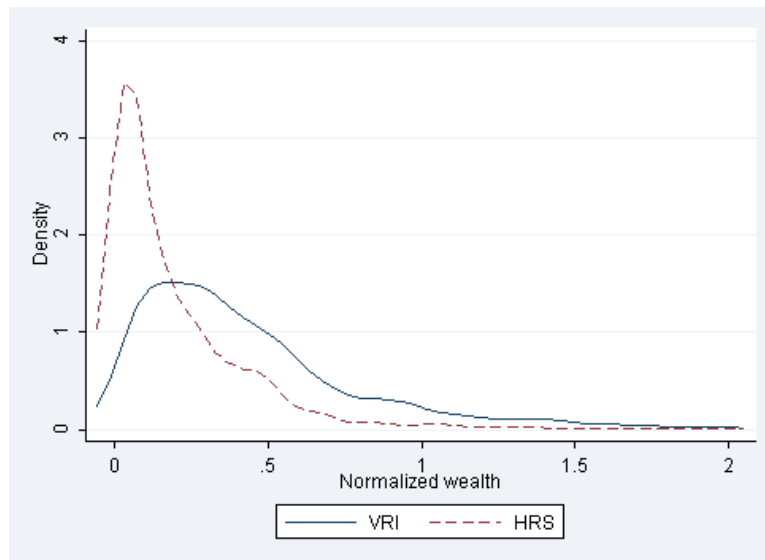
Figure 6. Size of Relocation Required to Have Asset Location Implied by Two Hypotheses (SCF)



Note: The curves show the kernel density estimations of the amount of relocation required from the actual asset location patterns to make them consistent with the patterns suggested by the two hypotheses. The blue curve corresponds to H1 and the red curve corresponds to H2.

Figure 7. Distribution of normalized financial wealth (kernel estimation)

A. VRI vs HRS



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

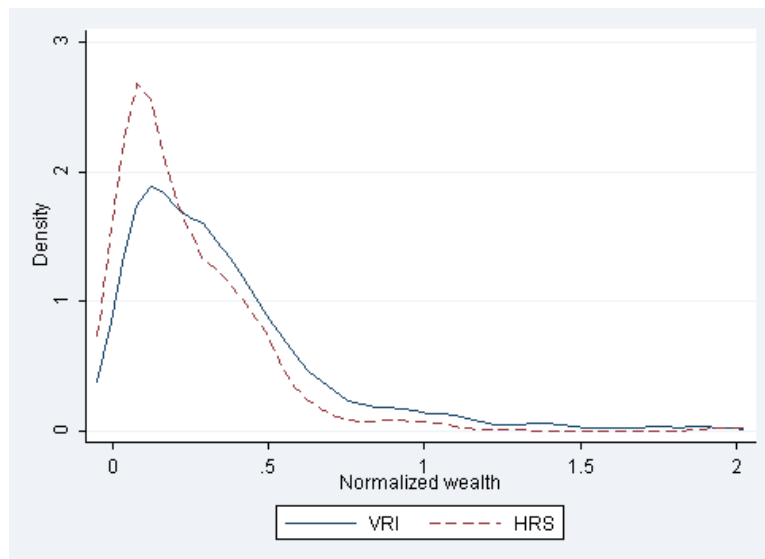
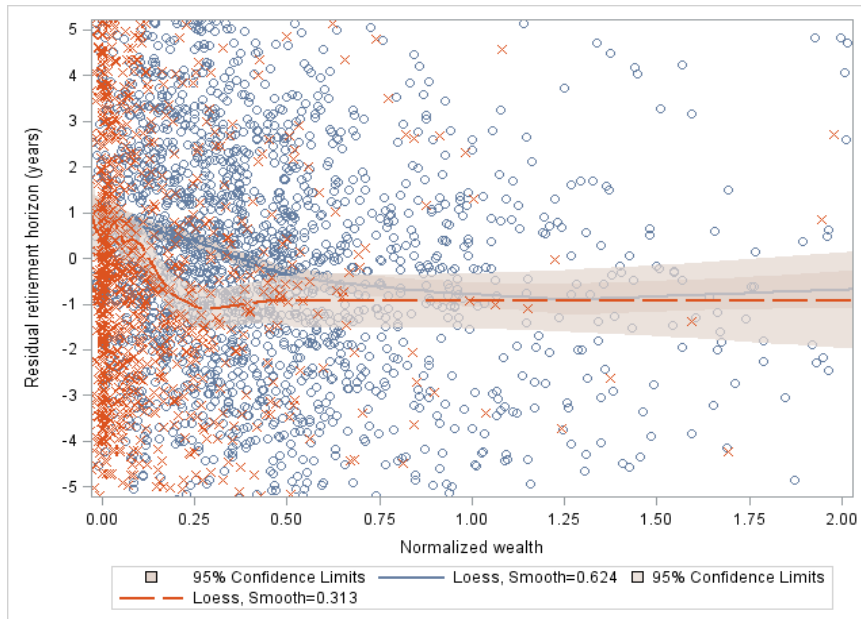


