Learning, ambiguity and life-cycle portfolio allocation

Learning, Ambiguity and Life-Cycle Portfolio Allocation*

Claudio Campanale

July 29, 2009

*Claudio Campanale, Departamento de Fundamentos del Análisis Económico, Universidad de Alicante, Campus San Vicente del Raspeig, 03690, Alicante, Spain. Phone: +34 965903614 ext. 3262. E-mail: claudio@merlin.fae.ua.es.

I wish to thank the editor (Gianluca Violante), an anonymous associate editor and referee, Samuel Bentolila, Paolo Ghirardato, Francisco Gomes, Dirk Krueger, Antonio Mele, Filippo Taddei and seminar participants at the University of Vigo, the University of Turin - Collegio Carlo Alberto, the University of Alicante and the participants to the Econometric Society North American Meetings 2008, the CEA 2009 and SED 2009 for discussion and suggestions. I also wish to thank the Ministerio de Educación y Ciencia proyecto SEJ 2007-62656 and IVIE for financial support and CeRP for generous hospitality during the development of this project. Any remaining errors or inconsistencies are entirely my responsibility.

Abstract

In the present paper I develop a life-cycle portfolio choice model where agents perceive stock returns to be ambiguous and are ambiguity averse. As in Epstein and Schneider (2005) part of the ambiguity vanishes over time as a consequence of learning over observed returns. The model shows that ambiguity alone can rationalize moderate stock market participation rates and conditional shares with reasonable participation costs but has strongly counterfactual implications for conditional allocations to stocks by age and wealth. When learning is allowed, conditional shares over the life-cycle are instead aligned with the empirical evidence and patterns of stock holdings over the wealth distribution get closer to the data.

Keywords: Portfolio choice, life-cycle, ambiguity, learning

JEL codes: G11, D91, H55

1 Introduction

The last decade has witnessed a substantial surge of academic interest in the problem of households' financial decisions, perhaps triggered by the increased importance of personal savings for retirement consumption that is taking place in response to the debate about downsizing the role of unfunded social security. A number of empirical facts have been documented regarding in particular the stockholding behavior of households. These include the fact that participation rates, even though increasing over the years, are still at about half of the population and the moderate share allocated to stocks by participants. It has also been documented that the share of financial wealth allocated to stocks is increasing in wealth and roughly constant or moderately increasing in age.[1] Equally important has been the development of theoretical models that, based on a workhorse of modern macroeconomics, that is, the precautionary savings model, have tried to explore the same issue. The current paper joins this latter line of research by exploring the role of a class of non standard preferences in the context of the model cited above.

More specifically, in this paper I present a model of life-cycle portfolio choice where agents perceive the return to one of the assets to be ambiguous and are averse to ambiguity. As in Epstein and Schneider (2005) ambiguity can be reduced over time through learning. The basic framework of the model

[1] Among the papers that have uncovered the patterns of household financial behavior are Ameriks and Zeldes (2004), Bertaut and Starr-McCluer (2000), Curcuru et al. (2004) and Heaton and Lucas (2000) for the US. The book of Guiso et al. (2001) documented the same facts for a number of other industrialized countries as well and the work by Calvet, Campbell and Sodini (2007) has gone in much greater details to document stock-holding behavior among Swedish households.

is otherwise standard: agents have finite life and receive a stochastic earnings stream during working life, followed by a constant pension benefit in retirement. Agents cannot insure against earnings uncertainty, thus use savings as a self-insurance instrument. Beside that they save for the other usual reasons, that is, to finance consumption during retirement, to insure against uncertainty about the length of life and to leave a bequest. Saving can occur through two assets, a risk-free bond and a risky stock, and exogenous no borrowing and no short sale constraints are imposed. Trading in the stock requires payment of a fixed per period cost. Where the model departs from the basic framework is in the way agents perceive the stock return process. In this regard the model assumes that agents perceive the stock return process to be ambiguous, that is, they think they cannot know the exact distribution governing that process but think it lies in some set of distributions. Agents are averse to ambiguity according to the max-min utility model developed by Gilboa and Schmeidler (1989) in a static framework and extended to a dynamic setting by Epstein and Schneider (2003). It is also assumed that the ambiguity present in the stock return process can be reduced through the observation of the realized returns and that stock market participants have an advantage at doing so.

The model is solved numerically and its properties are analyzed under a broad set of parameters. It is shown that ambiguity aversion alone can generate moderate participation rates and conditional shares without resorting to large participation costs and it does so by assuming a fairly reasonable amount of ambiguity in the stock return process. On the other hand the model with ambiguity but no learning shows two very counterfactual properties when we look deeper at its implications: stock shares for market participants are strongly declining

in both age and wealth which is at odd with the empirical evidence. When learning is introduced, the model, while retaining reasonable average participation rates and conditional shares, generates a life-cycle profile of conditional stock allocation that is slightly increasing but with little variation as in the data. It also displays constant stock shares over wealth levels, thus moving a step in the right direction towards matching the empirically observed increasing pattern. The intuition for these results is that under ambiguity — and with short-selling constraints — the equity premium that is relevant for the decision maker, henceforth called "worst case" equity premium, is the lowest given the distributions in the posterior set. With learning the set of posteriors shrinks over the life-cycle thus improving the "worst case" equity premium and inducing investors to hold a larger share of stocks in their portfolios as they age. Moreover, given the assumption that stock market participants have an advantage at learning, wealthier agents who have started to participate earlier will generally face a smaller set of posteriors, hence a higher "worst case" equity premium. This increases their demands of stocks, although quantitatively this does not go as far as allowing to fully match the wealth-share profile observed in the data.

