NBER WORKING PAPER SERIES

HOUSEHOLD STOCK MARKET BELIEFS AND LEARNING

Gábor Kézdi
Robert J. Willis

Working Paper 17614
http://www.nber.org/papers/w17614

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2011

Support from the National Institute of Aging (PO1 AG026571 and RO3 AG29469) is acknowledged. The authors thank Andras Fulop, Krisztina Molnar, Mathew Shapiro, Adam Szeidl, and seminar participants at Central European University, CERGE-EI Prague, NBER, the University of Michigan, the University of Munich, the EEA/ESEM 2008 meetings, and the First Jackson Hole Conference on Subjective Probabilities for their comments. Peter Hudomiet provided excellent research assistantship and many valuable comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Household Stock Market Beliefs and Learning
Gábor Kézdi and Robert J. Willis
NBER Working Paper No. 17614
November 2011
JEL No. C31,D12,D8

ABSTRACT

This paper characterizes heterogeneity of the beliefs of American households about future stock market returns, provides an explanation for that heterogeneity and establishes its relationship to stock holding behavior. We find substantial belief heterogeneity that is puzzling since households can observe the same publicly available information about the stock market. We propose a simple learning model where agents can invest in the acquisition of financial knowledge. Differential incentives to learn about the returns process can explain heterogeneity in beliefs. We check this explanation by using data on beliefs elicited as subjective probabilities and a rich set of other variables from the Health and Retirement Study. Both descriptive statistics and estimated relevant heterogeneity of the structural parameters provide support for our explanation. People with higher lifetime earnings, higher education, higher cognitive abilities, defined contribution as opposed to defined benefit pension plans, for example, possess beliefs that are considerably closer to what historical time series would imply. Our results also suggest that a substantial part of the reduced form relationship between stock holding and household characteristics is due to differences in beliefs. Our methodological contribution is estimating relevant heterogeneity of structural belief parameters from noisy survey answers to probability questions.

Gábor Kézdi
Department of Economics
Central European University
9 Nador St, Budapest, Hungary
kezdig@ceu.hu

Robert J. Willis
3254 ISR
University of Michigan
P. O. Box 1248
426 Thompson Street
Ann Arbor, MI 48106
and NBER
rjwillis@isr.umich.edu

1 Introduction

Beliefs about stock market returns are important determinants of households' investment behavior. Recent research has established the strong relationship between beliefs and stock-holding, and it also documented substantial heterogeneity in those beliefs (Vissing-Jorgensen, 2004; Dominitz and Manski, 2005 and 2007; Amromin and Sharpe, 2006). This heterogeneity is puzzling since stock returns are publicly observable, and all of its history as well as many analyses are public information. Understanding the source of heterogeneity is important to understand heterogeneity in household finances, which is substantial (Campbell, 2006).

The goal of this paper is to characterize heterogeneity of the stock market beliefs of American households, understand the sources of that heterogeneity, and establish its relation to household portfolios. Our substantive contribution is to provide a more systematic account of the heterogeneity than the previous literature and relate it to a relatively simple explanation. Our methodological contribution is to estimate structural belief parameters from noisy survey answers to probability questions of the type advocated by Manski (2004).

We hypothesize that heterogeneity is the result of differences in learning histories, which are in turn caused by differences in returns to and costs of learning (as well as in initial conditions). People learn about finance in general and the stochastic process of stock market returns in particular. The value of learning is proportional to savings, but the costs are fixed. As a result, people with higher earnings prospects should learn more than people with lower income prospects, especially if social security provides enough retirement income for the latter. Differences in the costs of learning and differences in general attitudes may also be heterogenous, creating additional heterogeneity in learning outcomes. Initial conditions matter, too. Those with very low expectations will be less likely to learn and will see their beliefs unchanged. In the end, those who learn will revise their initial beliefs to be more precise, closer to what historical series would imply, and learning makes beliefs less heterogenous. This is a human capital explanation applied to financial knowledge (as in Delavande, Rohwedder and Willis, 2008). It is also an application of the information choice theory of Veldkamp (2011).

We characterize beliefs by the subjective mean and subjective standard deviation of the one year ahead log return on the stock market index. These are unobserved variables that we relate to observed answers to two survey questions: one about the probability that the stock

market return would be positive and the other one about the probability that the returns would be 10 per cent or more. Our structural estimation model accounts for survey response error due to rounding, potential inattention, and the unwillingness or inability to make the necessary effort to give precise answers. A subset of the respondents in our sample answered the same pair of questions twice in the survey, about half an hour apart, which allows us to calibrate the moments of survey noise in a direct way.

We verify the implications of the learning theory by empirical evidence on stock market beliefs using a sample of 55 to 64 years old respondents of the Health and Retirement Study. Our sample consists of people who are at the peak of their asset accumulation process, and their beliefs and household portfolios are the result of their learning and investment history. We first show correlations and OLS regressions using observable answers to probability questions. Then we estimate the structural econometric model and estimate the theoretically interesting belief parameters conditional on the survey answers (analogously to the prediction of individual risk tolerance by Kimball, Sahm and Shapiro, 2009). Our structural model separates survey noise from relevant heterogeneity.

Our estimates show that respondents have low expectations and high perceived risk on average and substantial heterogeneity in expectations. Results from both the simple and the more structural analysis support the learning explanation. People who had stronger incentives to learn in the past indeed possess beliefs that are consistent with more learning. In particular, people with higher lifetime earnings, higher education, higher cognitive abilities, defined contribution as opposed to defined benefit pension plans, and those who are more optimistic and less uncertain about things in general have stock market beliefs that are less heterogeneous, somewhat less uncertain and considerably closer in levels to what historical time series would imply. Our results also show that the people who did not have incentives to learn are very pessimistic about stock market returns.

On top of the small literature on beliefs, many papers have looked at reduced-form associations of stock market participation with demography, education and wealth (Ameriks and Zeldes, 2004; Guiso, Haliassos, and Jappelli, 2002), cognitive capacity (Christelis, Jappelli and Padula, 2010), health (Rosen and Wu, 2003), or social interactions (Guiso, Sapienza and Zingales, 2004; Hong, Kubik and Stein, 2004). The results of our theoretical explanation and our empirical investigation are all in line with the results of that literature. They also suggest that part of those reduced form associations may operate through differential

incentives for learning about attainable stock returns.

The rest of the paper is structured the following way. Section 2 contains a brief characterization of stock market beliefs. Section 3 summarizes the setup and the most important implications of a simple theoretical model of household portfolio choices with learning. We then describe our data as briefly as possible in Section 4, and move on to descriptive evidence on the probability answers themselves in Section 5. Section 6 covers the estimation of the structural parameters of beliefs and their association with stockholding and the right hand-side variables. The last part concludes. Four appendices present the details of our investigation. Appendix A contains the formal structure of our theoretical model and its results. Appendix B shows more details of our data, descriptive statistics and results from linear regressions on observables. Appendix C contains the details of the structural estimation model, and Appendix D contains detailed estimation results and robustness checks.

2 Characterizing stock market beliefs

We assume that people believe that yearly log returns are i.i.d. and normally distributed. Throughout the paper we denote the mean of log returns as μ and the standard deviation as σ. For example, $\mu = 0.1$ means that the mean return is approximately ten per cent. At yearly frequency, the i.i.d. normal assumption for log returns lines up well with historical data available respondents to the 2002 wave of the survey we analyze. In the period of 1945 to 2002, yearly log nominal returns of the Dow Jones index were characterized by a mean of $\mu = 0.07$ and a standard deviation of $\sigma = 0.15$. Different windows can give lower and higher values of μ, and the value of σ is remarkably stable.

Under the i.i.d. lognormality assumption, the beliefs of individual i about the stock market returns are fully characterized by her beliefs about the mean and the standard deviation, and we denote those subjective beliefs by $\tilde{\mu}_i$ and $\tilde{\sigma}_i$. We define $\tilde{\mu}_i$ and $\tilde{\sigma}_i$ as the parameters that would characterize individual beliefs in investment situations. The goal of this paper is to characterize heterogeneity in $\tilde{\mu}_i$ and $\tilde{\sigma}_i$, understand the sources of that heterogeneity, and establish its relationship to heterogeneity in household portfolios.

$\tilde{\mu}_i$ and $\tilde{\sigma}_i$ are unobserved in our data. Instead, we observe answers to probability questions. In the larger part of the sample that we use to show descriptive statistics, one question was asked. This question (p_0) asked what the respondent thought the probability is that the

market will go up. In the sample that we use for the structural analysis, we have answers to another probability question as well (p_{10}), about the probability that the market will go up by at least 10 per cent. The questions themselves were phrased the following way.

p_0 question: By next year at this time, what is the percent chance that mutual fund shares invested in blue chip stocks like those in the Dow Jones Industrial Average will be worth more than they are today?

p_{10} question: By next year at this time, what is the chance they will have grown by 10 percent or more? [1]

When answers to both p_0 and p_{10} are available, identifying the mean and standard deviation of log returns from the two probabilities is relatively straightforward under the normality assumption. Let R denote one year ahead gross returns, which is a random variable with $\ln R \sim N(\mu, \sigma^2)$. In principle, one can relate these probabilities to the parameters of the lognormal distribution in a straightforward way. Let heterogeneity be denoted by an i index, the subjective nature of the probabilities by the tilde, and let stars denote theoretically correct probabilities derived from subjective beliefs; actual survey answers may be different, see later. Then,

$$\begin{aligned} \tilde{p}^*_{0i} &= \tilde{P}_i[R \geq 1] = \tilde{P}_i[\ln R \geq 0] = \Phi\left(\frac{\tilde{\mu}_i}{\tilde{\sigma}_i}\right) \\ \tilde{p}^*_{10i} &= \tilde{P}_i[R \geq 1.1] \approx \tilde{P}_i[\ln R \geq 0.1] = \Phi\left(\frac{\tilde{\mu}_i - 0.1}{\tilde{\sigma}_i}\right) \end{aligned}$$

Observing $\tilde{p}^*_{0i}$ and $\tilde{p}^*_{10i}$ would allow for a simple computation of $\tilde{\mu}_i$ and $\tilde{\sigma}_i$ by making use of the inverse normal c.d.f. Higher $\tilde{\mu}_i$ corresponds to higher probabilities, while higher $\tilde{\sigma}_i$ pushes the argument of Φ toward zero and thus pushes both probabilities towards 0.5.

In order to see the correspondence between $\tilde{\mu}_i$ and $\tilde{\sigma}_i$ on the one hand and $\tilde{p}^*_{0i}$ and $\tilde{p}^*_{10i}$ on the other hand in more intuitive ways, Figure 1 shows three probability distribution functions together with vertical lines at the cutoff points of 0 and 0.1 log returns that correspond to the p_0 and p_{10} questions. The continuous line shows a p.d.f. with historical

[1] Note that the wording of the questions ("will be worth more") is somewhat vague. We interpret it as nominal returns without taking inflation, taxes or investment costs into consideration. If financially more sophisticated people have higher and more precise expectations, and, at the same time, they are more likely to think in real and/or after-tax terms, we shall underestimate heterogeneity in beliefs and its relation to variables that are related to financial sophistication.

moments between 1945 and 2002 ($\mu = 0.07$ and $\sigma = 0.15$). The dashed line corresponds to a mean-preserving spread (higher perceived risk), and the dotted line corresponds to a lower mean (more pessimistic beliefs).

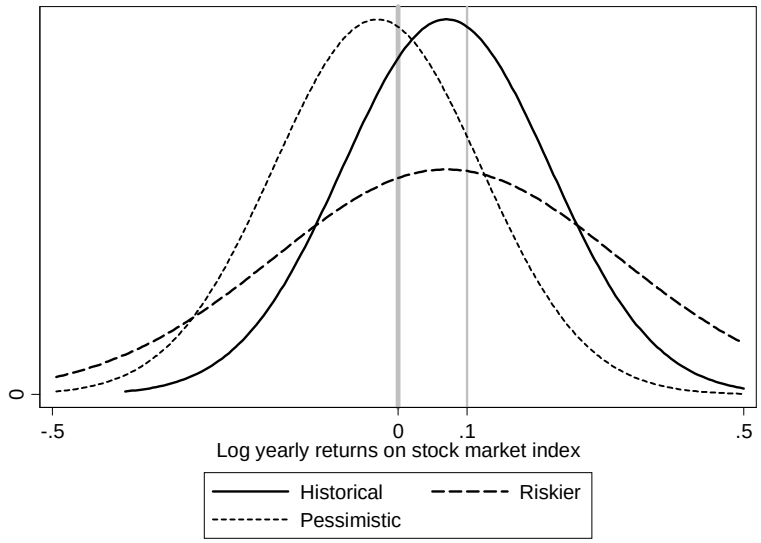


Figure 1. Examples for probability densities of normally distributed log returns, with the cutoff points for p_0 and p_{10}

$\tilde{p}^*_{0i}$ and $\tilde{p}^*_{10i}$ are equal to the area to the right of the corresponding bars at 0 and 0.1 log returns, respectively. The series of post-war returns to 2002 corresponds to $p^*_0 = 0.68$, $p^*_{10} = 0.42$ and $p^*_0 - p^*_{10} = 0.26$.

Holding risk constant, more pessimistic beliefs result in smaller values of $\tilde{p}^*_{0i}$ and $\tilde{p}^*_{10i}$. We can therefore think of the answer to the p_0 (or the p_{10}) questions as proxy variables for the perceived level of returns. A mean-preserving spread leads to smaller area between the two vertical bars, which equals the difference $\tilde{p}^*_{0i} - \tilde{p}^*_{10i}$. The difference between the two answers may thus serve as a proxy for the inverse of perceived risks.

These proxies are not clean, though. The effect of risk on the probabilities can be ambiguous: higher risk corresponds to a smaller area to the right of a cutoff point if the mean is to the right (as for cutoff 0 when comparing the solid and the dashed curves), but it corresponds to a larger area if the mean is to the left (as for cutoff 0.1). Optimism/pessimism affects the difference between the probabilities, too, in ambiguous ways. For example, optimism decreases the difference if the mean is shifted outside the interval between the two bars

from within the bars (as is the case for the dotted curve here), but the effect is the opposite if the mean is shifted towards to the middle of the interval. Simultaneous heterogeneity in the mean and the variance can lead to more complicated heterogeneity in the level and the difference of $\tilde{p}^*_{0i}$ and $\tilde{p}^*_{10i}$.

Observing $\tilde{p}^*_{0i}$ and $\tilde{p}^*_{10i}$ would identify $\tilde{\mu}_i$ and $\tilde{\sigma}_i$ at the individual level. Instead of $\tilde{p}^*_{0i}$ and $\tilde{p}^*_{10i}$, however, we are likely to observe something else, as answers to the probability questions contain substantial noise with a complicated structure.

$\tilde{\mu}_i$ and $\tilde{\sigma}_i$ are the parameters that are relevant in investment situations. There are, however, strong theoretical reasons to believe that people's answers to the probability questions are not equal to the p^* transformations of these parameters. There is little time to answer the questions, and, beyond a spirit of cooperation, there are no incentives to get the answers right. It is therefore better to look at actual answers as "guesses" for what the p^* values may be, given recollections of $\tilde{\mu}_i$ and $\tilde{\sigma}_i$.

The data reveals answer patterns that strongly support this view, and we shall document that later. Some of these answer patterns make computing the p^* values impossible. All of the answer patterns indicate that actual answers are noisy transformations of relevant beliefs. Our structural econometric model will address these problems.

3 Heterogeneity in beliefs and learning

In this section we briefly summarize the explanation we propose for the heterogeneity in stock market beliefs. It is in the spirit of the human capital literature and its application to financial knowledge (as in Delavande, Rohwedder and Willis, 2008), embedded in a standard life-cycle model with borrowing constraints. It can also be viewed as an application of the information choice theory of Veldkamp (2011). Appendix A contains a formal model, its numerical solution and its most important comparative static results.

Heterogeneity in stock market beliefs is the result of differences in learning histories, which differences are in turn caused by differences in returns to and costs of learning (and, potentially, differences in initial conditions). Suppose that individuals live for three periods: young adulthood, mature adulthood and old age. Young people are endowed with some initial sets of beliefs about the mean and the standard deviation of log returns ($\tilde{\mu}$ and $\tilde{\sigma}$), and they share the belief that returns are i.i.d. lognormal.

In young adulthood, people work, earn wages, consume and save, but are subject to borrowing constraints. Mature adults look the same except that wages are considerably higher. In old age, people receive pension benefits and earn no wage. Importantly, pensions are from a defined-benefit system such as Social Security, and pension benefits are a concave function of lifetime earnings. As a result, people who earn below a certain threshold do not have an incentive to save for retirement, people who earn above that threshold do save, and their saving rate depends on their lifetime earnings. Savings can be held in risk-free bank accounts or invested in risky stocks. Borrowing and short sales are not allowed, making the share of stocks in per period savings between zero and one. The returns on stocks are governed by the i.i.d. lognormal process with historical moments, regardless of individuals' beliefs. But whether and how much an individual chooses to invest into stocks in any period depends on her beliefs about those moments at the beginning of the period.

The essence of the model is the possibility to learn about the parameters of the returns process. Learning is the result of a choice. If people choose to learn, they can observe a long historical series of log returns and update their subjective beliefs in a standard Bayesian fashion. The results are a posterior mean that is closer to the historical average and a smaller posterior variance. Learning is more general than updating beliefs: it is understanding the ways in which investment works. As a result, earning can increase the attainable expected return and/or reduce the risk of their portfolio on top of the effects of learning on beliefs about stock market returns. If they choose to learn, people have to pay a fixed cost. People may also learn in a more passive fashion: If they have high enough earnings in young age, they may invest those into stocks, and observing the returns will allow for updating beliefs.

The implications of this model are straightforward. The value of learning is proportional to intended savings, but the costs are fixed. Those who have higher lifetime earnings or lower learning costs will be more likely to learn. Similarly, if we allow for heterogeneity in initial belief endowments, higher initial subjective mean and lower initial variance will also lead to higher propensity to learn. Higher risk tolerance and more patience also lead to higher propensity to learn.

These results have important implications for the empirical analysis of heterogeneous stock market beliefs and household portfolios. Heterogeneity in lifetime earnings reflects heterogeneity in general human capital which, in turn, is the result of differences in the costs of and the benefits to investment based on probability beliefs about future returns. (Becker,

1964,1993; Willis, 1986; Card, 1998). If stock market beliefs are result of investment into a specific form of human capital, all personal characteristics that are related to lifetime earnings will also be related to beliefs as well, even conditional on lifetime earnings. This is another channel through which earnings and household finances are related, on top of the more traditional argument for the role of background risk (emphasized by, e.g., Viceira, 2001).

4 Data

In this section we give a brief overview of our sample and discuss the definition of the variables we use in the analysis. The Appendix B contains additional information; figures and tables in the Appendix B are labeled as B1, B2 etc.

We use data from the Health and Retirement Study (HRS), a large biannual panel household survey that follows older Americans (see NIA, 2007, for review). The HRS is representative of the American population 50 years of age or older, and their households. HRS has had a number of probability questions from 1992 on. It added questions on stock market beliefs in 2002. Besides subjective probabilities, HRS collects data on the amount and structure of savings, including tax-sheltered accounts such as 401(k), a rich set of demographic variables, and measures of cognitive functioning. In addition, retrospective earnings data from W2 tax forms are linked to a large subset of the HRS respondents for long time periods (the latter data are available in a secure data use setting). For the descriptive analysis in this paper, we use data from four waves of HRS, from 2002 through 2008[2]; for the structural analysis we use data from 2002 only.

We restricted the sample to people who were 55 to 64 years of age and whose spouse was also in that age range. The age restriction has both a theoretical and a practical reason. Households in this age group are around the end of the wealth accumulation phase of the life cycle but have not yet started decumulating their wealth. The cross-section of these households allows us to analyze heterogeneity in the results of learning and investment histories. The practical reason for the age restriction is the availability of retrospective

[2]Hudomiet, Kézdi and Willis (2011) show that shortly after the fall of Lehman Brothers in September 2008 stock market beliefs of households changed substantially and in an unusual way. For this reason we decided to drop interviews that were made after September 2008 in this paper.

earnings data from administrative sources, an important variable in the analysis. Sample sizes are in table B1 in the Appendix B.

In 2002, the HRS asked the p_0 and the p_{10} questions, while in 2004 and 2006 only the p_0 questions. In 2008, the p_0 question was accompanied by a second question with eight randomized threshold values ranging from a decrease of 40 per cent or more to and increase of 40 per cent or more. Hudomiet, Kézdi and Willis (2011) use these probability variables from HRS 2008 to look at the effect of the crash of the stock market on households' beliefs. In this paper we use answers to the p_0 questions from all four survey waves and the p_{10} question from 2002.[3]

The 2002 wave of the HRS includes an "experimental module" with additional subjective probability questions about stock market returns. About five per cent of the respondents were randomly assigned to answer the questions in this module. Among others, the module included questions on p_0 and p_{10} once more. Typically, people answered the experimental module about 30 minutes and 60 questions after they answered the original p_0 and p_{10} questions. This small subsample allows for a direct analysis of measurement error in the probability answers, in the spirit of the test-retest reliability studies in the survey measurement literature.

Stockholding is measured at the level of households. In the HRS households are asked whether they had investments in stocks or mutual funds. If "yes," we call people in these households "stockholders outside retirement accounts". The survey asks about retirement accounts as well and the fraction of stocks in those (the latter in a simplified way until 2006). Persons who lived in households in which someone had stocks or mutual fund investments in retirement accounts are labelled "stockholders in retirement accounts." The union of these two sets is labelled "stockholders."

The fraction of stockholders is 51 per cent in 2002. Conditional on stockholding, the share of stocks in portfolios held outside retirement accounts is 59 per cent, and it is 80 per cent on retirement accounts. Stockholding status declines between 2002 and 2008 and so does the fraction of stocks in the portfolio conditional on stockholding. The likelihood of being a stockholder increases in wealth (both total net wealth and financial wealth). Conditional on stockholding, the share of stocks in the portfolio seems unrelated to wealth. Tables B2 and

[3]The varying thresholds for the second probability question in HRS 2008 introduce econometric complications that we do not address in this paper.