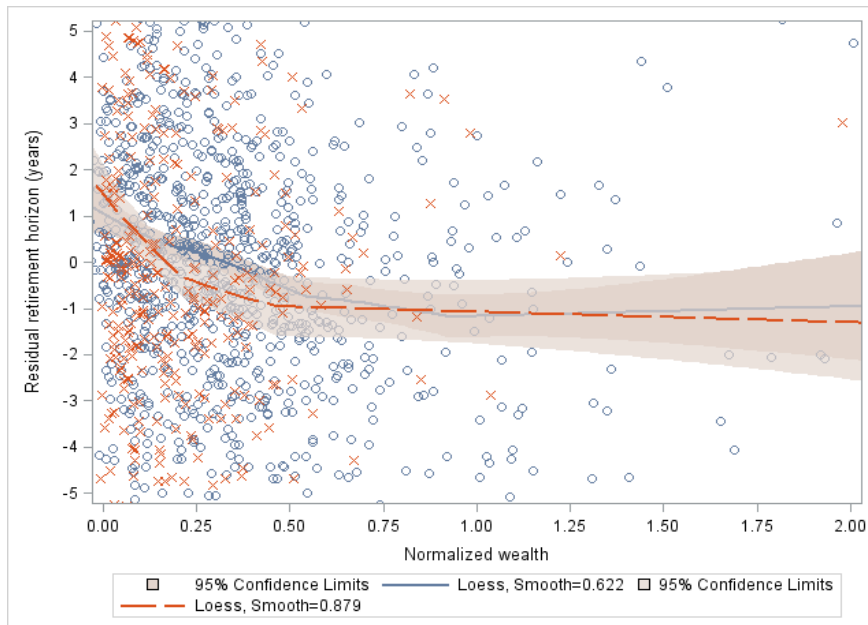
Figure 8. Retirement horizon versus normalized financial wealth: LOESS

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).

The Wealth of Wealthholders

September 2023

Appendix A. The VRI Sample Design

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 IRA accounts (either at Vanguard or elsewhere). For this age group, we sample the types in the proportion they appear in the population.

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).

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.

The Wealth of Wealthholders

September 2023

Appendix B. Account Sequence Example

Section 3.1 of the main text explained the structure of the wealth section of the survey in detail. In this appendix, we show actual screenshots 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-B1). She answers how many accounts she has for each type using a drop-down menu (Figure A-B2) and then gives each of the accounts a nickname (Figure A-B3). The survey shows the summary of responses so far (Figure A-B4) 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-B5 is one example). After the loop, the survey displays a summary table of account balances as well as a total (Figure A-B6). 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-B7). In this example, the respondent chooses to correct balances, indicates that she wants to correct the balance for “Roth IRA” (Figure A-B8), and then corrects the balance for that account (Figure A-B9). 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-B10).

The survey then asks two follow-up questions for each account: stock share (Figure A-B11) and whether that account is held at Vanguard (Figure A-B12). 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-B13). The respondent can print this summary page as a record.

Figure A-B1. 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-B2. 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-B3. 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.

