Who Saves?

The Role of Long-run Earnings and Life Events on Retirement Wealth

Deborah Foster

Michigan State University
Department of Economics
101 Marshall-Adams Hall
East Lansing, MI 48824-1038
dfoster@msu.edu

November 2007

Abstract

In recent years numerous studies have examined individual retirement saving behavior. An important paper by Venti and Wise (2001) investigates how much of the observed variation in retirement wealth is attributable to differences in household lifetime earnings and life events. Their analysis suggests that these factors have little effect, but their data suffer from several important shortcomings. We reexamine the role of long-run earnings and life events using the Panel Study of Income Dynamics (PSID) and find that these factors account for over five times more of the variation in retirement wealth than reported by Venti and Wise. Overall, almost half of the observed variation in retirement wealth is explained by long-run earnings and life events.

* I would like to thank Steven Haider, Paul Menchik, and Stacy Dickert-Conlin for their guidance and advice. I also greatly appreciate the comments and suggestions from seminar participants at Michigan State University and the University of Michigan PSID/CDS.

1. Introduction

The looming retirement of the baby boomers, the debate over the future of Social Security, and high-profile pension bankruptcies and reform have all heightened interest in personal retirement savings in America. One of the important questions concerns who is and isn't saving, and an extensive literature has emerged in which researchers investigate the correlates of wealth. For example, wealth differentials have been identified by race, gender, and family structure.

Venti and Wise (2001) addresses the issue of who saves by focusing on the ability to save. They use the Health and Retirement Survey (HRS) to determine how much of the observed variation in retirement wealth can be attributed to factors that affect the amount of money that can be allocated to savings. Specifically, they examine the role of household lifetime earnings and life events (including inheritances, children, marital status, and health) on wealth at retirement. Their analysis suggests that these factors have little effect, leading them to conclude that "the bulk of the dispersion must be attributed to differences in the amount that households choose to save" (57). In short, they attribute differences in wealth accumulation to different saving preferences, not to differences in the ability to save.

Venti and Wise is one of the first papers to closely study the role of life events, but their data suffer from several cruicial shortcomings. First, their measure of lifetime earnings comes from the Social Security Administration (SSA) earnings reports that are top-coded. Venti and Wise report that over 25% of the sample exceeded the earnings cap in 1971, but only 4.8% of their sample was affected in 1991 (30). Even though the portion of the sample affected by the earnings cap declines, it is still an important problem. High-earners typically have more wealth, and those are the exact individuals whose lifetime income measure is affected by the censoring.

Additionally, the demographic information used to measure life events is based on the HRS baseline survey completed in 1992, when the primary respondents are between the ages of 51 and 61. Although some retrospective information is included, little is known about the participants' lives prior to the survey.

The goal of this paper is to reexamine the role of long-run earnings and life events on retirement savings using the Panel Study of Income Dynamics (PSID). The survey includes uncensored income information that allows us to examine the income-wealth relationship more precisely. The panel structure enables us to observe households over an extended period of time and thus examine both the occurrence and timing of life events. Our findings indicate that the effects of income and life events are over five times greater than previously estimated. Although such findings are suggestive, this analysis suffers from the same difficulties in interpretation as Venti and Wise's analysis – the descriptive analysis cannot untangle the extent to which the life events have a causal effect on wealth accumulation or instead reflect underlying preference differences.

The remainder of the paper is organized as follows. Section 2 includes a brief review of the relevant literature. In Section 3 we discuss the data used in the analysis (with further detail available in the Data Appendix). Results are found in Section 4, and we summarize and discuss our findings in Section 5.

2. Literature Review

Wealth inequality in the United States is well-documented. Indeed, wealth is much more unevenly distributed than income.[1] To gain insight into why these disparities exist, researchers

[1] For example, see Hurst, Luoh, and Stafford (1998), Wolff (1998), and Cagetti and De Nardi (2005).

have examined the correlates of wealth. Such studies are typically cross-sectional and highlight differences in wealth holdings across different demographic characteristics.

For example, Blau and Graham (1990) report that young black families only have about 18% of the wealth held by young white families. Possible explanations include differences in income, inheritances, and asset allocation.[2] Gender differences are also evident. Studies such as Schmidt and Sevak (2006) document that women have less wealth than men. The gender wealth gap is often attributed to differences in income and the fact that women are more risk-averse and therefore invest differently than their male counterparts.[3]

Lupton and Smith (2003) examine wealth differences by marital status. They find that married couples not only have more per capita wealth than divorced individuals, but the duration of marital status matters, too; those married longer have more wealth, and those divorced longer have less. Along similar lines, Wilmoth and Koso (2002) find that wealth is affected not only by marital status, but by the sequence of marital events over a person's lifetime.

Researchers have also examined the effect of children on wealth. Schmidt and Sevak (2006) find that younger children do not affect wealth levels, but households with children between the ages of 18 and 24 have less wealth. They argue that these findings are consistent with parents saving for their children's education. On the other hand, Lupton and Smith (2003) do not find a significant relationship between children and wealth regardless of children's age.

The focus of our analysis is a study by Venti and Wise (2001) that includes many of the factors discussed above to determine how much of the overall variation in retirement wealth can be explained by lifetime earnings, a variety of life events, and investment choice. They find that, even for households with similar lifetime earnings, substantial variation in retirement wealth

---

[2] See Menchik and Jianakopolos (1997), Altonji, Doraszelski, and Segal (2000), and Gittleman and Wolff (2000) for further detail.
[3] Refer to Bajtelsmit and Bernasek (1996) for a survey of this evidence.

exists. Next, they control for age, marital status, children, inheritances, and health. The results show that these factors only explain about 4% of the observed variation. The effect of investment choice (about 8%) is also very small. From this, they conclude that the primary cause of wealth dispersion is differences in preferences – some choose to save while young and accumulate extensive wealth while others do not.

Hurst (2006) also considers a wide range of demographic characteristics and life events in his analysis. In particular, he includes income, unemployment, health, marital status, family size, race, education, occupation, and region in a regression to predict retirement wealth for households in the PSID. He finds that wealth accumulation varies, even for households that are similar across all of those dimensions.

3. Data

A brief description of the data and variables used in the analysis are provided below. Please see the Data Appendix for further detail.

3.1 Our Sample

The data are from the PSID. The longitudinal survey, conducted by the Survey Research Center at the University of Michigan Institute for Social Research, was administered annually from 1968 to 1997. Beginning in 1999, it switched to a biennial schedule. The most recent data available are from 2005. In all, there are 34 waves of data spanning a period of 38 years.

The sample is composed of males (and their households) that meet two selection criteria. The first requirement is that we have a measure of retirement wealth. Specifically, the household must complete a wealth supplement when the male is 55 (plus or minus 2) years old. Age fiftyfive is younger than the standard retirement age, but the earlier wealth measure mitigates problems arising from early retirees who have already begun to deplete their accumulated wealth while still allowing us to observe the majority of wealth accumulation.[4] The second criterion is that the household is observed twenty years prior to the measure of retirement wealth. This enables us to calculate a measure of long-run earnings. The final sample consists of 944 households.[5]

By construction, sample members are compiled from the 1989, 1994, 1999, 2003, and 2005 wealth supplements.[6] In the remainder of the paper, we will refer to those households whose retirement wealth measure comes from the same wealth supplement as cohorts.