B3 and figures B1 through B4 in the Appendix B show the details.

One of the most important variables is a proxy of lifetime earnings. The variable is defined as the cpi-adjusted mean earnings of households with individuals between age 40 and 55 based on the W-2 tax forms. The variable is from confidential data and is not available for part of the sample, which needed imputed values. Other right hand-side variables include standard demographics (age, gender, single or couple, years of education race and ethnicity) and wealth (measured in categories, separately for total net wealth and financial wealth).

Cognitive functioning is measured by the four short tests included in HRS (immediate word recall, delayed word recall, serial 7s (successively subtracting seven from one hundred) and dementia screening questions). We use the first factor of the four aggregate scores for each individual between 1992 and 2000. McArdle, Fisher and Kadlec (2007) argue that the first factor of these tests measures episodic memory.

We use three measures for general optimism/pessimism and one measure for general uncertainty as personal attitudes. Each of these measures is based on survey answers prior to the 2002 wave of the HRS. The first optimism variable is a dummy denoting positive errors in predicting sunny weather. HRS 1994 and 2000 included a "warm-up" question to the series of subjective probability questions about the probability that the day following the interview would be sunny. We obtained realized weather data for the day in question at the zip-code location of the interview, and we regressed the probability answer on sunny hours (their fraction to hours of daylight). The residual of this regression can be interpreted as a forecast error. The variable we use is a dummy indicating whether the respondent's average forecast error was positive on both of the two surveys. The use of the answers to the HRS sunshine question as a measure of optimism was first proposed by Basset and Lumsdaine (1999).

The second optimism variable is the individual's assessment of the likelihood that a major recession would occur the near future. The question was asked in HRS 1992, 1996 and 1998, and the measure we use is the average of those answers. This variable appears in the survey well before the stock market answers and is likely to reflect general pessimism about the economy. The third variable is a score created from the nine-item psychological depression tests administered to the respondents in all waves of the HRS between 1992 and 2000. This test lists symptoms of psychological depression, and we use the score as a measure of time-invariant general pessimism.

The measure for general uncertainty is the fraction of fifty per cent answers to all probability questions (except for the stock market questions) given by the individual in all of the surveys from year 1992 to 2002. The idea behind this measure is that a person's propensity to give 50-50 answers in many different domains indicates uncertainty in general. This variable is very similar to the one used in Hill, Perry and Willis (2005) and Sahm (2007).

The right hand-side variables include a proxy for risk tolerance for HRS respondents estimated by Kimball, Sahm and Shapiro (2008) from answers to hypothetical gambles over lifetime earnings in HRS 1992 to 2002. Using these measures, Sahm (2007) found a significant positive relationship between risk tolerance and stockholding in a larger sample of HRS respondents.

5 Descriptive analysis

Before turning to a more structural analysis, we show results from descriptive statistics and simple linear regressions using the answers to the probability questions. We first document survey noise in the probability answers; then we characterize observed heterogeneity in those answers; finally, we show that the probability answers predict stockholding in ways that are consistent with portfolio choice theory.

The answers to the stock market probability questions contain substantial noise. Tables B2 through B4 through B10 and figure B5 in the Appendix B show the detailed statistics.

95 per cent of the p_0 answers are rounded to ten or 25 or 75 per cent. Focal values at 50 per cent account for an especially large part of all answers. In the American context, the answer "fifty-fifty" to such a probability question may be interpreted as a synonym for "I don't know."[4] At the same time, 50 per cent is a frequent response to probability questions in Europe as well (Hurd, Rohwedder and Winter, 2005). The rounding in p_0 and p_{10} is typical for survey probability answers; see Manski (2004) for examples.

Many respondents give the same answer to p_0 and p_{10} that, taken at face value, would

[4]Beginning in 2006, HRS has asked a follow-up question to respondents who answer the p_0 question with an answer of "50" to distinguish between those who believe that the stock market is equally likely to go up or down in the coming year from those who are "just unsure" about the probability. About two-thirds answer that they are unsure. See Bruine de Bruin and Carman (2011) for a more detailed analysis of the 50 per cent responses.

imply infinitely large standard deviations of log returns. Rounding would allow for finite (but large) standard deviations to give that pattern. Some respondents give $p_0 < p_{10}$, which does not conform the laws of probability. It may be that these respondents do not understand probabilities at all. It is also possible that these answers reflect inattention to one or both questions. Empirical evidence is in line with the latter interpretation.

The most direct evidence on survey noise comes from comparing answers to the p_0 and p_{10} questions in the core questionnaire and the experimental module. When the randomly selected small subset of the respondents were asked to answer the same probability questions once again during the same interview about half hour later, most gave different answers.

Perhaps surprisingly, all three noise features (rounding, apparent violations of the laws of probability, and test-retest noise) seem largely random (see tables B6 through B10 in the appendix). The prevalence of these answer patterns are not related to stockholding or cognitive capacity. There are some weak associations between rounding and education, and the propensity to give the same answer to p_0 and p_{10} and education, lifetime earnings and wealth. Some demographic characteristics are also weakly predictive but no clear pattern emerges. The cross-sectional distribution of the probability answers in the experimental module is very similar to the cross-sectional distribution of the probability answers in the core questionnaire. The absolute difference between the core and module answers is unrelated to any observable variable.

Having established noise in the probability answers, we turn to relevant heterogeneity in them. The goal is to show variation in the probability answers across groups of respondents that, according to our argument, should have had different incentives for learning and thus should have different beliefs.

We focus on four statistics: the sample average of p_0 $(\bar{p}_0)$, the variance of p_0 in the sample $(V(p_{0i}))$, the average difference between p_0 and p_{10} $(\bar{p}_0 - \bar{p}_{10})$ and the fraction of missing p_0 answers. These statistics are computed using waves 2002 through 2008 of HRS, except for $(\bar{p}_0 - \bar{p}_{10})$, which is computed for 2002 only as p_{10} is not available in later years.

$\bar{p}_0$ can be thought of as a proxy for the mean level of stock market beliefs: higher values correspond to more optimistic beliefs, and the closer $\bar{p}_0$ is to 0.68 (or 0.61 for more recent years before 2002) the closer the level of beliefs is to what historical returns would imply. $V(p_{0i})$ is a measure of cross-sectional heterogeneity in expected stock returns, also called disagreement in the finance literature (Hong and Stein, 2007). $(\bar{p}_0 - \bar{p}_{10})$ is an inverse proxy

for perceived risk: the larger the difference the lower risk is attributed to stock returns. The fraction of missing p_0 answers is a proxy for ignorance, which can be thought of as extreme uncertainty about stock returns.

Table 1 shows the descriptive statistics by lifetime earnings, father's occupation, education, cognitive capacity, risk tolerance and stockholding status. Those with higher lifetime earnings, education and cognitive capacity should have beliefs that reflect past learning because of stronger incentives to learn actively, both through its costs and benefits. Defined contribution (DC) pensions create higher incentives for learning than defined benefit (DB) pensions. Those with fathers who were managers or professionals grew up in families that were more likely to be exposed to stockholding or had higher levels of financial knowledge. Father's occupation is, of course, related to lifetime earnings as well, through intergenerational income links. Risk tolerance is also likely to be related to stock market beliefs both through passive learning (higher levels of risk tolerance lead to stockholding at least in case of favorable beliefs) and active learning (by increasing expected benefits). Finally, those who hold stocks towards the end of their active career have stock market beliefs that reflect past learning; either passive learning through earlier stockholding or active learning.

Learning should lead to beliefs that are characterized by levels closer to historical average, lower perceived risk, lower levels of ignorance. In addition, groups whose members learned more should be characterized by lower levels of disagreement. Translated to the proxy variables in Table 1, these would imply $\bar{p}_0$ closer to 0.68, lower $V(p_{0i})$, higher $\bar{p}_0 - \bar{p}_{10}$ (and closer to 0.26) and lower fraction of missing p_0 answers.

Table 1. Descriptive statistics of the subjective probability answers to the stock market returns questions. HRS 2002 through 2008.

	$\bar{p}_0$	$V(p_{0i})$	$\bar{p}_0 - \bar{p}_{10}$	Fraction missing p_0
Top 25 per cent of lifetime earnings	0.56	0.067	0.113	0.03
Bottom 25 per cent of lifetime earnings	0.44	0.079	0.061	0.26
Education college or more	0.56	0.062	0.123	0.06
Education high school or less	0.45	0.074	0.062	0.23
Has DC pension (top 25% lifetime earnings)	0.60	0.059	0.148	0.02
Has DB pension (top 25% lifetime earnings)	0.55	0.064	0.137	0.03
Top 25 per cent of cognitive capacity	0.53	0.063	0.116	011
Bottom 25 per cent of cognitive capacity	0.42	0.082	0.053	0.31
Father was manager or professional	0.55	0.064	0.109	0.10
Father had other occupation	0.50	0.072	0.084	0.15
Top 25 per cent of risk tolerance	0.51	0.070	0.095	0.16
Bottom 25 per cent of risk tolerance	0.45	0.078	0.073	0.16
Stockholder	0.55	0.063	0.107	0.06
Not stockholder	0.45	0.074	0.063	0.24
Entire sample	0.50	0.071	0.086	0.16
Total number of observations	11,259	11,259	3,532	13,408

Sample: Health and Retirement Study, waves 2002, 4, 6 and 8 ($\bar{p}_0 - \bar{p}_{10}$ is from HRS 2002 only).

Respondents of age 55 through 64 with a spouse of the same age range (and singles)

p_0 is the answer to the probability of positive returns on stock markets by following year

The results are all consistent with the predictions of the learning model. Individuals with high lifetime earnings, DC pension plans, high levels of education, high cognitive capacity, high risk tolerance or who grew up in families that were exposed to stockholding (more likely if the father was manager or professional) have beliefs that reflect learning more than the beliefs of their complementary groups (non-stockholders, those with low lifetime earnings, DB pensions, low education, low cognitive capacity, low risk tolerance, non-managerial or non-professional father). Their beliefs are closer to historical probabilities, which also means more optimistic beliefs and lower perceived risk. There is less disagreement about stock returns in these groups, and there is less ignorance measured by the prevalence of missing

answers.

The figures also imply that expectations are low, disagreement is substantial and perceived risks are high. Note however that the probability answers match historical frequencies well in combined groups for whom learning incentives should be the highest. College educated people with top 25 per cent lifetime earnings and DC pension plans have an average p_0 of 0.64, average $p_0 - p_{10}$ at 0.17, zero per cent missing answers and very little disagreement. Nevertheless, low expectations and high perceived risk among the general population, and among non-stockholders in particular, is remarkable.

Table B11 in the Appendix B shows substantial variation of beliefs by the year of interview. Interestingly, average beliefs of stockholders exhibit remarkable stability over the years, and much of the cross-year variation is due to non-stockholders. The same is true for missing probability answers. Conversely, much of the cross-year variation in disagreement comes from stockholders.

OLS regressions reveal partial correlations that are very similar to the simple inter-group differences in Table 1 above (see table B12 in the Appendix B). The belief-specific right hand-side variables predict stock market probabilities in expected ways, too. Sunshine optimism is positively related to the level of p_0 answers, while past beliefs about economic recession and depressive symptoms are negatively related. The propensity to have given fifty-fifty answers in the past is strongly negatively related to the difference between p_0 and p_{10}, indicating strong positive correlation with perceived stock market risk. There are strong differences among demographic groups as well, even after holding the other variables constant. Women, singles and African Americans give probability answers that indicate more pessimistic beliefs, higher perceived risks, and they and Hispanics are more likely to give missing answers.

The probability answers predict stockholding, as documented by OLS regressions in table B13 of the Appendix B. We estimated two separate linear regressions for stockholding, one for the probability of nonzero stock-market based assets in the household portfolio, $\Pr(s_i > 0)$ and one for the share of such assets if nonzero $E[s_i | s_i > 0]$. The two types of regressions allow for looking at the relation of beliefs and stockholding at the extensive margin and the intensive margin separately. This specification does not "handle" selection into stockholding but it is the simplest way to look at the two margins.[5] For each left hand-side variable,

[5] Credible identification of a selection model would require exlcusion restrictions in the second regression, i.e. instruments that affect stockholding at the extensive margin but not the intensive margin. In principle,

we estimated one regression on the entire sample (with the appropriate age restriction) that includes p_{0i} and the dummy for missing p_{0i}, and another one on the HRS 2002 sample that includes $p_{0i} - p_{10i}$.

The results are all consistent with the role of beliefs in portfolio choice. Stockholding is strongly positively related to p_0 answers, negatively related to the propensity to give a missing p_0 answer, and it is positively related to $p_0 - p_{10}$, indicating a negative relation to perceived risk. Conditional on stockholding, the fraction of stock-market based assets in household portfolios are positively related to p_0 answers and $p_0 - p_{10}$, the latter again indicating a negative correlation with perceived risk. These results are strong because they are conditional on lifetime earnings, education, cognitive capacity and demographics. The results at the intensive margin are all the more remarkable because only beliefs and education have significant coefficients. Controlling for detailed measures of household wealth decreases the coefficients by half, but most remain significant. Wealth in this age group is endogenous as it is the result of savings and investment history, and thus these latter results are likely biased downward in magnitude.

The descriptive statistics and the linear regression results are consistent with the hypothesis that stock market beliefs are results of learning over the lifetime and predict stockholding. These results are robust in the sense that they are free from additional econometric assumptions. At the same time, they yield estimates that are hard to interpret, for two reasons. First, the probability answers and their simple transformations may be affected by heterogeneity both in the subjective mean ($\tilde{\mu}$) and the subjective standard deviation ($\tilde{\sigma}$). Second, measurement error is likely to distort the observed probability answers and thus the descriptive statistics derived from them. The next section presents a more structural measurement model that deals with these problems.

6 Structural analysis

We develop a structural measurement model to estimate heterogeneity in the relevant belief variables and handle survey noise. The model relates the latent belief variables $(\tilde{\mu}_i, \tilde{\sigma}_i)$ to

one would need variation in fixed costs that is exogenous to anything that affects investment choices that lead to variation in the fraction of stocks. We argue that fixed costs are mostly related to learning, the results of which naturally affect all investment choices. Valid instruments are thus hard to find in this case.

the observed answers to the probability questions (p_{0i}, p_{10i}). It accommodates the observed noise features in the data and our intuition about the way people answer difficult survey questions.

Our estimation strategy is structural in the sense that it focuses on the theoretically relevant parameters and the relevant heterogeneity in those parameters (net of survey noise). In particular, we estimate the moments of the distribution of $\tilde{\mu}_i$ and $\tilde{\sigma}_i$ in the population and in various subpopulations (analogously to Table 1 above), and we investigate the role of heterogeneity of $\tilde{\mu}_i$ and $\tilde{\sigma}_i$ in heterogeneity of stockholding.

The assumption of i.i.d. normal log returns implies that individual beliefs are fully characterized by $\tilde{\mu}_i$ and $\tilde{\sigma}_i$. As we noted earlier, probability judgments elicited in a survey situation are likely to be very different from those that enter into a real life investment decision. The survey response to a probability question takes less than thirty seconds, on average, and there are practically no incentives to get the answers right.[6] We model the difference in two steps. The first introduces survey noise, and the second step introduces rounding.

Noise is modeled as mean-zero additive components to the index $\tilde{\mu}/\tilde{\sigma}$ that enters the probabilities p_0 and p_{10}. The noise components, denoted by v_0 and v_{10}, are assumed to be jointly normal and potentially correlated. Let p_{0i}^{br} and p_{10i}^{br} denote hypothetical "before rounding" answers so that the observed answers p_{0i} and p_{10i} may be rounded versions of the former. Conditional on a draw of the noise variables, these hypothetical survey answers are then the following:

$$p_{0i}^{br} = \Phi\left(\frac{\tilde{\mu}_i}{\tilde{\sigma}_i} + v_{0i}\right) \tag{1}$$

$$p_{10i}^{br} = \Phi\left(\frac{\tilde{\mu}_i - 0.1}{\tilde{\sigma}_i} + v_{10i}\right) \tag{2}$$

$$\begin{bmatrix} v_{0i} \\ v_{10i} \end{bmatrix} \sim N\left(0, \sigma_v^2 \begin{bmatrix} 1 & \rho_v \\ \rho_v & 1 \end{bmatrix}\right) \tag{3}$$

The noise components are assumed to be independent of any relevant heterogeneity, which is consistent with the randomness of the test-retest error and the near-randomness of the

[6] At the same time, it is important to stress that more than a half century of survey research has shown that data from properly designed and executed sample surveys can be used to make valid inferences to population characteristics despite short response times and lack of incentive to tell the truth (or to lie).

other noise features. The bivariate nature of the noise accommodates answers of $p_{0i} < p_{10i}$ if that phenomenon is due to inattention on the survey (which, as noted earlier, is supported by the near-randomness of its prevalence). The correlation coefficient between v_0 and v_{10} is related to average inattention. $\rho_v = 1$ would mean that all respondents answer questions p_0 and p_{10} with the same noise, which would not allow for answers like $p_{0i} < p_{10i}$. At the other extreme, $\rho_v = 0$ would mean that all respondents forget their previous answers completely. The true value of ρ_v is likely to be in-between.

We identify moments of the noise process (σ_v^2, ρ_v) by making use of answers from the experimental module. Recall that a small subset of the respondents answered the same probability questions once more, in an experimental module. We assume that noise components in the core and module answers are independent.[7] Comparing answers to the core and experimental module questions identifies the noise variance σ_v^2. Conditional on σ_v^2, joint moments of the p_{0i} and p_{10i} answers identify the correlation ρ_v. The Appendix C contains the details of identification and the calibration results. Our preferred estimates are $\sigma_v = 0.95$ and $\rho_v = 0.42$ or 0.61 (the latter depending on whether covariates are entered or not).

Answers to the probability questions may differ from the hypothetical "before-rounding" probabilities p^{br} because of rounding. We accommodate rounding by an interval response model. An answer within a pre-specified interval can correspond to any probability p^{br} within that interval. Round numbers are in the middle of those intervals, which are defined in an exogenous fashion and are assumed to be the same for all respondents.

Formally, the vector of survey answers (p_{0i}, p_{10i}) is in the quadrant Q_{kl} if the vector of hypothetical probabilities p_{ij}^{br} is in that quadrant:

$$\begin{pmatrix} p_{0i} \\ p_{10i} \end{pmatrix} \in \mathbf{Q}_{kl} \Leftrightarrow \begin{pmatrix} \Phi\left(\frac{\tilde{\mu}_i}{\tilde{\sigma}_i} + v_{0i}\right) \\ \Phi\left(\frac{\tilde{\mu}_i - 0.1}{\tilde{\sigma}_i} + v_{10i}\right) \end{pmatrix} \in \mathbf{Q}_{kl} \tag{4}$$

$$\mathbf{Q}_{kl} = \begin{pmatrix} [q_k, q_{k+1}) \\ [q_l, q_{l+1}) \end{pmatrix} \tag{5}$$

In the implemented model, the intervals are defined, in percentage terms, as $[0, 5)$,

[7] This is probably a lower bound to the noise variance, because any "noise" that would be specific to the entire survey situation but would not affect investment decisions (e.g. the experience of a bad day) would affect the "core" and "module" answers in similar ways and would not be measured by the test-retest difference.

$[5, 15)$, $[15, 25)$, ..., $[95, 100]$. These intervals allow for rounding to the nearest ten. The interval response model is the simplest way of accommodating rounding that is compatible with the guesswork of calculating probabilities.

With additional assumptions on their cross-sectional distribution, this model allows for estimating moments of the relevant heterogeneity in $\tilde{\mu}_i$ and $\tilde{\sigma}_i$. We assume that $\tilde{\mu}_i$ is normally distributed and $\tilde{\sigma}_i$ follows a two-point distribution. We estimate the conditional mean of the normal distribution and the probability of the low point conditional on right hand-side variables. In the benchmark model we estimate the variance of the normal distribution and the high value of $\tilde{\sigma}$ unconditionally, and we set the low value of $\tilde{\sigma}$ to the historical standard deviation of log returns. As a robustness check, we explore models in which the low value of $\tilde{\sigma}$ is estimated as well.

$$\tilde{\mu}_i = \alpha + \beta'_\mu x_i + \gamma'_\mu z_{\mu i} + u_{\mu i}, \quad u_{\mu i} \sim N\left(0, \sigma^2_{u\mu}\right) \tag{6}$$

$$\tilde{\sigma}_i \in \{\tilde{\sigma}_{low}, \tilde{\sigma}_{high}\} \tag{7}$$

$$\Pr\left(\tilde{\sigma}_i = \tilde{\sigma}_{low}\right) = \Phi\left(\beta'_\sigma x_i + \gamma'_\sigma z_{\sigma i}\right) \tag{8}$$

In the equations, z_μ is the vector of optimism variables (positive error in forecasting sunshine, low expectations about the economy in the past, symptoms of clinical depression), z_σ is the proxy variable of person-specific uncertainty (the fraction of fifty-fifty answers to all probability answers in the past), and x_i denotes the vector of all other right hand-side variables.[8] α stands for the constant in the equation for $\tilde{\mu}$. We estimated it as a vector, by allowing for a different constant for $\tilde{\sigma}_i = \tilde{\sigma}_{low}$ versus $\tilde{\sigma}_i = \tilde{\sigma}_{high}$, which allows for a correlation between $\tilde{\mu}_i$ and $\tilde{\sigma}_i$.

