Version 4.4a, July 8, 2002
Presented at NBER Summer Institute on Aging, July 29, 2002

Who Becomes A Stockholder? Expectations, Subjective Uncertainty, and Asset Allocation

Preliminary. Do not quote.

Gábor Kézdi and Robert J. Willis
University of Michigan

Abstract. We develop a model of portfolio selection with subjective uncertainty and learning in order to explain why some people hold stocks while others don't. We model heterogeneity in information directly, which is an alternative to the existing explanations that emphasized heterogeneity in transaction costs of investment. We plan to calibrate the model to survey data (when available) on people's perception about the distribution of stock market returns. Our approach aslo leads to a model of learning with new implications such as zero optimal risky assets, or ex post correlation of uncorrelated labor income and optimal portfolio composition. It also points to two factors in probabilistic thinking that should have a major impact on stock ownership. These are the level and the precision of expectations. We construct proxy measures for the two parameters from the 1992-2000 waves of the Health and Retirement Study (HRS). We use a large battery of the subjective probability questions administered in each wave of HRS to construct an overall "index of optimism" (the correlated factor between all subjective probabilities) and "index of precision" (the fraction of nonfocal probability answers, following Lillard and Willis, 2001). We also construct measures for how people forecast the weather, their cognitive capacity, wealth, and basic demographics. Our results indicate that stock ownership and the probability of becoming a stockholder are strongly positively correlated with the indices of the level and precision of expectations. Interpretation of the former is quite challenging and further research is needed to understand its full content.

The research herein was performed pursuant to a grant from the U.S. Social Security Administration (SSA) to the Michigan Retirement Research Center. Additional support was received by the second author from NIA grant P0-1 AG10179. Data from the Health and Retirement Study used in this paper were funded by the National Institute on Aging with additional support from SSA.

When looking at portfolio choice, much of the theoretical finance literature has focused on how people should behave given their preferences. Considerably less is known about how people behave in reality. As part of the optimal choice, normative finance theory prescribes strategies for using the available information the best possible way. In contrast, a positive approach should ask what information people actually have when they make portolio choices, how they use that information, and how they acquire it. In this paper we aim at contributing to the latter approach. In particular, we try to understand how heterogeneity in the perception of uncertain outcomes translates into heterogeneity in portfolio choices, and how informational heterogeneity arises in the first place.

Our paper is related to recent developments in behavioral finance. Andrei Shleifer (2000) has shown that financial markets may not be efficient if there are agents with “erroneous” beliefs about the distribution of market returns. These agents he calls noise traders. Here we try to understand how these beliefs affect behavior and how they arise. Contrary to the behavioral approach, however, we stay in an expected utility framework. Conditional on their beliefs everyone behaves optimally, and beliefs are formed by learning in an optimal way. The model we develop is quite simple but it leads to interesting implications such as the inseparability of wealth and optimal portfolio structure even when preferences are isoelastic and asset returns are i.i.d. Throughout the paper we focus on the ownership margin. That is, we ask why some people hold certain assets while others do not. We show that, among other things, imprecise beliefs can lead to zero risky assets in the portfolio. Following Shliefer’s terminology, these agents we may call “noise non-traders”.

Another problem our paper helps to understand is the “stockholding puzzle”. The usual explanation for why some people hold stocks while others don’t is heterogeneity in transaction costs of investment (see, e.g., Halliosos and Bertaut, 1995; Bertaut, 1998; or Vissing-Jørgensen, 2001). Usually, these costs are defined in a broad way. Among other things, people’s imperfect knowledge about stock returns has been viewed as part of the costs of investment. In this paper we take a different approach. While acknowledging that some kind of fixed transaction costs are likely to affect investment decisions, we model heterogeneity in information in a more direct way.

There are two main advantages of this approach. First, we can look at survey measures of people's perception about the distribution of stock market returns and calibrate our model to get optimal portfolios. The data are not available yet, but we expect a lower and less heterogenous optimal share of risky assets than implied by the traditional approach (which would take historical data and use the same estimated process for everybody). Secondly, we can introduce learning into the portfolio selection problem. We show that the perspective of future learning can lead to zero risky assets in the portfolio in a setup where no learning would always imply a positive share. We have also started to model heterogeneity in learning, from exogenous factors (such as memory) and endogenous reasons (higher expected labor earnings makes it more worthwhile to acquire information). The extension helps to connect cognitive capacities to observed portfolio allocation. It may also help in understanding the ex post correlation of labor income (uncorrelated with stock returns) and portfolio allocation: people with higher expected earnings may find it worth acquiring more information and have a better understanding of the stock returns process, which may increase the share of stocks in their portfolio.

In the second part of the paper we examine the determinants of stock ownership using longitudinal data from the Health and Retirement Study (HRS). The major empirical novelty of our paper is that we relate the theoretical parameters to data on subjective measures of preferences and expectations from HRS. These include survey measures of risk aversion and measures of two aspects of probabilistic beliefs: the dispersion and the level of expectations ("precision" and "optimism"). For constructing these measures, we use a large battery of subjective expectations questions of HRS. Lillard and Willis (2001) have shown that one can interpret the propensity to give nonfocal answers to all subjective probability questions (measured by the fraction answers other than 0, 50, or 100 percent) as a proxy for the individuals' general precision when facing uncertainty. We use that measure as one of our measures for precision. Another measure we construct by looking at individuals' weather forecast (the probability that tomorrow will be a sunny day) and realized weather. The inverse of the absolute error people make is our second proxy for their general precision while dealing with uncertain events.

We use three measures for the level of people's expectations. One is their expectation about economic growth (the negative of the probability of a depression), while the two others are again general measures. The first one is an indicator whether they were overly optimistic in their weather forecast, on the assumption that sunny days are positive outcomes.[1] The second one is a common component of individuals' expectations through all domains. Basset and Lumsdaine (2001) have shown that answers to seemingly unrelated questions are significantly correlated. They interpret this phenomenon as a fixed effect that affects all of an individual's expectations. We take their idea one step further and extract this common component by factor analysis. Without any restrictions, the constructed index is significantly positively correlated with positive events and negatively correlated with negative events.

Perhaps the most intriguing result in our paper is the strong predictive power of the index of general optimism. While we label this variable "optimism," at this point we are quite agnostic about its content. It may reflect cognitive bias as well as optimistic beliefs that are justified by the individual's private information. Its significant positive correlation with both the actual sunny day answers and the optimistic weather forecast error indicate that it reflects "genuine" optimism, at least in part. On the other hand, we find strong evidence that people who give more optimistic answers to probability questions in general are healthier, wealthier, and more educated. This suggests that they are in part justified in having higher expectations. Although we relate the general index to expectations in our theoretical model, this second interpretation may have very little to do with probabilistic thinking. Instead, it may reflect a combination of lucky events in the past, general abilities, or past investments in human capital. At this point, we are not able to separate these effects, nor can we tell how much is cognitive bias and how much is "fundamentals" in the index. We leave these problems for future research.

Note that except for expectations about economic growth, none of our measures is related to stock returns. Rather, they reflect some person-specific components in how people form expectations over uncertain events.[2] In light of this we find especially remarkable the fact that all of the above measures predict stock ownership. Except for the weather forecast indices, they remain strong predictors even after controlling for demographics, wealth, and cognition. Risk aversion has no predictive power probably because it is measured with especially large error. Besides improving the predictive power of models of stockholding, expectations over economic growth, general optimism, and general precision seem to explain a significant part of the effects of other variables. The predictive power of education, race, and wealth are substantially reduced

[1] The idea of using the HRS sunny day question for measuring optimism was first introduced by Lumsdaine (1999).

when the expectations variables are also included. If we interpret these measures as proxies for probabilistic thinking, the results indicate that part of what used to be regarded as fixed costs of investment are indeed operating through people's expectations. Alternatively, they may be better indicators of cognitive abilities, wealth, health, and human capital than the conventional measures. The results are very robust. They hold for ownership as well as buying and selling stocks; they are robust to a very flexible specification for wealth effects; and they are remarkably stable across time.

The paper is organized the following way. The first part presents the theoretical model of portfolio choice. The second part discusses the various measures we use in explaining household choices. The third section presents the empirical model and the results, and the last part concludes.

1. Portfolio Selection With Subjective Uncertainty and Bayesian Learning

In this section we present a model of portfolio selection under subjective beliefs, uncertainty, and learning. Our goal is to introduce a simple analytical framework to facilitate our empirical investigations. In particular, we would like to relate stock ownership to expectations about asset returns and the precision of those expectations. We keep things as simple as possible. The model is an application of the well-known continuous time portfolio choice model of Merton (1969), augmented with subjective beliefs and Bayesian updating as derived by Gennotte (1986) and Brennan (1998). We allow for heterogeneity and look at the results from the angle of stock ownership, an application not considered in the previous literature.

1.1 Portfolio selection with known parameters

Consider an individual who saves for retirement. For simplicity, assume that at time 0 she has wealth W_0 to invest and she wants to maximize the expected utility of W_T, her wealth when she retires at some predetermined time T. Assume that she has a conventional constant relative risk aversion (CRRA) utility function with γ being the parameter of relative risk aversion. She can choose between investing into the risk-free asset with known instantaneous

---

[2] HRS 2002 contains questions about subjective expectations of stock market returns and more detailed questions in an experimental module. In the future we intend to use that data to tie the empirical work more closely to theory.

rate of return r and one risky asset. The instantaneous rate of return of the risky asset, denoted by dS/S, is assumed to follow a Brownian motion with constant mean μ and variance σ^2. For the time being we assume that μ and σ^2 are known constants. The investment decision consists of choosing an optimal fraction of wealth invested into the risky asset for each time t between 0 and T, which we denote by α_t.

The equation of motion for the instantaneous return to the risky asset is given by

$$\frac{dS}{S} = \mu\, dt + \sigma\, dz\ , \tag{1}$$

where dz is the increment to a standard Wiener process. This is a continuous time generalization of a random walk with drift, where the instantaneous drift is μ and the variance is σ^2. Throughout the analysis we assume that the investor knows the random walk nature of the process and that its parameters are constant. For now we also assume that she also knows the parameters themselves, an assumption we will relax later.

With fraction α_t of wealth W_t invested into the risky asset at each time t, wealth also follows a geometric Brownian motion given by

$$\frac{dW}{W} = \left(r + \alpha_t\left(\mu - r\right)\right) dt + \alpha_t\, \sigma\, dz\ , \tag{2}$$

where r is the known instantaneous rate of return on the risk-free asset.

Subject to this budget constraint, the investor's problem is

$$\max_{\alpha_t} E_t \frac{W_T^{1-\gamma}}{1-\gamma}\ . \tag{3}$$

Assuming that $\gamma > 1$,[3] the standard solution to this problem (Merton, 1969) is a constant fraction of wealth invested into the risky asset

$$\alpha^* = \frac{\mu - r}{\gamma\sigma^2}. \tag{4}$$

The optimal share invested into stocks is increasing in its mean return, decreasing in the return of the risk-free asset, and decreasing in the variance and the degree of risk aversion.

[3] If the coefficient of relative risk aversion is smaller than one and $\mu > r$, it is optimal to hold the entire portfolio in the risky asset. The reason is that if $\gamma < 1$ the concavity of the utility function is not sufficiently strong to offset the convex relationship between terminal wealth and the rate of return caused by compounding.

1.2. Stockholder Puzzles

The simple and elegant result in (4) comes at the cost of being at odds with a number of empirical regularities. In particular, people in the world of the model choose the optimal portfolio at time zero and never change it, implying that no one becomes a stockholder at a later date or sells off an initial holding. In addition, α^* is always positive if the expected return is higher than the rate on the risk-free asset ($\mu > r$) Given empirical evidence that stockmarket returns are high –indeed, so high as to create the "equity premium puzzle" (Mehra and Prescott, 1985)–the theory predicts that everyone, no matter how risk averse, should hold a positive fraction of the risky asset in their portfolio. Moreover, the "Tobin separation theorem" (Tobin, 1958) suggests that the composition of the optimal portfolio should be independent of the optimal level of wealth.