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

Figure A-B4. 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-B5. 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-B6. 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	<input type="radio"/>	<input type="radio"/>
IRA 2: Roth IRA	No response provided		
Employer-sponsored retirement plan (401(k), 403(b), 457, etc.) 1: Retirement	\$400,000	<input type="radio"/>	<input type="radio"/>
Checking account 1: Chase	\$15,000	<input type="radio"/>	<input type="radio"/>
Mutual fund account (other than money market) 1: Vanguard	\$275,000	<input type="radio"/>	<input type="radio"/>
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-B7. 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-B8. 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	<input type="checkbox"/>
IRA 2: Roth IRA	No response provided	<input checked="" type="checkbox"/>
Employer-sponsored retirement plan (401(k), 403(b), 457, etc.) 1: Retirement	\$400,000	<input type="checkbox"/>
Checking account 1: Chase	\$15,000	<input type="checkbox"/>
Mutual fund account (other than money market) 1: Vanguard	\$275,000	<input type="checkbox"/>

Figure A-B9. 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-B10. 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	<input checked="" type="radio"/>	<input type="radio"/>
IRA 2: Roth IRA	\$150,000	<input checked="" type="radio"/>	<input type="radio"/>
Employer-sponsored retirement plan (401(k), 403(b), 457, etc.) 1: Retirement	\$400,000	<input checked="" type="radio"/>	<input type="radio"/>
Checking account 1: Chase	\$15,000	<input type="radio"/>	<input checked="" type="radio"/>
Mutual fund account (other than money market) 1: Vanguard	\$275,000	<input checked="" type="radio"/>	<input type="radio"/>
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-B11. 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-B12. Which Accounts at Vanguard

Which accounts are currently held at Vanguard?

Please select one for each row.

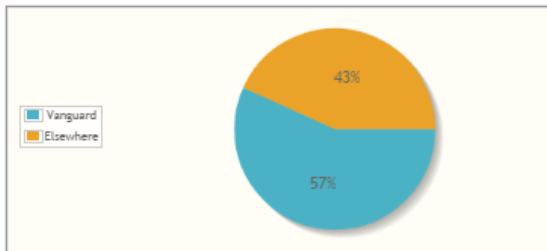
Accounts	Held at Vanguard	
	Yes	No
IRA 1: Rollover IRA	<input checked="" type="radio"/>	<input type="radio"/>
IRA 2: Roth IRA	<input checked="" type="radio"/>	<input type="radio"/>
Employer-sponsored retirement plan (401(k), 403(b), 457, etc.) 1: Retirement	<input type="radio"/>	<input checked="" type="radio"/>
Mutual fund account (other than money market) 1: Vanguard	<input checked="" type="radio"/>	<input type="radio"/>

Figure A-B13. Summary Table and Charts

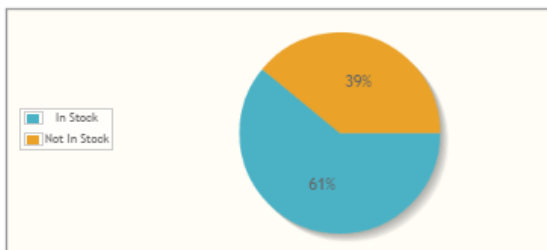
Thank you very much for providing this detailed information about your financial assets. Before continuing with the final sections of the survey, we would like to present you with a summary of your assets. You may wish to print this for your own use:

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

September 2023

Appendix C. Definition of concepts

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

Total financial wealth. 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 the 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.)