We estimated the model by Maximum Likelihood. The details of the likelihood function are in the Appendix C. We estimated the model with and without the right hand-side variables. The details of the parameter estimates are in table D1 of the Appendix D.[9]

[8] Excluding z_σ from the equation of $\tilde{\mu}$ and excluding z_μ from the equation of $\tilde{\sigma}$ are motivated by the fact that they do not influence the belief proxies in the simple linear regressions (Appendix B, table B12). These exclusions may in principle be important in identifying the association of $\tilde{\mu}_i$ and $\tilde{\sigma}_i$ on the one hand and stockholding on the other hand in the stockholding equations below. However, the inclusion of z_σ and z_μ in all equations (including the stockholding equations) does not change any of the results, see the additional estimates presented in the Online Appendix D, tables D9 through D11.

[9] There we show additional results for the restricted sample of financial respondents and for a more

Table 2 shows the estimates of the most important unconditional moments of the relevant heterogeneity in stock market beliefs.

Table 2. Relevant heterogeneity in stock market beliefs. Estimates from the structural model

	Model w/o covariates		Model with covariates	
	Point estimate	SE*	Point estimate	SE*
Population average of $\tilde{\mu}$	-0.066	0.018	-0.050	0.015
Population standard deviation of $\tilde{\mu}$	0.197	0.019	0.218	0.027
Population average of $\tilde{\sigma}$	0.576	0.077	0.532	0.091

*Bootstrap standard errors

Sample: HRS 2002, 55 to 64 years old financial respondents (partner is also 55 to 64)

The results imply low expectations and high perceived risks on average and substantial heterogeneity in expectations. The population moment estimates are similar whether covariates are used or not used in the estimation, but the model with covariates indicates somewhat less pessimistic expectations and lower perceived risk. According to the estimates with covariates, the population average of $\tilde{\mu}_i$ is negative 5 per cent. The population standard deviation of $\tilde{\mu}$ is 22 per cent, indicating that over 40 per cent of the population has positive expectations, and almost 30 per cent have expectations at or above the historical average of 0.07. The population average of the perceived standard deviation is over 50 per cent, to be compared to the historical standard deviation of 15 per cent.

Perhaps even more interesting are the moments conditional on the covariates. Instead of interpreting the estimated coefficients of the model, we show estimated moments of $\tilde{\mu}_i$ and $\tilde{\sigma}_i$ in various groups.

The moments within subgroups are based on predicted values of $\tilde{\mu}_i$ and $\tilde{\sigma}_i$, which are conditional on observed survey answers p_{0i} and p_{10i} and the other covariates in the estimation model. The predicted subjective mean and subjective standard deviation are denoted by $\hat{\mu}_i$ and $\hat{\sigma}_i$, and they come from the following conditional expectations (where hats on the expectation operator mean estimates):

flexible way of estimating heterogeneity in $\tilde{\sigma}_i$. Those results are qualitatively very similar to our benchmark esitmates, except that heterogeneity in $\tilde{\sigma}_i$ is less successfully pinned down in some cases.

$$\hat{\mu}_i = \widehat{\mathbb{E}}\left[\tilde{\mu}_i | x_i, z_{\mu i}, (p_{0i}, p_{10i}) \in \mathbf{Q}_{kl}\right] \quad (9)$$

$$\hat{\sigma}_i = \widehat{\mathbb{E}}\left[\tilde{\sigma}_i | x_i, z_{\sigma i}, (p_{0i}, p_{10i}) \in \mathbf{Q}_{kl}\right] \quad (10)$$

The conditional expectations are relatively straightforward to compute by Bayes' Rule after the results of the structural model that specifies the full distributions for $\tilde{\mu}_i$ and $\tilde{\sigma}_i$. The predicted $\hat{\mu}_i$ and $\hat{\sigma}_i$ are then the sample analogues to those. The details of the derivation are in the Appendix C.

This prediction method is analogous to the prediction of risk tolerance based on survey answers to hypothetical gambles by Kimball, Sahm and Shapiro (2008). The predicted values are different from the true values, creating measurement error in the variables. The measurement error is one of prediction error. It has zero mean and thus leads to an unbiased estimate of the population mean, but it leads to an underestimation of the population standard deviation (because the predicted values are less dispersed than the true values). Using $\hat{\mu}_i$ and $\hat{\sigma}_i$ on the right-hand side of a regression leads to consistent estimates as long as all the covariates that are used in the predictions are also entered in the regression in question. The standard errors in this regression are inconsistent, though, and thus bootstrap standard errors are advised. If one uses the $\hat{\mu}_i$ and $\hat{\sigma}_i$ in regressions that have different covariates from the ones used in the prediction equations, OLS is inconsistent and a more sophisticated GMM procedure is appropriate (see Kimball, Sahm and Shapiro, 2008, for more details).

Table 3 shows the average predicted belief variables in various groups in the sample. The table is analogous to table 1 in the descriptive analysis, except that table 3 does not have a measure of disagreement (because that is not estimated well by the dispersion of $\hat{\mu}_i$), and it does not repeat the fraction of missing answers.

Table 3. Estimated mean of the structural parameters of stock market beliefs in various subpopulations. HRS 2002

	Average $\hat{\mu}_i$	Average $\hat{\sigma}_i$
Top 25 per cent of lifetime earnings	0.065	0.542
Bottom 25 per cent of lifetime earnings	−0.070	0.540
Education college or more	0.041	0.539
Education high school or less	−0.094	0.536
Has DC pension (top 25% lifetime earnings)	0.072	0.536
Has DB pension (top 25% lifetime earnings)	0.051	0.573
Top 25 per cent of cognitive capacity	0.023	0.551
Bottom 25 per cent of cognitive capacity	−0.138	0.517
Father was manager or professional	0.030	0.519
Father had other occupation	−0.049	0.550
Top 25 per cent of risk tolerance	0.008	0.515
Bottom 25 per cent of risk tolerance	−0.141	0.569
Financial respondent in couple	0.031	0.512
Non-financial respondent in couple	−0.051	0.558
Entire sample	−0.041	0.539
Total number of observations	3,314	3,314

Sample: Health and Retirement Study, wave 2002. Respondents of age 55 through 64 (partner also 55-64)

$\hat{\mu}_i$ and $\hat{\sigma}_i$: subjective mean and subjective standard deviation of the one-year ahead stock return, predicted value

The inter-group differences in average expectations are large. Those with top 25 per cent lifetime earnings or top 25 per cent cognitive capacity believe that expected stock returns are 15 percentage points higher than those with bottom 25 per cent earnings or cognitive capacity. College educated respondents believe that expected returns are 13 percentage points higher than those with high school education or less. Average $\hat{\mu}$ is still below the historical average of 0.07 in these categories, but in combined categories it exceeds that (it is 0.11 for college educated respondents with top 25 per cent lifetime earnings and a DC pension plan). The differences by pension plan, the father's occupation, risk tolerance and financial respondent status are sometimes smaller but still substantial. These differences are all in line with the predictions of the learning model. On average, individuals who had

higher incentives to learn believe that expected stock market returns are positive and closer to historical evidence than other individuals.

The estimated intergroup differences in perceived risk are smaller and do not always have the expected direction, because our model is less successful in capturing heterogeneity in $\tilde{\sigma}$. Nevertheless, in most of the cases when the difference is substantial, the results are in line with the predictions of the learning model: those who had higher incentives to learn believe that risks are lower and closer to historical evidence.

Finally, we examine the association of predicted individual beliefs with stockholding. Stockholding is specified as a two-tier hurdle model. The extensive margin (whether a household holds any stock-market based assets at all) is a probit, and the intensive margin (how much it holds if it holds any) is a truncated regression. Subjective beliefs are entered in the right-hand side of these equations in the form of their predicted values $\hat{\mu}_i$ and $\hat{\sigma}_i$, in additive ways.

$$\Pr(s_i = 0) = \Phi(\alpha_1 \hat{\mu}_i + \beta_1 \hat{\sigma}_i + \delta_1' x_i) \quad (11)$$

$$s_i^* = \alpha_2 \hat{\mu}_i + \beta_2 \hat{\sigma}_i + \delta_2' x_i + u_{2i} \quad (12)$$

We estimated both models in two ways, first without the stock market belief estimates and second with them. Apart from the belief-specific variables, the same right hand-side variables are included in the structural model as in the stockholding equations. As a result the coefficients of the belief variables are consistently estimated (see our discussion above and also in Kimball, Sahm and Shapiro, 2008). Because of inconsistency of the analytical standard errors, we present bootstrap standard errors that are re-sampled at the level of households (in order to allow for within-household correlations that are obviously strong because of the common left hand-side variable).

Besides the coefficients of the belief variables, it is also interesting to see whether and to what extent coefficients of the other variables change with the inclusion of the belief variables. Table 4 shows the most important results from the stockholding equation (11). Table D2 in the Appendix D contains all estimates.

Expected stock market returns ($\hat{\mu}_i$) are strongly predictive of the probability of stockholding and the share of stocks in household portfolios. Individuals who believe that stock

market returns are higher by one percentage point live in households that are 0.7 percentage points more likely to own stocks, and if they own stocks, the share of stocks among their financial assets is 0.3 percentage points higher. The estimated correlation of perceived risk, $\hat{\sigma}$, is not significant in the equations.

Other right hand-side variables have strong associations with the probability of stockholding, and most are of the expected sign. They are, however, at most weakly predictive of the share of stocks in household portfolios, which makes the strong predictive power of expected returns even more remarkable.

Inclusion of the stock market beliefs decreases the association of stockholding and the other right hand-side variables. The association with education and cognitive capacity are cut by a third. Single men and women are significantly less likely to hold stocks than couples in the reduced form but not if we condition on their stock market beliefs. Females in couples are of course just as likely to hold stocks as the reference group of coupled men because stockholding is defined at the household level. Their beliefs are, however, a lot less optimistic, and that's why, conditional on their beliefs, they should be more likely to hold stocks according to the second model. African Americans are 23 percentage points less likely to hold stocks in the first model (conditional on all the other right hand-side variables), and the difference drops to 18 percentage points if beliefs are also controlled. The difference between Hispanics and non-Hispanics are 23 percentage points (again conditional on the other right hand-side variables), and is unaffected by the inclusion of beliefs.[10]

Overall, our findings suggest that those people who should learn about returns in the stock market do learn and, given their beliefs, those people who should invest do invest. It is important to emphasize that the estimated coefficients in Table 4 do not capture the causal effect of beliefs about stock returns on stock holding. Rather, as our theoretical model emphasizes, beliefs are the product of a process of learning (or failure to learn) that takes place over the life cycle with many feedback loops between observations of market returns, evolution of earnings and wealth and investments in learning.

[10]For robustness checks, we re-estimated all models with the financial respondents only, as well as with alternative specifications for the heterogeneity in $\tilde{\sigma}_i$. The results, shown in the appendix tables D3 through D10 in the Appendix D, are very similar to those presented above.

Table 4. Subjective stock market beliefs and stockholding at the extensive margin.

	$\Pr(s_i > 0)$, partial effects		$E(s_i \mid s_i > 0)$	
	(1)	(2)	(3)	(4)
$\hat{\mu}_i$		0.734		0.302
		(0.088)**		(0.110)**
$\hat{\sigma}_i$		−0.050		0.144
		(0.108)		(0.131)
Log lifetime earnings	0.041	0.034	0.000	−0.003
	(0.010)**	(0.010)**	(0.012)	(0.012)
Education	0.033	0.023	0.011	0.008
	(0.004)**	(0.004)**	(0.05)**	(0.005)
Cognitive capacity	0.069	0.044	0.002	−0.009
	(0.012)**	(0.012)**	(0.015)	(0.015)
Log risk tolerance	0.023	−0.039	0.048	0.030
	(0.028)	(0.029)	(0.028)*	(0.029)
Single female	−0.110	0.021	−0.023	0.009
	(0.023)**	(0.032)	(0.029)	(0.037)
Single male	−0.094	−0.031	−0.022	−0.005
	(0.029)**	(0.031)	(0.037)	(0.038)
Female in couple	−0.003	0.087	0.010	0.029
	(0.016)	(0.024)**	(0.017)	(0.025)
African American	−0.233	−0.181	0.034	0.057
	(0.030)**	(0.030)**	(0.046)	(0.053)
Hispanic	−0.229	−0.225	−0.008	−0.004
	(0.044)**	(0.043)**	(0.066)	(0.064)
Other variables	YES	YES	YES	YES

Probit models (1) and (2); truncated regression models (3) and (4).

Clustered standard errors in parentheses; bootstrapped for models (2).and (4)

** significant at 1%; * significant at 5%

Sample: HRS 2002, 55 to 64 years old financial respondents (partner is also 55 to 64)

7 Conclusions

Using survey data on subjective probabilities and a rich set of personal characteristics, this paper estimates heterogeneity in stock market beliefs and proposes an explanation for the source of that heterogeneity. We show descriptive evidence and develop a structural measurement model to capture the theoretically important belief parameters and separate survey noise from relevant heterogeneity. We provide detailed evidence on survey noise and the measurement model accommodates all the noise features we document. The results are consistent with our proposed explanation for heterogeneity in stock market beliefs. They also reinforce previous results about the predictive power of beliefs on stockholding.

Our results establish the importance of belief heterogeneity in household finances. They show that survey answers to probability questions can be helpful in characterizing individual beliefs, but their analysis should recognize the importance of survey noise. Our econometric model is a simple but sensible attempt to deal with measurement error that may be a useful reference for further research in this direction.

Our structural estimation results on the subjective mean of stock returns are relatively strong, while our results on the subjective standard deviation are weaker. We explored different models with different assumptions about the form of heterogeneity in the subjective standard deviation, and the results were always qualitatively similar. It is possible that answers to more probability questions or probability questions that are defined for more distant horizons would result in stronger identification in the presence of substantial survey noise.

So, what can we learn from these results? First, the HRS survey data is consistent with a model in which beliefs about the stock market depend on financial knowledge and the acquisition of financial knowledge is costly. Although our results emphasize the importance of beliefs, on a cautionary note, they also suggest that the strong correlation between beliefs and stock market participation in the HRS and other surveys cannot be interpreted as a causal relationship. Second, our results in some ways support the recent emphasis on finding ways to improve financial literacy as potentially useful policy to help people prepare for retirement. It would be useful to know more than we do about the mechanisms by which people acquire financial knowledge. Our model suggests that feedback effects through learning by doing may have large cumulative effects in the long run. Thus, policies that

encourage participation in stockholding at a small scale early in the life cycle may motivate people both to improve their knowledge of risks and returns and to increase their level of saving.

References

[1] Ameriks, John, and Stephen P. Zeldes (2004), "How Do Household Portfolio Shares Vary With Age? " Working Paper, Columbia Business School.

[2] Amromin, Gene and Steven A. Sharpe (2006), "From the horse's mouth: Gauging conditional expected stock returns from investor surveys." Working paper, Federal Reserve Bank of Chicago

[3] Bassett, William F. and Robin L. Lumsdaine (1999), "Outlook, Outcomes and Optimism." *Unpublished manuscript.*

[4] Becker, Gary S. (1964, 1993, 3rd ed.), *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education.* Chicago, University of Chicago Press

[5] Bruine de Bruin, Wandi and C. G. Carman. (2011), "Measuring risk perceptions: What does the excessive use of 50% mean?" *Medical Decision Making*, forthcoming

[6] Campbell, John Y. (2006), "Household Finance." *The Journal of Finance* 61(4), 1553-1604.

[7] Card, David (1999), "The causal effect of education on earnings," in: O. Ashenfelter & D. Card (ed.), *Handbook of Labor Economics*, edition 1, volume 3, chapter 30, pages 1801-1863 Elsevier

[8] Christelis, Dimitrios, Tullio Jappelli and Mario Padula (2010), "Cognitive Abilities and Portfolio Choice." *European Economic Review*, 54(1), 18-38.

[9] Delavande A., S. Rohwedder and R. J. Willis (2008), "Preparation for Retirement, Financial Literacy and Cognitive Resources," Michigan Retirement Research Center Working Paper No. 2008-190.

[10] Dominitz, J. and C. Manski (2005), "Measuring and Interpreting Expectations of Equity Returns" *NBER Working Paper*, 11313

[11] Dominitz, J. and C. Manski (2007), "Expected Equity Returns and Portfolio Choice: Evidence from the Health and Retirement Study" *Journal of the European Economic Association*, 5(2-3), 369-379.

[12] Guiso, Luigi, Michalis Haliassos, Tullio Jappelli (2002), *Household Portfolios.* Cambridge: MIT Press.

[13] Guiso, Luigi, Paola Sapienza, and Luigi Zingales (2004), "The Role of Social Capital in Financial Development," *American Economic Review* 94, 526-556.

[14] Hill, Daniel, M. Perry and R. J. Willis (2005), "Estimating Knightian Uncertainty from Probability Questions on the HRS." *Presented at World Congress of the Econometric Society*, London, August 19-24, 2005.

[15] Hong, Harrison, Jeffrey D. Kubik, and Jeremy C. Stein (2004), "Social Interaction and Stock Market Participation," *Journal of Finance,* 59, 137-63.

[16] Hong, Harrison and Jeremy C. Stein (2007), "Disagreement and the Stock Market." *Journal of Economic Perspectives*, 21(2), 109-128.

[17] Hudomiet, P, G. Kezdi and R. J. Willis (2011), "Stock Market Crash and Expectations of American Households", *Journal of Applied Econometrics*, Vol. 26, No. 3, 2011, pp. 393-415.

[18] Hurd, M. and K. McGarry (1995), "Evaluation of the Subjective Probabilities of Survival in the Health and Retirement Study," *Journal of Human Resources*, 30, S268-S292.

[19] Hurd, M., S. Rohwedder and J. Winter (2005), "Subjective Probabilities of Survival: An International Comparison," RAND manuscript.

[20] Hurd, M, M. van Rooij, and J. Winter (2008) "Stock market expectations of Dutch households." *Journal of Applied Econometrics*, Vol. 26, No. 3, 2011, pp. 416-436.

[21] Inst. Econ. Res., Univ. Munich

[22] Juster, T. and R. Suzman (1995), "An Overview of the Health and Retirement Study," *Journal of Human Resources*, 30, S7-S56.

[23] Kimball, M. S., C. R. Sahm and M. D. Shapiro (2008), "Imputing Risk Tolerance from Survey Responses." *Journal of the American Statistical Association* 103 (Sept 2008) 1028-1038

[24] Manski, C. (2004), "Measuring Expectations," *Econometrica*, 72, 1329-1376.

[25] McArdle, John J., Fisher, Gwenith G., Kadlec, Kelly M. (2007) "Latent Variable Analyses of Age Trends of Cognition in the Health and Retirement Study, 1992-2004" *Psychology and Aging*. 22:3 p.525-545

[26] National Institute of Aging. (2007), *Growing Old in America. The Health and Retirement Study*. NIH Publications 07- 5757, Washington D.C.

[27] Rosen, Harvey S. and Stephen Wu (2004), "Portfolio Choice and Health Status," *Journal of Financial Economics,* 72, 457-84.

[28] Sahm, Claudia R. (2007), "Stability of Risk Preference." Mimeo, University of Michigan.

[29] Veldkamp, Laura (2011), *Information Choice in Macroeconomics and Finance*. Princeton University Press

[30] Viceira, L.M. (2001), "Optimal portfolio choice for long-horizon investors with nontradeable labor income," *Journal of Finance*, 56(2), 433-470.

[31] Vissing-Jorgensen, Annette (2004), "Perspectives on Behavioral Finance: Does Irrationality Disappear with Wealth? Evidence from Expectations and Actions," in: M. Gertler and K. Rogoff eds., *The NBER Macroeconomics Annual 2003*, Cambridge: MIT Press.

[32] Willis, Robert J. (1986) "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions" In O. Ashenfelter and R. Layard, eds. *Handbook of Labor Economics*, (Amsterdam: North Holland, 1987): 525 602.

Appendices

A Details of the theoretical model

In this Appendix we present and solve a simple three-period life-cycle model of consumption and saving with risky assets, heterogeneous beliefs about the parameters of the distribution of returns, and potential learning about those parameters. The model is built on the "small scale" model of Haliassos and Michaels (2002), and we add to it elements that are connected to ideas in the human capital literature (e.g., Becker, 1964), its application to the acquisition of financial knowledge (Delavande, Rohwedder and Willis, 2008), and the theory of information choice (Veldkamp, 2011).

Consider individual i who lives for three periods. Period 1 contains her young active years (e.g., age 20 through 40), period 2 her active years in mature age (e.g., age 40 through 60), and in period 3 she is retired. In periods $t = 1, 2$ she receives labor income Y_{it}. In period 3 she receives pension benefits that are a function of her previous earnings $Y_{i3} = \pi(Y_{i1}, Y_{i2})$. We abstract away from taxes and non-labor income other than pensions. Importantly, pensions are from a defined-benefits-type system such as Social Security, and pension benefits are a concave function of lifetime earnings.[11] As a result, people who earn above certain threshold have an incentive to save for retirement, and the saving rate may depend on lifetime earnings. For simplicity, we assume that there is no uncertainty in earnings, the retirement age, pension benefits or the length of life.

In each period, individual i can save. Savings can be invested in bank accounts (B_{it}) that yield a fixed gross interest R_f or in equity (S_{it}) that yield stochastic potential return $R_t \sim iid \log N(\mu, \sigma)$. The individual cannot borrow or short the risky asset so $B_{it} \geq 0$ and $S_{it} \geq 0$. Risky returns R_t are defined as potential returns in the sense that the actual returns individual i can earn come at a discount of τ so that effective returns are $R_t^e = R_t e^{-\tau} \sim \log N(\mu - \tau, \sigma)$. The idea here is that R_t denotes the yearly gain on an ideal portfolio of risky assets. In this paper we assume that the ideal portfolio is the stock market index fund, and therefore realizations of R_t are the realized returns on the stock market index. We assume

[11] The Social Security benefit formula is very concave indeed. It starts with defining average monthly earnings from the lifetime earnings history, in which months without earnings count as zero. Benefits are 90 per cent of that average up to a relatively low level of earnnings; earnings in the middle range are transformed into benefits by a 32 per cent factor; and a factor of 15 per cent is used for high levels of earnings.

that individuals earn less than the return on such an ideal portfolio because of proportional transaction costs (therefore the notation τ) and sub-optimal choice of underlying assets.