3.2 Measuring Wealth

The wealth measure is total household wealth and comes from the PSID wealth supplement.[7] The survey's measure of total wealth is composed of eight asset categories: business/farm, checking/savings, debt, real estate, stock, vehicles, home equity, and other savings. The wealth measure does not include public or private pension wealth. All values are converted to 2005 dollars using the Bureau of Economic Analysis' Personal Consumption Expenditures (PCE) price index.

---

[4] Use of age 55 also makes our results more readily comparable to those of Venti and Wise whose sample has an average age of 55.4 years.

[5] The results presented here exclude households in the SEO sub-sample. The analysis was also performed including those sample members with the PSID-provided family weights, and the results (available in the Data Appendix) are not substantially different than those presented here.

[6] The 1984 supplement is excluded by the requirement that we observe households 20 years prior to the wealth measure because the PSID did not begin until 1968. The 2001 supplement is not used because all but two of those households are picked-up in the 1999 and 2003 supplements.

[7] Households with married adults should save more since they must finance retirement for two people instead of one. Some research (such as Wilmoth and Koso 2002) analyzes wealth per person instead of household wealth, and poverty research frequently applies family equivalence measures to adjust for family size. However, the use of equivalence scales unadjusted for income differences is also problematic (Aaberge and Melby 1998). We therefore follow the traditional wealth literature and use the unadjusted household as the basis of analysis.

Self-reported wealth information is subject to severe non-response. Because the total wealth measure is comprised of several different categories, failure to report any single component means that a total wealth measure is not available for that household. Higher-wealth households have complex holdings that can be difficult to value and are therefore more likely not to be reported. If the non-response is not random (as in this situation) then estimates based on the data will be biased.

The PSID wealth supplement addresses this issue with unfolding brackets.[8] Individuals are first asked to report the dollar amount of a given asset. If the respondent does not state a specific dollar amount then the interviewer follows-up with a series of questions (along the lines of "Would it amount to $50,000 or more?") to determine the appropriate bracket (range of values). Based on those responses, the PSID uses a three-level hot deck procedure to impute the missing values.[9]

3.3 Measuring Long-run Earnings

Ideally, we would include lifetime earnings in the analysis. Given that the PSID only collects contemporaneous income, we are restricted to a measure of long-run earnings with an inherent trade-off – it gets closer to actual lifetime earnings as we include more years, but at the same time, the sample gets smaller.

Table 1 displays this trade-off. The individual cells indicate the correlation between long-run earnings using the indicated age range and long-run earnings based on the 20 years prior to the retirement measure (which is, on average, ages 35-54). Sample sizes are noted in brackets. The effect of including additional years at the beginning of the age range is evident

[8] See Juster and Smith (1997) for details on how brackets remedy non-response bias.
[9] Details regarding the PSID imputation process are provided in the PSID data documentation.

when going from right to left across a row; moving down a column, one can see the effect of adding more years to the end of the age span. It is immediately obvious that increasing the age range greatly reduces sample size. For example, for a starting age of 35, increasing the ending age just five years from 54 to 59 decreases the sample size from 944 to 540. That extra five years eliminates over 40% of the sample, but the measure including those extra years is highly correlated (0.9616) with the original measure. Based on this analysis, we conclude that long-run earnings for ages 35-54 is highly correlated with lifetime earnings, but still retains a large sample size.

The PSID variable used to create the long-run earnings measure is total labor income. It includes wages/salaries, bonuses, overtime, tips, commissions, professional practice or trade, market gardening, additional job income, and miscellaneous labor income. Since retirement decisions are typically made at the household level, the variable of interest is *household* long-run earnings. To that end, the spouse's total labor income is included when applicable. All income is converted to 2005 dollars using the PCE and discounted using a 2% discount rate. Our results are not sensitive to the discount factor used.

Because the survey switched to a biennial schedule in 1997, income data are not available for all years. For the affected sample members, income for the missing years is calculated as the simple average of income in the previous and following year. To adjust for missing income data in years when the survey was conducted (which is different from reported income of zero) we use the average of non-missing household income as the long-run earnings measure.

3.4 Measuring Life Events

The life events considered here are the same as those included in Venti and Wise: inheritance, children, marital status, and health. The difference is, due to the panel structure of the PSID, we can analyze when these events occurred. Specifically, we divide the life event variables into 3 time periods corresponding to approximate ages 35-39, 40-49, and 50-54. That is roughly events occurring during the male's 30s, 40s, and 50s.

Inheritance data are from a question included throughout the PSID survey as to whether any lump payments were received. It includes both inheritances and settlements from an insurance company or lawsuit. The number of children living in the household at the time of the survey (regardless of biological relationship with other household members) is collected in the PSID, and we use the maximum number reported for each age range. The marital status variables included are indicators for those who remain married throughout, get married, or whose marriage ends (through divorce, separation, or widowhood) during the relevant period. Our measure of health comes from a series of questions about work missed due to own illness or illness of a family member.

4. Results

The results are organized as follows. First, we reexamine long-run earnings, life events, and demographic characteristics as analyzed by Venti and Wise and discuss the differences in our results and possible reasons for them. We then consider two additional factors - earnings volatility and pre-existing wealth.

Following Venti and Wise, we focus on how well long-run earnings and life events explain the observed variation in retirement wealth. Before discussing the specific measures considered, a word of caution is required. While measures of goodness-of-fit are appropriate for this analysis, their shortcomings should not be overlooked. They do not measure the quality of the model, and they are within-sample measures only (not population estimates).

There are several common measures of goodness-of-fit including R-squared, adjusted R-squared, and root mean square error (RMSE). The R-squared is the percentage of variation in the dependent variable that is explained by the independent variables. The adjusted R-squared is similar, but it makes a degrees-of-freedom adjustment which effectively penalizes the inclusion of an additional explanatory variable. Because of that alteration, the adjusted R-squared should no longer be interpreted as a percentage like the R-squared. An alternative measure is the RMSE. Unlike the R-squared and adjusted R-squared, this measure is based on the residuals and thus expresses the amount of variation *not* attributable to the regressors. Like the adjusted R-squared, the RMSE includes a degrees-of-freedom adjustment. The RMSE is linked to the R-squared and adjusted R-squared through the identity that the total sum of squares (SST) is equal to the explained sum of squares (SSE) plus the sum of squared residuals (SSR). The R-squared is the ratio of SSE/SST, and the percentage change in RMSE is a function of SSR/SST. Thus, as the R-squared and adjusted R-squared increase, the RMSE decreases. Whether we look at the increase in one of the R-squared measures or the percent decrease in the RMSE, we are essentially capturing the same thing.