Web-survey eligibility. 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 has 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

September 2023

Appendix D. 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-D1 compares the 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 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-D2 and A-D3 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-D4 and A-D5. 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-D6 compares education, health, and marital status across samples. Tables A-D7, A-D8, and A-D9 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 it 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-D10 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-D1. 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-D2. Effect of Imposing VRI Sampling Screens: Income distribution

		Mean	Percentiles				
			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-D3. 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-D4. 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-D5. 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-D6. 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-D7. 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-D8. Effect of Imposing VRI Sampling Screens: Fraction with Very Good or Excellent Health 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
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-D9. Effect of Imposing VRI Sampling Screens: Fraction Married or Partnered 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
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-D10. Effect of Imposing VRI Sampling Screens: Retirement Rate 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	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

September 2023

Appendix E. Estimating Retirement/Wealth Relationship

HRS sample. Table A-E1 shows how many observations we lose in the HRS by imposing each additional condition on the 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 earner satisfies the age condition, some are retired while some have dual main earners. 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-E1, 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 6, 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-E2 shows the distribution of newly defined normalized wealth and Figure A-E3 shows the new LOESS estimates. For both figures, Panel A is for the entire sample used in Section 6. Panel B is for the employer-sponsored subsets.

Figure A-E2A shows that the VRI sample still has higher replacement rates, though the gap is less stark than in Figure 7A. The VRI has many observations in the range between 1 and

¹ Some earners 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.

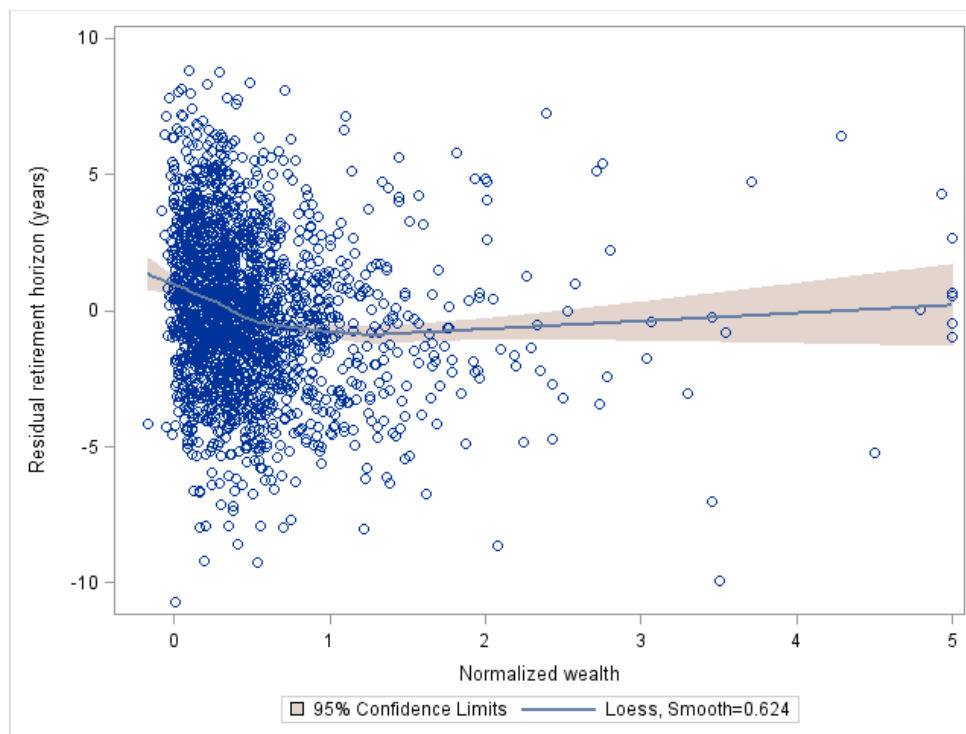
2, while for the HRS, most of the observations have normalized wealth smaller than 1. The LOESS estimate (Figure A-E3A) shows basically the same relationship as the baseline model (Figure 8A). With the VRI sample, we can estimate a negative and statistically significant relationship for a wider range (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-E2B). Figure A-E3B 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-E1. HRS Sample Size for Retirement Horizon Analysis: Effect of Each Condition

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

Figure A-E1. Retirement horizon versus normalized financial wealth: LOESS (full range of data)