Importantly, we assume that the survey questions in HRS (p_0 and p_{10}) ask about potential returns R_t, but individual investment decisions are based on effective returns R_t^e. R_t is a random variable but the factor τ is not. Realizations of potential returns R_t are common across all individuals. However, individuals may differ in their beliefs about the parameters of the distribution of R_t (but they all think it is i.i.d. lognormal). Individual beliefs about the parameters are denoted by tilde over the greek letters denoting true parameters, e.g., $\tilde{\mu}$.

In the beginning of period 1, individual i is endowed with a set of beliefs about the parameters of the distribution of potential returns R_t. Beliefs about parameter σ^2 are assumed to be the same for everybody (this setup is the same as the one used by Brennan, 1998). At the same time, there is uncertainty about μ with heterogeneous beliefs. Individuals have some belief $\tilde{\mu}$ but they know that they don't know the true μ. We refer to incomplete knowledge about μ as uncertainty and model it by a prior distribution of μ, which is normal, centered around $\tilde{\mu}_{i1}$, and its variance $\tilde{\sigma}^2_{\mu i1}$ is potentially heterogenous, too.

In this setup, the distribution of log gross potential returns is perceived as normal with mean $\tilde{\mu}_{i1} = \tilde{E}_{it}[\mu]$ (the individual-specific mean of the random variable μ) and variance $\tilde{\sigma}^2_{i1} = \tilde{\sigma}^2_{\mu i1} + \sigma^2$ (the reduced-form variance is the sum of variance due to individual-specific parameter uncertainty and fixed variance due to risk). When individual i makes the portfolio choice decision in period 1, she thinks that risky returns follow a lognormal distribution with parameters $\left(\tilde{\mu}_{i1}, \tilde{\sigma}^2_{i1}\right)$. Heterogeneity in period 1 beliefs is predetermined by differences in what people may learn at home or in school, or differences in personality (degree of general optimism and general uncertainty). If they do not learn more about the returns, individuals enter period 2 with the same beliefs: $\tilde{\mu}_{i2} = \tilde{\mu}_{i1}$ and $\tilde{\sigma}^2_{i2} = \tilde{\sigma}^2_{i1}$. However, their beliefs can change as results of two kinds of learning.

The first kind is mechanical learning, or passive learning following the terminology of Veldkamp (2011). If an individual invests in S_{i1}, the realized returns will make her change her beliefs by Bayesian updating. Since the length of period 1 is unity, and the realized

returns are R_1 for everyone, the results of passive learning are the Bayesian posteriors

$$\tilde{\mu}_{i2} = \frac{\sigma^2 \tilde{\mu}_{1i} + \tilde{\sigma}^2_{\mu i1} R_1}{\sigma^2 + \tilde{\sigma}^2_{\mu i1}} \tag{13}$$

$$\tilde{\sigma}^2_{\mu i2} = \left(\frac{1}{\tilde{\sigma}^2_{\mu i1}} + \frac{1}{\sigma^2} \right)^{-1} = \frac{\sigma^2 \tilde{\sigma}^2_{\mu i1}}{\sigma^2 + \tilde{\sigma}^2_{\mu i1}} \tag{14}$$

$$\tilde{\sigma}^2_{i2} = \tilde{\sigma}^2_{\mu i2} + \sigma^2 \tag{15}$$

As a result of passive learning, individuals update their $\tilde{\mu}$ in the direction of the realized stock market returns in period 1, and their uncertainty decreases.

The second kind of learning is active learning (again, following the terminology of Veldkamp, 2011). Individuals can invest in learning even if they do not invest in period 1. Also, those who are investors in period 1 may learn more than simply observing the returns they realize. Active learning is an investment in one's financial knowledge, which is a form of human capital. Many insights of the large literature on investment into human capital may apply to active learning (see, e.g., Becker, 1964).

Active learning affects attainable returns in two ways. The first is Bayesian updating of one's beliefs about μ and σ. The investor can update her beliefs by observing a history of past returns, where the length of the history is h_i. While h_i should be a decision variable in general, we abstract away from that in this simple model and fix it to h. We set $h > 1$ in order to reflect a longer horizon than available in mechanical learning. With history length h and observed average stock market returns $\bar{R}_h$ the result of active learning is the Bayesian posterior distribution

$$\tilde{\mu}_{i2} = \frac{\sigma^2 \tilde{\mu}_{1i} + h \tilde{\sigma}^2_{\mu i1} \bar{R}_h}{\sigma^2 + h \tilde{\sigma}^2_{\mu i1}} \tag{16}$$

$$\tilde{\sigma}^2_{\mu i2} = \left(\frac{1}{\tilde{\sigma}^2_{\mu i1}} + \frac{h}{\sigma^2} \right)^{-1} = \frac{\sigma^2 \tilde{\sigma}^2_{\mu i1}}{\sigma^2 + h \tilde{\sigma}^2_{\mu i1}} \tag{17}$$

$$\tilde{\sigma}^2_{i2} = \tilde{\sigma}^2_{\mu i2} + \sigma^2 \tag{18}$$

Similarly to passive learning, individuals update their $\tilde{\mu}$ in the direction of the realized returns in the observed time horizon, and their uncertainty decreases. Those are ex post results of learning. Active learning is a choice based on results that are expected ex ante. Ex ante, individuals do not know in which direction their $\tilde{\mu}$ will be updated. In particular, they do not expect their mean to change after learning. But they know that learning will

decrease their uncertainty.[12]

The second aspect of learning affects individual transaction costs τ that discount potential returns. Recall that although potential returns are R_t, individuals can attain returns of $R_t e^{-\tau}$ on their investment S_{it}. By active learning, we assume that individuals can decrease their transaction cost τ. For simplicity, we assume that active learning leads to $\tau = 0$ so that active learners can expect to realize (and do realize) R_t on their investment S_{it}.

Active learning is an investment. We assume that its two aspects are bundled so that those who choose to learn will see their beliefs updated as in (16) through (18) and their transaction costs τ reduced (to zero in this simple setup). Active learning entails individual-specific costs of D_i that are to be paid in period 1. Note a key aspect of this investment setup: while the benefits to active learning are related to the amount to invest into the risky assets, the costs are not. This aspect will drive many of our most important implications.

Combining all the ingredients outlined above, the decision problem of the individual can be formulated the following way.

$$
\begin{aligned}
EU_i &= \sum_{t=1}^{3} \beta^{t-1} Eu\left(C_{it}\right) && (19)\\
u\left(C_{it}\right) &= \frac{1}{1-\gamma} C_{it}^{1-\gamma} && (20)\\
X_{it} &\geq C_{it} + B_{it} + S_{it} + f_{it} + D_{it} && (21)\\
f_{it} &= f \times 1\left(S_{it} > 0\right) && (22)\\
D_{i1} &= D \text{ if active learning in period 1} && (23)\\
D_{it} &= 0 \text{ otherwise} && (24)\\
X_{it} &= S_{i(t-1)}\left(R_t e^{-\tau_{it}} - R_f\right) + \left(B_{i(t-1)} + S_{i(t-1)}\right) R_f + Y_{it} && (25)\\
B_{it} &\geq 0, \quad S_{it} \geq 0 && (26)
\end{aligned}
$$

The utility function in (19) is standard expected utility, C_{it} is consumption, β is the discount factor. The instantaneous utility function is CRRA, and γ is the parameter for risk aversion and the inverse of the intertemporal elasticity of substitution at the same time. The budget line in (21) states that the sum of investments B_{it} and S_{it} (bonds and stocks,

[12] In this setup, the decrease in uncertainty is a deterministic function of h because of the simplistic assumption of known σ^2. But uncertainty decreases in h in richer setups as well as long as the observed returns are from a stationary distribution.

respectively), the fixed costs of investment (f_{it}), and the cost of active learning (D_{it}) cannot exceed cash on hand (X_{it}). Fixed costs need to be paid if one invests in the risky assets, and their role is to prevent very small investments. D needs to be paid if one invests in active learning. In our setup the only time people may invest into active learning is period 1 (no one wants to save in period 3, and thus it is never optimal to learn later than period 1). Equation (25) describes the equation of motion for cash on hand. In the beginning of every period t earnings (Y_{it}) are received, and the returns on previous period $(t-1)$ investments are collected. In case of stocks, these are net returns that include proportional transaction costs τ. Equation (26) states the nonnegativity constraints that make borrowing and short selling impossible.

This model is relatively simple, but it does not yield to analytical solutions. In order to get the implications for our empirical investigation, we simulated out the policy function. The model can be solved with backward induction. In the third period the optimal behavior is trivial. There is only one state variable, X_{i3} (cash-on-hand in period 3) and one control variable C_{i3} (consumption). The optimal policy is to consume everything, $C_{i3} = X_{i3}$. The second period is more complex. There are four state variables: X_{i2}, D_{i1} (whether the individual had active learning in period 1) and the belief parameters $\tilde{\mu}_{i2}$ and $\tilde{\sigma}^2_{i2}$. The two control variables are B_{i2} and S_{i2} which then imply C_{i2}. The optimal second period policy function and the implied value function can be computed by simulation. We computed the optimal decision for a large number of grid points on the state variables and then we used cubic splines to approximate the functions for their entire domains. In the first period there is no state variable, but there are three control variables: B_{i1}, S_{i2} and D_{i1}.

We solved the model for a large number of different parameter values. Some parameters values were borrowed from the literature such as:

$$\begin{aligned} \gamma &= 3 \\ \beta &= 0.97 \\ R_f &= e^{0.02} \\ \mu &= 0.07 \\ \sigma &= 0.15 \end{aligned}$$

The second set of variables are the wage variables. The heterogeneity of lifetime income and its link to learning and investment is our primary focus so we computed the optimal

policy function for a large set of wage values. We generated a distribution of earnings that resembles the observed distribution. As a benchmark, we set the ratio of Y_{i2} to Y_{i1} to 2 (so that $Y_{i2} = 2Y_{i1}$). The distribution of earnings is set to lognormal in both the first and the second period (or generation). We have set the 5^{th} percentile of the Y_{i1} to be 0.4 and the 95^{th} percentile of the Y_{i1} to be 2.2. This way the population average of Y_{i1} is normalized to roughly 1.

For the third period income we used a simplified social security formula. First we computed the average indexed monthly earnings (AIME) as the average of the prior wages, $PIA_i = 0.5\,(Y_{i2} + Y_{i1})$. Then we have chosen two bendpoints, Q_1 and Q_2. The third period social security income was defined as

$$Y_{i3} = \begin{cases} 0.9PIA_i & \text{if} \quad PIA_i \leq Q_1 \\ 0.9Q_1 + 0.32\,(PIA_i - Q_1) & \text{if} \quad Q_1 < PIA_i \leq Q_2 \\ 0.9Q_1 + 0.32\,(Q_2 - Q_1) + 0.15\,(PIA_i - Q_3) & \text{if} \quad Q_2 < PIA_i \end{cases}$$

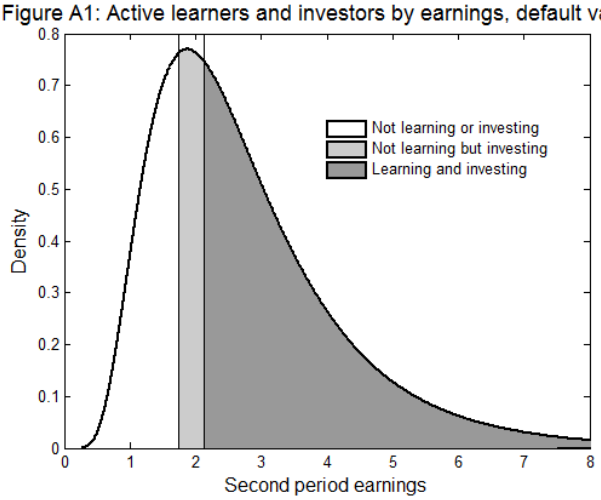
The bendpoints were chosen to be $Q_1 = 1.25$ (approximately the 40th percentile) and $Q_2 = 2.65$ (approximately the 90th percentile).

The rest of the parameters, due to lack of consensus about their values, had to be calibrated differently. We have chosen basic values that made the results interesting, and we have run sensitivity analyses to see how the results change as we move away from these values. The default values of these parameters were:

$$\begin{aligned} \tau &= 0.025 \\ f &= 0.06 \\ D &= 0.02 \\ \tilde{\sigma}^2_{\mu 1} &= 0.15 \\ h &= 3 \end{aligned}$$

Perhaps the most important but also a rather straightforward result of the model is that an increase in lifetime earnings leads to an increased propensity to learn and to invest. In the setup here, the only source of heterogeneity is in earnings. Figure A1 illustrates the results using our default parameter values. There is a first threshold value of second period earnings (~ 1.75) below which nobody learns and nobody participates on the stock market. Between this and a second threshold value (~ 2.13) people participate on the stock market but they do

Figure A1: Active learners and investors by earnings, default values



not acquire financial knowledge. People whose second period earnings are above the second threshold, and consequently who had the most incentive to save, both learn actively and participate on the stock market. The pattern that, other things equal, the lowest earners do not learn and do not invest, the middle income people do not learn but invest and the rich both invest and learn is universal in this model, but the two threshold values can coincide in which case all investors are knowledgeable.

This relationship between lifetime earnings and learning is due to saving motives in this model. Expected benefits of learning are increasing in the level of period-2 savings. Ceteris paribus, those who intend to save less will see lower benefits to learning than those who intend to save more. Since intended period-2 savings are increasing in lifetime earnings, expected benefits to learning are increasing in lifetime earnings as well. At the same time, the costs of investment, D, aren't directly related to the amount to invest. As a result, the likelihood of learning and investing is increasing in lifetime earnings.

In a richer and more realistic setting learning costs would also be heterogeneous. In reality, learning costs are likely to be negatively correlated with earnings. Heterogeneity in lifetime earnings reflects heterogeneity in general human capital (Becker, 1964). Heterogeneity in human capital is the result of differences in the costs as well as the benefits to investment into human capital (Willis, 1986, and Card, 1998). Those costs include general skills and family background, which likely play important roles in determining costs of learning about stock returns, too. Therefore, those who have higher lifetime earnings because of higher levels of human capital are also likely to face lower learning costs of stock returns.

This amplifies the positive relationship between learning and earnings.

The 8 panels in Figures A2 show additional comparative static results. Each figure shows the fraction of individuals who choose to learn in period 1 and the fraction of individuals holding stocks in period 2**.** These fractions are calculated using the simulated distribution of earnings as described above.

The results are very intuitive. Panel a) of Figure A2 shows that as the cost of learning increases, the fraction of people who choose to learn decreases. The increasing learnings costs make active learning less beneficial but that does not necessarily discourage stock market participation. As long as learning costs are sufficiently high to begin with, a further increase in it would only make people participate on the stock market without learning. This is the case on Figure A1, where an increase in learning cost leads to a monotonic decrease in active learning, but that does not fully translate into lower stockholding above some level of learning costs.

Panel b) shows a reverse picture. As the period-1 expectation of the mean of log potential returns increases, stockholding in period 2 increases dramatically, and the probability of learning increases as well up to a point. Above some expected potential returns, some people can acquire sufficiently high effective returns in the second period even without financial knowledge, but they choose not to invest in knowledge in the first period when they are relatively poor.

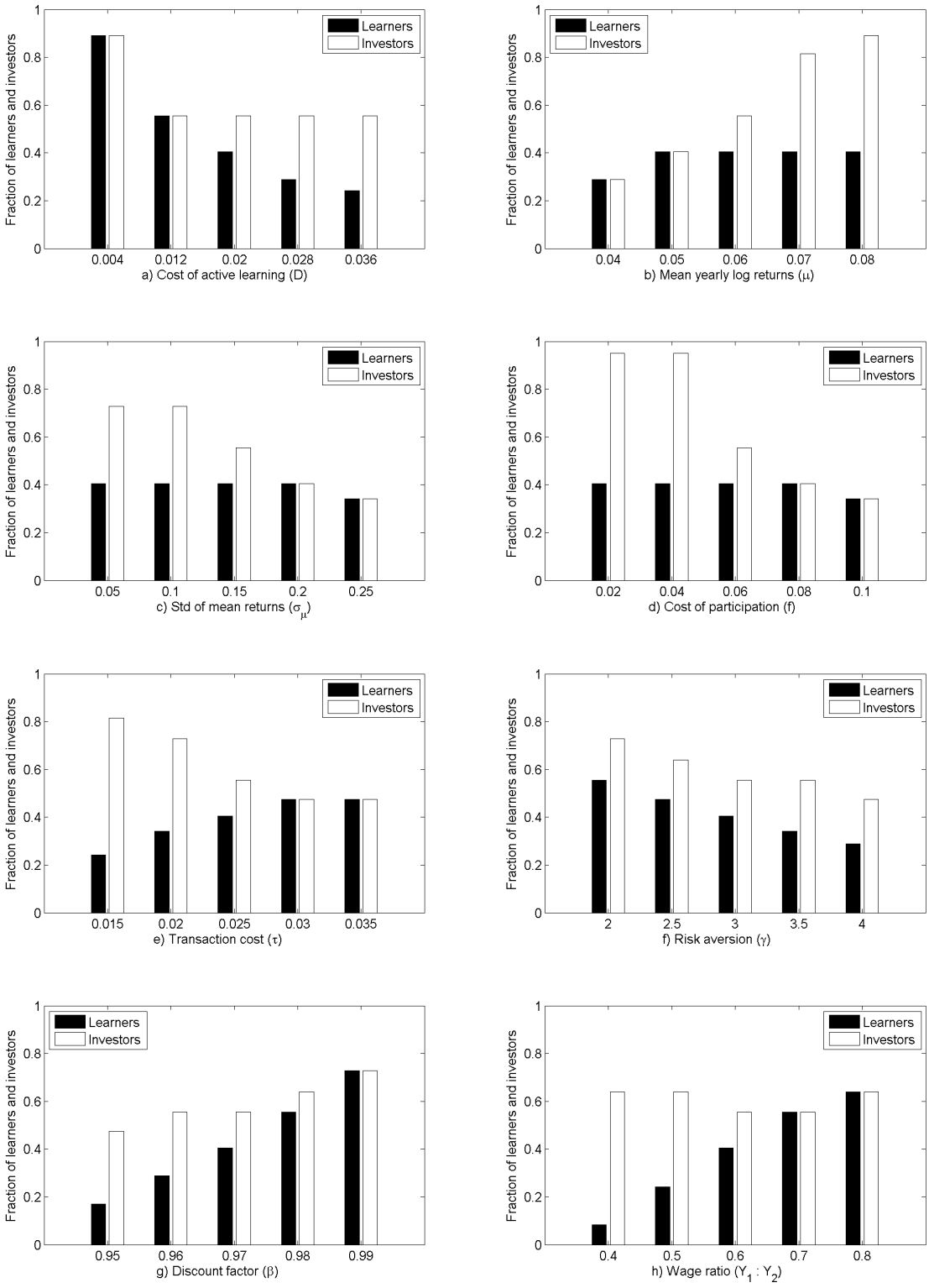
Panel c) shows that parameter uncertainty (i.e., uncertainty about μ) is negatively related to stock market participation and weakly negatively related to learning. The expected value of a lognormally distributed variable positively depends on the variance of its logarithm[13]. This panel nets out this effect. In order to show the pure effect of increasing uncertainty, we have imposed a mean-preserving spread such that expected returns are the same in all five cases. This result shows that, in this setup, the prospect of decreased uncertainty is not an important motive for learning. The expectation of gains is the important motive.[14] Panel d) shows that increasing the fixed costs of participation leads to decreasing stockholding, and albeit in a much less pronounced way, it also leads to less learning.

Panel e) shows the effect of increasing transaction costs τ, the discount from potential

[13] $E\left(R_{it}\right) = \exp\left(\tilde{\mu}_{it} + 0.5\tilde{\sigma}_{it}^2\right)$

[14] In case we do not make the adjustment of the mean log return the dependence between uncertainty and learning vanishes completely, and the dependence between uncertainty and participation becomes very weakly positive.

Figure A2: Fraction of active learners and second period stock market participants by different parameter values



returns. Increasing this discount decreases effective attainable returns conditional on R_t, but it increases the expected gains from active learning. The results imply that the effects on both learning and on stockholding in period 2 are substantial, but the two effects go in the opposite direction. Higher discount makes participation without financial knowledge less beneficial. Some of these people would leave the market, but some would decide to learn and stay on the market.

Panel f) shows a strong and monotonic negative relationship between risk aversion on the one hand, and learning and subsequent stockholding on the other hand. Higher risk aversion leads to a smaller fraction of savings put into stocks, ceteris paribus, which decreases the value of learning about stock returns (especially since the primary effect of such learning is increased expected returns and not decreased risk). Panel g) shows that increased patience increases stockholding and learning as well. It is partly because more patient individuals plan to save more, and partly because they are more willing to pay the costs of learning in period 1 for its expected benefits in period 2.

Finally, panel h) shows that as the age-earnings profile gets flatter (period 1 earnings increase at the cost of period 2 earnings), the probability of learning increases and the effect on stock market participation is rather ambiguous. In order to net out wealth effects lifetime earnings are kept constant in all five specifications and only the ratio of first and second period wage is changing. A flatter wage profile makes any investment in the first period more likely as the marginal value of consumption loss in period 1 decreases. There are two opposing effects on stock market participation. First, a flatter wage profile decreases second period earnings, which makes people less likely to participate on the stock market. Second, if the earnings profile is sufficiently flat, the increasing number of financially knowledgeable people would push stock market participation up. In this particular setup, the first effect dominates at very steep profiles, and the second effect dominates at very flat profiles. In general, it is not evident which of the two effects is stronger.