We elect to follow Venti and Wise and use the RMSE. The findings are similar regardless of which measure is used, but using RMSE makes our estimates comparable to those of Venti and Wise.

Because the focus here is on overall explanatory power, individual coefficient estimates are not included in the text of the paper. However, coefficients and standard errors for the overall model are reported in the Data Appendix.

4.1 The Role of Long-run Earnings

To alleviate problems arising from the use of censored income data, Venti and Wise rank their sample by lifetime earnings and divide it into deciles. They concede that results for the lowest and highest deciles are less reliable but state that "the ranking by Social Security earnings represents a good approximation to a ranking based on actual total earnings, and that thus the deciles are a good approximation to actual lifetime earnings deciles" (30).

Because we have uncensored data, it is possible to examine the effectiveness of their correction method. We begin by applying the SSA's annual earnings maximums to the annual earnings information collected in the PSID. If the reported earnings are below the earnings maximum, actual earnings are used. When actual earnings exceed that maximum, we substitute the maximum amount for actual earnings. Once this adjustment is made for each year of income for both the male and, when applicable, his wife, we create a new variable called Social Security-adjusted long-run earnings. As with the original long-run earnings variable, it is the average of annual earnings for the twenty years prior to the retirement wealth measure, which is approximately ages 35-54.

Each household is assigned to a long-run earnings decile two times. The first is based on censored earnings data, and the second is from the uncensored earnings data. Table 2 provides a cross-tabulation of the censored and uncensored deciles. If the SSA maximum had no effect (meaning households are assigned to the same decile whether censored or uncensored data are used), all households would be along the diagonal. It is evident looking at Table 2 that the effect of censored data is not restricted to the last decile as Venti and Wise assume. Instead, the effect is spread primarily between deciles 4-10. More than 52% of households in the censored fifth decile change groups, and at least 70% of the seventh through ninth deciles move. Overall, 49% of the sample is assigned to a different decile when the SSA maximum is imposed, and 20% of those move more than one decile.[10]

Table 3 compares long-run earnings and the distribution of retirement wealth for censored and uncensored deciles.[11] Looking first at median long-run earnings, censored and uncensored income measures are similar for the lower deciles, but large differences emerge in the upper deciles. The difference between median income for deciles using censored and uncensored data increases from 10% of the censored data for the fifth decile up to 41% for the tenth decile. We see a similar pattern for median wealth; the lower deciles are relatively similar, but the upper deciles differ considerably.

Because one purpose of the analysis is to determine how much of the observed variation in wealth is attributable to income, differences in the distribution of wealth within deciles are even more important than differences in median wealth levels across deciles. As the coefficient of variation column demonstrates, the within-decile variation is considerably different for the censored and uncensored data. For all but 3 deciles, the degree of variation in wealth to be explained by income is overstated when censored income is used.

The effect of censored income data is further evident in Table 4. The first two rows verify that the subsequent findings are not driven by inherent differences between the HRS and PSID samples; Venti and Wise report that lifetime earnings reduce the RMSE by 5.05%, and the

---

[10] See the Data Appendix for a similar analysis and discussion using long-run earnings quintiles.
[11] See the Data Appendix for a comparison of Venti and Wise's HRS wealth distribution with the PSID wealth distribution.

PSID censored long-run earnings yield a decrease of 1.57%. Using long-run earnings deciles based on uncensored annual income instead of censored data reduces the RMSE by 5.35%. Removing the SSA maximums more than doubles the explanatory power of long-run earnings. Use of the long-run earnings measure in place of decile indicators more than doubles the percent reduction in RMSE. When uncensored long-run earnings and four polynomial terms are included in the regression, the residual standard deviation is reduced by 25.98%.[12] Therefore, differences in long-run earnings alone explain almost 26% of the variation in retirement wealth. These results represent a huge departure from Venti and Wise's estimate of 5%.

Table 2 provides some insight to the causes for this large difference. We know from previous literature that wealth in the U.S. is concentrated at the top. In Table 2 it is clear that those high-wealth households are also the ones most affected by the censored data. Indeed, 25 of the 94 households in the uncensored tenth decile are included in censored deciles five through eight. Use of censored income data inappropriately allocates some of the highest wealth households to lower long-run earnings deciles, obfuscating the correlation between wealth and long-run earnings.

Recall, our sample differs from that of Venti and Wise because it is composed of five different cohorts. Each cohort was exposed to a different macroeconomic environment which may affect their wealth accumulation. To control for any macroeconomic effects, we add a series of cohort indicators to the regression in the final row of Table 4. This specification is used as the baseline regression for the remainder of the analysis.

[12] If we instead use the R-squared or adjusted R-squared to measure explanatory power, long-run earnings explain approximately 37% of the observed variation.

4.2 The Role of Life Events

Venti and Wise identify inheritances, children, marital status, and health as factors that could affect retirement savings, but their findings indicate that these life events explain at most 4% of the observed variation in wealth. However, as was the case with long-run earnings, the data used in their analysis limit the dependability of the results. The demographic information used to measure life events is based on the HRS baseline survey completed in 1992 when the primary respondents are between the ages of 51 and 61. Although some retrospective information is included, little is known about the participants' lives prior to the survey. The panel structure of the PSID allows us to observe households over an extended period of time and therefore control for the timing of these life events. As discussed in Section 3, we do so by controlling for events that occurred when the male was roughly in his 30s, 40s, and 50s.

Table 5 highlights the effect of the improved life event measures. The first row reports Venti and Wise's finding of 4.03% of variation explained by life events. The next row shows the results when life event variables at age 55 (our closest approximation to Venti and Wise) are used.[13] Replication of Venti and Wise's analysis with the PSID sample yields very similar results. However, when we exploit the panel structure of the PSID and employ age-varying life events, the explanatory power of life events increases dramatically (from 6% to 22%).

We can also see the combined effect of income and life events in Table 5. The total percent reduction in RMSE when both factors are included is 48.4%. Thus, we find that the total explanatory power of long-run earnings and life events is over five times greater than Venti and Wise's estimate of 9.08%.

---

[13] Venti and Wise's variables are: inheritances received before 1980, between 1980 and 1988, and after 1988, current marital status, number of children, and current health rating. We use inheritances received between ages 35-39, 40-49, and 50-55, current marital status, maximum number of children ever reported to the PSID between the ages of 35 and 55, and a current health rating.

The lower panel of Table 5 shows the RMSE and incremental effect for each life event individually. Due to missing values for some variables, the sample size is different for each factor. We therefore run the baseline regression for that sub-sample and then add the indicated life event.

Although we do not want to focus on the individual regressions too much since it is likely that the life event variables are correlated with one another, considering each variable one at a time does allow us to see which of the life events seem to have a larger impact. The factor with the single largest effect is marital status. This is not surprising given previous research, but it is remarkable that we get such a large effect using a relatively simple measure. The smallest effect comes from health and is most likely due to difficulties in measuring health.