Microeconomic data on stockholding as well as conventional financial counseling and advice contradict these implications for households engaged in long term saving for retirement. (See Campbell and Visceira, 2002, for an excellent survey of the literature.). We illustrate some of the empirical patterns of stock holding using five waves of longitudinal data from households of Health and Retirement Study (HRS) participants born in 1941-51 who were age 51-61 at the first wave in 1992. (The data are described in more detail below in Section 2.) Table 1 shows that only about one-third of households own stocks directly.[4] The table also shows that there was modest growth in this fraction between 1992 and 2000 from 32 percent to 36.5 percent as this cohort was approaching retirement during a period of high stock market returns. Table 2 shows a large variation in stockholding by education but growth in ownership at every education level. The fraction of respondents with less than high school grew from 11.0 to 13.6 percent between 1992 and 2000 compared with growth from 55 percent to 60 percent among the college educated. Finally, Table 3 shows evidence of considerable mobility in stock ownership from wave to wave. Between 1992 and 1994, for example, 11.2 of households who did not own stocks in 1992

[4] In addition to direct ownership, households may hold stocks indirectly in IRAs and Keogh's, 401k accounts or defined pension plans. They might also hold other forms of risky assets such as business assets. However, even under a broader definition of risky assets a large fraction of households are non-owners. In this paper, we restrict our focus to the direct ownership of assets.

became owners by 1994 while 22.0 percent of owners became non-owners. These patterns show considerable stability over time.

1.3 Portfolio Selection with Subjective Uncertainty

As noted in the Introduction, heterogeneity in transaction costs has been used to explain heterogeneity in stock ownership. Researchers usually interpret those costs in a very broad sense. In particular, they include the costs of acquiring and processing information. We take a different approach here, using a model by Brennan (1998) which augments the Merton model by considering individuals who are uncertain about the parameters of process governing stock returns and who learn through Bayesian updating . By modeling heterogeneity in beliefs of the returns process directly, we show that Brennan's model can explain ownership differences without differences in transaction costs. More generally, our goal is to introduce a simple analytical framework to facilitate our empirical investigations. In particular, we would like to relate stock ownership to expectations about asset returns and the precision of those expectations.

Suppose that the investor does not know all the parameters of the returns process (μ and σ^2). To keep things simple, assume that she has a one-point belief about σ^2 but a distribution over μ. In particular, at time zero, she thinks that μ is drawn from a Normal distribution with mean m_0 and variance v_0. v_0 represents subjective uncertainty of those expectations. The inverse of v_0 is often called the precision of the beliefs. We will use the concept of precision in this sense, and will also refer to v_0 as the degree of prior imprecision.

With time she observes the realized returns, regardless of having invested into stocks or not. Based on the observed series, she continuously updates her belief about the distribution of μ.[5] The updated μ (conditional on the realization) is Normal, with parameters m_t and v_t, the former being a diffusion process itself while the latter is a deterministic function of time:

$$dm = \frac{v_t}{\sigma^2}\left(\frac{dS}{S} - m_t\, dt\right) \tag{5}$$

$$dv = -\frac{v_t}{\sigma^2}\, dt \ , \tag{6}$$

[5] We assume that she does not update her beliefs about σ^2. If the price path induced by the Brownian motion is continuous, it is possible to estimate σ^2 over any short interval of data, perhaps justifying treating σ^2 as a known parameter.

where (6) can be solved to get

$$v_t = v_0 \exp\left(-\frac{t}{\sigma^2}\right). \tag{7}$$

These results were first derived by Lipster and Shiryayev (1978) and were used by Gennotte (1986) and Brennan (1998) for the portfolio selection problem.

Since the perceived parameters of the stochastic process of stock returns (and therefore wealth) change over time, the optimal fraction invested in risky assets will also vary with time. However, the investor's problem can be separated into first updating the parameters and then making the choice based on the posterior. As Brennan (1998) shows, the solution to this new problem is quite complicated. On the other hand, it can be represented by a closed form function ψ of γ and m. The optimal rule, then, is given by

$$\alpha_t^* = \frac{(m_t - r)}{\gamma\sigma^2} + \frac{v_t\psi(m_t, \gamma)}{\gamma\sigma^2}. \tag{8}$$

Note that the first term on the right hand side of (8) is the conventional expression for the optimal portfolio share from Merton's model where m_t the expected rate of return on stocks, given current information. The second term in (8) represents what Brennan calls an "intertemporal hedging demand" for the risky asset that arises from subjective uncertainty about the "true" rate of return.[6] Brennan shows that sign of the hedging demand depends on the degree of risk aversion. Specifically, ψ is positive for $0 < \gamma < 1$ (mild risk aversion), zero for $\gamma = 1$ (logarithmic utility), and negative for $\gamma > 1$ (strong risk aversion).

1.4 Heterogeneity and the ownership margin

Although Brennan does not emphasize this, it is possible that parameter uncertainty and potential learning may be of sufficient importance that strongly risk averse persons with $\gamma > 1$ may choose to hold no stocks, even if the expected return on the risky assets (m_t) exceeds the risk-free return. If sufficiently strong parameter uncertainty (large v_t) is accompanied by

[6] Note that hedging demand goes to zero as t goes to infinity since, from (7), uncertainty about the rate of return, v_t, disappears asymptotically as data on the history of stock prices increases.

sufficiently strong risk aversion (leading to negative ψ), the expression in (8) can become nonpositive. Ruling out short sales so that α_t^* is non-negative, strong risk aversion and sufficient parameter uncertainty may therefore lead to a zero fraction invested into the risky asset.

Since this implication is central to the concerns of this paper, it is worth developing the intuition underlying it. To do so, we consider an extremely simple model in which an individual with CRRA utility seeks to maximize retirement wealth, which is consumed at the end of the final period. The return on the risky asset for person i in period t, m_{it}, can be either $\mu+\delta_{it}$ or $\mu-\delta_{it}$, with equal probability. Assume that the decisionmaker lives for two periods and maximizes her wealth at the end of period two (W_2) by choosing the share of risky assets in periods 1 and 2. Crucially, assume that she can reoptimize the portfolio after period 1. Thus, let her value function be

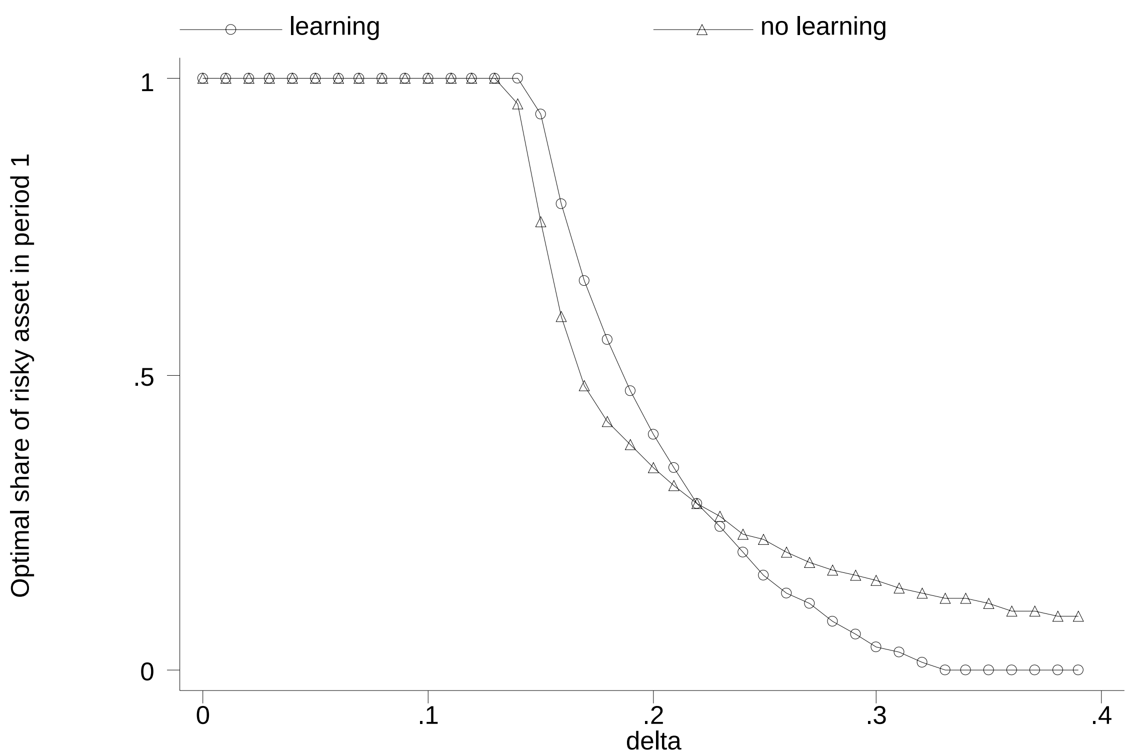
$$J\left(W_0\right)=\max_{s_1,s_2} Eu\left(W_2\right)$$

subject to the constraints

$$\begin{aligned}
&u\left(W_{i2}\right)=\frac{W_{i2}^{\;1-\gamma}}{1-\gamma},\\
&W_0 = given,\\
&W_t = W_{t-1}\left(1-\alpha_t\right)\left(1+r\right)+W_{t-1}\alpha_t\left(1+m_{it}\right),\quad t=1,2,\\
&\Pr\left(m_{it}=\mu+\delta_{it}\right)=0.5,\quad \Pr\left(m_{it}=\mu-\delta_{it}\right)=0.5.
\end{aligned}$$

We consider two cases: one in which δ_{it} is *i.i.d.* so that observing m_{i1} in period 1 does not tell anything about the future. In this case, there is no role for learning. In the second case, $\delta_{i1}=\delta_{i2}=\delta_i$ so that observing m_{i1} leaves no uncertainty for period 2. This is an extreme case of learning. The question we ask is how the possibility of learning affects the optimal allocation in period 1 before learning takes place. The problem is solved by backward induction.

Optimal allocation in period 1 takes into account the best choice in period 2. Figure 1 shows an example for the optimal share of the risky asset in period 1 by possible values of δ in the two cases where we assume r=0.05, μ=0.10 and γ=2. A larger δ corresponds to larger uncertainty, therefore the share of the risky asset is nonincreasing in δ. The picture for larger values of γ are similar, with steeper curves which start to decline at lower values of δ.. The two curves cross at some low optimal share. The possibility of learning makes it worth holding more risky assets in period 1 above that point, and less below it. For large enough δ, the share of risky assets reaches zero if learning is expected in period 2, and it stays positive no matter how large the risk if learning is not possible.



Why does the potential for learning lead to the possibility that it is optimal not to own any stock even if transaction costs are zero? To answer this question, it is useful first to review why a positive share will always be optimal if utility is CRRA, the expected return is greater than the risk free rate, and there is no learning. The reason is that, in the neighborhood of $\alpha = 0$, the investor has eliminated all but an arbitrarily small amount of risk from her portfolio. Since risk aversion is a second-order phenomenon, the investor should place at least some small amount of the risky asset in her optimal portfolio if its expected rate of return exceeds the risk free rate.

The situation is different if learning can occur and the investor can alter her portfolio in light of new information. The assumption that learning is both possible and probable implies that the investor will not eliminate risk from terminal wealth even if she chooses a totally "riskless" portfolio in the initial period (i.e., chooses $\alpha_1 = 0$). The reason is that if sufficiently good news about stock returns occurs in the future, such an investor knows that she will then buy some stock because the expected rate of return has increased and risk has been reduced while, if bad news occurs, she will choose to remain fully invested in the safe asset. Thus, from the perspective of period 1, uncertainty about the return parameter implies that the investor's terminal wealth is stochastic even if she currently holds only safe assets. A risk averse person will attempt to minimize risk, but with parameter uncertainty and future learning, it is not optimal to eliminate all risk because of the option value of acting on new information. Hence, sufficiently risk averse persons people will be at a corner with $\alpha_1^* = 0$. Note that this is true whether the the source of the new information is changes in stock market prices, as in Brennan's model, or any other information that affects the individual's subjective beliefs about long term stock market returns such as news of the latest accounting scandal, the next big thing on the Internet or the anticipation of hot stock tips from Uncle Harry who is scheduled to visit next week.