References

[1] Becker, Gary S. (1964), *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education.* Chicago, University of Chicago Press

[2] Brennan, Michael (1998), "The role of learning in dynamic portfolio decisions." *European Finance Review*, 1:295–306.

[3] Card, David (1999), "The causal effect of education on earnings," in: O. Ashenfelter & D. Card (ed.), *Handbook of Labor Economics*, edition 1, volume 3, chapter 30, pages 1801-1863 Elsevier

[4] Delavande A., S. Rohwedder and R. J. Willis (2008), "Preparation for Retirement, Financial Literacy and Cognitive Resources," Michigan Retirement Research Center Working Paper No. 2008-190.

[5] Haliassos, Michael, and Alexander Michaelides (2002), "Calibration and Computation of Household Portfolio Models." In: Luigi Guiso, Michael Haliassos and Tulio Japelli, eds., *Household Portfolios.* Cambridge: M.I.T. Press

[6] Veldkamp, Laura (2011), *Information Choice in Macroeconomics and Finance.* Princeton University Press. Forthcoming.

[7] Willis, Robert J. (1987), "Wage determinants: A survey and reinterpretation of human capital earnings functions": in: O. Ashenfelter & R. Layard (ed.), *Handbook of Labor Economics*, edition 1, volume 1, chapter 10, pages 525-602. Elsevier.

B Data, descriptive statistics and detailed evidence on noise and information in the probability answers

B.1 Sample and stockholding

Table B1. Sample size

	Age 55-64[a]	Other respondents[b]
HRS 2002	4,056	12,074
HRS 2004	3,676	14,651
HRS 2006	3,182	14,027
HRS 2008 (before Sep)	2,512	11,161

[a]Individuals of age 55 to 64 and whose spouse is of age 55 to 64 as well (or have no spouse)

Table B2. Fraction of stockholders in the sample

	Stockholders		
	outside retirement acc.[a]	in retirement acc.[b]	All
HRS 2002	0.37	0.33	0.51
HRS 2004	0.34	0.32	0.49
HRS 2006	0.29	0.29	0.45
HRS 2008 (before Sep)	0.26	0.29	0.42

[a] Have investments in stocks or mutual funds outside retirement accounts

[b] Have stocks or mutual funds within retirement accounts

Table B3. Share of stocks in the portfolio among stockholders in the sample

	Stockholders		
	outside retirement acc.[a]	in retirement acc.[b]	All
HRS 2002	0.59	0.79	0.56
HRS 2004	0.58	0.75	0.56
HRS 2006	0.53	0.81	0.56
HRS 2008 (before Sep)	0.51	0.78	0.53

[a] Have investments in stocks or mutual funds outside retirement accounts

[b] Have stocks or mutual funds within retirement accounts

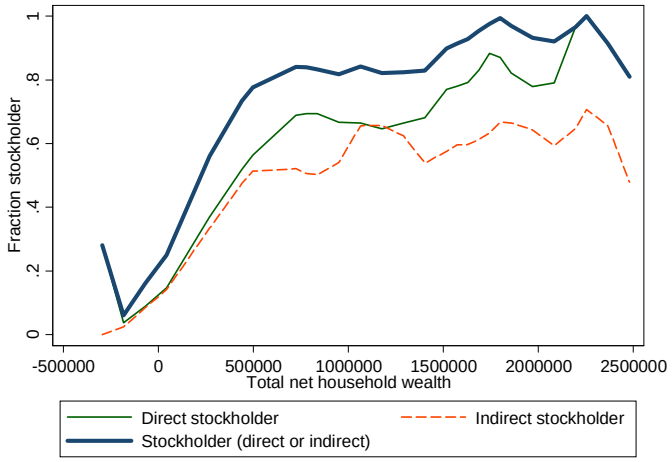


Figure B1. Fraction of stockholders and total net wealth. HRS 2002.

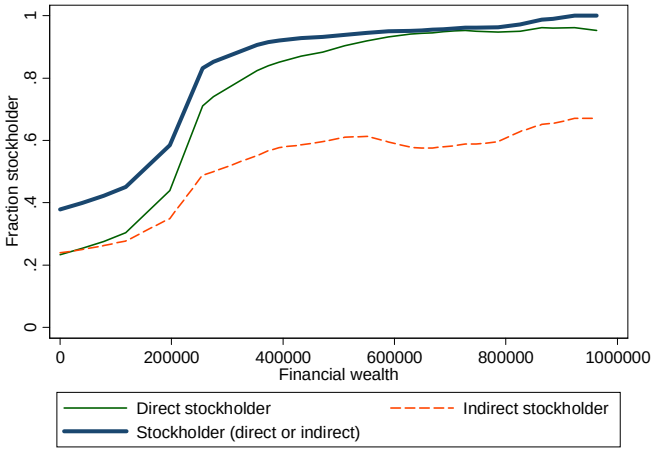


Figure B2. Fraction of stockholders and financial wealth. HRS 2002.

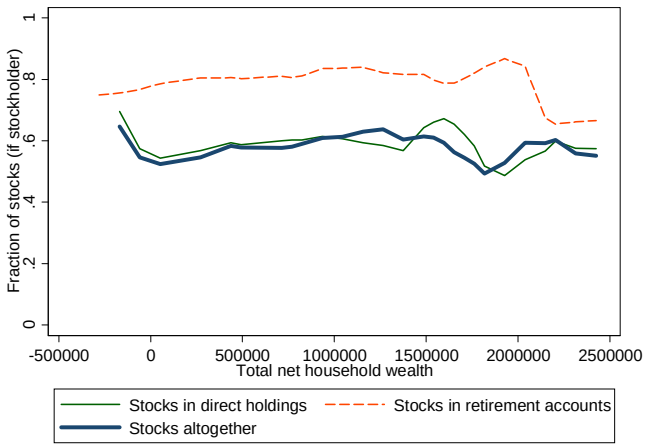


Figure B3. Share of stocks in the portfolio of stockholders, and total net wealth. HRS 2002.

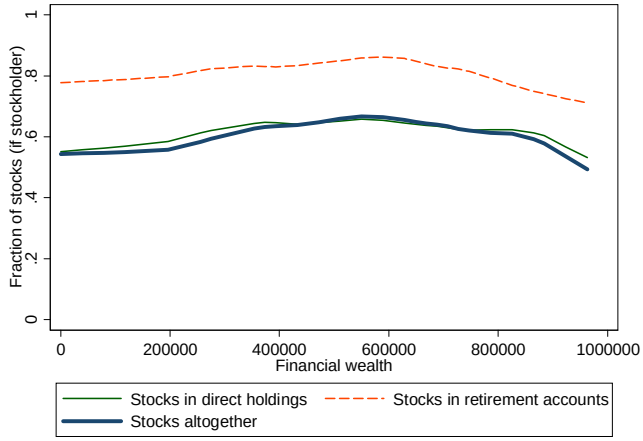


Figure B4. Share of stocks in the portfolio of stockholders, and financial wealth. HRS 2002.

B.2 The proxy variable for lifetime earnings

The source of the lifetime earnings data is the Detailed and the Summary Earnings Records (DER and SER) derived from the Master Earnings File (MEF) of the Social Security Administration that is linked to HRS. For details about the MEF and the linking procedure see Olsen and Hudson (2009) and the documents on the HRS website.[15]

The DER data is derived from the W-2 forms filed by employers to the Internal Revenue Service each year, and it is available from 1978 onward. The SER data is available since 1951, but it contains information only on jobs covered by social security and income up to the taxable maximum. In principle the DER data is superior to the SER as it covers more jobs and it provides more precise information on high income people whose earnings records are capped in SER but uncapped in DER. Therefore we gave priority to the DER data and we only used the SER in exceptional cases described below.

The main issue of the linking procedure is that HRS needed to acquire written consent from sample members in order to get the administrative information on them. HRS made a lot of effort to increase the participation rate, but it remained below 100 percent. Generally HRS has a relatively good coverage rate for earnings before 1992 (slightly above 80 percent) and moderately good coverage rate for earnings afterwards (around 60 percent). Below we provide precise numbers about the attrition rate for our target sample, which will be higher than these numbers. HRS asked for consent in each wave, but in some waves only people with prior consent were asked. Before 2006 the consent covered years up to the interview year, but since 2006 the consent covers future years as well. The consequence is that, as of now, the coverage rates are typically higher for earlier waves (people had more chance to provide consent), but in the future this difference will diminish. Another problem beyond coverage rate is selection. There is evidence that giving consent is not random. Men, the educated, the rich and minorities are underrepresented in the merged sample. See the text for details about how we handle this problem.

Our primary sample is a ten year cohort, people 55-64 years old in 2002 and whose spouse

[15] There are two relatively detailed documents under the data section at hrsonline.isr.umich.edu that can be accessed after free registration. Note that the social security data is not public, and thus only these documents are available but not the data. The website also provides detailed information about how to get permission to use the restricted data.

is in the same age range, too.[16] In some specifications we look at people in the same age range in 2004, 2006 and 2008 as well. As Table B4 shows our target sample size is 4056 in 2002, 3672 in 2004, 3174 in 2006 and 2506 in 2008. Hudomiet, Kézdi and Willis (2011) show that shortly after the fall of Lehman Brothers in September 2008 stock market beliefs of households changed substantially and in an unusual way. For this reason we decided to drop interviews that were made after September 2008 in this paper.

The earnings data we created is the average CPI-adjusted earnings in a 15 year period, between age 40 and 55.[17] The earliest year we used is 1978, which is for earnings at age 40 for people who were 64 in 2002 ($2002 - 64 + 40 = 1978$). The DER data in principle is available from 1978, but the version of the data stored at HRS only covers years 1980 onward. The HRS staff claims that there were some technical problems with the 1978 and the 1979 DER data, and therefore they decided against merging it to HRS. Therefore all the 1978 and 1979 earnings information is coming from the SER. Another issue happened in 1998-2000. HRS first acquired only the DER data until 1997, and then it acquired the SER data until 1999. Therefore for people who stopped giving consent after 2000, we only have SER information for their earnings from 1998 and 1999.

Table B4 shows the quality of the social security earnings information in HRS. As we can see we had no information about the earnings of 612 people in 2002 (15 percent). This number is similar in later waves as well, but due to the falling sample size the ratio of missing values is increasing. Among those who provided some information the majority did so for all the 15 years we needed for our lifetime earnings proxy. The nature of the data is such that missing years can only happen at the end of the period and only for those who stopped giving consent to HRS to collect the earnings data on them. In 2002 we have all the necessary years for 2733 people, we have 10-14 years of information for 590 people and less than 10 years for 121 people. The corresponding numbers for later waves are smaller in level but very similar as a percentage. Here the decision we made was to disregard the earnings data for everyone for whom we have less than 10 years of information, and use the available years for imputation for those who only have 5 or less missing years.

[16] People who are at least 55 years old, but they haven't turned 65 yet.

[17] People who are at least 40 years old but haven't turned 55 yet.

Table B4. Social Security earnings availablility in our target samples

	2002	2004	2006	2008*
Target sample size	4056	3672	3174	2506
All 15 years available	2733	2521	1974	1337
10-14 years available	590	376	405	469
1-9 years available	121	91	64	44
no SSA information	612	684	731	656

* Interviews made prior to October 2008

For confidentiality reasons HRS top-coded all the earnings variables. For people whose earnings were above $250,000 in a given year, we only have interval information, where the intervals are $250,000-$299,999; $300,000-$499,999 and $500,000 and above. Topcoded responses were imputed with a procedure described below. HRS also rounded earnings below $250,000 to the closest multiple of $100, with the exception of $0-$49, where we can differentiate between a true $0 and a $1-$49 value.

The DER data contains five earnings variables:

1. Total compensation: This variable amounts to the sum of the Box 1 values on each W-2 forms submitted on behalf of a person by all his employers. Total compensation includes wages, bonuses, non-cash payments and tips[18]. Total compensation typically does not include deferred payments such as contributions to a 401k plans, but certain plans are included. This variables is uncapped, meaning that high income vales are not censored, only topcoded.

2. Social security earnings: This variables is derived from the Box 3 values of the corresponding W-2 forms. There are two major differences between this variable and total compensation. The first difference is that social security earnings contain information on deferred compensation as well. The second difference is that this variable is capped at the taxable maximum. The taxable maximum was changing year by year. In 2002 it was $80,400, for example, meaning that any earnings beyond this amount are missing.

3. Medicare earnings: This variable is based on the Box 5 values of the W-2 forms. Medicare earnings are almost identical to social security earnings. The main difference is between the taxable maximums used for the two measures. Before 1991 the

[18] Only tips that the employee reported to the employer. Allocated tips are not part of Box 1.

medicare and the social security caps were identical. Since 1994 there is no limit on the taxable earnings for medicare, and between 1991 and 1993 the difference between the medicare and the social security taxable maximums were diverging.

4. FICA taxable self employment earnings: This variable is based on Form 1040 Schedule SE reported by the self employed to IRS. The variable is capped at the same amounts as the social security earnings.

5. Medicare taxable self employment earnings: This variable is almost identical to the previous, but here the less restrictive medicare caps are used.

The SER data contains only one variable which is the sum of all his wage, salary and self employment income. Similarly to social security DER earnings the variable has information only on jobs covered by social security and contains capped values at the social security taxable maximum.

The correlation between these variables are generally very high, but they are not identical. In principle the best quality data is the post 1994 values of the medicare earnings which is uncapped and it also contains information on deferred compensation. The decision we made was the following. First we took the maximum of the total compensation, the social security, and the medicare earnings. In case the maximum was capped or topcoded, we imputed a value with a procedure described below. Second, we took the maximum of the FICA and the medicare taxable self employment earnings. Again, if the maximum was capped or topcoded, we used imputation. Third, we added the employment and the self-employment values. Fourth, we compared this sum to the SER data and took the maximum. After this procedure we had an almost complete person-year-earnings dataset.[19] The final step was the imputation of the remaining missing values.

We needed to impute earnings in three cases. The first is topcoded and rounded responses; the second is for people who stopped giving consent to HRS and therefore their earnings are missing for their last years; and the third is for capped earnings values. Out of these three only the second one affected many respondents (590 in 2002), topcoding and capping were less severe issues.

[19] One technical issue was that missing values and zero earnings were hard to distinguish in the DER data, but it was precisely stored in the version of the SER data HRS provided.

Topcoded and rounded responses were imputed in a very simple way. For the $250,000-$299,999 interval we imputed $270,000; for the $300,000-$499,999 interval we imputed $370,000; for the $500,000 and above interval we imputed $710,000, and for the $1-$49 interval we imputed $40. Other rounded responses were not imputed, we used the rounded values. The values we used were motivated by interval regressions for the logarithm of earnings. If one assumes log-normally distributed earnings, estimates an interval regression, and computes the conditional expected value of a given interval, then he gets numbers that are very close to the values we used. We estimated models with and without flexible time trends in earnings and with and without basic demographic variables such as gender, age and education, and the resulting conditional expectations were always very close to these numbers.

For people whose last earnings values were missing we used their earlier earnings for imputation. As described above, we only have people in our sample with at least ten years of information and thus maximum 5 years of missing earnings. We saw two possibilities for imputation. We could either impute the mean of earlier wages, or we could put more weight on recent years. We have found that many people in our sample had notable fluctuations in their earnings so we decided to use the second approach. First we identified the last four valid earning values for each missing value. Second, we adjusted all the four values with the cpi to get an initial guess for the missing earnings. Then we averaged these values with relative weights $1/t$ if the initial guess was based on earnings t years before the missing response. For example let us say that for a given person we only had earnings up to 1991, and we wanted to impute values for 1992-1994. Let us take the 1994 value. We took the person's earnings from 1988,1989,1990 and 1991, adjusted these values with the cpi, and averaged them with relative weights $1/6, 1/5, 1/4$ and $1/3$.

Capped earning values were imputed in a very similar fashion to the previous. Recall that capping applied to people whose earnings were higher than the taxable maximum. As the taxable maximum increased over time we decided to use the next four earnings values instead of the last four. Moreover, when an initial guess turned out to be lower than the taxable maximum, we replaced the guess with the taxable maximum. When the final guess was equal to the taxable maximum (when all the four initial guesses were lower than that) we imputed 110 percent of the taxable maximum. Another problem was that sometime we had less than four initial guesses, in which case we used as many as we had. When there was no initial guess at all, we again imputed 110 percent of the taxable maximum.

Finally, we simply imputed the sample mean for all the missing observations. The last two rows of Table B4 shows that the number of imputed observations were 733 in 2002 and similar in magnitude in the later waves as well. In the regression analyses, we entered a dummy variable for missing (and therefore imputed) earnings data.

B.3 Noise in the probability answers

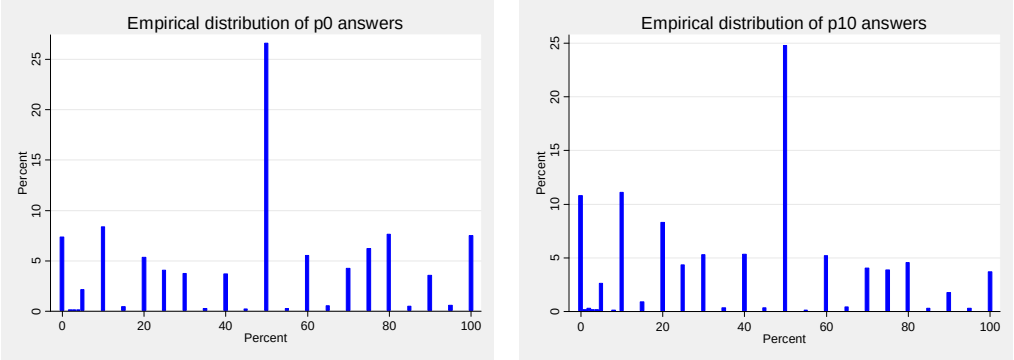


Figure B5. The distribution of reported subjective probabilities of a gain of the stock market (p_0) and the 10 per cent or larger gain (p_{10}). HRS 2002, estimation sample ($n = 2969$)

Table B4. Patterns of survey noise in the core questionnaire

Fraction of responses where	HRS 2002	HRS 2004	HRS 2006	HRS 2008
$p_0 = 0.5$	0.238	0.262	0.239	0.254
$p_0 = 0.0$ or $p\,0 = 1.0$	0.119	0.077	0.073	0.062
p_0 rounded other ten per cent	0.509	0.512	0.559	0.539
p_0 rounded 25% or 75%	0.088	0.096	0.082	0.096
p_0 not round number	0.047	0.054	0.047	0.048
Total	1.000	1.000	1.000	1.000
$p_0 > p_{10}$	0.425			
$p_0 = p_{10}$	0.439			
$p_0 < p_{10}$	0.136			
Total	1.000			

Table B5. Direct evidence on survey noise: Test-retest comparisons using core questionnaire and experimental module answers to the same probability questions from HRS 2002

	p_0	p_{10}
Mean answer in core questionnaire	0.486	0.396
Mean answer in module	0.479	0.334
Difference (core minus module)	0.007	0.063
Standard dev. in core	0.290	0.272
Standard dev. in module	0.272	0.303
Difference (core minus module)	0.018	−0.031
Fraction who gave the same answer in core and module	0.273	0.179
Absolute value of difference between core and module	0.231	0.240
Correlation core and module answers	0.467	0.356

Table B6. The propensity to give round answer to the to p_0 question (multiple of 10% or 25% or 75%)

OLS regression results for the noise patterns in HRS 2002-2008

LHS variable: p0 answer rounded (dummy). HRS 2002-2008						
Stockholder dummy	-0.004					
	[0.005]					
Log lifetime earnings		-0.001				
		[0.003]				
Education			-0.002			
			[0.001]**			
Cognition				0.002		
				[0.004]		
Single female					0.013	
					[0.007]	
Single male					0.011	
					[0.009]	
Female in couple					0.008	
					[0.007]	
Age					-0.001	
					[0.001]	
Black					-0.020	
					[0.009]*	
Hispanic					0.014	
					[0.009]	
Father manager/professional					-0.013	
					[0.008]	
Log risk tolerance					0.001	
					[0.011]	
Wealth non-positive						-0.004
						[0.011]
Wealth in middle						-0.007
						[0.007]
Wealth high						-0.011
						[0.009]
Fin. wealth zero						0.000
						[0.010]
Fin. wealth in middle						-0.004
						[0.008]
Fin. wealth high						-0.005
						[0.009]
Dummies for p0 categories	YES	YES	YES	YES	YES	YES
Observations	11,113	11,113	11,113	11,113	11,112	11,113
R-squared	0.09	0.09	0.09	0.10	0.10	0.06
F-test statistic for shown coeffs					1.27	1.01
p-value					0.257	0.418

Robust standard errors in brackets. * significant at 5%; ** significant at 1%

Mean fill for missing cognition, father's occ, risk tolearnce variables; dummies for missing values included

Table B7. The propensity to give the same answer to p_0 and p_{10}

OLS regression results for the noise patterns in HRS 2002

LHS variable: p_0=p_{10} (dummy). HRS 2002						
Stockholder dummy	-0.016					
	[0.020]					
Log lifetime earnings		-0.020				
		[0.009]*				
Education			-0.008			
			[0.004]*			
Cognition				-0.007		
				[0.011]		
Single female					0.015	
					[0.026]	
Single male					0.047	
					[0.034]	
Female in couple					0.046	
					[0.023]*	
Age					0.011	
					[0.004]*	
Black					0.018	
					[0.028]	
Hispanic					0.007	
					[0.043]	
Father manager/professional					-0.032	
					[0.027]	
Log risk tolerance					0.001	
					[0.029]	
Wealth non-positive						0.088
						[0.044]*
Wealth in middle						0.003
						[0.029]
Wealth high						-0.034
						[0.034]
Fin. wealth zero						-0.045
						[0.037]
Fin. wealth in middle						0.021
						[0.030]
Fin. wealth high						-0.017
						[0.035]
Dummies for p0 categories	YES	YES	YES	YES	YES	YES
Observations	3,520	3,520	3,520	3,520	3,519	3,520
R-squared	0.06	0.06	0.06	0.07	0.06	0.00
F-test statistic for shown coeffs					2.31	0.73
p-value					0.019	0.623