Although not the focus of this analysis, the coefficient estimates (reported in the Data Appendix) warrant mention. The interaction term of inheritances received in the 30s and long-run earnings is statistically significant, and it is jointly significant with inheritances received in the 30s. Mean long-run earnings in the sample are $85,507, so the coefficients indicate that the correlation of inheritances is small for households with average earnings, positive for those with below-average earnings, and negative for those with higher-than-average earnings. The coefficients for inheritances in the 40s and 50s are not statistically significant, but they do suggest that inheritances received in the 40s increase wealth for those with average earnings, and inheritances in the 50s have virtually no effect for households with average long-run earnings. An F-test indicates that the inheritance variables are jointly significant at the 5% level.

The children-income interaction terms are significant for all three age groups, and the individual and income-interacted terms are jointly significant for all ages. For those with average earnings, the presence of children in the household when in their 30s decreases wealth by about $50,000, but having kids in their 40s increases wealth by over $70,000. Those who still have children in the household in their 50s have significantly less wealth ($110,000 for those with average earnings).

As suggested by the large decrease in RMSE and verified by an F-test, the marital status variables are highly significant. The effect of being married consistently from 35-39 or 50-54 is not statistically significant, but for those with average earnings, being married consistently in the 40s increases wealth by over $250,000. The other large effect is for those whose marriage ends during their 30s. Those with above average earnings have more wealth, but for men with lower earnings, the effect is negative and large in magnitude. The health variables for the 30s are jointly significant at the 10% level, but their effect is relatively small. Overall, the health variables are not jointly significant.

4.3 Race and Education

Venti and Wise initially exclude race and education from their analysis because those variables are likely to be correlated with saving preferences. After finding little effect of long-run earnings, life events, and investment choice, they add both education and race to the regression. Their results indicate that these variables explain very little of the variation in retirement wealth.

Results including education and race in our baseline regression are in Table 6. The individual effect of both education and race is negligible. Briefly returning to the coefficient estimates in the Data Appendix, we see that none of the race and education variables are statistically significant at the 10% level. This is interesting given the documented racial wealth gap discussed in the literature review.

4.4 Other Financial Factors

Our reexamination of the correlates of wealth strongly suggests that the ability to save is more important than Venti and Wise's findings imply. Additional data available in the PSID allow us to delve even deeper. First, we use the uncensored annual income to consider earnings volatility. Next, we exploit the repetition of the PSID wealth supplement to examine the relationship between retirement wealth and wealth holdings earlier in life. The results including these financial factors are in Table 6.

4.4.1 Earnings Volatility

Thus far, our analysis has focused on the relationship between wealth and total earnings. However, it is possible that the timing of income is also important. For example, the precautionary savings model suggests that, due to uncertainty about the future, households increase saving in the current period to insure themselves against unforeseen shocks (such as job loss or medical bills) in the future.[14] It seems likely that the precautionary savings motive would be greater for households with more variation in annual earnings than for those with more stable earnings. As a descriptive way of considering this, we add earnings volatility to the regression analysis.[15]

We use the standard deviation in annual household income to represent earnings volatility. As shown in Table 6, this variable decreases the RMSE by 5.3%. We then add the coefficient of variation to the regression to allow the effect to vary by income level, but it has no effect. Since earnings volatility is likely related to macroeconomic factors, it seems possible that the effect

---

[14] See Kazarosian (1997) for empirical evidence of the precautionary savings motive.

[15] In addition to earnings volatility, the timing of income may also be important. For instance, it makes sense that those who earn most of their money earlier in life would be able to accumulate more retirement wealth than others. To account for this possibility, we add a variable for the percentage of total long-run income earned in the first five years to the regression. The results are not reported, but the variable is not statistically different from zero and has a negligible effect on the RMSE. The same is true for variables representing the percentage of income earned in the first ten years and last five years.

will vary across cohorts. To allow for this, we interact the standard deviation of earnings with the cohort indicator dummy variables. Doing so decreases the RMSE by an additional 8.4 percentage points.

The coefficient estimates (reported in the Data Appendix) are consistent with a precautionary savings motive; higher variation in annual earnings increases wealth. The variable is also highly significant. The income-adjusted term is not statistically significant.

4.4.2 Pre-existing Wealth

The PSID administered the wealth supplement approximately every 5 years beginning in 1984. It is therefore possible to observe household wealth 20 years prior to the measure of retirement wealth for a portion of the sample. Table 6 displays the results including this pre-existing wealth in the baseline regression. The incremental effect of this variable is over 4%. Wealth 20 years earlier has almost as much explanatory power as inheritances and children (see Table 5). These results suggest that some households are able to accumulate substantial wealth early in life, and those are the very households that enter retirement with greater wealth. As with all of our results, it is possible that the driver is differences in tastes instead of ability, but it is doubtful that differences in taste would have a large effect at such a young age. It is more likely that differences so early in life are due to differences in the ability to save. For instance, individuals who start their working career with educational loans have less ability to save than those whose parents paid for college. These findings indicate that a closer look at wealth accumulation earlier in life could add further insight to the issue.

5. Conclusion and Discussion

We utilize the PSID to reexamine the correlation between retirement wealth and both household long-run earnings and a variety of life events. Previous research by Venti and Wise suggests that the effect of these variables is minimal. Indeed, they conclude that long-run earnings explain about 5% of the variation and only an additional 4% is attributable collectively to inheritances, children, marital status, and health. Data limitations, namely the use of censored income data and one-time measures of life events taken later in life, cast doubt on the dependability of their findings.

We perform an analysis similar to that of Venti and Wise and find strikingly different results. In contrast to the meager 5% of variation they find is attributable to long-run earnings, our analysis with uncensored income data indicates that household long-run earnings explain over 25% of the observed variation in wealth. We find that the percent of wealth variation explained by life events is 22% (compared to Venti and Wise's estimate of 4%). Overall, our findings suggest that the explanatory power of long-run earnings and life events is over five times greater than previous estimates imply.

When long-run earnings, life events, and earnings volatility are included jointly, the reduction in RMSE is 56.7%. These results indicate that the ability to save is important. This conclusion is in stark contrast to Venti and Wise's finding that the variation in wealth is almost entirely attributable to preferences.

The shortcomings of our analysis should not be ignored. The requirement to observe the same individual over 20 years yields a much smaller sample than that used by Venti and Wise and makes it harder to distinguish specific attributes. Like Venti and Wise, we also lack an ideal measure of health. Use of an age-varying health measure represents an improvement over the one-time measure at age 55, but it is doubtful that the measure is able to fully capture the effect of health.

Although the magnitude of our results is striking, this analysis suffers from the same difficulties in interpretation as Venti and Wise's study. The goal is to identify the ability to save and its effect on retirement wealth, but our results are in no way causal. A careful examination of these correlations is insightful, but this descriptive analysis cannot eliminate other factors such as tastes and preferences.