1.4 Heterogenous learning

We can extend the model to allow for heterogenous (and possibly endogenous) learning the following way. Instead of observing δ, the investor tries to estimate it using a number of signals. This way there remains uncertainty in period 2 because estimates have errors, but the uncertainty is reduced. The better the estimate the smaller period 2 uncertainty. The number and quality of signals affects the quality (bias and precision) of the estimate. Assume that all signals are informative, equally noisy and can be described by

$$\delta_{ij} = \delta + v_{ij}, \quad E\left(v_{ij}\right) = 0, \quad Var\left(v_{ij}\right) = \sigma_v,$$

where i denotes the individual and j corresponds to each and every signal. After observing N_i signals, individual i estimates δ by simply taking a sample average:

$$\hat{\delta}_i = \frac{1}{N_i}\sum_{j=1}^{N_i}\delta_{ij} \ .$$

The distribution of estimates among individuals is normal with mean δ and variance $\sigma_i^2 = \sigma_\nu^2 / (N_i - 1)$. Apart from randomness, differences in learning are results of heterogeneity in the number of signals N_i. The more signals one observes the better (more precise) are the estimates.

The effects of cognitive capacities such as memory can be modeled through the sample size: $N_i = N(x_i)$, where the x_i are the personal characteristics that affect the number of signals individual i observes. We can also think of N_i as a choice variable (possibly still affected by cognitive capacities): people can buy a sample of signals. In such a model they do so in order to form a better view about the parameters of the return process (here δ). We can model the role of the exogenous factors x_i as factors affecting the cost of a sample. This way we introduced another decision into our problem: the decision maker has to choose an optimal sample size N_i^*, by contrasting its costs to its expected benefits.

We have not worked out the solution in this version of the paper. The endogenous learning model is probably more interesting if we allow for uncertain non-asset income (earnings) to play a role. In such a model, people with higher expected earnings in period 2 may find it more beneficial to have more precise estimates and therefore have an incentive to buy more signals. Better estimates have a positive effect on the optimal share of risky assets in period 2. Therefore such a model would imply an ex post correlation between income and asset allocation even if labor income is uncorrelated with stock returns.

The learning models may help to understand recent stock market phenomena. The story of the 1990's ("the rise of the equity culture," Poterba, 2001) can be interpreted as the result of a series of unusually high signals. This made some people invest in stocks who would not have done so if signals had been less encouraging. Some argue that people who started investing into stocks started to learn more quickly, which reduced their transaction costs and encouraged further investments, creating the emergence of the "equity culture". In our model this phenomenon will probably arise if learning (the amount of information) is larger for stockholders.

2. Measuring Parameter Heterogeneity

The empirical novelty of this paper lies in the various survey measures we use to proxy for the heterogeneous parameters that determine stock ownership. In this section we describe those measures in detail. We consider four types of measures: those that try to capture heterogeneity in prior expectations; those that we think are related to prior precision (inverse of σ_i^2); risk preferences (γ); and general cognitive abilities, including education. This last group is not part of our theoretical model yet. While the model allows for heterogeneous (if not updated) beliefs about the variance, we do not try to measure that here. The focus of our paper is on prior expectations and the precision of beliefs. We define various measures for them but one is common to all: they are based on the expectations questions of the Health and Retirement Study (HRS). The HRS is a large household panel with detailed information on cognition, expectations, and asset ownership. In this paper, we use data on the initial HRS cohort of 12,670 persons born in 1931-41 (plus spouses) who were first surveyed in 1992 and have been resurveyed every two years. The latest available data is from the "early release" for 2000. See Juster and Suzman (1995) and Willis (1999) for more detailed descriptions of the HRS studies.

First, we describe our measures for expectations, next we turn to measures of precision of beliefs. The third subsection describes the measure of risk aversion, and the fourth subsection defines the cognitive measures we use. The last subsection looks at how all of these measures are related.[7]

2.1 Expectations

We use three measures to proxy prior expectations. One is specific to the expected growth of the economy and is therefore more directly connected to returns on the stock market, while the two others are general measures of expectations about different kinds of events.[8]

[7] Note that in most cases we do not try to relate the magnitude of the measures to the theoretical variables. The reason is that we do not have explicit measures for expectations about stock market returns, and the measures we use are therefore not related directly to the parameters. Instead, they serve as proxies for them with artificial units of measurement. The only quantitative comparisons we make are to show which measures are the most powerful in predicting behavior.

[8] HRS 2002 will include an experimental module on expectations about the stock market. Those data will allow us to capture expectations and their precision in a more direct way. In addition, they may help us capturing heterogeneity in beliefs about the variance.

Each wave of HRS contains a large number of questions about expectations about various future events.[9] HRS asks the following question in each survey: "What do you think are the chances that the U.S. economy will experience a major depression sometime during the next 10 years or so?" For ease of interpretation, we use the probability of the complementary event, that is one minus the subjective probability of a depression. For each period we examine, we use the beginning of period answer to this question as a proxy for prior expectation. Since in 1998 the question was asked only for those who were not interviewed before,[10] we use 1996 values when explaining changes from 1998 to 2000. Table 5 shows the summary statistics of the variables (the subjective probability of no depression), scaled to be between 0 and 1. On average, expectations increased significantly between 1992 and 1994 but did not change much between 1994 and 1996. This is finding is consistent with the substantial increase in stock ownership between 1992 and 1994 but not between 1994 and 1996. The variance stayed by and large the same. Through the years, expectations were positively correlated: those who were more optimistic than average in 1992 were likely to stay so.

Besides expectations about economic growth, we are interested in high expectations, or optimism in general. Bassett and Lumsdaine (2001) show that some people have systematically higher and others systematically lower expectations about future events in the HRS. Expectations are correlated across events that do not seem to be related at first sight. They argue that these person-specific "fixed effects" reflect some otherwise unmeasured heterogeneity. We take their approach one step further, by extracting the common component and using it for our analysis. It is important to note that we capture this person-specific component in a mechanical way, without any assumptions about the importance of the different domains of expectations and the correlation between them. We have extracted the correlation between all subjective probability questions by factor analysis. Since not all probability questions were asked from everybody, the set of questions that identify the factor may be different for different individuals.[11] We have generated factors of optimism for all waves together and for each wave separately. Most of the analysis will focus on the all-waves measure, while the year-by-year

[9] HRS started in 1992, and it interviews people in every two years. Data from 1992, 1994, 1996, and 1998 are publicly available, and an early release of the 2000 survey can also be used.
[10] Same was true for HRS 2000.
[11] Technically, we solved the missing value problem by standardizing all variables and replacing missing values by zero (the standardized mean). Note that we standardized first and imputed the missing values second so that the variance of the transformed variables is a decreasing function of the number of missing cases.

index we will use for robustness checks only. The information content of the all-waves measure may be different because not everybody is part of every sample. The factor analysis basically created an artificial variable with mean zero and variance 1 according to the equation

$$a_{ij} = \lambda_j f_i + \omega_{ij} \,. \tag{9}$$

i is the index for individuals, *j* is the index for the probability question (in the different waves), the *a* are the standardized answers to the probability questions, and the λ_j are the "factor loadings". f_i is the common unobserved component for individual *i* across all probability questions, while ω_{ij} contains the components that are unique to the particular question. These equations can be interpreted as linear regressions for each question *j*, where regression constants are zero because everything is standardized, and we estimate f_i on top of λ_j. We used the principal factor method that chooses the solution that predicts the original covariance matrix best.[12] At the end $\hat{f}_i$, the artificially created variable captures how individual *i*'s expectations differ from the average, through all probability questions, and we call it the *index of optimism*.

Table 6 contains the pairwise correlation of the index of optimism with the individual questions. Note that we created the index without restricting the sign and magnitude of these correlations. Quite remarkably, the correlations between the index and the different answers are very stable across waves. Moreover, the factor is positively correlated with positive events and negatively correlated with negative events. This is also true for general events like the probability of sunshine or inflation (except for the questions about Social Security), which are outside the control of the individuals and don't contain private information. This is the reason why we labeled the extracted factor "index of optimism".

Although we call it optimism, we are quite agnostic about the content of the latent variable we proxy with this index. In principle, optimism can reflect cognitive bias in expectations or expectations that are brighter but justified by the individuals' information. This latter may occur in the context of very general events, too: people who know the economic history of the U.S. better may think that there is less chance of a double-digit inflation in the near future. In this second interpretation, the index of optimism conveys information about some

[12] The covariance matrix of the expectation variables is not equal to their correlation matrix because their variance is decreased by the missing observations. This way questions with more valid observations were given a larger weight than those with less valid observations.

“fundamental” heterogeneity across people.[13] The index is strongly correlated with the survival probabilities, but correlations with the probability of giving help, leaving inheritance or that income will keep up with inflation are also substantial. It is conceivable that we capture health effects. It is also possible that we capture wealth effects. The reverse relationship with the Social Security questions suggest that being well informed might be part of the story (more informed people may think that it is more likely to become less generous), although those effects are small. The results support the fundamental heterogeneity story, but the positive correlation with sunshine suggests that cognitive bias may also play a role. People with higher value of the optimism index have higher expectations than others probably because they are healthier, wealthier and wiser, but partly just because they see things brighter.

We have tried to explore the content our index of optimism by examining its relationship with other observable variables, especially those related to health, wealth, and demographics. Table 7 shows the mean of the optimism factor by categories of self rated health, for each wave of HRS. Recall that the overall mean is 0, and the variance is 1. The results indicate that our factor is strongly positively related to how well people feel. Table 8 presents regression results where the left-hand side variable is the optimism factor, and the right-hand side variables are usual demographic covariates, religion, and wealth (total net worth in 1992). In a separate equation we also enter a factor we created from the self-rated health using all waves (larger values of which mean better health). Education, age, wealth and the factor of health are standardized, the others are binary variables. Recall that the left-hand side variable is also standardized. The results show that more educated, female, younger, and white people are more “optimistic.” The same is true slightly for Catholics, and strongly for Jews. These results weaken somewhat when subjective health is entered into the regression. The race coefficients basically drop to zero, the effect of education drops by a third, and the effect of age drops by two thirds. Subjective health becomes the most important determinant of the latent variable, and wealth is also a significant predictor. These results support our hypothesis that the optimism factor reflects something “real”: people who answer more positively the subjective probability questions in general are healthier, wealthier, and somewhat more educated. They are also more likely to be female and Jewish, which is less straightforward to interpret.

[13] In principle, the index of optimism could contain information about “fundamental” differences if a positive bias can actually feed back to behavior. Unfortunately, the direction of the impact is not clear as one could imagine both

In terms of our theoretical model, we relate the index of optimism to prior expectations about the mean return. Note, however, that it can capture elements that may correlate with learning abilities (better knowledge about the economy) or the incentive to learn (life expectancy). We leave these possibilities for future research.

We also construct an alternative measure for general optimism, one that might be closer to the cognitive bias interpretation. For this measure, we use answers to the 1994 and 2000 questions about the probability of a sunny day the day after the interview. We match the respondents with the closest weather station and use actual weather data for the day in question.[14] We then regress the subjective probability variable on the observed fraction of sunny hours (number of registered sunny hours divided by daylight for the day in the year at the given latitude). A positive residuum corresponds to an optimistic forecast: the respondent thought that it would be sunnier than it turned out to be. The distribution of these errors is slightly skewed with some 57 percent of the respondents giving a sunnier forecast than the actual weather. We use the sign of this (reverse) forecast error as an indicator variable as another measure of optimism. We also created an average measure for the sign from the two observations we have (1994 and 2000): it is zero if the forecast was overly pessimistic both years, one if it was overly optimistic both years, and 0.5 if the two years gave a different sign. Our rationale for having this alternative measure is that forecast errors are probably more closely related to cognitive bias, whereas the general optimism factor probably reflects more fundamental heterogeneity.