Robust standard errors in brackets. * significant at 5%; ** significant at 1%

Mean fill for missing cognition, father'socc, risk tolearnce variables; dummies for missing values included t

Table B8. The propensity to give smaller answer to p_0 than p_{10}

OLS regression results for the noise patterns in HRS 2002

LHS variable: $p_0<p_{10}$ (dummy). HRS 2002						
Stockholder dummy	-0.005					
	[0.014]					
Log lifetime earnings		0.009				
		[0.006]				
Education			-0.002			
			[0.003]			
Cognition				-0.007		
				[0.008]		
Single female					-0.019	
					[0.019]	
Single male					-0.032	
					[0.022]	
Female in couple					-0.029	
					[0.017]	
Age					-0.006	
					[0.003]*	
Black					0.035	
					[0.022]	
Hispanic					0.069	
					[0.035]*	
Father manager/professional					-0.003	
					[0.018]	
Log risk tolerance					-0.024	
					[0.021]	
Wealth non-positive						-0.032
						[0.031]
Wealth in middle						-0.014
						[0.021]
Wealth high						0.002
						[0.023]
Fin. wealth zero						-0.005
						[0.028]
Fin. wealth in middle						-0.018
						[0.022]
Fin. wealth high						-0.016
						[0.024]
Dummies for p0 categories	NO	NO	NO	NO	NO	NO
Observations	3,520	3,520	3,520	3,520	3,519	3,520
R-squared	0.06	0.06	0.06	0.06	0.07	0.06
F-test statistic for shown coeffs					1.56	0.62
p-value					0.141	0.713

Robust standard errors in brackets. * significant at 5%; ** significant at 1%

Mean fill for missing cognition, father'socc, risk tolearnce variables; dummies for missing values included t

Table B9. Absolute value of the difference between p_0 in the core and p_0 in the module

OLS regression results for the noise patterns in HRS 2002

LHS variable: $\lvert p_0 - p_{0\_module} \rvert$. HRS 2002						
Stockholder dummy	0.044					
	[0.037]					
Log lifetime earnings		-0.006				
		[0.017]				
Education			0.003			
			[0.007]			
Cognition				0.034		
				[0.020]		
Single female					0.006	
					[0.061]	
Single male					-0.056	
					[0.073]	
Female in couple					-0.041	
					[0.052]	
Age					-0.006	
					[0.007]	
Black					-0.087	
					[0.046]	
Hispanic					-0.025	
					[0.127]	
Father manager/professional					0.003	
					[0.056]	
Log risk tolerance					-0.028	
					[0.069]	
Wealth non-positive						-0.009
						[0.055]
Wealth in middle						0.047
						[0.060]
Wealth high						-0.007
						[0.065]
Fin. wealth zero						0.027
						[0.081]
Fin. wealth in middle						0.044
						[0.057]
Fin. wealth high						0.024
						[0.063]
Dummies for p0 categories	NO	NO	NO	NO	NO	NO
Observations	205	205	205	205	205	205
R-squared	0.01	0.00	0.00	0.01	0.02	0.02
F-test statistic for shown coeffs					1.26	1.19
p-value					0.272	0.314

Robust standard errors in brackets. * significant at 5%; ** significant at 1%

Mean fill for missing cognition, father'socc, risk tolearnce variables; dummies for missing values included t

Table B10. Absolute value of the difference between p_{10} in the core and p_{10} in the module

OLS regression results for the noise patterns in HRS 2002

LHS variable: $\lvert p_{10} - p_{10\_module} \rvert$. HRS 2002						
Stockholder dummy	0.025					
	[0.036]					
Log lifetime earnings		0.003				
		[0.016]				
Education			-0.005			
			[0.008]			
Cognition				0.027		
				[0.020]		
Single female					-0.022	
					[0.059]	
Single male					-0.026	
					[0.062]	
Female in couple					-0.043	
					[0.048]	
Age					-0.009	
					[0.008]	
Black					0.012	
					[0.048]	
Hispanic					0.200	
					[0.124]	
Father manager/professional					0.093	
					[0.060]	
Log risk tolerance					0.075	
					[0.061]	
Wealth non-positive						0.051
						[0.063]
Wealth in middle						0.052
						[0.056]
Wealth high						0.039
						[0.064]
Fin. wealth zero						0.146
						[0.094]
Fin. wealth in middle						0.031
						[0.051]
Fin. wealth high						0.017
						[0.059]
Dummies for p0 categories	NO	NO	NO	NO	NO	NO
Observations	196	196	196	196	196	196
R-squared	0.00	0.00	0.00	0.01	0.06	0.03
F-test statistic for shown coeffs					1.26	1.19
p-value					0.272	0.314

Robust standard errors in brackets. * significant at 5%; ** significant at 1%

Mean fill for missing cognition, father'socc, risk tolearnce variables; dummies for missing values included t

B.4 Relevant heterogeneity in the probability answers

Table B11. Descriptive statistics of the subjective probability answers to the stock market returns questions by survey wave. HRS 2002 through 2008.

	$\bar{p}_0$	$V(p_{0i})$	$\bar{p}_0 - \bar{p}_{10}$	Fraction missing p_0
All respondents				
2002	0.48	0.081	0.088	0.18
2004	0.52	0.068		0.12
2006	0.51	0.068		0.19
2008 (before September)	0.50	0.067	0.104	0.15
Stockholders				
2002	0.56	0.081	0.113	0.06
2004	0.58	0.056		0.03
2006	0.57	0.056		0.05
2008 (before September)	0.56	0.055	0.129	0.04
Not stockholders				
2002	0.40	0.081	0.064	0.27
2004	0.46	0.072		0.20
2006	0.46	0.072		0.28
2008 (before September)	0.46	0.071	0.086	0.21

Sample: Health and Retirement Study, waves 2002, 4, 6 and 8 ($\bar{p}_0 - \bar{p}_{10}$ is from HRS 2002 only).

Respondents of age 55 through 64 with a spouse of the same age range (and singles)

p_0 is the answer to the probability of positive returns on stock markets by following year

Table B12. OLS regression results for the stock market probability answers.

Panel 1: Without wealth on the right-hand side

	p0	resid square	p0 - p10	missing p0
Log lifetime earnings	0.010	0.000	0.003	0.032
	[0.004]*	[0.001]	[0.004]	[0.003]**
DB pension	-0.009	-0.002	0.009	-0.012
	[0.008]	[0.002]	[0.011]	[0.008]
DC pension	0.016	-0.004	0.010	-0.023
	[0.007]*	[0.002]	[0.010]	[0.008]**
Education	0.008	-0.001	0.005	-0.015
	[0.002]**	[0.000]**	[0.002]*	[0.002]**
Cognition	0.024	-0.007	0.007	-0.037
	[0.004]**	[0.001]**	[0.006]	[0.005]**
Financial respondent	0.029	0.006	0.000	0.164
	[0.009]**	[0.003]*	0.000	[0.011]**
Log risk tolerance	0.044	0.002	0.022	-0.033
	[0.012]**	[0.004]	[0.014]	[0.012]**
Single female	-0.082	-0.004	-0.045	0.092
	[0.009]**	[0.003]	[0.012]**	[0.010]**
Single male	-0.032	0.000	-0.018	0.038
	[0.012]**	[0.004]	[0.015]	[0.012]**
Female in couple	-0.065	-0.007	-0.023	0.071
	[0.008]**	[0.002]**	[0.011]*	[0.009]**
Age	-0.001	0.001	0.002	0.013
	[0.001]	[0.000]	[0.002]	[0.002]**
Black	-0.048	0.006	-0.019	0.022
	[0.009]**	[0.003]	[0.011]	[0.011]*
Hispanic	0.001	0.005	-0.029	0.112
	[0.013]	[0.004]	[0.017]	[0.016]**
Father manager/professional	0.019	0.001	0.018	0.004
	[0.009]*	[0.003]	[0.012]	[0.009]
Sunny day optimism	0.016	0.000	-0.006	0.004
	[0.007]*	[0.002]	[0.008]	[0.007]
Economic pessimism	-0.092	-0.002	0.000	-0.003
	[0.014]**	[0.004]	[0.017]	[0.014]
Depressive symptomes	-0.010	0.003	0.004	0.014
	[0.004]**	[0.001]*	[0.004]	[0.004]**
Fraction fifty answers	-0.041	-0.124	-0.193	-0.069
	[0.036]	[0.011]**	[0.045]**	[0.040]
Dummies for missing varaibles	YES	YES	YES	YES
Dummies for years	YES	YES	YES	YES
Observations	9131	9131	3323	10887
R-squared	0.1	0.05	0.03	0.23

Robust standard errors in brackets. * significant at 5%; ** significant at 1%

Table B12. OLS regression results for the stock market probability answers.

Panel 2: Wealth included on the right-hand side

	p0	resid square	p0 - p10	missing p0
Log lifetime earnings	0.007	0.000	0.001	0.035
	[0.004]	[0.001]	[0.004]	[0.003]**
DB pension	-0.010	-0.001	0.010	-0.008
	[0.008]	[0.002]	[0.011]	[0.008]
DC pension	0.017	-0.003	0.011	-0.022
	[0.007]*	[0.002]	[0.010]	[0.008]**
Education	0.005	-0.001	0.003	-0.012
	[0.002]**	[0.000]*	[0.002]	[0.002]**
Cognition	0.019	-0.007	0.006	-0.031
	[0.004]**	[0.001]**	[0.006]	[0.005]**
Financial respondent	0.033	0.006	0.000	0.156
	[0.009]**	[0.003]*	0.000	[0.011]**
Log risk tolerance	0.042	0.002	0.021	-0.031
	[0.012]**	[0.004]	[0.014]	[0.012]*
Single female	-0.068	-0.005	-0.038	0.078
	[0.009]**	[0.003]	[0.012]**	[0.010]**
Single male	-0.019	-0.001	-0.011	0.024
	[0.012]	[0.004]	[0.015]	[0.012]*
Female in couple	-0.068	-0.007	-0.026	0.074
	[0.008]**	[0.002]**	[0.011]*	[0.009]**
Age	-0.001	0.001	0.001	0.014
	[0.001]	[0.000]	[0.002]	[0.002]**
Black	-0.031	0.005	-0.010	0.001
	[0.009]**	[0.003]	[0.011]	[0.011]
Hispanic	0.013	0.005	-0.022	0.095
	[0.013]	[0.004]	[0.018]	[0.016]**
Father manager/professional	0.014	0.002	0.014	0.007
	[0.009]	[0.003]	[0.013]	[0.009]
Sunny day optimism	0.015	0.000	-0.006	0.004
	[0.007]*	[0.002]	[0.008]	[0.007]
Economic pessimism	-0.081	-0.003	0.008	-0.011
	[0.014]**	[0.004]	[0.017]	[0.014]
Depressive symptomes	-0.007	0.002	0.006	0.010
	[0.004]	[0.001]*	[0.004]	[0.004]*
Fraction fifty answers	-0.045	-0.116	-0.191	-0.044
	[0.036]	[0.011]**	[0.045]**	[0.040]
Wealth non-positive	0.017	0.002	-0.010	-0.014
	[0.015]	[0.005]	[0.016]	[0.016]
Wealth in middle	0.026	-0.002	0.002	-0.029
	[0.008]**	[0.003]	[0.011]	[0.010]**
Wealth high	0.049	0.000	0.025	-0.038
	[0.010]**	[0.003]	[0.014]	[0.012]**
Fin. wealth zero	-0.021	0.003	0.002	0.056
	[0.012]	[0.004]	[0.014]	[0.015]**
Fin. wealth in middle	0.025	-0.003	0.014	-0.041
	[0.008]**	[0.003]	[0.012]	[0.010]**
Fin. wealth high	0.046	-0.001	0.028	-0.043
	[0.010]**	[0.003]	[0.014]*	[0.011]**
Dummies for missing varaibles	YES	YES	YES	YES
Dummies for years	YES	YES	YES	YES
Observations	9131	9131	3323	10887
R-squared	0.11	0.05	0.04	0.24

Robust standard errors in brackets. * significant at 5%; ** significant at 1%

Table B13. OLS regression results for the stock market probability answers.

Panel 1: Without wealth or belief-specific variables on the right-hand side

	Stockholding		Share of stocks if stockholder	
	HRS 2002-8	HRS 2002	HRS 2002-8	HRS 2002
p0	0.222		0.077	
	[0.019]**		[0.020]**	
p0 missing	-0.140		0.005	
	[0.012]**		[0.021]	
p0 - p10		0.094	0.000	0.043
		[0.033]**	0.000	[0.030]
Log lifetime earnings	0.042	0.046	0.005	0.008
	[0.007]**	[0.008]**	[0.008]	[0.009]
DB pension	0.040	0.021	-0.023	-0.005
	[0.014]**	[0.021]	[0.012]	[0.018]
DC pension	0.019	0.040	0.000	0.003
	[0.013]	[0.020]*	[0.011]	[0.017]
Education	0.027	0.025	0.006	0.005
	[0.002]**	[0.003]**	[0.003]*	[0.004]
Cognition	0.055	0.063	0.001	0.001
	[0.006]**	[0.009]**	[0.008]	[0.011]
Financial respondent	-0.078	-0.159	0.001	0.001
	[0.010]**	[0.019]**	[0.011]	[0.037]
Log risk tolerance	0.038	0.037	0.024	0.031
	[0.021]	[0.024]	[0.021]	[0.023]
Single female	-0.096	-0.129	-0.001	-0.006
	[0.014]**	[0.019]**	[0.015]	[0.021]
Single male	-0.115	-0.130	0.022	-0.016
	[0.018]**	[0.025]**	[0.020]	[0.027]
Female in couple	0.029	0.011	0.007	0.018
	[0.008]**	[0.012]	[0.007]	[0.011]
Age	0.003	-0.001	0.002	0.009
	[0.002]	[0.003]	[0.002]	[0.003]**
Black	-0.196	-0.209	0.005	0.022
	[0.016]**	[0.022]**	[0.023]	[0.035]
Hispanic	-0.136	-0.172	0.012	0.018
	[0.019]**	[0.027]**	[0.031]	[0.047]
Father manager/professional	0.095	0.069	0.001	0.006
	[0.017]**	[0.022]**	[0.014]	[0.018]
Dummies for missing varaibles	YES	YES	YES	YES
Dummies for years	YES	YES	YES	YES
Observations	10901	4055	4850	1876
R-squared	0.27	0.26	0.01	0.01

Robust standard errors in brackets. * significant at 5%; ** significant at 1%

Table B13. OLS regression results for the stock market probability answers.

Panel 2: Wealth and belief-specific variables are included on the right-hand side

	Stockholding		Share of stocks if stockholder	
	HRS 2002-8	HRS 2002	HRS 2002-8	HRS 2002
p0	0.109		0.070	
	[0.017]**		[0.020]**	
p0 missing	-0.079		0.005	
	[0.011]**		[0.020]	
p0 - p10		0.026	0.000	0.044
		[0.030]	0.000	[0.030]
Log lifetime earnings	0.017	0.016	0.003	0.006
	[0.005]**	[0.006]*	[0.008]	[0.008]
DB pension	0.022	0.005	-0.019	-0.001
	[0.012]	[0.019]	[0.012]	[0.018]
DC pension	0.020	0.041	0.002	0.006
	[0.011]	[0.018]*	[0.011]	[0.017]
Education	0.007	0.002	0.005	0.004
	[0.002]**	[0.003]	[0.003]	[0.004]
Cognition	0.020	0.029	0.002	-0.001
	[0.005]**	[0.008]**	[0.008]	[0.011]
Financial respondent	-0.039	-0.068	-0.001	0.001
	[0.008]**	[0.016]**	[0.011]	[0.037]
Log risk tolerance	0.031	0.030	0.021	0.027
	[0.018]	[0.022]	[0.021]	[0.023]
Single female	0.005	-0.025	0.003	0.001
	[0.012]	[0.018]	[0.015]	[0.022]
Single male	-0.024	-0.035	0.027	-0.017
	[0.016]	[0.023]	[0.020]	[0.027]
Female in couple	0.009	-0.009	0.009	0.017
	[0.007]	[0.010]	[0.007]	[0.011]
Age	-0.001	-0.005	0.002	0.008
	[0.002]	[0.003]	[0.002]	[0.003]*
Black	-0.100	-0.095	0.005	0.010
	[0.014]**	[0.021]**	[0.022]	[0.036]
Hispanic	-0.074	-0.092	-0.006	-0.007
	[0.017]**	[0.025]**	[0.032]	[0.047]
Father manager/professional	0.053	0.027	-0.006	-0.004
	[0.015]**	[0.019]	[0.014]	[0.018]
Sunny day optimism	0.005	0.029	0.008	-0.016
	[0.010]	[0.013]*	[0.012]	[0.015]
Economic pessimism	-0.106	-0.082	-0.061	-0.089
	[0.018]**	[0.026]**	[0.023]**	[0.033]**
Depressive symptomes	-0.008	-0.012	0.011	0.009
	[0.005]	[0.007]	[0.007]	[0.010]
Fraction fifty answers	-0.092	-0.076	-0.149	-0.072
	[0.053]	[0.073]	[0.066]*	[0.089]
Wealth non-positive	-0.005	-0.011	0.083	0.180
	[0.012]	[0.020]	[0.050]	[0.077]*
Wealth in middle	0.138	0.134	0.023	0.011
	[0.016]**	[0.024]**	[0.021]	[0.031]
Wealth high	0.279	0.278	0.069	0.045
	[0.020]**	[0.031]**	[0.023]**	[0.035]
Fin. wealth zero	-0.006	-0.016	0.134	0.168
	[0.012]	[0.020]	[0.045]**	[0.067]*
Fin. wealth in middle	0.203	0.232	-0.113	-0.073
	[0.016]**	[0.025]**	[0.021]**	[0.032]*
Fin. wealth high	0.352	0.379	-0.099	-0.043
	[0.020]**	[0.031]**	[0.022]**	[0.035]
Dummies for missing varaibles	YES	YES	YES	YES
Dummies for years	YES	YES	YES	YES
Observations	10887	4054	4848	1876
R-squared	0.42	0.43	0.04	0.04

References

[1] Olsen, Anya and Russell Hudson (2009), "Social Security Administration's Master Earnings File: Background Information." *Social Security Bulletin.* 69(3).

C Details of the structural econometric model

C.1 Deriving the likelihood function

The hypothetical "before rounding" survey answers are the following:

$$p_{0i}^{br} = \Phi\left(\frac{\tilde{\mu}_i}{\tilde{\sigma}_i} + v_{0i}\right) \tag{27}$$

$$p_{10i}^{br} = \Phi\left(\frac{\tilde{\mu}_i - 0.1}{\tilde{\sigma}_i} + v_{10i}\right) \tag{28}$$

Observed probability answers are modeled as interval responses:

$$\begin{pmatrix} p_{0i} \\ p_{10i} \end{pmatrix} \in \mathbf{Q}_{kl} \Leftrightarrow \begin{pmatrix} \Phi\left(\frac{\tilde{\mu}_i}{\tilde{\sigma}_i} + v_{0i}\right) \\ \Phi\left(\frac{\tilde{\mu}_i - 0.1}{\tilde{\sigma}_i} + v_{10i}\right) \end{pmatrix} \in \mathbf{Q}_{kl} \tag{29}$$

$$\mathbf{Q}_{kl} = \begin{pmatrix} [q_k, q_{k+1}) \\ [q_l, q_{l+1}) \end{pmatrix} \tag{30}$$

Expressing the event in scalar terms makes it clear how we can invert the standard normal c.d.f. and get algebraic expressions in terms of the latent variables $\tilde{\mu}_i$, $\tilde{\sigma}_i$ and v_{0i} and v_{10i}.

$$p_{0i} \in [q_k, q_{k+1}) \Leftrightarrow q_k \leq \Phi\left(\frac{\tilde{\mu}_i}{\tilde{\sigma}_i} + v_{0i}\right) < q_{k+1} \tag{31}$$

$$\Leftrightarrow \frac{\tilde{\mu}_i}{\tilde{\sigma}_i} + v_{0i} \in \left[\Phi^{-1}(q_k), \Phi^{-1}(q_{k+1})\right]$$

$$p_{10i} \in [q_l, q_{l+1}) \tag{32}$$

$$\Leftrightarrow \frac{\tilde{\mu}_i}{\tilde{\sigma}_i} + v_{10i} \in \left[\Phi^{-1}(q_l) + \frac{0.1}{\tilde{\sigma}_i}, \Phi^{-1}(q_{l+1}) + \frac{0.1}{\tilde{\sigma}_i}\right] \tag{33}$$

We need distributional assumptions on the random variables to close the econometric model and make it suitable for Maximum Likelihood estimation. First of all, we assume that conditional on observables the individual mean, $\tilde{\mu}_i$ is distributed normally:

$$\tilde{\mu}_i = \beta_\mu' x_{\mu i} + \gamma_\mu' z_{\mu i} + u_{\mu i}$$

$$u_{\mu i} \sim N(0, V(u_\mu))$$

Second, we assume that individual uncertainty, $\tilde{\sigma}_i$ can take two values $\tilde{\sigma}_i \in \{\tilde{\sigma}_{low}, \tilde{\sigma}_{high}\}$ where $\tilde{\sigma}_{low}$ is the low value corresponding to certain people and $\tilde{\sigma}_{high}$ is the high value for

uncertain ones. These two cut points can be estimated, but sometimes we set $\tilde{\sigma}_{low} = 0.15$ which is the historical standard deviation of yearly log-returns. Whether someone has high or low uncertainty is a probit:

$$\begin{aligned} \tilde{\sigma}_i &= \begin{cases} \tilde{\sigma}_{low} & \text{if} \quad \beta'_\sigma x_{\sigma i} + \gamma'_\sigma z_{\sigma i} + u_{\sigma i} \geq 0 \\ \tilde{\sigma}_{high} & \text{if} \qquad\qquad \text{otherwise} \end{cases} \\ u_{\sigma i} &\sim N(0,1) \end{aligned}$$

Third, we assume that the noise components, v_{0i} and v_{10i} follow a bivariate normal distribution:

$$\begin{bmatrix} v_{0i} \\ v_{10i} \end{bmatrix} \sim N\left(0, \sigma_v^2 \begin{bmatrix} 1 & \rho_v \\ \rho_v & 1 \end{bmatrix}\right) \tag{34}$$

Lastly, we assume that $u_{\mu i}, u_{\sigma i}$ and $\mathbf{v}_i \equiv (v_{0i},\ v_{10i})'$ are mutually independent.