Data Appendix

A1. PSID Variables

The PSID's measure of total wealth includes business/farm, checking/savings, debt, real estate, stock, vehicles, home equity, and other savings. The value collected for business/farm is how much the individual would realize if it were sold and all debts on it paid off. Checking/savings amounts include money in checking or savings accounts, money market funds, CDs, government savings bonds, and treasury bills. Items included as debt are credit card balances, student loans, medical or legal bills, and loans from relatives. Mortgages and vehicle loans are explicitly excluded. The real estate category contains real estate other than the main home, land, rental real estate, or money owed on a land contract. Again, the reported value is the amount realized if sold and all debts paid. Stock also includes mutual funds and investment trusts. Prior to 1994, it included IRAs and private annuities. In 1994, IRAs and private annuities were specifically excluded from the category, and beginning in 1999, they were reported separately. Vehicle wealth is the value, net of any amounts still owed, of all cars, trucks, motor homes, trailers, and boats. Home equity is equity on the primary residence. Other savings includes bond funds, cash value of a life insurance policy, a valuable collection for investment purposes, and rights in a trust or estate.

An important component of wealth, pensions, is not included in the PSID total wealth measure. A pension supplement was completed in 1999 that attempts to value pension benefits for both heads and spouses. For defined contribution (DC) pensions, that is relatively easy to do. Individuals can usually give an estimate of their account balance. Then, a life expectancy table can be used to convert that dollar amount to an annuity value. Unfortunately, defined benefit (DB) pensions are much harder to value because they are generally based on a complex formula that varies across employers and typically includes age, years of service, and final salary. Due to their complexity, individuals often do not know the precise formula for their benefits, much less an estimate of the value. This is particularly problematic for our sample whose working years were in the 70s and 80s. Over that period the percentage of DB pensions was decreasing, but it remained between 20-30% of total pensions (EBRI 2003).

Use of the PSID pension information would further restrict the sample size since the supplement was not administered until 1999. Even more importantly, the value-added from this information is questionable. For instance, one question asks whether the individual can estimate what his pension benefits will be. In 2005, only 713 respondents said yes, while 1,555 responded no. Of those who responded yes, there is no way to know how accurate those estimates really are. We therefore do not include pension amounts in our analysis.

Non-labor income is not included in long-run earnings. Items such as rental property income and interest earned on financial assets represent returns on previous investments. The goal here is not to analyze the quality of investments or their rate of return. Including these sources in the income measure would incorporate different rates of return and potentially distort the analysis.

For the sample members with missing data due to the biennial survey schedule, we use information reported in the next available survey. For example, we use the demographic data from the 1999 survey for both 1998 and 1999. Any exceptions to that general rule are noted below.

The inheritance variable is an indicator of whether a lump payment of at least $10,000 was received in the given year. The PSID collects information about the amount of the inheritance, but that amount was capped at $10,000 for survey years 1968-1982. For consistency, we do not include any amounts in our analysis.

A complication arises in the later years when the survey changes to a biennial schedule. Fortunately, a question was added beginning with the 1999 survey which asks if gifts or inheritances worth a total of $10,000 or more were received in the last five years.[16] For each gift/inheritance reported, the year in which it was received is recorded. We therefore construct variables for gifts/inheritances received in each year and allocate all reported amounts based on the year it was received.

As explained above, we consider events that occur roughly during the head's 30s, 40s, and 50s. To that end, we create a separate variable for each of those age ranges that is equal to one if an inheritance of at least $10,000 was received in any of those years. To allow the effect of inheritances to vary by earnings level, we interact each of the age-range variables with long-run earnings.

The PSID collects the number of children living in the household at the time of the survey. This includes all individuals under the age of 18 regardless of their biological relationship with other members of the household. As those children grow up and move out, they are no longer considered part of that household. To create a variable for children during the male's 30s, 40s, and 50s, we take the maximum number of children reported over the time period and then interact that with a variable indicating whether or not the household has any children. Again, we interact the variables for the 30s, 40s, and 50s with long-run earnings to allow the effect of children to be different for those with different earnings.

There are two sources for marital status in the PSID. Each survey asks about current marital status, plus there is a marital history file which covers the individual's entire marital

[16] The 2003 and 2005 surveys ask only about the previous two years.

history. We use the current report as our primary source and supplement any missing values with the marital history report. Marital status is categorized as never married, married, widowed, divorced, or separated.

The relationship between wealth and marital status is complex.[17] In an effort to keep the analysis relatively simple but still capture the important dynamics, we create variables for those who remain married throughout, get married, or whose marriage ends during the relevant period. The base group is therefore those who are consistently unmarried (including those who are never married, widowed, divorced, or separated).

Marital status variables are again created for the three different age ranges and are interacted with dummy variables taking a value of one if the marital status applies to the individual. As with all life event variables, each of the age range marital status variables is interacted with long-run earnings.

The "married" variable indicates that the individual's status is married in each year. The "get married" variable means that the individual's marital status changed from either never married, widowed, divorced, or separated to married within the time frame considered. The "end marriage" variable represents those who transition from married to divorced, separated, or widowed during the period. By construction, it is possible for the same person to be included in multiple categories. For instance, someone who was married at age 35, got divorced at 36, and then remarried at 39 would be included in both the "get divorced" and "get married" groups in his 30s.

Like Venti and Wise, we have limited information on health. They use the individual's current health status reported at the time of the survey when respondents are approximately 55 years old. The PSID variable most similar to that comes from a question that asks the individual

[17] For more details, see Wilmoth and Koso (2002).

to rate his or her health on a scale of one to five. This question was added to the survey in 1984 and is therefore available at age 55 for most of our sample. However, as Venti and Wise recognize, a subjective health rating given later in life is not necessarily an indicator of health over the person's lifetime.

The PSID provides a better alternative. The survey includes a series of questions about missing work due to illness. The questions are: Did you miss any work in [the previous year] because you were sick, how much work did you miss [because you were sick], did you miss any work in [the previous year] because someone else in the family was sick, and how much work did you miss [because someone else in the family was sick]?[18] Although it is not a direct measure of health, these questions do give us an indication of the individual's health during their lifetime. Also, because the survey asks about missing work for family members, we have at least some information about the health of children and spouses when applicable.

We use this information to create a variable equal to the total number of work weeks missed due to either own illness or that of a family member. The number of weeks is already calculated for all years except 1999, 2001, and 2003. For those three surveys, the amount of work missed is reported in days, weeks, or months and we converted all amounts to weeks. That variable is then interacted with a dummy variable indicating if any work was missed. To create the variables for the 30s, 40s, and 50s, that interaction term is summed for the relevant years. Each of the age range variables is interacted with long-run earnings.

Additional variables considered in the analysis are education and race. The PSID collects the number of years of education completed to date. We use the highest level of education reported for the individual as our measure of his educational attainment.

[18] For survey years 1968-1975, work missed due to own illness and work missed to illness of another family member were combined.

Race is included in each survey. However, because race should not change for a given individual, we only need to collect it once. For consistency, we use the race reported in the last survey used for each individual.[19] Possible race categories are white, black, Hispanic, and other.

[19] The one exception is the second cohort. Race was not collected in the 1994 survey, so we use race from the 1993 survey for those individuals.