2.2 Subjective uncertainty

Our strategy to measure prior subjective uncertainty is based on the assumption that uncertainty about future stock returns is correlated with uncertainty about other future events. Moreover, we treat this general uncertainty component as a fixed individual trait that stays constant over time (at least for the eight years of the survey). One way to formalize this is to assume that prior subjective uncertainty (σ_i^2) can be decomposed into two parts: one that is

---

a negative feedback (carelessness) and a positive one (positive attitude).

[14] The matching was based on the latitude and longitude (of the center of the Zip-code area) of the respondents' residence and the latitude and longitude of the weather station. The number of land-based weather stations was cut substantially in the late 1990s, leaving 78 stations for 2000 out of the 152 that existed in 1994. The matching produced good results in 1994 but less so in 2000. The average distance to the closest weather station was about 50

specific to the event (σ_{ui}^2) and one that is common across all events for any individual ($\sigma_{\delta i}^2$). This second component may be thought of as some kind of a cognitive trait. In that interpretation, some people understand uncertainty better than others. In our case, we assume that σ_i^2, individual *i*'s subjective parameter variance about stock returns, is the sum of an individual and domain-specific variance term and a variance term that is fixed for each individual through all domains:

$$\sigma^2{}_i = \sigma^2{}_{ui} + \sigma^2{}_{\delta i}\,. \tag{10}$$

Our two measures are proxies for the common variance term $\sigma^2{}_{\delta i}$.

The first measure we use is the fraction of nonfocal probability answers, where focal answers are 0, 50-50, or 100 percent. Lillard and Willis (2001) show that if people form a subjective distribution about the probabilities and report the mode of this distribution when asked in a survey, the fraction of focal answers directly measures $\sigma^2_{\delta_i}$.[15] Even if we do not rely on the "modal choice" hypothesis of survey response, this measure is intuitively appealing in that answers of zero, fifty or a hundred percent probably reflect a very crude understanding of probabilities. Table 9 describes the fraction of exact (nonfocal) probability answers for our sample. A higher value of this index reflects higher degree of precision over all the subjective probability questions in HRS.

As an alternative measure of general subjective uncertainty, we also use the absolute magnitude of the 1994 and 2000 forecast error from the sunny day question when compared to actual weather based on the regression described above. The absolute error varies between 0 and 0.77, with mean 0.25 and standard deviation 0.15 (these numbers are basically the same for the two years). We also created an average absolute error from the two waves. In order to make interpretation easier, we recoded the variables in such a way that a larger value means higher precision (we subtracted the variables from their maximum). We label this variable "precision in weather forecast".

---

miles, with 90 percent of the people being within 100 miles. The same numbers for 2000 are 130 and 300 miles, respectively.

[15] Lillard and Willis present evidence confirming the internal validity of this interpretation of focal responses to the probability questions. They argue on theoretical grounds that individuals who are more uncertain (i.e., have less precise beliefs about probabilities) will tend to behave more risk aversely. Consistent with this hypothesis, they find the fraction of focal answers is negatively related to the presence of risky assets in a household's portfolio and to the rate of growth of the portfolio's value over time.

2.3 Risk aversion

HRS measures risk aversion based on responses to hypothetical gambles over lifetime earnings. Barsky, Juster, Kimball, and Shapiro (1997) show that those measures contain reasonable information about risk preferences. HRS 1992 asks the questions from the whole sample. In 1994, 1998, and 2000 the risk preference questions were administered for subsamples of 6-8% of the HRS age eligible. There was no such question in 1996. Table 10 describes the measures in four risk aversion category for the age eligible, in the four waves of HRS where the question was asked. The distribution is stable over time, with around 60 percent of people showing very strong risk aversion. They would turn down a gamble that would double their lifetime earnings or could result in a 20% decrease, both with a fifty percent chance. The distribution is even across the three other categories.

Based on the 1992 and 1994 surveys, Barsky et al. (1997) also document that the risk preference measure is quite noisy. Although we do not present the corresponding figures, comparing the 1994, 1996, and 1998 measures also supports the presence of substantial measurement error. Despite the noise, Barsky et al. show that the measure predicts risky behavior such as smoking, heavy drinking, or not having health or life insurance, after controlling for demographics. They also show that the measure predicts stock ownership.

2.4 Cognition

One factor that probably affects learning is cognitive abilities. For measuring learning abilities, we use education and the various cognitive measures of HRS. In the context of our current setup, these measures can be a thought of as proxies for prior parameter uncertainty: those with weaker memory or analytical and numerical abilities might have learned less so they have more diffuse priors at the beginning of our time period.

We use all cognitive test scores in HRS: items from the WAIS IQ test contained in HRS 1992 (they ask respondents to define the relationship of two different things like an orange and a banana, or praise and punishment); immediate and delayed word recall tests from all waves; counting back by seven from 100; and screening questions for dementia 1996, 1998, and 2000. The latter include questions about the date (day, month, and year), the day of the week, the President and Vice President of the U.S., and naming two things after hearing their definition (scissors and cactus). We extracted one common factor in each of the four groups (IQ, word recall, counting back by sevens, dementia screen) for all survey waves together, and we also created a factor from all cognitive questions. Table 11 shows the correlation of the different factors with each other. All factors are positively correlated, and the overall factor shows a strong correlation with the individual factors. Table 12 shows the correlation of the overall cognition factor with each of the items in each wave. The correlation is always positive and statistically significant. The results suggest that the overall cognition factor captures both memory properties (word recall) and numerical abilities (counting back by sevens). Its relationship to IQ type questions and the dementia control variables is weaker, though. Therefore, we shall use the domain-specific factors separately in our analysis.

2.5 Relationship of the different measures

Before we turn to how the above-defined measures predict who becomes a stockholder, we examine how they are related to each other. Recall that all variables are defined so that they reflect higher expectation/optimism, higher precision (lower uncertainty), more risk tolerance, and higher cognitive abilities. For the weather forecast error variables we used the two-wave averages.

Table 13 shows the correlation between the different measures. Risk tolerance is uncorrelated with everything else, but all the measures of optimism and precision are positively correlated with each other and with cognition. This suggests a cognitive interpretation of the measures, including the optimism factor and the index of exact probability answers. The weather forecast errors behave the way we would expect if the optimism factor contained elements of cognitive bias or the weather forecast was influenced by general well being. The sign of the error is more correlated with the optimism factor than the index of exact probability answers, while the (inverse of the) magnitude of the error behaves the opposite way. At the same time, the small correlations for both of the weather variables probably reflect a large amount of noise in them.

3. Estimation

We estimate three models: the probability of being a stockholder in the baseline survey (1992), the probability of becoming a stockholder between two survey years conditional on not being one at the beginning of the period, and the probability of selling all stocks between two survey years. The results we present focus on stock ownership outside retirement accounts; all qualitative results hold for broader definitions, too.

3.1 The empirical models

We estimate linear probability models for easier interpretation. Probit and logit counterparts give essentially the same results. We estimate the following regression:

$$s_{it} = \alpha_0 + \alpha_m ' m_{it} + \alpha_x ' x_{it} + u_{it} \, , \tag{11}$$

$$\Delta^+ s_{it} = \beta_0 + \beta_m ' m_{it} + \beta_x ' x_{it} + v_{it} \, , \tag{12}$$

$$\Delta^- s_{it} = \gamma_0 + \gamma_m ' m_{it} + \gamma_x ' x_{it} + w_{it} \, , \tag{13}$$

where s_{it} denotes stock ownership (0 or 1), $\Delta^+ s_{it} = s_{it+1} - s_{it}$ (0 or 1) is and indicator for becoming a stockholder conditional on no stocks at the beginning of the time period, and $\Delta^- s_{it} = s_{it+1} - s_{it}$ (0 or -1) is an indicator for selling all stocks conditional on having stocks at the beginning of the time period.

The x_i are demographic variables (age, coupleness, gender if single, education, race), together with the initial (t) level of total wealth. m_{it} is the vector of subjective measures: three measures of expectations (economic growth, the index of optimism, and optimistic weather forecast); two measures of precision of beliefs (index of precision taken over all probability variables and inverse of the absolute weather forecast error); and the measure risk aversion. *m* also contains measures of cognition. These variables are time-invariant except for expectations over economic growth and total wealth, both of which reflect the situation at the beginning, *t*. When we replace the time-independent variables such as the indices of optimism and precision by measures for each year (*t*), the results remain qualitatively the same but the magnitudes drop. This is consistent with wave-by-wave indices being a more noisy measure for the same fixed latent variable.

We estimate (11)-(13) for each time period and also for a pooled sample of all waves. The pooled data is an unbalanced panel of individuals. In accordance with the descriptive analysis in Section 1, the ownership is assigned to each member of couples. The standard errors are estimated by allowing for clustering at the household level, both across individuals and for the pooled sample, across time. The right hand side variables of major interest are signed in such a way that we predict all of them to be positive. We use standardized values of all variables except for the binary ones (optimistic weather forecast, single female, couple, race variables).

3.2 Results

Before presenting the estimates for the regressions specified above, we look at pairwise correlations of the subjective measures with each left-hand side variable. The purpose of this exercise is to see whether the variables we constructed predict stock ownership as hypothesized. Table 14 contains the results. Except for risk aversion, all measures are correlated with the stockholding variables in the way we expect. These correlations are highly significant, stable across time, and often substantial in magnitude.

Table 15 presents the main results for equations (11)-(13) estimated on the pooled sample of the five survey waves (1992 to 2000). For each left-hand side variable, the first column shows results where no subjective measures are included, while the second column presents the results from the full specification. Since the right-hand variables are either binary or standardized to have mean 0 and variance 1, the magnitudes are directly comparable.

The results are similar for the three different left-hand side variables, with a clear ranking in predictive power. Ownership shows the strongest correlations, selling out all stocks the weakest, and becoming a stockholder is somewhere in-between. This is consistent with the hypothesis that stock ownership is measured with noise, and therefore first differences are noisier than levels. Selling out is the noisiest because the upward trends in the 1990's resulted in a higher signal-to-noise ratio for buying than selling. The baseline results show by and large what we know from the previous literature: couples are significantly more likely to hold stocks than singles; same is true for whites, followed by Hispanics, African Americans being the least likely stockholders; and education and total wealth are also strong predictors.

All subjective measures and cognitive variables have the right sign, except risk tolerance, but not all are significant. Expectations over economic growth is significant but not large in magnitude. General optimism has a very strong effect, stronger than education and wealth in general. The Lillard-Willis index of precision is also significant and so are some of the cognition variables, above all the IQ-type scores. Risk aversion has no power whatsoever, and the weather forecast variables also lose significance when entered with the other measures. Entering the subjective measures increases the predictive power of the model by 15-20 percent, and they also decrease the effect of race, education, and wealth. This suggests that what previous research identified as fixed transaction costs of investment do contain informational elements indeed. Even our quite crude measures explain a significant part of their covariance with stock ownership, even though they are not related to stock market returns at all.

3.3 Robustness of the results

Table 16 and 17 present the results year by year. All qualitative results hold for each time period separately. In fact, even the magnitudes are quite stable. That is not so surprising when the highly serially correlated ownership is on the left-hand side, but it is a little more so for the transition variables. In addition to the year-by-year runs, we have performed three other robustness checks. Table 18 shows the results from sample for the baseline results to be the same as for the full specification. They are essentially the same as in Table 15. In table 19 we show what happens when we enter wealth with a piecewise linear spline by its 10 deciles. All other effects become weaker, including education, race, and all subjective measures. The qualitative results, however, are unchanged. In Table 20 we include the index of general optimism and precision to refer to t0 only (as opposed to the whole 1992-2000 time period). Again, the results are qualitatively the same, but the effects of the variables in question are smaller in magnitude. This is consistent with the hypothesis that they are noisier measures of the same personal characteristics. Other robustness checks remain to be done. These include alternative measures of stock ownership, and a more detailed index for optimism constructed by domains of expectations.

As we indicated earlier, the left-hand side variables are probably measured with considerable noise, especially in equations (12)-(13). It is quite possible that this measurement error is negatively correlated with cognitive capacities: people with worse memory and classification skills are more likely to make mistakes. That would introduce a bias in the cognition-related coefficients in the two transition regressions. The bias is probably downward when the left-hand-side variable is buying, and upward when selling (because the latter is coded –1, 0). The reason is that part of the coefficients would pick up transition due to reporting error differentials in the two interviews, which is presumably negatively correlated with the relevant variables. The direct measures of cognition are not much different in the two transition equations, although the point estimates are somewhat larger for the selling than the buying equations. This indicates that the bias we worry about is there but it is probably not too large.