A. VRI



B. HRS

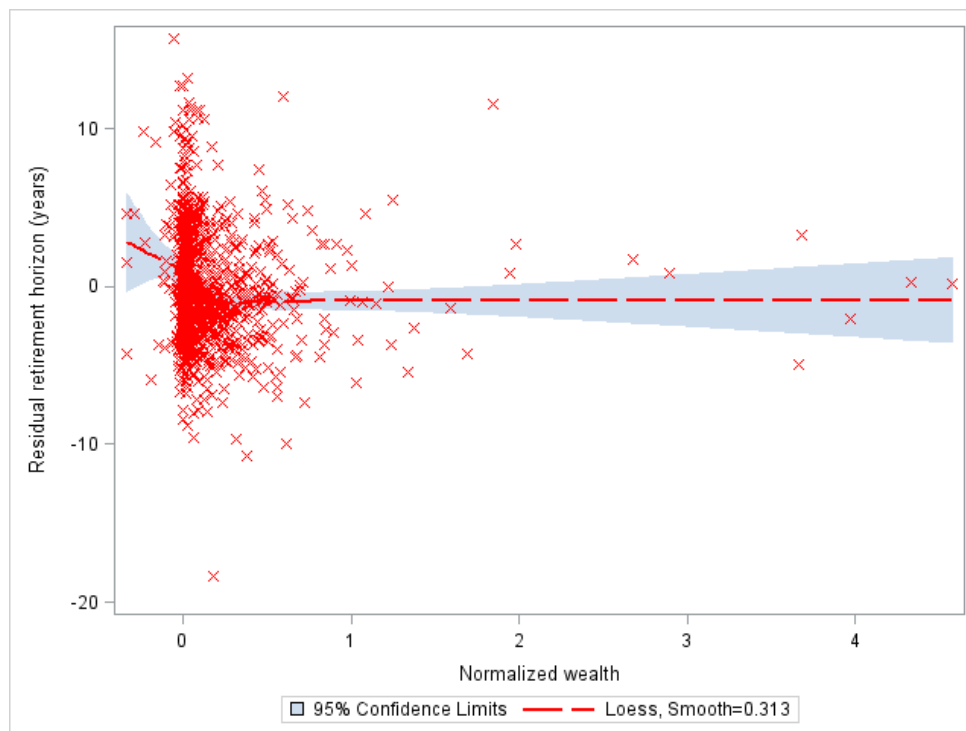
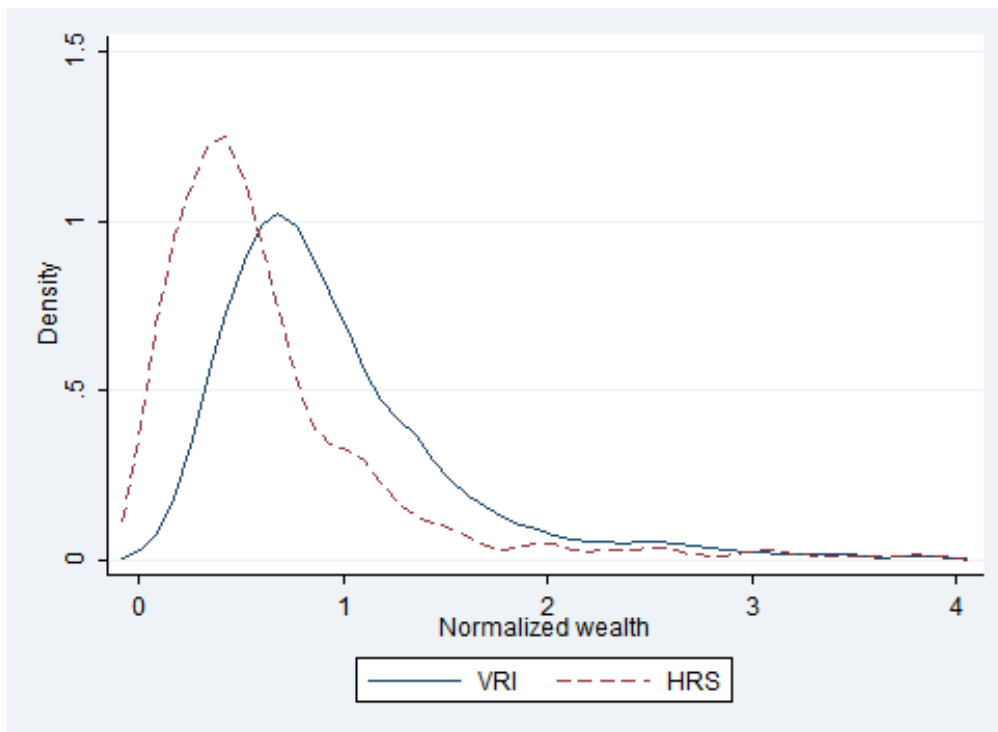