A2. Regression Coefficients

Table A1: Overall Regression Results

Variable	Coefficient	Standard Error	t-statistic
Long-run Earnings	6.56	5.53	1.19
Long-run Earnings^2	-1.30E-04	4.30E-05	-3.04
Long-run Earnings^3	4.61E-10	1.27E-10	3.64
Long-run Earnings^4	-4.44E-16	1.10E-16	-4.03
Cohort 2	70,073.01	90,078.22	0.78
Cohort 3	14,397.58	82,878.23	0.17
Cohort 4	-38,116.65	81,169.76	-0.47
Cohort 5	-103,268.30	97,439.50	-1.06
Inheritance in 30s	348,144.80	192,470.70	1.81
Inheritance in 30s*Long-run Earnings	-4.54	1.64	-2.76
Inheritance in 40s	206,562.00	150,031.90	1.38
Inheritance in 40s*Long-run Earnings	-0.87	1.43	-0.61
Inheritance in 50s	222,878.30	161,233.10	1.38
Inheritance in 50s*Long-run Earnings	-2.27	1.62	-1.41
Children in 30s	130,225.80	66,424.47	1.96
Children in 30s*Long-run Earnings	-2.11	0.81	-2.62
Children in 40s	-101,886.00	73,633.51	-1.38
Children in 40s*Long-run Earnings	2.05	0.88	2.33
Chidlren in 50s	98,733.86	66,844.20	1.48
Chidlren in 50s*Long-run Earnings	-2.45	0.72	-3.41
Married in 30s	808,756.60	507,622.50	1.59
Married in 30s*Long-run Earnings	-9.18	6.26	-1.46
Married in 40s	-2,100,093.00	349,129.60	-6.02
Married in 40s*Long-run Earnings	27.81	3.83	7.26
Married in 50s	810,425.80	462,516.40	1.75
Married in 50s*Long-run Earnings	-9.45	5.56	-1.70
Get Married in 30s	929,100.70	506,981.80	1.83
Get Married in 30s*Long-run Earnings	-12.64	6.26	-2.02
Get Married in 40s	-604,210.90	503,660.20	-1.20
Get Married in 40s*Long-run Earnings	9.46	5.88	1.61
Get Married in 50s	-53,254.87	401,380.50	-0.13
Get Married in 50s*Long-run Earnings	-0.31	4.95	-0.06
End Marriage in 30s	-1,121,914.00	279,345.90	-4.02
End Marriage in 30s*Long-run Earnings	14.50	3.26	4.45
End Marriage in 40s	-981,752.00	530,221.80	-1.85
End Marriage in 40s*Long-run Earnings	12.43	6.52	1.91
End Marriage in 50s	379,260.30	570,949.90	0.66
End Marriage in 50s*Long-run Earnings	-7.92	6.98	-1.14
Missed Work in 30s	-9,408.79	7,765.24	-1.21
Missed Work in 30s*Long-run Earnings	0.24	0.12	2.02
Missed Work in 40s	-1,476.85	4,792.58	-0.31
Missed Work in 40s*Long-run Earnings	0.02	0.06	0.38
Missed Work in 50s	1,054.50	7,228.51	0.15
Missed Work in 50s*Long-run Earnings	0.05	0.10	0.48
Black	-102,602.80	110,565.60	-0.93
Hispanic	-472,740.80	298,870.10	-1.58
Other Race	233,948.90	195,913.90	1.19
Education	-5,694.99	12,675.72	-0.45
Std. Dev. Of Earnings	19.29	2.60	7.42
Std. Dev. Of Earnings/Mean Earnings	-15,549.97	174,533.60	-0.09
Constant	87,182.46	304,778.00	0.29
Observations			806
R-squared			0.618

Note: Dependent variable is retirement wealth.
Source: Author's calculations from PSID

A3. Alternative Analyses

Censored v. Uncensored Long-run Earnings

When examining the effect of censored income data, sample size will affect the results. There will be fewer observations in each decile for smaller samples, making it more likely that observations will change categories. Since our sample size is smaller than Venti and Wise's, we also follow the same procedure with quintiles. As shown in Table A3, we still find that a substantial number of observations (257) change categories and that the effect is largely concentrated among the upper quintiles.

Table A2: Comparison of Long-run Earnings Quintiles with Censored and Uncensored Data

		Uncensored Quintile				
		1	2	3	4	5
Censored Quintile	1	181	7	-	1	-
	2	8	157	19	3	2
	3	-	25	120	24	20
	4	-	-	50	101	38
	5	-	-	-	60	128

Note: Sample size is 944
Source: Author's calculations from PSID

SEO Sub-sample with Weights

The PSID initial sample of approximately 4,800 households comes from two sources. The first 1,872 households are from the sample used in the 1966-1967 Survey of Economic Opportunity, which was a Census study to analyze the effect of the War on Poverty. That sub-sample is composed primarily of low-income households and is referred to as the SEO sample. To make the sample nationally representative, a cross-section sample of 2,930 households was added from the Survey Research Center's national sampling frame. That segment of the sample is called the SRC.

In each year of the survey, the PSID provides weights to appropriately combine the two sub-samples and make the sample nationally-representative. However, given the structure of our sample in which households are combined from different waves of the survey, it is unclear which weights should be used. We therefore elect to present the results excluding the SEO sample members. However, we also performed the same analysis using the full sample. The table below includes results using the 1989 family weights for all households as well as those using family weights for the last year the male is observed in our sample (which is approximately age 55). Overall, the central finding remains unchanged: substantially more of the variation in retirement wealth is attributable to long-run earnings and life events when better measures are employed.

Table A3: Regression Goodness of Fit with PSID Family Weights

	Percent Reduction in RMSE	
Control Variables	1989 Family Weights	Age 55 Family Weights
Long-run Earnings	27.60%	25.56%
Life Events	14.46%	7.33%
Percent of Variation Attributable to Long-run Earnings and Life Events	**42.06%**	**32.88%**

Notes:

[1] Long-run Earnings regression is retirement wealth on long-run earnings, long-run earnings polynomial terms, and cohort indicator dummy variables.

[2] Life events include inheritance, children, marital status and health. All life event variables are interacted with long-run earnings.

[3] Sample size for 1989 family weights is 1,123; sample size for Age 55 family weights is 1,106.

Source: Author's calculations from PSID

Excluding Home Equity

The PSID total wealth measure used in the main analysis includes home equity, and this is consistent with the wealth measure in Venti and Wise's analysis. However, the retirement wealth literature contains extensive discussion as to whether it is correct to include home equity in retirement wealth (for instance, see Gale 1997). We therefore complete the same analysis using a measure of total wealth that excludes home equity. The results, shown in Table A4 below, verify that the explanatory power of long-run earnings and life events is quite similar for the two wealth measures.