4. Discussion and conclusions

Our results strongly support the role of prior expectations and the precision of beliefs as major determinants of who becomes a stockholder, the major implication of the theory of portfolio selection with learning. Expectations about economic growth matter, but more important is what we call general optimism. The relative weakness of the former effect partly reflects the noisy nature of our survey measure, but the strength of latter one is an intriguing result in itself. In order to have a better explanation, we plan to investigate the content of our index of optimism in detail in the future. The index of precision (fraction of nonfocal answers to all expectations questions) also predicts who becomes a stockholder. This result provides further support for Lillard and Willis (2001), who argue that this index is a useful measure of how precise people’s beliefs are in general.

The strength of our results are somewhat surprising in the light that they do not rely on any direct subjective information about stock market returns. Our measure that is closest to what returns people expect on the stock market is how likely they think the whole economy would slip into a depression. Expectations and precision of beliefs over other, apparently not related events seem to correlate strongly with those about asset returns.

HRS 2002 contains survey measures of perceived stock market returns. It asks questions not only about the expected rate of return but whole distribution. When the data are available, we will use them to estimate the optimal share for each individual. We then will contrast this to their actual stockholding status. This way, we can assign an implicit threshold for not holding stocks for each respondent. If we recalibrate the model to the past distribution of stock returns, we can compare the level and heterogeneity of thresholds implied by the two strategies. We expect that allowing for heterogeneity in beliefs largely reduces both the level and variance of the thresholds.

The theoretical model of portfolio choice with subjective uncertainty and learning offers implications about what kind of heterogeneity should matter in determining who does and who does not become a stockholder.

References

Barsky, Robert B., F. Thomas Juster, Miles S. Kimball, and Matthew D. Shapiro (1997), “Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study,” *Quarterly Journal of Economics*, 537-579.

Bartaut, C. 1998. "Stockholding Behavior of U.S. Households: Evidence from the 1983-1989 Survey of Consumer Finances." *Review of Economics and Statistics*, 80(2) 263-275.

Bassett, W. F., and R. L. Lumsdaine (1999), “Outlook, Outcomes and Optimism” unpublished manuscript, Brown University.

Bassett, W. F., and R. L. Lumsdaine (2001), “Probability Limits,” *Journal of Human Resources*, 36(2), 327-363.

Brennan, M. J. (1998), “The Role of Learning in Dynamic Portfolio Decisions,” *European Finance Review* 1: 295-306.

Campbell, John Y. and Viceira, Luis M. (2002), *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*. Oxford University Press.

Genotte, G. (1986), “Optimal Portfolio Choice under Incomplete Information,” *Journal of Finance*, 61, 733-749.

Haliassos, M. and C. Bertaut. (1995) "Why Do So Few Hold Stocks?" *Economic Journal*, 1110-1129.

Juster, F. T. and R. Suzman (1995) “An Overview of the Health and Retirement Study,” *The Journal of Human Resources*, 30, S7-S56.

Lipster, R.S. and A. N. Shirayayev (1978), *Statistics of Random Processes*, Springer-Verlag, New York.

Lillard, Lee A. and Willis, Robert J. (2001), “Cognition and Wealth: The Importance of Probabilistic Thinking.” Presented at the Third Annual Joint Conference for the Retirement Research Consortium “Making Hard Choices About Retirement,” May 17-18, 2001, Washington, DC.

Mehra, Rajnish and Edward C. Prescott, (1985), "The Equity Premium: A Puzzle," *Journal of Monetary Economics* 15:145-161.

Merton, R. C. (1969), “Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case,” *The Review of Economics and Statistics*, 51(3), 247-257.

Poterba, J. M. (2001), “The Rise of the ‘Equity Culture’: U.S. Stockownership Patterns, 1989-1998.” *Mimeo.*

Shliefer, Andrei (2000) *Inefficient Markets: An Introduction to Behavioral Finance*. Oxford University Press.

Tobin, James (1958). "Liquidity Behavior as Behavior Toward Risk" *Review of Economic Studies*, No. 67, 65-86.

Vissing-Jørgensen, Annette. 2000. "Non-Financial Income and Participation Cost Structures." Department of Economics, University of Chicago.

Willis, Robert J. (1999) "Theory Confronts Data: How the HRS Is Shaped by Economics of Aging and How the Economics of Aging Will Be Shaped by HRS," *Labour Economics*, 6(2), 119-145.

Table 1. Fraction in stockholder households. HRS age eligibles (born 1931-41).

Year	Outside retirement accounts
1992	32.0
1994	35.9
1996	36.0
1998	35.7
2000	36.5

Table 2. Stock ownership outside retirement accounts by education groups
HRS cohort, person weights

year	<12	12	13-15	16 or more	all
1992	11.0	30.6	38.9	55.1	32.0
1994	13.9	35.7	44.2	55.4	35.9
1996	13.5	34.5	44.4	58.7	36.0
1998	12.3	33.7	45.0	58.9	35.7
2000	13.6	33.5	45.0	60.1	36.5

Table 3. Changes in stock ownership outside retirement accounts
HRS age eligible sample, person weights

	1994			
1992	no	yes	total	sample size
no	57.2	9.9	67.1	
yes	6.8	26.1	32.9	
total	64.0	36.0	100.0	8,787

	1996			
1994	no	yes	total	sample size
no	56.0	7.4	63.4	
yes	7.3	29.4	36.6	
total	63.3	36.8	100.0	7,954

	1998			
1996	no	yes	total	sample size
no	56.0	7.3	63.2	
yes	8.1	28.7	36.8	
total	64.0	36.0	100.0	7,532

	2000			
1998	no	yes	total	sample size
no	55.3	8.4	63.7	
yes	8.3	28.1	36.3	
total	63.6	36.4	100.0	7,080

Table 5. Probability of no economic depression (one minus the probability of an economic depression in the near future). Summary statistics.
HRS cohort. Weighted by t0 person weight.

	mean	**std**	**obs**	**min**	**max**
1992	0.45	0.26	9086	0	1
1994	0.62	0.28	7878	0	1
1996	0.61	0.28	7427	0	1
average	0.54	0.22	9086	0	1

Correlation

	1992	**1994**	**1996**	**average**
1992	1.00			
1994	0.34	1.00		
1996	0.32	0.43	1.00	

Table 6. Correlation of the optimism factor with the expectation questions. p-values in parentheses, number of observations below.

Question	1992	1994	1996	1998	2000
Sunny day		0.162			0.094
		(0.00)			(0.00)
		10517			8655
Income will keep up with inflation			0.435	0.445	0.458
			(0.00)	(0.00)	(0.00)
			9779	8976	8368
Will leave inheritance (>$10K)		0.553	0.560	0.580	0.573
		(0.00)	(0.00)	(0.00)	(0.00)
		10462	9951	9195	8597
Will leave inheritance (>$100K)		0.540	0.591	0.609	0.610
		(0.00)	(0.00)	(0.00)	(0.00)
		8384	8034	7596	7172
Will leave inheritance (any)			0.236	0.237	0.239
			(0.00)	(0.00)	(0.00)
			1845	1489	1320
Will receive inheritance		0.459	0.472	0.439	0.392
		(0.00)	(0.00)	(0.00)	(0.00)
		10479	9965	9243	8657
Will lose job	-0.161	-0.188	-0.160	-0.148	-0.154
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	6439	5255	4516	3833	3178
Would find another job if lost current one	0.124	0.174	0.191	0.257	0.234
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	6446	5245	4521	3838	3173
Will work sometime if not working		0.180	0.201	0.165	0.182
		(0.00)	(0.00)	(0.00)	(0.00)
		4029	4299	4417	4627
Will work past age 62	0.075	0.137	0.133	0.169	0.114
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	7496	7130	4518	3391	2318
Will work past age 65	0.118	0.170	0.172	0.233	0.215
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	5512	5131	3163	2559	1983
Health will limit work activity	-0.263	-0.278	-0.237	-0.237	-0.237
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	7635	7809	5479	4617	3825
Will find a job if looking			0.142	0.142	0.213
			(0.00)	(0.00)	(0.00)
			653	491	350

Table 6, cont.

Will move in the next 2 years				0.057	0.064
				(0.01)	(0.00)
				2211	3103
Will live to be 75	0.461	0.512	0.553	0.585	0.564
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	11734	9665	8823	6714	5866
Will live to be 85	0.383	0.420	0.440	0.492	0.507
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	10939	9543	8934	6277	7931
Will give financial help to someone	0.118	0.482	0.506	0.539	0.526
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	11837	10478	9935	9184	8559
Will receive financial help		0.122	0.117	0.095	0.086
		(0.00)	(0.00)	(0.00)	(0.00)
		10519	9957	9220	8601
Will go to a nursing home		-0.179	0.008	-0.108	-0.150
		(0.00)	(0.60)	(0.00)	(0.00)
		2993	4349	2117	2971
Probability of 2-digit inflation	-0.122	-0.027	-0.049	-0.080	-0.095
	(0.00)	0.006	(0.00)	(0.00)	(0.00)
	11601	9953	9311	8380	7860
Prob. of economic depression	-0.146	-0.136	-0.134		
	(0.00)	(0.00)	(0.00)		
	11701	10122	9524		
Social Security will be less generous	0.054		0.060		
	(0.00)		(0.00)		
	11734		9711		
Social Security will be more generous	-0.046				
	(0.00)				
	11752				

Table 7. Self rated health and the optimism factor (means by category)

	1992	1994	1996	1998	2000
Excellent	0.56	0.65	0.69	0.74	0.79
Very good	0.21	0.29	0.28	0.42	0.38
Good	-0.16	-0.15	-0.16	-0.05	-0.10
Fair	-0.57	-0.62	-0.62	-0.54	-0.56
Poor	-0.76	-0.88	-0.95	-0.88	-0.85
All	0.00	0.00	0.00	0.00	0.00
Anova R^2	0.19	0.23	0.23	0.22	0.22

Table 8. Regression results: predictors of the index of optimism

Mean of LHS variable	*0.000*	*0.000*
Age [a]	-0.073	-0.036
	(8.91)**	(4.93)**
Female [b]	0.180	0.178
	(11.08)**	(12.36)**
Education [a]	0.337	0.191
	(37.84)**	(22.99)**
Black [b]	-0.152	-0.008
	(6.96)**	(0.40)
Hispanic [b]	-0.151	-0.047
	(4.85)**	(1.71)
Catholic[b]	0.078	0.034
	(4.11)**	(2.00)*
Jewish [b]	0.165	0.187
	(2.68)**	(3.43)**
Total wealth in 1992[a]	0.201	0.156
	(24.62)**	(21.48)**
Self-rated health [a]		0.426
		(57.28)**
Constant	-0.068	-0.091
	(5.25)**	(7.91)**
Observations	12,002	12,002
R-squared	0.23	0.39

t-statistics in parentheses. Standard errors are robust to heteroskedasticity.

* significant at 5% level; ** significant at 1% level
[a] standardized RHS variable: mean=0, std=1
[b] binary variable

Table 9. Fraction of exact answers. Summary statistics.

Variable	# obs	Mean	Std. D	Min	Max
Fraction exact, 1992	11,879	0.51	0.25	0.00	1.00
Fraction exact, 1994	10,635	0.41	0.22	0.00	1.00
Fraction exact, 1996	10,099	0.39	0.23	0.00	1.00
Fraction exact, 1998	9,548	0.38	0.23	0.00	1.00
Fraction exact, 2000	8,950	0.41	0.22	0.00	1.00
Fraction exact, avg.	7,460	0.42	0.17	0.00	0.94

Correlations
(obs=7460)

	1992	1994	1996	1998	2000	avg
Fraction exact, 1992	1.00					
Fraction exact, 1994	0.34	1.00				
Fraction exact, 1996	0.33	0.49	1.00			
Fraction exact, 1998	0.29	0.43	0.46	1.00		
Fraction exact, 2000	0.27	0.41	0.44	0.45	1.00	
Fraction exact, avg.	0.59	0.75	0.78	0.73	0.72	1.00

Table 10. Distribution of respondents in four risk preference categories (percent)
HRS age eligibles, person weights.