Figure A-E2. Distribution of normalized financial wealth (including future DB pension and SS income)

A. VRI vs HRS



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

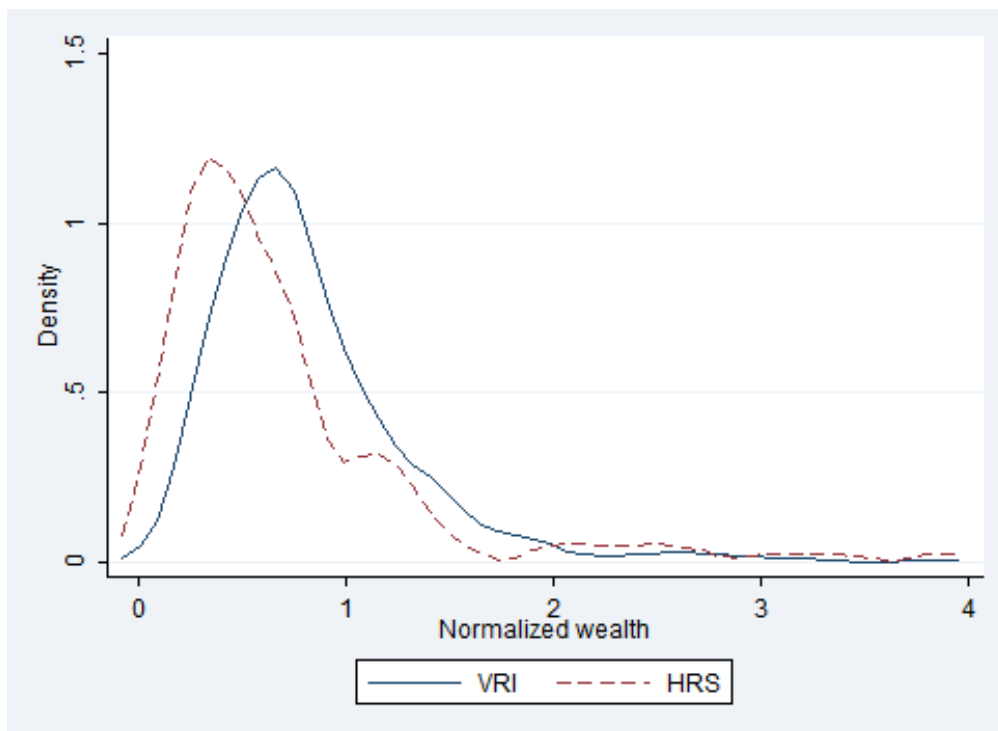
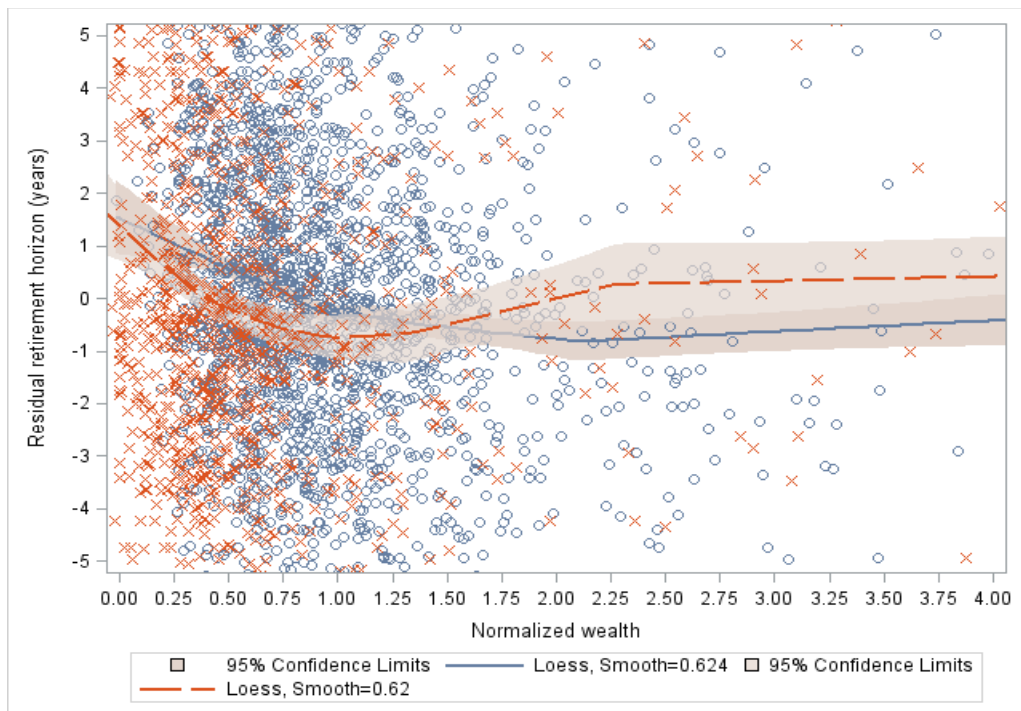