Table A4: Regression Goodness of Fit Excluing Home Equity

Specification	Incremental Effect
Long-run Earnings	23.76%
Life Events	20.78%
Total	44.54%

Note: Specification of regressions is identical to that presented in the main text of the paper.
Source: Author's calculations from PSID

A4. Wealth-Earnings Distribution Comparison

Table A5 below compares median wealth by long-run earnings decile for the HRS sample used in Venti and Wise (2001) and the PSID sample used in this analysis. The only clear difference between the two wealth measures is that the Venti and Wise wealth amounts include pensions. In an effort to make the two samples more comparable, we subtracted the median pension wealth by decile from the respective median wealth measure.

After adjusting for pensions, PSID wealth is considerably larger. Although it is not possible to trace the exact cause for this difference, it is most likely due to the fact that Venti and

Wise include women in their sample. As discussed in the literature review, the wealth gap between men and women is substantial. The larger wealth in our sample may also be contributing to our larger results.

Table A5: Sample Comparison of Median Wealth by Long-run Earnings Deciles

Long-run Earnings Decile	Median Wealth			
	Venti & Wise	V&W without Pensions	Censored PSID	Uncensored PSID
1	$5,000	$5,000	$28,750	$26,320
2	34,429	34,429	66,332	65,346
3	52,803	52,803	85,298	85,298
4	82,620	82,620	185,008	156,263
5	105,166	101,166	187,941	136,904
6	144,188	130,153	159,522	202,207
7	189,832	156,039	233,331	214,766
8	221,692	180,884	255,636	287,453
9	305,536	247,536	360,100	376,950
10	387,609	304,350	375,003	646,225

Notes:

[1] Venti & Wise wealth includes business equity, personal financial assets, real estate, personal retirement assets (including IRA and 401(k) balances), vehicles, home equity, and pensions; PSID wealth includes business/farm equity, checking/savings accounts, real estate, private annuity/IRA, stock, vehicle, home equity, and other personal savings.

[2] V&W without Pensions is calculated as the original Venti & Wise median wealth less the median value of traditional pensions for each decile.

[3] PSID wealth values are given in 1992 dollars to be more comparable with Venti and Wise.

Sources: Venti and Wise (2001) and author's calculations from PSID

References

Aaberge, Rolf and Ingrid Melby. 1998. "The Sensitivity of Income Inequality to Choice of Equivalence Scales." *Review of Income and Wealth*, 44(4): 565-596.

Altonji, Joesph G., Ulrich Doraszelski, and Lewis Segal. 2000. "Black/white differences in wealth." *Economic Perspectives*, Federal Reserve Bank of Chicago, Winter 2000.

Banks, James, Richard Blundell, and Sarah Tanner. 1998. "Is There a Retirement Savings Puzzle?" *American Economic Review*, 88(4): 769-788.

Bajtelsmit, Vickie L. and Alexandra Bernasek. 1996. "Why Do Women Invest Differently Than Men?" *Financial Counseling and Planning*, 7: 1-10.

Blau, Francine D. and John W. Graham. 1990. "Black-White Differences in Wealth and Asset Composition." *Quarterly Journal of Economics*, 105(2): 321-339.

Cagetti, Marco and Mariacristina De Nardi. 2005. "Wealth inequality: data and models." Federal Reserve Bank of Chicago Working Paper 2005-10.

EBRI research and education team. 2003. "EBRI Research Highlights: Retirement Benefits." *EBRI Special Report SR-42*, Issue Brief Number 258.

Gale, William G. 1997. "The Aging of America: Will the Baby Boom Be Ready for Retirement?" *The Brookings Review*, 15(3): 4-9.

Gittleman, Mary and Edward N. Wolff. 2000. "Racial Wealth Disparities: Is the Gap Closing?" Working Paper.

Golderberger, Arthur S. 1991. *A Course in Econometrics*, Harvard University Press.

Haider, Steven J. and Melvin Stephens, Jr. 2006. "Is There a Retirement-Consumption Puzzle? Evidence Using Subjective Retirement Expectations." *Review of Economics and Statistics*, 89(2): 247-264.

Hamermesh, Daniel S. 1984. "Consumption During Retirement: The Missing Link in the Life Cycle." *Review of Economics and Statistics*, 66(1): 1-7.

Hurd, Michael D. 1990. "Research on the Elderly: Economic Status, Retirement, and Consumption and Saving." *Journal of Economic Literature*, 2 (June): 565-637.

Hurd, Michael and Susann Rohwedder. 2003. "The Retirement-Consumption Puzzle: Anticipated and Actual Declines in Spending at Retirement." NBER Working Paper 9586.

Hurst, Erik. 2006. "Grasshoppers, Ants and Pre-Retirement Wealth: A Test of Permanent Income Consumers." NBER Working Paper 10098.

Hurst, Erik, Ming-Ching Luoh, and Frank Stafford. 1998. "Wealth Dynamics of American Families: 1984-94." *Brookings Papers on Economic Activity*, 1998(1): 267-337.

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(440): 1268-1278.

Kazarosian, Mark. 1997. "Precautionary Savings – A Panel Study." *The Review of Economics and Statistics*, 79(2): 241-247.

Kotlikoff, Laurence J. and Lawrence H. Summers. 1981. "The Role of Intergenerational Transfers in Aggregate Capital Accumulation." *The Journal of Political Economy*, 89(4): 706-732.

Lupton, Joseph P. and James P. Smith. 2003. "Marriage, Assets, and Savings." In *Marriage and the Economy: Theory and Evidence from Advanced Industrial Societies*, edited by Shoshana A. Grossbard-Shechtman, 129-152. New York: Cambridge University Press.

Menchik, Paul L. and Nancy Ammon Jianakopolos. 1997. "Black-White Wealth Inequality: Is Inheritance the Reason?" *Economic Inquiry* 35(April 1997): 428-442.

Schmidt, Lucie and Purvi Sevak. 2006. "Gender, Marriage, and Asset Accumulation in the United States." *Feminist Economics,* 12(1-2): 139-166.

Venti, Steven F. 2006. "Choice, Behavior, and Retirement Saving." In *The Oxford Handbook of Pensions and Retirement Income*, edited by Gordon L. Clark, Alicia H. Munnell, and J. Michael Orszag. Oxford: Oxford University Press.

Venti, Steven F. and David A. Wise. 2001. "Choice, Chance, and Wealth Dispersion at Retirement." In *Aging Issues in the United States and Japan*, edited by Seiritsu Ogura, Toshiaki Tachibanaki, and David A. Wise, 25-64. Chicago: The University of Chicago Press.

Venti, Steven F. and David A. Wise. 1999. "Lifetime Earnings, Saving Choices, and Wealth at Retirement." In *Wealth, Work, and Health: Innovations in Survey Management in the Social Sciences*, edited by James P. Smith and Robert J. Willis, 87-120. Ann Arbor: The University of Michigan Press.

Venti, Steven F. and David A. Wise. 1998. "The Cause of Wealth Dispersion at Retirement: Choice or Chance?" *The American Economic Review,* Papers and Proceedings of the Hundred and Tenth Annual Meeting of the American Economic Association, 88(2): 185-191.