Risk preference categories	1992	1994	1998	2000
I. very strong risk aversion ($\gamma > 4$)	64.7	63.2	58.2	64.1
II. strong risk aversion ($4 > \gamma > 2$)	12.0	12.9	16.2	14.4
III. weak risk aversion ($2 > \gamma > 1$)	10.6	13.2	9.6	8.5
IV. very weak risk aversion ($1 > \gamma > 0$)	12.7	10.7	16.0	13.0
All	100.0	100.0	100.0	100.0
Number of observations	9,089	591	628	760

Table 11. Cognition: correlation between different factors

	Memory	IQ	Countig back by 7	Dementia control
Memory	1.00			
IQ	0.36	1.00		
Countig back by 7	0.58	0.38	1.00	
Dementia control	0.36	0.33	0.47	1.00

Table 12. Cognition: correlation between the overall factor and the different items
HRS cohort. p-values in parentheses.

Question	**1992**	**1994**	**1996**	**1998**	**2000**	***IQ-type questions***	**1992**
Immediate word recall	0.60	0.64	0.71	0.74	0.72	Orange & banana	0.38
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)		(0.00)
Delayed word recall	0.57	0.61	0.69	0.73	0.73	Table & chair	0.30
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)		(0.00)
Counting back by sevens			0.65	0.67	0.68	Eye & ear	0.44
			(0.00)	(0.00)	(0.00)		(0.00)
Date month			0.21	0.20	0.24	Egg & seed	0.33
			(0.00)	(0.00)	(0.00)		(0.00)
Date day			0.22	0.24	0.24	Air & water	0.27
			(0.00)	(0.00)	(0.00)		(0.00)
Date year			0.23	0.26	0.20	Fly & tree	0.30
			(0.00)	(0.00)	(0.00)		(0.00)
Day of week			0.15	0.14	0.15	Praise & punishment	0.25
			(0.00)	(0.00)	(0.00)		(0.00)
Scissors			0.11	0.10	0.12		
			(0.00)	(0.00)	(0.00)		
Cactus			0.43	0.39	0.42		
			(0.00)	(0.00)	(0.00)		
President			0.30	0.29	0.29		
			(0.00)	(0.00)	(0.00)		
Vice president			0.40	0.44	0.46		
			(0.00)	(0.00)	(0.00)		
Counting back from 20			0.25	0.27	0.24		
			(0.00)	(0.00)	(0.00)		
Counting back from 86			0.37	0.36	0.36		
			(0.00)	(0.00)	(0.00)		

Table 13. Correlation of time-invariant subjective measures of expectations, precision of beliefs,
risk aversion, and cognition.
HRS cohort. p-values in parentheses, number of observations below.

	Index of optimism	Optimism in weather	Index of precision	Precision in weather	Measure of risk tolerance[b]
Optimism in weather	0.151 (0.00) 10239				
Index of precision	0.143 (0.00) 11879	0.027 (0.01) 10077			
Precision in weather	0.096 (0.00) 10239	0.201 (0.00) 10239	0.179 (0.00) 10077		
Measure of risk tolerance[b]	0.010 (0.27) 11603	0.027 (0.01) 9885	0.044 (0.00) 11594	0.009 (0.38) 9885	
Cognition: memory	0.279 (0.00) 12175	0.091 (0.00) 10239	0.121 (0.00) 11879	0.085 (0.00) 10239	0.005 (0.62) 11611
Cognition: IQ	0.324 (0.00) 12175	0.101 (0.00) 10239	0.235 (0.00) 11879	0.145 (0.00) 10239	0.012 (0.20) 11611
Cognition: sevens	0.295 (0.00) 12175	0.070 (0.00) 10239	0.120 (0.00) 11879	0.146 (0.00) 10239	0.012 (0.19) 11611
Cognition: dementia	0.265 (0.00) 12175	0.069 (0.00) 10239	0.121 (0.00) 11879	0.118 (0.00) 10239	0.028 (0.03) 11611
Overall cognition	0.359 (0.00) 12175	0.106 (0.00) 10239	0.168 (0.00) 11879	0.140 (0.00) 10239	0.007 (0.47) 11611

[b] Inverse of the estimated γ.

Table 14. Pairwise correlation of outcomes and subjective measures.
p-values below the coefficients.

Panel A. LHS: stock ownership outside retirement accounts (0 or 1)

	1992	1994	1996	1998	2000	pooled	ever	always
economic growth	0.106	0.102	0.081	0.084	0.089	0.093	0.126	0.098
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
general optimism	0.307	0.337	0.363	0.355	0.378	0.347	0.404	0.279
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
optimistic weather forecast	0.066	0.063	0.054	0.062	0.054	0.060	0.069	0.042
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
general precision	0.113	0.143	0.148	0.160	0.158	0.143	0.170	0.124
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
precision in weather forecast	0.068	0.083	0.088	0.083	0.078	0.080	0.095	0.055
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
risk tolerance	0.022	0.004	0.002	0.002	0.003	0.005	0.002	0.017
	0.02	0.66	0.87	0.88	0.77	0.83	0.83	0.12

Panel B. LHS: becoming a stockholder if not one at t0 (0 or 1)

	1992-1994	1991994-6	1996-1998	1998-2000	pooled
economic growth	0.074	0.038	0.047	0.051	0.054
	0.00	0.00	0.00	0.00	0.00
general optimism	0.216	0.210	0.204	0.256	0.221
	0.00	0.00	0.00	0.00	0.00
optimistic weather forecast	0.027	0.025	0.050	0.036	0.034
	0.03	0.05	0.00	0.01	0.00
general precision	0.080	0.096	0.098	0.100	0.092
	0.00	0.00	0.00	0.00	0.00
precision in weather forecast	0.052	0.044	0.045	0.056	0.049
	0.00	0.00	0.00	0.00	0.00
risk tolerance	0.013	0.011	0.009	-0.010	0.007
	0.27	0.38	0.47	0.45	0.29

Panel C. LHS: selling all stocks if stockholder at t0 (-1 or 0)

	1992-1994	1994-1996	1996-1998	1998-2000	pooled
economic growth	0.037	0.012	0.041	0.051	0.035
	0.04	0.51	0.02	0.01	0.00
general optimism	0.136	0.162	0.098	0.151	0.136
	0.00	0.00	0.00	0.00	0.00
optimistic weather forecast	0.029	0.011	0.014	-0.011	0.011
	0.10	0.52	0.46	0.58	0.21
general precision	0.075	0.036	0.065	0.042	0.055
	0.00	0.04	0.00	0.03	0.00
precision in weather forecast	0.059	0.043	0.012	0.003	0.031
	0.00	0.02	0.53	0.88	0.00
risk tolerance	0.009	-0.003	-0.019	0.007	-0.002

	0.61	0.86	0.29	0.71	0.86

Table 15. Main regression results (OLS). Pooled sample.

	LHS: owhership (0 or 1)		**LHS: buy (0 or 1)**		**LHS: sell (0 or –1)**	
Mean LHS variarble	**0.343**	**0.343**	**0.123**	**0.123**	**-0.218**	**-0.218**
couple[b]	0.115	0.081	0.051	0.036	0.033	0.026
	(11.33)**	(8.62)**	(8.85)**	(6.31)**	(1.99)*	(1.58)
single female[b]	0.018	0.013	0.009	0.005	0.013	0.016
	(1.21)	(0.91)	(0.99)	(0.57)	(0.52)	(0.62)
age[a]	0.020	0.032	0.000	0.005	0.016	0.024
	(4.74)**	(7.50)**	(0.02	(1.63)	(3.10)**	(4.40)**
black[a]	-0.186	-0.139	-0.061	-0.048	-0.167	-0.138
	(17.30)**	(12.85)**	(9.71)**	(7.28)**	(5.66)**	(4.64)**
hispanic[b]	-0.130	-0.111	-0.048	-0.043	-0.093	-0.073
	(9.66)**	(8.78)**	(6.62)**	(6.01)**	(2.47)*	(1.95)
education[a]	***0.129***	***0.058***	***0.050***	***0.021***	***0.068***	***0.032***
	(22.59)**	***(10.38)*****	***(16.18)*****	***(6.43)*****	***(9.71)*****	***(4.04)*****
net worth[a]	0.071	0.057	0.067	0.047	0.011	0.008
	(3.59)**	(3.64)**	(6.02)**	(4.68)**	(2.84)**	(2.28)*
economic growth[a]		0.011		0.004		0.002
		(3.66)**		(1.98)*		(0.48)
general optimism[a]		***0.080***		***0.041***		***0.036***
		(16.31)**		***(13.98)*****		***(6.61)*****
optimistic weather forecast[b]		0.004		-0.005		-0.006
		(0.46)		(0.83)		(0.51)
general precision[a]		***0.031***		***0.016***		***0.011***
		(6.97)**		***(5.06)*****		***(2.06)****
precision in weather forecast[a]		0.000		0.000		0.007
		(0.06)		(0.18)		(1.34)
risk tolerance[a]		0.002		-0.001		-0.007
		(0.45)		(0.51)		(1.67)
cognition: memory[a]		0.018		0.005		0.008
		(3.60)**		(1.61)		(1.24)
cognition: iq[a]		0.020		0.010		0.014
		(4.38)**		(3.28)**		(2.38)*
cognition: numbers[a]		0.015		0.004		0.013
		(3.46)**		(1.50)		(1.96)
cognition: dementia control[a]		0.001		0.000		0.010
		(0.33)		(0.19)		(1.07)
Constant	0.274	0.286	0.117	0.134	-0.274	-0.292
	(30.18)**	(27.73)**	(19.97)**	(19.03)**	(17.28)**	(16.21)**
Observations	42,137	42,137	21,539	21,539	11,277	11,277
R-squared	0.16	0.20	0.06	0.08	0.03	0.04

t-statistics in parentheses. Standard errors are robust to heteroskedasticity and clustering within households.
* significant at 5% level; ** significant at 1% level
[a] standardized RHS variable: mean=0, std=1. [b] binary variable (average positive weather error can also be 0.5)

Table 16. Regression results: year by year. Baseline.