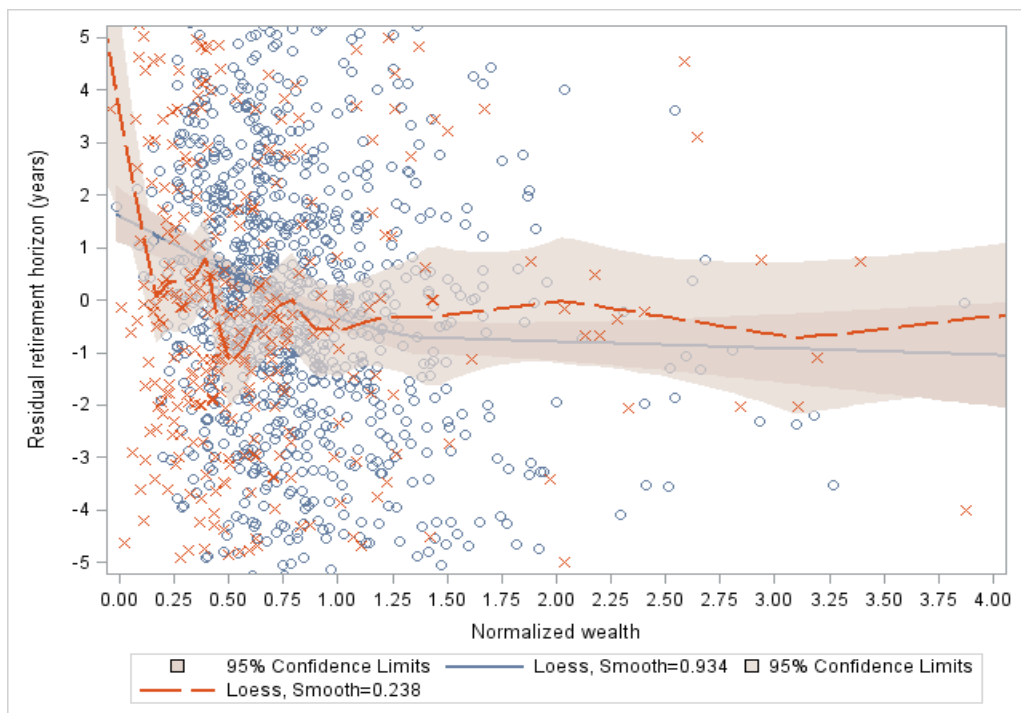
Figure A-E3. Retirement horizon versus normalized financial wealth: LOESS
(Normalized wealth including future DB pension and SS income)

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).