Let $\tilde{\boldsymbol{\eta}}_i$ denote the vector of beliefs, $\mathbf{v}$ the vector of noise terms, $\mathbf{p}_i$ the vector of probability answers and let the parameter vector $\boldsymbol{\theta}$ denote all parameters of the conditional density and $\mathbf{X}_i$ denote the vector of all right hand-side variables. The quadrant $\mathbf{Q}_{kl}$ that contains the observed probability answers is defined above.

$$\begin{aligned} \tilde{\boldsymbol{\eta}}_i &\equiv (\tilde{\mu}_i,\ \tilde{\sigma}_i)' \\ \mathbf{v}_i &\equiv (v_{0i},\ v_{10i})' \\ \mathbf{p}_i &\equiv (p_{0i},\ p_{10i})' \\ \boldsymbol{\theta} &= \left(\beta'_\mu, \gamma'_\mu, \tilde{\sigma}_{low}, \tilde{\sigma}_{high}, \beta'_\sigma, \gamma'_\sigma,\ V(u_\mu),\ \sigma_v^2,\ \rho_v\right)' \\ \mathbf{X}_i &\equiv \left(x'_i,\ z'_{\mu i},\ z'_{\sigma i}\right) \end{aligned}$$

Then the event described by (31) and (32) can be summarized as

$$\mathbf{p}_i \in \mathbf{Q}_{kl} | \tilde{\boldsymbol{\eta}}_i, \mathbf{v}_i$$

The individual (conditional) likelihood is the probability of observing that event conditional on observables.

$$\ell_i \equiv \ell\left(\mathbf{p}_i | \mathbf{X}_i;\ \boldsymbol{\theta}\right) = \Pr\left(\mathbf{p}_i \in \mathbf{Q}_{kl} \mid \mathbf{X}_i;\ \boldsymbol{\theta}\right)$$

It is worth expanding the likelihood by conditioning on $\tilde{\sigma}_i$

$$\begin{aligned}
\ell_i &= \ell\left(\mathbf{p}_i|\mathbf{X}_i;\ \boldsymbol{\theta},\tilde{\sigma}_i = \tilde{\sigma}_{low}\right)\Pr\left(\tilde{\sigma}_i = \tilde{\sigma}_{low}|\mathbf{X}_i;\ \boldsymbol{\theta}\right) \\
&\quad +\ell\left(\mathbf{p}_i|\mathbf{X}_i;\ \boldsymbol{\theta},\tilde{\sigma}_i = \tilde{\sigma}_{high}\right)\Pr\left(\tilde{\sigma}_i = \tilde{\sigma}_{high}|\mathbf{X}_i;\ \boldsymbol{\theta}\right) \\
&= \ell\left(\mathbf{p}_i|\mathbf{X}_i;\ \boldsymbol{\theta},\tilde{\sigma}_i = \tilde{\sigma}_{low}\right)\Phi\left(\beta'_\sigma x_{\sigma i} + \gamma'_\sigma z_{\sigma i}\right) \\
&\quad +\ell\left(\mathbf{p}_i|\mathbf{X}_i;\ \boldsymbol{\theta},\tilde{\sigma}_i = \tilde{\sigma}_{high}\right)\left(1 - \Phi\left(\beta'_\sigma x_{\sigma i} + \gamma'_\sigma z_{\sigma i}\right)\right)
\end{aligned}$$

Thus it is enough to find an expression for $\ell\left(\mathbf{p}_i|\mathbf{X}_i;\ \boldsymbol{\theta},\tilde{\sigma}_i\right)$. Let us denote $w_{0i} \equiv \frac{u_{\mu i}}{\tilde{\sigma}_i} + v_{0i}$ and $w_{10i} \equiv \frac{u_{\mu i}}{\tilde{\sigma}_i} + v_{10i}$. The conditional likelihood is

$$\begin{aligned}
\ell\left(\mathbf{p}_i|\mathbf{X}_i;\ \boldsymbol{\theta},\tilde{\sigma}_i\right) &= \Pr\left(\left[\begin{array}{c} p_{0i} \\ p_{10i} \end{array}\right] \in \left[\begin{array}{c} [q_k, q_{k+1}) \\ [q_l, q_{l+1}) \end{array}\right] \middle| \mathbf{X}_i;\ \boldsymbol{\theta},\tilde{\sigma}_i\right) \\
&= \Pr\left(\left[\begin{array}{c} w_{0i} \\ w_{10i} \end{array}\right] \in \left[\begin{array}{c} \left[w_{0i}^k, w_{0i}^{k+1}\right) \\ \left[w_{10i}^l, w_{10i}^{l+1}\right) \end{array}\right] \middle| \mathbf{X}_i;\ \boldsymbol{\theta},\tilde{\sigma}_i\right)
\end{aligned}$$

with the notation $w_{0i}^{k+1} \equiv \Phi^{-1}\left(q_{k+1}\right) - \beta'_\mu x_{\mu i} + \gamma'_\mu z_{\mu i}$, $w_{0i}^k \equiv \Phi^{-1}\left(q_k\right) - \beta'_\mu x_{\mu i} + \gamma'_\mu z_{\mu i}$, $w_{10i}^{l+1} \equiv \Phi^{-1}\left(q_{l+1}\right) - \beta'_\mu x_{\mu i} + \gamma'_\mu z_{\mu i} + \frac{0.1}{\tilde{\sigma}_i}$ and $w_{10i}^l \equiv \Phi^{-1}\left(q_l\right) - \beta'_\mu x_{\mu i} + \gamma'_\mu z_{\mu i} + \frac{0.1}{\tilde{\sigma}_i}$

$\mathbf{w}_i \equiv \left(w_{0i},\ w_{10i}\right)'$ has a centered bivariate normal distribution and thus the likelihood can be expressed from the bivariate normal c.d.f.

$$\begin{aligned}
\ell\left(\mathbf{p}_i|\mathbf{X}_i;\ \boldsymbol{\theta},\tilde{\sigma}_i\right) &= Binorm\left(w_{0i}^{k+1}, w_{10i}^{l+1}, \mathbf{C}_i\right) + Binorm\left(w_{0i}^k, w_{10i}^l, \mathbf{C}_i\right) \\
&\quad -Binorm\left(w_{0i}^k, w_{10i}^{l+1}, \mathbf{C}_i\right) - Binorm\left(w_{0i}^{k+1}, w_{10i}^l, \mathbf{C}_i\right) \qquad (35)
\end{aligned}$$

with $\mathbf{C}_i$ representing the variance-covariance matrix of $\mathbf{w}_i|\mathbf{X}_i;\ \boldsymbol{\theta},\tilde{\sigma}_i$

$$\mathbf{C}_i = \left[\begin{array}{cc} \frac{V(u_\mu)}{\tilde{\sigma}_i^2} + \sigma_v^2 & \frac{V(u_\mu)}{\tilde{\sigma}_i^2} + \rho_v\sigma_v^2 \\ \frac{V(u_\mu)}{\tilde{\sigma}_i^2} + \rho_v\sigma_v^2 & \frac{V(u_\mu)}{\tilde{\sigma}_i^2} + \sigma_v^2 \end{array}\right]$$

As the bivariate normal distribution is available in standard econometric packages such as Stata 11 the likelihood can be evaluated using (35).

C.2 Expected $\tilde{\mu}$ and $\tilde{\sigma}$ conditional on the probability answers

The goal is to get

$$\hat{\boldsymbol{\eta}}_i = \widehat{\mathbb{E}}\left[\tilde{\boldsymbol{\eta}}_i|\mathbf{p}_i \in \mathbf{Q}_{kl}\right]$$

Start from the individual likelihood (31) and (32). These describe the probability of the probability answers falling in a certain interval, conditional on $\tilde{\mu}_i$ and $\tilde{\sigma}_i$. In the parsimonious notation, (31) and (32) describe the event $\mathbf{p}_i \in \mathbf{Q}_{kl}|\tilde{\boldsymbol{\eta}}_i$. Our question is the reverse: it is the density (and then the expectation) of $\tilde{\mu}_i$ and $\tilde{\sigma}_i$ conditional on the probability answer:

$$\mathbb{E}\left[\tilde{\boldsymbol{\eta}}_i|\mathbf{p}_i \in \mathbf{Q}_{kl}\right] = \int \tilde{\boldsymbol{\eta}}_i \times f\left(\tilde{\boldsymbol{\eta}}_i|\mathbf{p}_i \in \mathbf{Q}_{kl}\right) d\left(\tilde{\boldsymbol{\eta}}_i\right)$$

By Bayes' theorem,

$$f\left(\tilde{\boldsymbol{\eta}}_i|\mathbf{p}_i \in \mathbf{Q}_{kl}\right) = \frac{\Pr\left(\mathbf{p}_i \in \mathbf{Q}_{kl}|\tilde{\boldsymbol{\eta}}_i\right) \times f\left(\tilde{\boldsymbol{\eta}}_i\right)}{\Pr\left(\mathbf{p}_i \in \mathbf{Q}_{kl}\right)} = \frac{\Pr\left(\mathbf{p}_i \in \mathbf{Q}_{kl}|\tilde{\boldsymbol{\eta}}_i\right)}{l_i} \times f\left(\tilde{\boldsymbol{\eta}}_i\right)$$

so that

$$\hat{\boldsymbol{\eta}}_i = \mathbb{E}\left[\tilde{\boldsymbol{\eta}}_i|\mathbf{p}_i \in \mathbf{Q}_{kl}\right] = \int \tilde{\boldsymbol{\eta}}_i \times \frac{\Pr\left(\mathbf{p}_i \in \mathbf{Q}_{kl}|\tilde{\boldsymbol{\eta}}_i\right)}{l_i} \times f\left(\tilde{\boldsymbol{\eta}}_i\right) d\left(\tilde{\boldsymbol{\eta}}_i\right) \tag{36}$$

The only unknown part is $\Pr\left(\mathbf{p}_i \in \mathbf{Q}_{kl}|\tilde{\boldsymbol{\eta}}_i\right)$. It can be computed similarly to the likelihood function.

$$\begin{aligned}
\Pr\left(\mathbf{p}_i \in \mathbf{Q}_{kl}|\tilde{\boldsymbol{\eta}}_i\right) &= \Pr\left(\left[\begin{array}{c} p_{0i} \\ p_{10i} \end{array}\right] \in \left[\begin{array}{c} [q_k, q_{k+1}) \\ [q_l, q_{l+1}) \end{array}\right] \middle| \mathbf{X}_i;\ \boldsymbol{\theta}, \tilde{\sigma}_i, \tilde{\mu}_i\right) \\
&= \Pr\left(\left[\begin{array}{c} v_{0i} \\ v_{10i} \end{array}\right] \in \left[\begin{array}{c} \left[v_{0i}^k, v_{0i}^{k+1}\right) \\ \left[v_{10i}^l, v_{10i}^{l+1}\right) \end{array}\right] \middle| \mathbf{X}_i;\ \boldsymbol{\theta}, \tilde{\sigma}_i, \tilde{\mu}_i\right)
\end{aligned}$$

where $v_{0i}^{k+1} \equiv \Phi^{-1}\left(q_{k+1}\right) - \frac{\tilde{\mu}_i}{\tilde{\sigma}_i}$, $v_{0i}^{k} \equiv \Phi^{-1}\left(q_{k}\right) - \frac{\tilde{\mu}_i}{\tilde{\sigma}_i}$, $v_{10i}^{l+1} \equiv \Phi^{-1}\left(q_{l+1}\right) - \frac{\tilde{\mu}_i}{\tilde{\sigma}_i} + \frac{0.1}{\tilde{\sigma}_i}$ and $v_{10i}^{l} \equiv \Phi^{-1}\left(q_{l}\right) - \frac{\tilde{\mu}_i}{\tilde{\sigma}_i} + \frac{0.1}{\tilde{\sigma}_i}$. Note that this is different from the analogous formula in the likelihood function because, at this stage we "know" $(\tilde{\mu}_i, \tilde{\sigma}_i)$. (In practice, we simulate it out using the estimated parameters which completely specify its distribution.)

The probability in question is again the probability mass over a rectangle:

$$\begin{aligned}
\Pr\left(\mathbf{p}_i \in \mathbf{Q}_i|\tilde{\boldsymbol{\eta}}_i\right) &= Binorm\left(v_{0i}^{k+1}, v_{10i}^{l+1}, \mathbf{D}\right) + Binorm\left(v_{0i}^{k}, v_{10i}^{l}, \mathbf{D}\right) \\
&\quad - Binorm\left(v_{0i}^{k}, v_{10i}^{l+1}, \mathbf{D}\right) - Binorm\left(v_{0i}^{k+1}, v_{10i}^{l}, \mathbf{D}\right)
\end{aligned} \tag{37}$$

with covariance matrix $\mathbf{D}$ from (34) so that $\mathbf{D} = \sigma_v^2 \left[\begin{array}{cc} 1 & \rho_v \\ \rho_v & 1 \end{array}\right]$.

Having all elements in (36) the integration can be approximated by simulation. With drawing M simulation draws $\tilde{\boldsymbol{\eta}}_{i,s}$ from the distribution of $\tilde{\boldsymbol{\eta}}_i$ the approximation can be written as

$$\begin{aligned}\hat{\boldsymbol{\eta}}_i &= \int \tilde{\boldsymbol{\eta}}_i \times \frac{\Pr\left(\mathbf{p}_i \in \mathbf{Q}_{kl} | \tilde{\boldsymbol{\eta}}_i\right)}{l_i} \times f\left(\tilde{\boldsymbol{\eta}}_i\right) d\left(\tilde{\boldsymbol{\eta}}_i\right) \\ &\approx \frac{1}{K_i} \sum_{s=1}^{M} \tilde{\boldsymbol{\eta}}_{i,s} \times \frac{\Pr\left(\mathbf{p}_i \in \mathbf{Q}_{kl} | \tilde{\boldsymbol{\eta}}_{i,s}\right)}{l_i}\end{aligned}$$

where K_i is a normalization factor:

$$K_i = \sum_{s=1}^{M} \frac{\Pr\left(\mathbf{p}_i \in \mathbf{Q}_i | \tilde{\boldsymbol{\eta}}_{i,s}\right)}{l_i}$$

C.3 Estimating the variance and correlation of survey noise

The goal of this exercise is to estimate moments of the noise distribution so that we can calibrate those in the estimation. We are interested in σ_v^2 and ρ_v. In this simple exercise, we make use of the probability answers in the core questionnaire (p_{0i}, p_{10i}) and the probability answers in the experimental module (p_{M0i}, p_{M10i}), and we ignore rounding.

The hypothetical "before rounding" survey answers are, conditional on the noise variables, the following

$$\begin{aligned}p_{0i}^{br} &= \Phi\left(\frac{\tilde{\mu}_i}{\tilde{\sigma}_i} + v_{0i}\right) \\ p_{10i}^{br} &= \Phi\left(\frac{\tilde{\mu}_i - 0.1}{\tilde{\sigma}_i} + v_{10i}\right) \\ p_{M0i}^{br} &= \Phi\left(\frac{\tilde{\mu}_i}{\tilde{\sigma}_i} + v_{M0i}\right) \\ p_{M10i}^{br} &= \Phi\left(\frac{\tilde{\mu}_i - 0.1}{\tilde{\sigma}_i} + v_{M10i}\right)\end{aligned}$$

As a result, we have that $\Phi^{-1}\left(p_{0i}^{br}\right) = \frac{\tilde{\mu}_i}{\tilde{\sigma}_i} + v_{0i}$, $\Phi^{-1}\left(p_{10i}^{br}\right) = \frac{\tilde{\mu}_i - 0.1}{\tilde{\sigma}_i} + v_{10i}$, $\Phi^{-1}\left(p_{M0i}^{br}\right) = \frac{\tilde{\mu}_i}{\tilde{\sigma}_i} + v_{M0i}$, and $\Phi^{-1}\left(p_{M10i}^{br}\right) = \frac{\tilde{\mu}_i - 0.1}{\tilde{\sigma}_i} + v_{M10i}$.

By assumption, the noise components are jointly normally distributed, and they are uncorrelated across core questionnaire and the experimental module.

$$\begin{bmatrix} v_{0i} \\ v_{10i} \\ v_{M0i} \\ v_{M10} \end{bmatrix} \sim N\left(\mathbf{0}, \sigma_v^2 \begin{bmatrix} 1 & & & \\ \rho_v & 1 & & \\ 0 & 0 & 1 & \\ 0 & 0 & \rho_v & 1 \end{bmatrix}\right)$$

C.3.1 Moment conditions 1 and 2

Compare the inverse normal of the core and module answers to the same probability question (p_0 and p_{M0} or p_{10} and p_{M10}), and and take expectation of the squares:

$$\begin{aligned} E\left[\left\{\Phi^{-1}\left(p_{0i}^{br}\right) - \Phi^{-1}\left(p_{M0i}^{br}\right)\right\}^2\right] &= \left[\left\{\frac{\tilde{\mu}_i}{\tilde{\sigma}_i} + v_{0i} - \frac{\tilde{\mu}_i}{\tilde{\sigma}_i} - v_{M0i}\right\}^2\right] \quad (38) \\ &= E\left[\left(v_{0i} - v_{M0i}\right)^2\right] = 2\sigma_v^2 \\ E\left[\left\{\Phi^{-1}\left(p_{10i}^{br}\right) - \Phi^{-1}\left(p_{M10i}^{br}\right)\right\}^2\right] &= \left[\left\{\frac{\tilde{\mu}_i - 0.1}{\tilde{\sigma}_i} + v_{10i} - \frac{\tilde{\mu}_i - 0.1}{\tilde{\sigma}_i} - v_{M10i}\right\}^2\right] \quad (39) \\ &= E\left[\left(v_{10i} - v_{M10i}\right)^2\right] = 2\sigma_v^2 \end{aligned}$$

C.3.2 Moment conditions 3 and 4

Similar comparisons across questions (p_0 and p_{M10} or p_{10} and p_{M0}) yield moments that are similar to (38) and (39), but they also include the subjective beliefs about the standard deviation of stock market returns.

$$\begin{aligned} E\left[\left\{\Phi^{-1}\left(p_{0i}^{br}\right) - \Phi^{-1}\left(p_{M10i}^{br}\right)\right\}^2\right] &= \left[\left\{\frac{\tilde{\mu}_i}{\tilde{\sigma}_i} + v_{10i} - \frac{\tilde{\mu}_i - 0.1}{\tilde{\sigma}_i} - v_{M10i}\right\}^2\right] \quad (40) \\ &= E\left[\left(v_{0i} - v_{M10i} + \frac{0.1}{\tilde{\sigma}_i}\right)^2\right] = 2\sigma_v^2 + 0.01E\left[\frac{1}{\tilde{\sigma}_i^2}\right] \\ E\left[\left\{\Phi^{-1}\left(p_{10i}^{br}\right) - \Phi^{-1}\left(p_{M0i}^{br}\right)\right\}^2\right] &= \left[\left\{\frac{\tilde{\mu}_i - 0.1}{\tilde{\sigma}_i} + v_{10i} - \frac{\tilde{\mu}_i}{\tilde{\sigma}_i} - v_{M10i}\right\}^2\right] \quad (41) \\ &= E\left[\left(v_{0i} - v_{M10i} - \frac{0.1}{\tilde{\sigma}_i}\right)^2\right] = 2\sigma_v^2 + 0.01E\left[\frac{1}{\tilde{\sigma}_i^2}\right] \end{aligned}$$

C.3.3 Moment condition 5

Compare the adjacent probability answers in the core questionnaire (p_{0i}, p_{10i}) and take expectation of the squares:

$$\begin{aligned} E\left[\left\{\Phi^{-1}\left(p_{0i}^{br}\right)-\Phi^{-1}\left(p_{10i}^{br}\right)\right\}^2\right] &= E\left[\left\{\frac{\tilde{\mu}_i}{\tilde{\sigma}_i}+v_{0i}-\frac{\tilde{\mu}_i-0.1}{\tilde{\sigma}_i}-v_{10i}\right\}^2\right] \\ &= E\left[\left(v_{0i}-v_{10i}+\frac{0.1}{\tilde{\sigma}_i}\right)^2\right]=2\left(1-\rho_v\right)\sigma_v^2+0.01E\left[\frac{1}{\tilde{\sigma}_i^2}\right] \end{aligned} \tag{42}$$

In principle, one can do this for the answers from the experimental module, (p_{M0i}, p_{M10i}). Because of low number of observations in the experimental module, we do not make use of that comparison.

C.3.4 Estimation of ρ and σ_v^2 by Minimum Distance

In principle, this is a simple Minimum Distance problem with five moment conditions ((38) through (42)) in three parameters $\left(\sigma_v^2,\ \rho_v \text{ and } E\left[1/\tilde{\sigma}_i^2\right]\right)$. Of these three parameters, we are interested in two, $(\sigma_v^2, \text{ and } \rho_v)$.