Wilmoth, Janet and Gregor Koso. 2002. "Marital History and Wealth Outcomes." *Journal of Marriage and Family*, 64 (February 2002): 254-268.

Wolff, Edward N. 1998. "Recent Trends in the Size Distribution of Household Wealth." *Journal of Economic Perspectives*, 12(3): 131-150.

Table 1: Correlations between Long-run Earnings for Average Ages 35-54 and Select Age Ranges

		Average Starting Age			
		20	25	30	35
Average Ending Age	49	0.9518	0.9556	0.9660	0.9728
		[404]	[634]	[790]	[944]
	54		0.9832	0.9937	1.0000
			[634]	[790]	[944]
	59			0.9571	0.9616
				[386]	[540]
	64				0.9750
					[310]

Note: Sample size in brackets
Source: Author's calculations from PSID

Table 2: Comparison of Long-run Earnings Deciles with Censored and Uncensored Data

		Uncensored Long-run Earnings Decile									
		1	2	3	4	5	6	7	8	9	10
Censored Long-run Earnings Decile	1	91	3	-	-	-	-	-	-	-	-
	2	3	83	6	1	-	-	1	-	-	-
	3	-	8	76	7	1	1	-	-	1	-
	4	-	-	12	63	10	6	3	-	1	-
	5	-	-	-	24	46	10	-	6	3	6
	6	-	-	-	-	38	28	10	8	8	3
	7	-	-	-	-	-	46	20	10	11	8
	8	-	-	-	-	-	4	55	16	11	8
	9	-	-	-	-	-	-	6	48	17	23
	10	-	-	-	-	-	-	-	6	42	46

Note: Sample size is 944
Source: Author's calculations from PSID

Table 3: Household Long-run Earnings and Wealth by Decile

Censored Data

Lifetime Income Decile	Median Household Long-run Earnings	Wealth: 25th Percentile	Wealth: 50th Percentile	Wealth: 75th Percentile	Wealth: Coefficient of Variation
1	$24,088.19	$4,354.15	$29,584.51	$83,320.67	176.43
2	40,315.95	29,224.41	71,529.20	147,817.94	309.28
3	49,124.05	34,039.46	100,563.66	216,446.16	210.50
4	57,757.16	78,652.89	186,854.92	355,603.75	238.43
5	65,205.86	79,082.04	203,000.00	394,423.22	287.51
6	74,982.73	88,690.13	180,000.00	330,795.48	184.07
7	83,283.33	157,319.40	272,968.45	526,985.10	230.85
8	91,200.08	164,900.00	294,399.67	553,257.50	107.73
9	102,088.31	191,849.09	419,166.46	828,830.41	98.96
10	122,074.25	277,100.00	440,627.50	885,000.00	256.47

Uncensored Data

Lifetime Income Decile	Median Household Long-run Earnings	Wealth: 25th Percentile	Wealth: 50th Percentile	Wealth: 75th Percentile	Wealth: Coefficient of Variation
1	$24,882.15	$4,354.15	$29,450.68	$79,984.73	167.14
2	40,281.15	29,224.41	70,398.75	138,038.72	153.11
3	50,544.33	40,500.00	100,563.66	202,818.42	284.01
4	60,728.24	76,500.00	176,715.51	328,000.00	192.69
5	72,037.76	73,908.44	141,951.80	252,344.55	222.60
6	82,735.51	95,869.81	216,073.03	379,295.57	260.63
7	93,350.46	162,254.74	238,811.55	351,936.54	203.36
8	106,222.00	175,268.60	323,612.47	549,034.16	101.57
9	124,054.20	264,645.02	432,373.67	849,500.00	151.23
10	171,514.48	422,776.43	681,995.80	1,387,848.71	184.93

Notes:

[1] Total sample of 944 households which completed a wealth supplement when the male was near age 55 and were observed 20 years prior to the wealth measure.

[2] Household Long-run Earnings is the average of non-missing discounted male's real total labor earnings for 20 years prior to retirement wealth supplement and, when applicable, the wife's discounted real total labor earnings in those same years.

[3] Total wealth includes value of business/farm owned, checking/savings accounts, real estate, stock, vehicle, private annuity/IRA, home equity, and other personal savings.

[4] All dollar amounts are deflated to 2005 dollars.

Source: Author's calculations from PSID

Table 4: Regression Goodness of Fit by Long-run Earnings Measure

Specification	N	Percent Reduction in RMSE
Venti & Wise Lifetime Earnings Deciles	3,992	5.05%
PSID Censored Long-run Earnings Deciles	808	1.57%
PSID Uncensored Long-run Earnings Deciles		5.35%
PSID Uncensored Long-run Earnings		13.62%
PSID Uncensored Long-run Earnings with Polynomial Terms		
2nd		14.39%
3rd		21.08%
4th		25.98%
PSID Uncensored Long-run Earnings with Polynomial Terms and Cohort Indicators		26.03%

Note: Wealth is the dependent variable for all regressions.

Sources: Venti and Wise (2001) and author's calculations from PSID

Table 5: Regression Goodness of Fit by Life Events Measure

Specification	N	Percent Reduction in RMSE	Incremental Effect
Venti & Wise One-time Life Event Measures	3,992	9.08%	4.03%
PSID One-time Life Event Measures	808	31.98%	5.95%
PSID Age-varying Life Events Measures		48.39%	22.36%
Age-varying Life Events Added Individually			
Baseline Regression	842	26.00%	
Plus Inheritance		31.81%	5.81%
Baseline Regression	882	24.04%	
Plus Children		29.26%	5.22%
Baseline Regression	862	23.93%	
Plus Marital Status		44.50%	20.57%
Baseline Regression	877	23.95%	
Plus Health		24.26%	0.31%

Notes:

[1] Venti and Wise interact life event variables with earnings decile, so here life event variables are interacted with long-run earnings.

[2] Baseline regression is retirement wealth on long-run earnings, long-run earnings polynomial terms, and cohort indicator dummy variables.

Sources: Venti and Wise (2001) and author's calculations from PSID

Table 6: Regression Goodness of Fit Including Additional Factors

Control Variables	N	Percent Reduction in RMSE	Incremental Effect
Additional Demographics			
Baseline Regression	943	16.57%	
Plus Education		16.53%	-0.04%
Baseline Regression	933	16.84%	
Plus Race		16.78%	-0.06%
Earnings Volatility			
Baseline Regression	944	16.57%	
Plus Standard Deviation of Annual Income		21.84%	5.28%
Plus Standard Deviation of Annual Income Divided by Mean Long-run Earnings		21.80%	-0.04%
Plus Standard Deviation of Annual Income Interacted with Cohort Indicators		30.19%	8.39%
Pre-existing Wealth			
Baseline Regression	404	19.41%	
Plus Pre-existing Wealth		23.56%	4.15%

Note: Baseline regression is retirement wealth on long-run earnings, long-run earnings polynomial terms, and cohort indicator dummy variables.

Source: Author's calculations from PSID