Panel A. LHS: stock ownership outside retirement accounts (0 or 1)

	1992	**1994**	**1996**	**1998**	**2000**	**ever**	**always**
mean LHS	0.286	0.323	0.324	0.324	0.326	0.568	0.159
couple	0.096	0.098	0.104	0.116	0.117	0.157	0.044
	(9.47)**	(8.65)**	(9.17)**	(9.70)**	(9.63)**	(11.23)**	(4.54)**
single female	0.024	-0.003	0.007	0.033	0.009	0.057	0.011
	(1.62	(0.2)	(0.45)	(1.89)	(0.5)	(2.64)**	(0.69)
age[a]	0.018	0.022	0.015	0.015	0.014	0.016	0.021
	(4.20)**	(4.76)**	(3.10)**	(3.03)**	(2.73)**	(3.40)**	(4.89)**
black	-0.156	-0.172	-0.193	-0.193	-0.195	-0.248	-0.101
	(14.56)**	(14.36)**	(16.12)**	(15.15)**	(14.28)**	(15.31)**	(12.47)**
hispanic	-0.108	-0.125	-0.129	-0.133	-0.138	-0.232	-0.051
	(8.63)**	(8.81)**	(8.95)**	(9.16)**	(7.51)**	(12.54)**	(4.69)**
education[a]	0.109	0.103	0.112	0.129	0.127	0.154	0.074
	(23.14)**	(20.47)**	(20.76)**	(22.11)**	(12.64)**	(28.84)**	(15.19)**
net worth[a]	0.091	0.100	0.097	0.056	0.044	0.065	0.071
	(10.79)**	(9.52)**	(7.33)**	(1.92)	(1.87)	(8.89)**	(7.84)**
Constant	0.244	0.283	0.282	0.271	0.278	0.471	0.137
	(25.85)**	(27.32)**	(27.43)**	(25.61)**	(25.83)**	(35.44)**	(15.33)**
Observations	12,450	11,177	10,728	10,282	9,399	10,324	8,343
R-squared	0.17	0.18	0.19	0.17	0.16	0.24	0.12

t-statistics in parentheses. Standard errors are robust to heteroskedasticity and clustering within households.
* significant at 5% level; ** significant at 1% level
[a] standardized RHS variable: mean=0, std=1
[b] binary variable

Panel B. LHS: becoming a stockholder if not one at t0 (0 or 1)

	1992-4	**1994-6**	**1996-8**	**1998-2000**
mean LHS	0.131	0.103	0.108	0.117
couple	0.050	0.042	0.045	0.042
	(5.15)**	(4.70)**	(4.84)**	(4.08)**
single female	-0.002	-0.002	0.028	0.003
	(0.16)	(0.18)	(1.96)	(0.21)
age[a]	0.010	0.000	-0.007	-0.005
	(2.28)*	(0.01)	(1.69)	(1.03)
black	-0.072	-0.058	-0.044	-0.060
	(6.99)**	(6.24)**	(4.19)**	(5.51)**
hispanic	-0.066	-0.041	-0.038	-0.034
	(6.05)**	(3.96)**	(3.95)**	(2.93)**
education[a]	0.045	0.039	0.049	0.052
	(10.40)**	(9.00)**	(10.93)**	(9.85)**
net worth[a]	0.044	0.064	0.090	0.169
	(3.43)**	(3.89)**	(6.25)**	(2.79)**
Constant	0.131	0.108	0.113	0.131
	(13.56)**	(11.39)**	(12.41)**	(10.38)**
Observations	7,881	6,759	6,465	6,183
R-squared	0.05	0.06	0.07	0.07

t-statistics in parentheses. Standard errors are robust to heteroskedasticity and clustering within households.
* significant at 5% level; ** significant at 1% level
[a] standardized RHS variable: mean=0, std=1
[b] binary variable

Panel C. LHS: selling all stocks if stockholder at t0 (-1 or 0)

	1992-4	**1994-6**	**1996-8**	**1998-2000**
mean LHS	-0.221	-0.219	-0.225	-0.238
couple	0.016	0.036	0.034	0.027
	(0.53)	(1.25)	(1.19)	(0.99)
single female	-0.013	0.016	0.061	-0.027
	(0.29)	(0.37)	(1.46)	(0.61)
age[a]	0.018	0.008	0.009	0.020
	(2.09)*	(0.88)	(1.09)	(2.12)*
black	-0.225	-0.225	-0.084	-0.145
	(4.48)**	(4.54)**	(1.57	(2.80)**
hispanic	-0.067	-0.079	-0.099	-0.180
	(1.01)	(1.21)	(1.56)	(2.46)*
education[a]	0.056	0.076	0.063	0.081
	(4.59)**	(6.63)**	(5.03)**	(6.42)**
net worth[a]	0.005	0.021	0.014	0.006
	(0.78)	(3.49)**	(1.89)	(1.09)
Constant	-0.250	-0.281	-0.286	-0.291
	(8.42)**	(10.14)**	(10.27)**	(10.82)**
Observations	3,308	3,368	3,223	3,057
R-squared	0.03	0.04	0.02	0.03

t-statistics in parentheses. Standard errors are robust to heteroskedasticity and clustering within households.

* significant at 5% level; ** significant at 1% level
[a] standardized RHS variable: mean=0, std=1
[b] binary variable

Table 17. Regression results: year by year. Full specification.

Panel A. LHS: stock ownership outside retirement accounts (0 or 1)

	1992	1994	1996	1998	2000	ever	always
mean LHS	0.312	0.343	0.353	0.359	0.359	0.576	0.168
couple[b]	0.080	0.077	0.078	0.087	0.080	0.125	0.030
	(6.85)**	(6.20)**	(6.01)**	(6.50)**	(5.94)**	(8.52)**	(2.80)**
single female[b]	0.032	-0.005	-0.007	0.035	0.006	0.051	0.005
	(1.75)	(0.25)	(0.36)	(1.70)	(0.30)	(2.22)*	(0.29)
age[a]	0.033	0.035	0.033	0.024	0.029	0.028	0.031
	(6.41)**	(6.37)**	(5.90)**	(3.86)**	(4.44)**	(5.21)**	(6.31)**
black[a]	-0.128	-0.135	-0.145	-0.145	-0.134	-0.183	-0.081
	(9.76)**	(9.59)**	(9.83)**	(9.15)**	(8.32)**	(10.19)**	(8.46)**
hispanic[b]	-0.101	-0.115	-0.113	-0.115	-0.110	-0.198	-0.046
	(6.80)**	(7.19)**	(6.63)**	(6.44)**	(6.13)**	(9.72)**	(3.95)**
education[a]	0.064	0.046	0.046	0.065	0.066	0.075	0.041
	(10.19)**	(6.87)**	(6.16)**	(8.38)**	(8.22)**	(10.42)**	(6.58)**
net worth[a]	0.082	0.078	0.072	0.039	0.026	0.045	0.063
	(10.40)**	(7.98)**	(6.68)**	(1.97)*	(1.95	(7.08)**	(7.14)**
economic growth[a]	0.016	0.015	0.007	0.010	0.007	0.016	0.012
	(3.56)**	(3.21)**	(1.48	(1.90)	(1.40)	(3.33)**	(2.88)**
general optimism[a]	0.059	0.072	0.082	0.083	0.095	0.097	0.043
	(11.30)**	(13.17)**	(14.32)**	(13.80)**	(16.61)**	(18.21)**	(9.13)**
optimistic weather forecast[b]	0.014	0.010	-0.002	0.005	-0.011	-0.003	-0.005
	(1.25)	(0.85)	(0.18)	(0.34)	(0.77)	(0.20)	(0.40)
general precision[a]	0.024	0.032	0.031	0.035	0.034	0.036	0.020
	(4.45)**	(5.85)**	(5.13)**	(5.47)**	(5.01)**	(6.30)**	(3.78)**
precision in weather forecast[a]	-0.004	0.002	0.004	0.000	-0.001	0.003	-0.003
	(0.83)	(0.44)	(0.74)	(0.02)	(0.17)	(0.65)	(0.67)
risk tolerance[a]	0.008	0.004	0.000	-0.003	-0.005	-0.002	0.004
	(1.79)	(0.90)	(0.04)	(0.57)	(0.87)	(0.50)	(0.91)
cognition: memory[a]	0.018	0.016	0.023	0.016	0.014	0.005	0.018
	(2.96)**	(2.53)*	(3.36)**	(2.19)*	(1.82)	(0.80)	(3.25)**
cognition: iq[a]	0.015	0.016	0.021	0.021	0.027	0.021	0.008
	(2.72)**	(2.80)**	(3.43)**	(3.18)**	(4.04)**	(3.48)**	(1.64)
cognition: numbers[a]	0.013	0.015	0.014	0.017	0.017	0.016	0.011
	(2.45)*	(2.62)**	(2.41)*	(2.61)**	(2.63)**	(2.58)**	(2.35)*
cognition: dementia control[a]	0.000	0.009	0.002	-0.002	-0.005	0.018	-0.009
	(0.00)	(2.19)*	(0.50)	(0.31)	(0.99)	(3.47)**	(2.56)*
Constant	0.254	0.293	0.300	0.290	0.301	0.488	0.148
	(19.88)**	(21.89)**	(21.33)**	(19.74)**	(20.23)**	(31.07)**	(12.02)**
Observations	9,706	9,340	8,311	7,666	7,114	8,589	7,342
R-squared	0.20	0.21	0.21	0.20	0.21	0.28	0.15

t-statistics in parentheses. Standard errors are robust to heteroskedasticity and clustering within households.
* significant at 5% level; ** significant at 1% level
[a] standardized RHS variable: mean=0, std=1
[b] binary variable (average positive weather error can also be 0.5)

Panel B. LHS: becoming a stockholder if not one at t0 (0 or 1)

	1992-1994	1994-1996	1996-998	1998-2000
mean LHS	0.138	0.110	0.119	0.122
couple[b]	0.038	0.034	0.037	0.031
	(3.62)**	(3.40)**	(3.39)**	(2.72)**
single female[b]	-0.001	-0.005	0.028	0.002
	(0.09)	(0.37)	(1.57)	(0.13)
age[a]	0.015	0.006	-0.009	0.002
	(3.00)**	(1.24)	(1.73)	(0.31)
black[a]	-0.051	-0.048	-0.041	-0.050
	(4.22)**	(4.41)**	(3.19)**	(4.06)**
hispanic[b]	-0.060	-0.036	-0.035	-0.032
	(4.98)**	(2.98)**	(2.78)**	(2.21)*
education[a]	0.014	0.016	0.038	0.023
	(2.53)*	(2.94)**	(5.98)**	(3.40)**
net worth[a]	0.034	0.046	0.074	0.080
	(2.72)**	(2.91)**	(4.37)**	(1.75)
economic growth[a]	0.008	0.001	0.006	0.001
	(2.12)*	(0.26)	(1.59)	(0.16)
general optimism[a]	0.047	0.036	0.027	0.050
	(8.87)**	(6.81)**	(5.04)**	(8.41)**
optimistic weather forecast[b]	-0.008	-0.010	0.005	-0.005
	(0.71)	(0.84)	(0.42)	(0.39)
general precision[a]	0.014	0.015	0.016	0.018
	(2.63)**	(2.75)**	(2.60)**	(2.76)**
precision in weather forecast[a]	0.001	-0.002	-0.001	0.000
	(0.36)	(0.53)	(0.32)	(0.07)
risk tolerance[a]	-0.001	-0.002	-0.002	0.002
	(0.31)	(0.56)	(0.44)	(0.44)
cognition: memory[a]	0.004	0.011	0.001	0.005
	(0.63)	(1.91)	(0.16)	(0.76)
cognition: iq[a]	0.004	0.011	0.007	0.022
	(0.81)	(2.11)*	(1.18)	(3.49)**
cognition: numbers[a]	0.013	0.002	0.002	-0.004
	(2.64)**	(0.50)	(0.32)	(0.68)
cognition: dementia control[a]	0.009	-0.001	-0.004	-0.007
	(2.85)**	(0.32)	(1.17)	(1.70)
Constant	0.150	0.124	0.123	0.139
	(11.97)**	(9.96)**	(9.11)**	(9.66)**
Observations	6,596	5,575	4,883	4,485
R-squared	0.08	0.07	0.08	0.09

t-statistics in parentheses. Standard errors are robust to heteroskedasticity and clustering within households.
* significant at 5% level; ** significant at 1% level
[a] standardized RHS variable: mean=0, std=1
[b] binary variable (average positive weather error can also be 0.5)

Panel C. LHS: selling all stocks if stockholder at t0 (-1 or 0)