The first two moment conditions allow for a Minimum Distance estimation of σ_v^2, while the fifth moment, together with the third and the fourth moments, allows for a Minimum Distance estimation of ρ_v. To see the latter, consider the difference (40) − (42), and the difference (41) − (42) is, of course, analogous.

$$\begin{aligned} &E\left[\left\{\Phi^{-1}\left(p_{0i}^{br}\right)-\Phi^{-1}\left(p_{M10i}^{br}\right)\right\}^2\right] \\ &-E\left[\left\{\Phi^{-1}\left(p_{0i}^{br}\right)-\Phi^{-1}\left(p_{10i}^{br}\right)\right\}^2\right] &= 2\sigma_v^2+0.01E\left[\frac{1}{\tilde{\sigma}_i^2}\right]-2\left(1-\rho_v\right)\sigma_v^2-0.01E\left[\frac{1}{\tilde{\sigma}_i^2}\right] \\ &&= 2\rho_v\sigma_v^2 \end{aligned}$$

Unfortunately, we do not observe p_{0i}^{br} and p_{10i}^{br} only their survey response versions that are rounded versions for almost all respondents. We address rounding in the likelihood estimation is by interval regressions, which is consistent under any rounding model (as long as rounding is within the pre-defined intervals). In this simple exercise, we assume away rounding error and treat observed answers as if they were the hypothetical pre-rounding variables p_{0i}^{br} and p_{10i}^{br} .

However, an important practical consequence of rounding is the prevalence of answers at 0 and 1, and $\Phi^{-1}(p)$ is not defined for $p = 0$ or $p = 1$. In this simple exercise we opted for

an ad-hoc solution replacing $p = 0$ to $p = 0 + \varepsilon$ and $p = 1$ to $p = 1 - \varepsilon$, respectively. Various values for ε were considered (0.05, 0.025, 0.01, and 0.005), and we present the results as a function of those values.

We have four equations in two unknowns, with the first two and second two equations being symmetric in p_0 and p_{10} (or their counterparts in the experimental module). This symmetry implies that the optimum Minimum Distance estimator has identity weights under the structure of our model.

The estimation results are the following.

Table C1. Estimated variance and correlation of survey noise.

Results of the Minimum Distance exercise by values of the auxiliary parameter ε

	$\varepsilon = 0.050$	$\varepsilon = 0.025$	$\varepsilon = 0.010$	$\varepsilon = 0.005$
σ_v	0.96	1.05	1.17	1.26
ρ_v	0.61	0.61	0.61	0.60

These results should be viewed as very crude approximations because they ignore rounding (we substituted in the actual answers p for the hypothetical, before-rounding answers p^{br}) and because they handle boundary values in a very ad-hoc way. The estimates of ρ_v seem robust to our handling the boundary problem, but the estimates of σ_v are not.

C.3.5 Estimation of ρ and σ_v^2 by Minimum Distance with covariates

The likelihood function and the estimator for $(\hat{\mu}_i, \hat{\sigma}_i)$ conditions on observed covariates $(\mathbf{X}_i)$ as well as the observed answers to the stock market probability questions $(\mathbf{p}_i)$. The variance and correlation coefficient of the noise variables (v_{0i}, v_{10i}) may be different if conditioned on those covariates.[20]

In this subsection we present estimates of the noise parameters that use moment conditions conditional on covariates. In practice, we repeated the Minimum Distance exercise described above, but instead of the inverse of the observed (and ε-adjusted) variables $\Phi^{-1}(p_{0i})$ etc. we used their residuals after having regressed on all covariates $(\mathbf{X}_i)$. The results are in table C2.

[20] Tables B6 through B10 in the Online Appendix B show that the observed noise features are not strongly associated with covariates. That was the basis for our assumption of unbiased and homoskedastic noise. However, even those weak associations may result in a conditional noise variance that is somewhat smaller than the unconditional one, which may make a difference in the likelihood estimation procedure.

Table C2. Estimated variance and correlation of survey noise, conditional on covariates. Results of the Minimum Distance exercise by values of the auxiliary parameter ε

	$\varepsilon = 0.050$	$\varepsilon = 0.025$	$\varepsilon = 0.010$	$\varepsilon = 0.005$
σ_v	0.95	1.04	1.15	1.24
ρ_v	0.42	0.42	0.43	0.44

D Detailed estimation results from the structural econometric model

D.1 Detailed estimates from the benchmark model

Table D1. Detailed structural estimates 2-point distribution for $\tilde{\sigma}$, low fixed at 0.15.

	E[mu]	P[sig=low] probit coeff	E[mu]	P[sig=low] probit coeff	E[mu]	P[sig=low] probit coeff
Log lifetime earnings			0.005	-0.055	0.003	-0.053
			[0.009]	[0.048]	[0.010]	[0.058]
Education (years)			-0.046	-0.309	-0.042	-0.294
			[0.021]*	[0.262]	[0.021]*	[0.291]
Cognitive score			0.003	-0.019	0.005	-0.009
			[0.018]	[0.179]	[0.018]	[0.197]
DB pension plan			0.01	-0.01	0.005	-0.009
			[0.004]*	[0.032]	[0.004]	[0.035]
DC pension plan			0.02	-0.121	0.017	-0.112
			[0.014]	[0.101]	[0.014]	[0.118]
Financial respondent			0.036	0.073	0.035	0.057
			[0.020]	[0.193]	[0.020]	[0.207]
Log risk tolerance			0.088	0.137	0.087	0.11
			[0.027]**	[0.239]	[0.026]**	[0.276]
Single female			-0.158	-0.763	-0.145	-0.833
			[0.033]**	[0.234]**	[0.034]**	[0.257]**
Single male			-0.076	-0.28	-0.061	-0.361
			[0.030]*	[0.238]	[0.029]*	[0.268]
Female in couple			-0.108	-0.715	-0.112	-0.798
			[0.028]**	[0.235]**	[0.031]**	[0.305]**
Age			0.003	0.079	0.002	0.072
			[0.004]	[0.039]*	[0.004]	[0.045]
Black			-0.072	-0.184	-0.041	-0.174
			[0.034]*	[0.269]	[0.032]	[0.267]
Hispanic			0	0.029	0.021	0.036
			[0.039]	[0.297]	[0.039]	[0.331]
Father professional			0.028	0.202	0.018	0.166
			[0.021]	[0.189]	[0.020]	[0.217]
Missing lifetime earnings			0.015	0.519	0.005	0.558
			[0.020]	[0.170]**	[0.021]	[0.189]**
Missing risk tolerance			0.023	0.32	0.018	0.314
			[0.022]	[0.212]	[0.021]	[0.235]
Missing father occupation			0.005	0.291	0.006	0.312
			[0.025]	[0.205]	[0.024]	[0.222]
Non-positive wealth					-0.008	-0.623
					[0.120]	[1.107]
Medium wealth					0.015	-0.396
					[0.025]	[0.209]
Hugh wealth					0.068	-0.061
					[0.030]*	[0.296]
Zero financial wealth					0.02	0.272
					[0.040]	[0.353]
Medium financial wealth					0.024	-0.215
					[0.027]	[0.237]
High finacial wealth					0.049	-0.138
					[0.031]	[0.278]
Sunshine optimism			0.04		0.038	
			[0.016]*		[0.015]*	
Pessimism in economic outlook			-0.206		-0.182	
			[0.041]**		[0.038]**	
Depressive symptomes			-0.018		-0.011	
			[0.010]		[0.009]	
Missing sunshine			-0.034		-0.033	
			[0.031]		[0.030]	
Missing economic pessimism			-0.038		-0.036	
			[0.038]		[0.038]	
Fraction fifty answers				-3.724		
				[1.025]**		
Constant	-0.009	-0.716	-0.148	-3.871	-0.052	-0.76

Table D2. Detailed probit estimates 2-point distribution for $\tilde{\sigma}$, low fixed at 0.15

	Pr(S=1)		E(s|s>0)	
	Probit coefficients		Truncated regression coefficients	
mu_hat		2.361***		0.302***
		[0.257]		[0.099]
sigma_hat		-0.160		0.144
		[0.285]		[0.111]
Log lifetime earnings	0.128***	0.109***	0.000	-0.003
	[0.027]	[0.026]	[0.008]	[0.008]
Education (years)	0.102***	0.075***	0.011**	0.008
	[0.013]	[0.013]	[0.005]	[0.005]
Cognitive score	0.214***	0.143***	0.002	-0.009
	[0.035]	[0.035]	[0.014]	[0.014]
DB pension plan	0.017	0.108	-0.020	-0.012
	[0.066]	[0.069]	[0.024]	[0.024]
DC pension plan	0.097	0.079	0.006	0.004
	[0.064]	[0.065]	[0.021]	[0.021]
Financial respondent	-0.071**	-0.152***	-0.029***	-0.040***
	[0.035]	[0.038]	[0.010]	[0.011]
Log risk tolerance	0.072	-0.127	0.048*	0.030
	[0.082]	[0.086]	[0.028]	[0.029]
Single female	-0.344***	0.067	-0.023	0.009
	[0.072]	[0.093]	[0.029]	[0.035]
Single male	-0.292***	-0.098	-0.022	-0.005
	[0.087]	[0.092]	[0.034]	[0.035]
Female in couple	-0.011	0.280***	0.010	0.029
	[0.039]	[0.060]	[0.013]	[0.020]
Age	-0.013	-0.016	0.006	0.006*
	[0.010]	[0.010]	[0.004]	[0.004]
Black	-0.728***	-0.582***	0.034	0.057
	[0.087]	[0.090]	[0.045]	[0.046]
Hispanic	-0.715***	-0.723***	-0.008	-0.004
	[0.125]	[0.124]	[0.065]	[0.063]
Father professional	0.162**	0.094	0.002	-0.003
	[0.074]	[0.075]	[0.023]	[0.023]
Constant	-1.452**	-1.066	0.110	0.060
	[0.690]	[0.724]	[0.276]	[0.282]
Observations	3323	3323	974	974
Log likelihood	-1078	-1049	-129	-128

Standard errors are clustered at the household level

* significant at 5%; ** significant at 1%

D.2 Results for financial respondents

Table D3. Relevant heterogeneity in stock market beliefs. Estimates from the structural model Financial respondents, 2-point distribution for $\tilde{\sigma}$, low fixed at 0.15.

	Model w/o covariates		Model with covariates	
	Point estimate	SE*	Point estimate	SE*
Population average of $\tilde{\mu}$	−0.048	0.011	−0.038	0.056
Population standard deviation of $\tilde{\mu}$	0.147	0.010	0.202	0.094
Population average of $\tilde{\sigma}$	0.170	0.002	0.449	0.086

*Bootstrap standard errors

Sample: Health and Retirement Study, wave 2002. Financial respondents, age 55-64 (age of spouse also 55-64)

Table is analogous to table 2 in the main text

Table D4. Estimated mean of the structural parameters of stock market beliefs in various subpopulations. HRS 2002

Financial respondents, 2-point distribution for $\tilde{\sigma}$, low fixed at 0.15.

	Average $\hat{\mu}_i$	Average $\hat{\sigma}_i$
Top 25 per cent of lifetime earnings	0.134	0.434
Bottom 25 per cent of lifetime earnings	−0.078	0.481
Education college or more	0.061	0.461
Education high school or less	−0.074	0.469
Has DC pension (top 25% lifetime earnings)	0.140	0.437
Has DB pension (top 25% lifetime earnings)	0.092	0.471
Top 25 per cent of cognitive capacity	0.042	0.474
Bottom 25 per cent of cognitive capacity	−0.128	0.455
Father was manager or professional	0.047	0.442
Father had other occupation	−0.021	0.477
Top 25 per cent of risk tolerance	0.033	0.437
Bottom 25 per cent of risk tolerance	−0.141	0.504
Entire sample of financial respondents	−0.020	0.468
Total number of observations	2,313	2,313

Sample: Health and Retirement Study, wave 2002. Financial respondents, age 55-64 (age of spouse also 55-64)

$\hat{\mu}_i$ and $\hat{\sigma}_i$ are the subjective mean and subjective standard deviation of the one-year ahead stock return, predicted va

Table is analogous to table 3 in the main text

Table D5. Subjective stock market beliefs and stockholding at the extensive margin.

Financial respondents, 2-point distribution for $\tilde{\sigma}$, low fixed at 0.15.

	$\Pr(s_i > 0)$, partial effects		$E(s_i \mid s_i > 0)$	
	(1)	(2)	(3)	(4)
$\hat{\mu}_i$		0.862		0.355
		(0.097)**		(0.129)**
$\hat{\sigma}_i$		−0.158		0.170
		(0.115)		(0.146)
Log lifetime earnings	0.031	0.019	−0.003	−0.008
	(0.008)**	(0.007)**	(0.008)	(0.008)
Education	0.037	0.029	0.010	0.008
	(0.004)**	(0.004)**	(0.06)**	(0.007)
Cognitive capacity	0.060	0.039	−0.002	−0.013
	(0.013)**	(0.013)**	(0.018)	(0.018)
Log risk tolerance	0.015	−0.064	0.049	0.028
	(0.030)	(0.030)*	(0.035)	(0.036)
Single female	−0.132	0.025	−0.032	0.001
	(0.025)**	(0.034)	(0.031)	(0.042)
Single male	−0.108	−0.034	−0.028	−0.014
	(0.029)**	(0.031)	(0.032)	(0.038)
Female in couple	−0.026	0.101	−0.003	0.018
	(0.026)	(0.033)**	(0.032)	(0.039)
African American	−0.218	−0.155	0.036	0.058
	(0.027)**	(0.028)**	(0.044)	(0.045)
Hispanic	−0.210	−0.162	−0.004	0.002
	(0.043)**	(0.042)**	(0.067)	(0.064)
Other variables	YES	YES	YES	YES

Table analogous to table 4 in main text. Probit models (1) and (2); truncated regression models (3) and (4).

Sample: Health and Retirement Study, wave 2002. Financial respondents, age 55-64 (age of spouse also 55-64)

Clustered standard errors in parentheses; bootstrapped for models (2).and (4); ** significant at 1%; * significant at 5%

D.3 Results with freely estimated 2-point distributions for $\tilde{\sigma}$

Table D6. Relevant heterogeneity in stock market beliefs. Estimates from the structural model 2-point distribution for $\tilde{\sigma}$, low point estimated as well.

	Model w/o covariates		Model with covariates	
	Point estimate	SE*	Point estimate	SE*
Population average of $\tilde{\mu}$	−0.054	0.015	−0.036	0.010
Population standard deviation of $\tilde{\mu}$	0.210	0.027	0.178	0.013
Population average of $\tilde{\sigma}$	0.516	0.055	0.165	0.011

*Bootstrap standard errors

Sample: Health and Retirement Study, wave 2002. All respondents, age 55-64 (age of spouse also 55-64)

Table is analogous to table 2 in the main text

Table D7. Estimated mean of the structural parameters of stock market beliefs in various subpopulations. HRS 2002

2-point distribution for $\tilde{\sigma}$, low point estimated as well.

	Average $\hat{\mu}_i$	Average $\hat{\sigma}_i$
Top 25 per cent of lifetime earnings	0.703	0.164
Bottom 25 per cent of lifetime earnings	−0.004	0.164
Education college or more	0.054	0.164
Education high school or less	−0.041	0.164
Has DC pension (top 25% lifetime earnings)	0.071	0.164
Has DB pension (top 25% lifetime earnings)	0.068	0.165
Top 25 per cent of cognitive capacity	0.041	0.164
Bottom 25 per cent of cognitive capacity	−0.071	0.163
Father was manager or professional	0.042	0.164
Father had other occupation	−0.008	0.164
Top 25 per cent of risk tolerance	0.021	0.164
Bottom 25 per cent of risk tolerance	−0.055	0.164
Entire sample respondents	−0.036	0.164
Total number of observations	3,314	3,314

Sample: Health and Retirement Study, wave 2002. All respondents, age 55-64 (age of spouse also 55-64)

$\hat{\mu}_i$ and $\hat{\sigma}_i$ are the subjective mean and subjective standard deviation of the one-year ahead stock return, predicted va

Table is analogous to table 3 in the main text

Table D8. Subjective stock market beliefs and stockholding at the extensive margin. 2-point distribution for $\tilde{\sigma}$, low point estimated as well.

	$\Pr(s_i > 0)$, partial effects		$E(s_i \mid s_i > 0)$	
	(1)	(2)	(3)	(4)
$\hat{\mu}_i$		0.712		2.283
		(0.080)**		(0.265)**
$\hat{\sigma}_i$		−11.62		−37.2
		(13.95)		(44.7)
Log lifetime earnings	0.041	0.037	0.000	0.121
	(0.008)**	(0.008)**	(0.008)	(0.027)**
Education	0.033	0.026	0.011	0.080
	(0.004)**	(0.004)**	(0.05)**	(0.013)**
Cognitive capacity	0.069	0.054	0.002	0.176
	(0.011)**	(0.013)**	(0.014)	(0.037)
Log risk tolerance	0.023	−0.008	0.049	−0.027
	(0.026)	(0.027)*	(0.029)	(0.084)
Single female	−0.111	−0.033	−0.023	−0.107
	(0.027)**	(0.027)	(0.029)	(0.083)
Single male	−0.093	−0.051	−0.022	−0.164
	(0.029)**	(0.028)	(0.034)	(0.090)
Female in couple	−0.003	0.050	0.006	0.161
	(0.012)	(0.018)**	(0.013)	(0.056)**
African American	−0.233	−0.204	0.034	−0.655
	(0.027)**	(0.028)**	(0.045)	(0.089)**
Hispanic	−0.229	−0.227	−0.008	−0.729**
	(0.024)**	(0.038)**	(0.065)	(0.125)
Other variables	YES	YES	YES	YES

Table analogous to table 4 in main text. Probit models (1) and (2); truncated regression models (3) and (4).

Sample: Health and Retirement Study, wave 2002. All respondents, age 55-64 (age of spouse also 55-64)

Clustered standard errors in parentheses; bootstrapped for models (2).and (4); ** significant at 1%; * significant at 5%

D.4 Results with including all the belief-specific right hand-side variables (z) in all models with the other covariates (x)

Table D9. Relevant heterogeneity in stock market beliefs. Estimates from the structural model 2-point distribution for $\tilde{\sigma}$, low point fixed to 0.15. z_μ and z_σ are always included with x

	Model w/o covariates		Model with covariates	
	Point estimate	SE*	Point estimate	SE*
Population average of $\tilde{\mu}$	−0.066	0.018	−0.046	0.021
Population standard deviation of $\tilde{\mu}$	0.197	0.019	0.213	0.036
Population average of $\tilde{\sigma}$	0.576	0.077	0.524	0.089

*Bootstrap standard errors

Sample: HRS 2002, 55 to 64 years old financial respondents (partner is also 55 to 64)

Table is analogous to table 2 in the main text

Table D10. Estimated mean of the structural parameters of stock market beliefs in various subpopulations. HRS 2002

2-point distribution for $\tilde{\sigma}$, low point fixed to 0.15. z_μ and z_σ are always included with x

	Average $\hat{\mu}_i$	Average $\hat{\sigma}_i$
Top 25 per cent of lifetime earnings	0.065	0.546
Bottom 25 per cent of lifetime earnings	−0.092	0.535
Education college or more	0.042	0.539
Education high school or less	−0.090	0.532
Has DC pension (top 25% lifetime earnings)	0.073	0.540
Has DB pension (top 25% lifetime earnings)	0.049	0.576
Top 25 per cent of cognitive capacity	0.024	0.551
Bottom 25 per cent of cognitive capacity	−0.132	0.509
Father was manager or professional	0.030	0.520
Father had other occupation	−0.046	0.548
Top 25 per cent of risk tolerance	0.011	0.512
Bottom 25 per cent of risk tolerance	−0.142	0.562
Financial respondent in couple	0.032	0.511
Non-financial respondent in couple	−0.049	0.556
Entire sample	−0.038	0.537
Total number of observations	3,314	3,314

Sample: Health and Retirement Study, wave 2002. Respondents of age 55 through 64 (partner also 55-64)

$\hat{\mu}_i$ and $\hat{\sigma}_i$: subjective mean and subjective standard deviation of the one-year ahead stock return, predicted value

Table is analogous to table 3 in the main text

Table D11. Subjective stock market beliefs and stockholding at the extensive margin.

2-point distribution for $\tilde{\sigma}$, low point fixed to 0.15. z_μ and z_σ are always included with x

	$\Pr(s_i > 0)$, partial effects		$E(s_i \mid s_i > 0)$	
	(1)	(2)	(3)	(4)
$\hat{\mu}_i$		0.608		0.240
		(0.095)**		(0.117)*
$\hat{\sigma}_i$		−0.071		0.240
		(0.096)		(0.131)
Log lifetime earnings	0.037	0.033	−0.000	−0.003
	(0.008)**	(0.008)**	(0.009)	(0.012)
Education	0.030	0.023	0.009	0.007
	(0.004)**	(0.004)**	(0.05)	(0.005)
Cognitive capacity	0.054	0.042	−0.002	−0.011
	(0.011)**	(0.012)**	(0.015)	(0.015)
Log risk tolerance	0.026	−0.025	0.045	0.034
	(0.026)	(0.029)	(0.028)	(0.030)
Single female	−0.088	0.015	−0.017	0.010
	(0.023)**	(0.032)	(0.029)	(0.037)
Single male	−0.079	−0.030	−0.024	−0.021
	(0.029)**	(0.031)	(0.035)	(0.038)
Female in couple	0.005	0.078	0.017	0.018
	(0.013)	(0.020)**	(0.013)	(0.024)
African American	−0.233	−0.190	0.028	0.047
	(0.026)**	(0.030)**	(0.045)	(0.046)
Hispanic	−0.220	−0.216	−0.005	−0.006
	(0.039)**	(0.040)**	(0.063)	(0.061)
Other variables	YES	YES	YES	YES

Table analogous to table 4 in main text. Probit models (1) and (2); truncated regression models (3) and (4).

Sample: Health and Retirement Study, wave 2002. All respondents, age 55-64 (age of spouse also 55-64)

Clustered standard errors in parentheses; bootstrapped for models (2).and (4); ** significant at 1%; * significant at 5%