	1992-1994	1994-1996	1996-998	1998-2000
mean LHS	-0.216	-0.214	-0.214	-0.228
couple[b]	0.004	0.031	0.041	0.016
	(0.12)	(1.05)	(1.35)	(0.57)
single female[b]	-0.004	0.015	0.073	-0.034
	(0.08)	(0.33)	(1.64)	(0.73)
age[a]	0.030	0.020	0.013	0.032
	(3.19)**	(2.05)*	(1.28)	(2.97)**
black[a]	-0.197	-0.182	-0.058	-0.080
	(3.78)**	(3.58)**	(1.06)	(1.45)
hispanic[b]	-0.024	-0.049	-0.098	-0.136
	(0.35)	(0.74)	(1.38)	(1.82)
education[a]	0.017	0.039	0.029	0.043
	(1.17)	(2.83)**	(1.99)*	(2.82)**
net worth[a]	0.001	0.016	0.013	0.003
	(0.15)	(2.78)**	(2.13)*	(0.52)
economic growth[a]	0.001	-0.007	0.009	0.009
	(0.07)	(0.80)	(0.96)	(0.86)
general optimism[a]	0.045	0.040	0.017	0.041
	(4.54)**	(4.22)**	(1.78)	(4.03)**
optimistic weather forecast[b]	0.010	-0.028	0.007	-0.018
	(0.49)	(1.40)	(0.33)	(0.79)
general precision[a]	0.024	-0.003	0.017	0.002
	(2.66)**	(0.28)	(1.76)	(0.16)
precision in weather forecast[a]	0.018	0.009	0.001	-0.002
	(1.94)	(1.01)	(0.09)	(0.21)
risk tolerance[a]	-0.008	-0.006	-0.001	-0.013
	(1.03)	(0.79)	(0.16)	(1.51)
cognition: memory[a]	0.008	0.019	-0.001	0.005
	(0.79)	(1.91)	(0.11)	(0.41)
cognition: iq[a]	0.014	0.022	0.013	0.004
	(1.38)	(2.10)*	(1.16)	(0.37)
cognition: numbers[a]	-0.012	0.014	0.016	0.044
	(1.06)	(1.19)	(1.18)	(3.14)**
cognition: dementia control[a]	0.009	0.003	0.014	0.018
	(0.55)	(0.21)	(0.87)	(0.99)
Constant	-0.260	-0.286	-0.306	-0.307
	(7.72)**	(9.12)**	(9.07)**	(9.55)**
Observations	2,986	2,995	2,736	2,560
R-squared	0.04	0.06	0.03	0.05

t-statistics in parentheses. Standard errors are robust to heteroskedasticity and clustering within households.
* significant at 5% level; ** significant at 1% level
[a] standardized RHS variable: mean=0, std=1
[b] binary variable (average positive weather error can also be 0.5)

Table 18. Robustness of the pooled OLS results. Baseline results on the same sample as those from full specification

	LHS: owhership (0 or 1)		LHS: buy (0 or 1)		LHS: sell (0 or –1)	
Mean LHS variarble	**0.343**	**0.315**	**0.123**	**0.115**	**-0.218**	**-0.225**
couple[b]	0.115	0.106	0.051	0.046	0.033	0.030
	(11.33)**	(12.05)**	(8.85)**	(9.08)**	(1.99)*	(1.9)
single female[b]	0.018	0.014	0.009	0.007	0.013	0.010
	(1.21)	(1.16)	(0.99)	(0.92)	(0.52)	(0.40)
age[a]	0.020	0.017	0.000	0.000	0.016	0.014
	(4.74)**	(4.85)**	(0.02	(0.13)	(3.10)**	(2.91)**
black[a]	-0.186	-0.182	-0.061	-0.060	-0.167	-0.177
	(17.30)**	(19.84)**	(9.71)**	(10.87)**	(5.66)**	(6.16)**
hispanic[b]	-0.130	-0.126	-0.048	-0.047	-0.093	-0.104
	(9.66)**	(11.44)**	(6.62)**	(7.79)**	(2.47)*	(2.85)**
education[a]	***0.129***	***0.117***	***0.050***	***0.047***	***0.068***	***0.069***
	(22.59)**	***(23.79)*****	***(16.18)*****	***(18.45)*****	***(9.71)*****	***(10.42)*****
net worth[a]	0.071	0.077	0.067	0.066	0.011	0.011
	(3.59)**	(4.09)**	(6.02)**	(6.42)**	(2.84)**	(2.97)**
Constant	0.274	0.271	0.117	0.118	-0.274	-0.278
	(30.18)**	(34.31)**	(19.97)**	(22.22)**	(17.28)**	(18.20)**
Observations	42,137	54,036	21,539	27,288	11,277	12,956
R-squared	0.16	0.16	0.06	0.06	0.03	0.03

t-statistics in parentheses. Standard errors are robust to heteroskedasticity and clustering within households.
* significant at 5% level; ** significant at 1% level
[a] standardized RHS variable: mean=0, std=1
[b] binary variable (average positive weather error can also be 0.5)

Table 19. Robustness of the pooled OLS results. Total wealth entered as a linear spline (on its 10 deciles)

	LHS: owhership (0 or 1)		LHS: buy (0 or 1)		LHS: sell (0 or –1)	
Mean LHS variarble	**0.315**	**0.343**	**0.115**	**0.123**	**-0.225**	**-0.218**
couple[b]	0.032	0.027	0.027	0.025	0.000	0.001
	(4.45)**	(3.24)**	(5.47)**	(4.31)**	(0.02)	(0.07)
single female[b]	0.003	0.006	0.004	0.004	0.008	0.013
	(0.27)	(0.44)	(0.55)	(0.41)	(0.35)	(0.55)
age[a]	0.000	0.010	-0.004	0.000	0.004	0.012
	(0.00)	(2.45)*	(1.75)	(0.06)	(0.85)	(2.19)*
black[b]	-0.100	-0.082	-0.044	-0.038	-0.139	-0.111
	(12.56)**	(8.23)**	(7.99)**	(5.87)**	(4.87)**	(3.75)**
hispanic[b]	-0.084	-0.084	-0.039	-0.039	-0.089	-0.061
	(8.53)**	(7.02)**	(6.69)**	(5.57)**	(2.40)*	(1.62)
education[a]	***0.065***	***0.036***	***0.036***	***0.018***	***0.048***	***0.021***
	(18.31)**	***(7.24)*****	***(14.86)*****	***(5.52)*****	***(7.28)*****	***(2.76)*****
net worth (1)[a]	0.001	-0.001	-0.014	-0.012	-0.364	-0.409
	(0.06)	(0.03)	(0.59)	(0.45)	(3.52)**	(4.15)**
net worth (2)[a]	0.472	0.342	0.186	0.069	3.205	3.473

	(3.60)**	(2.26)*	(1.72)	(0.59)	(4.09)**	(4.34)**
net worth (3)[a]	1.044	0.994	0.623	0.508	-1.100	-1.233
	(5.70)**	(4.86)**	(3.92)**	(2.95)**	(1.91)	(2.08)*
net worth (4)[a]	-2.024	-1.646	-1.014	-0.548	1.836	1.894
	(6.48)**	(4.60)**	(3.37)**	(1.62)	(1.85)	(1.86)
net worth (5)[a]	2.432	1.869	1.330	0.740	-1.882	-1.852
	(6.55)**	(4.32)**	(3.24)**	(1.56)	(1.53)	(1.46)
net worth (6)[a]	2.057	1.962	0.380	0.275	1.746	1.353
	(5.46)**	(4.73)**	(0.91)	(0.59)	(2.16)*	(1.59)
net worth (7)[a]	1.546	1.446	0.666	0.487	0.313	0.436
	(5.66)**	(4.91)**	(2.06)*	(1.37)	(0.76)	(1.02)
net worth (8)[a]	0.654	0.562	0.065	0.002	0.306	0.306
	(4.25)**	(3.39)**	(0.32)	(0.01)	(1.69)	(1.67)
net worth (9)[a]	0.340	0.328	0.165	0.170	0.111	0.095
	(5.97)**	(5.42)**	(1.89)	(1.83)	(1.98)*	(1.67)
net worth (10)[a]	0.007	0.006	0.004	0.002	-0.005	-0.005
	(2.52)*	(2.41)*	(0.41)	(0.16)	(1.17	(1.12)
economic growth[a]		0.010		0.004		0.003
		(3.61)**		(1.92)		(0.61)
general optimism[a]		***0.031***		***0.032***		***0.018***
		(7.81)*		***(10.85)****		***(3.25)****
optimistic weather forecast[b]		0.006		-0.004		-0.005
		(0.67)		(0.63)		(0.46)
general precision[a]		***0.027***		***0.015***		***0.012***
		(6.50)*		***(4.82)****		***(2.33)****
precision in weather forecast[a]		-0.002		-0.001		0.006
		(0.52)		(0.44)		(1.16)
risk tolerance[a]		0.003		-0.001		-0.007
		(0.86)		(0.22)		(1.62)
cognition: memory[a]		0.010		0.003		0.006
		(2.28)*		(0.99)		(1.01)
cognition: iq[a]		0.016		0.009		0.014
		(3.88)**		(3.06)**		(2.45)*
cognition: numbers[a]		0.011		0.004		0.011
		(2.74)**		(1.36)		(1.71)
cognition: dementia control[a]		-0.001		0.000		0.008
		(0.36)		(0.24)		(0.89)
Constant	0.107	0.129	0.051	0.080	-0.666	-0.699
	(10.49)**	(8.08)**	(4.39)**	(5.62)**	(9.72)**	(10.29)**
Observations	54,036	42,137	27,288	21,539	12,956	11,277
R-squared	0.280	0.280	0.080	0.090	0.060	0.060

t-statistics in parentheses. Standard errors are robust to heteroskedasticity and clustering within households.
* significant at 5% level; ** significant at 1% level
[a] standardized RHS variable: mean=0, std=1
[b] binary variable (average positive weather error can also be 0.5)

Table 20. Robustness of the pooled OLS results. Optimism and precision entered as t0 measures (as opposed to all waves)

	LHS: owhership (0 or 1)		LHS: buy (0 or 1)		LHS: sell (0 or –1)	
Mean LHS variarble	**0.315**	**0.343**	**0.115**	**0.123**	**-0.225**	**-0.218**
couple[b]	0.106	0.099	0.046	0.046	0.030	0.034
	(12.05)**	(10.14)**	(9.08)**	(7.90)**	(1.9)	(2.07)*
single female[b]	0.014	0.023	0.007	0.009	0.010	0.019
	(1.16)	(1.56)	(0.92)	(1.07)	(0.40)	(0.76)
age[a]	0.017	0.030	0.000	0.003	0.014	0.020
	(4.85)**	(6.79)**	(0.13)	(1.08)	(2.91)**	(3.77)**
black[a]	-0.182	-0.144	-0.060	-0.047	-0.177	-0.143
	(19.84)**	(13.06)**	(10.87)**	(7.16)**	(6.16)**	(4.78)**
hispanic[b]	-0.126	-0.116	-0.047	-0.045	-0.104	-0.078
	(11.44)**	(8.98)**	(7.79)**	(6.39)**	(2.85)**	(2.10)*
education[a]	***0.117***	***0.074***	***0.047***	***0.029***	***0.069***	***0.042***
	(23.79)**	***(12.97)*****	***(18.45)*****	***(8.79)*****	***(10.42)*****	***(5.44)*****
net worth[a]	0.077	0.064	0.066	0.058	0.011	0.010
	(4.09)**	(3.63)**	(6.42)**	(5.44)**	(2.97)**	(2.70)**
economic growth[a]		0.015		0.006		0.005
		(4.91)**		(2.76)**		(0.95)
general optimism[a]		***0.048***		***0.021***		***0.018***
		(13.88)**		***(9.30)*****		***(3.86)*****
optimistic weather forecast[b]		0.012		0.000		-0.001
		(1.21)		(0.06)		(0.12)
general precision[a]		***0.022***		***0.008***		***0.002***
		(7.59)**		***(3.46)*****		***(0.56)***
precision in weather forecast[a]		0.001		0.001		0.008
		(0.38)		(0.40)		(1.53)
risk tolerance[a]		0.002		-0.001		-0.007
		(0.67)		(0.32)		(1.50)
cognition: memory[a]		0.023		0.007		0.010
		(4.48)**		(2.15)*		(1.67)
cognition: iq[a]		0.026		0.014		0.017
		(5.68)**		(4.49)**		(2.86)**
cognition: numbers[a]		0.019		0.006		0.016
		(4.37)**		(2.25)*		(2.30)*
cognition: dementia control[a]		0.003		0.001		0.011
		(0.73)		(0.49)		(1.18)
Constant	0.271	0.265	0.118	0.119	-0.278	-0.296
	(34.31)**	(24.90)**	(22.22)**	(16.99)**	(18.20)**	(16.30)**
Observations	54,036	41,652	27,288	21,416	12,956	11,243
R-squared	0.17	0.19	0.06	0.07	0.03	0.03

t-statistics in parentheses. Standard errors are robust to heteroskedasticity and clustering within households.
* significant at 5% level; ** significant at 1% level
[a] standardized RHS variable: mean=0, std=1
[b] binary variable (average positive weather error can also be 0.5)