The main contribution of the paper is to document the implications of ambiguity aversion and learning in an ambiguous environment for household life-cycle portfolio allocation and to show that these features may have an important role in explaining the observed pattern of household financial choices. In doing so it joins two very active lines of research. The first one is the literature on portfolio allocation in precautionary savings models. This literature was first explored by Heaton and Lucas (1997 and 2000) and Haliassos and Michaelides (2003) in an infinite horizon setting and by Campbell et al. (2001), Cocco, Gomes

and Maenhout (2005) and Gomes and Michaelides (2005) in a life-cycle setting. These papers documented the basic properties of this type of model and pointed out the difficulties it has to explain the low participation rates and conditional stock shares observed in the data, in some cases proposing possible solutions. More recently a number of papers and in particular the ones by Benzoni et al. (2007), Lynch and Tan (2008) and Wachter and Yogo (2008) have looked for explanations of patterns of household stock market investment over the life-cycle and over wealth levels. In particular Benzoni et al. (2007) assume that labor income and stock market returns are co-integrated and show that under this assumption human capital is more like a stock for young agents lowering their demand of equity. Lynch and Tan (2008) obtain a similar result by allowing for correlation between stock returns and labor income growth and volatility. Wachter and Yogo (2008) assume the existence of both basic and luxury goods and show that under this assumption conditional portfolio shares of stocks are increasing in wealth. Contrary to those papers the current one retains both the assumption that labor earnings and stock returns are uncorrelated and that the utility function is homothetic. The major departure from the more traditional framework lies in the fact that stock returns are ambiguous, agents are averse to ambiguity and that ambiguity lessens over time through learning. The second line of research to which this paper is related is the one that has studied the implications of model uncertainty in asset pricing and portfolio choice. Contributions in a dynamic framework go back to Epstein and Wang (1994). More recently Cao, Wang and Zhang (2005) explored the implications of heterogeneity in ambiguity aversion for stock market participation and the equilibrium equity premium in a static framework and Leippold, Trojani and

Vanini (2005) studied a dynamic Lucas-style exchange economy with both ambiguity and learning. While the three papers cited above used the max-min model of ambiguity aversion, Ju and Miao (2007) introduced the Klibanoff, Marinacci and Mukerji (2006) smooth ambiguity model in a dynamic endowment economy with learning about the hidden state and showed that the model can match a wide set of asset pricing facts. Model uncertainty has also been studied in the alternative framework of robust control of which two applications to asset pricing are Maenhout (2004) in an endowment economy and Cagetti et al. (2002) in a business cycle model. Examples of explorations of the role of ambiguity aversion in portfolio choice models are Garlappi, Uppal and Wang (2006) that use a static mean-variance approach and the same paper by Maenhout cited above. This latter paper is dynamic as the present one, however it omits labor income and uses the robust control approach.

The rest of the paper is organized as follows. In Section 2 I present the description of the model, in Section 3 I report the choice of parameters, in Section 4 I report the main findings of the analysis and finally in Section 5 some short conclusions are outlined. The paper is completed by two appendixes where a short but formal treatment of the learning model and a description of the numerical methods used to solve the model are provided.

2 The Model

2.1 Demography and Preferences

Time is discrete and the model period is assumed to be 1 year. Adult age is denoted with the letter t and can range from 1 to $T = 80$ years. Agents are

assumed to enter the model at age 20 so that real life age is equal to $t + 20$. Each agent faces an age changing conditional probability of surviving to the next period which will be denoted with π_t. Surviving agents work the first 45 years and retire afterwards.

Agents do not value leisure, hence they derive utility from the stream of consumption they enjoy during their life-time only. Utility over consumption is defined by a period utility index $u(c_t)$ which will be assumed to be of the standard iso-elastic form. Agents also derive utility from leaving a bequest; the bequest motive is of the so called *warm glow* form hence can be simply represented by a function $D(.)$ defined over terminal wealth.

In the economy there are two independent sources of uncertainty. The first one is determined by the stochastic process for labor earnings and it is standard in that I assume that agents know its distribution. This process will be described in a later subsection. The second one is the process for stock returns. Following Epstein and Schneider (2005) it is assumed that this process is i.i.d. and that agents perceive it as ambiguous. In other words they assume that stock returns may be drawn from a whole family of distributions and even if they can learn from past observations of realized returns, they can never shrink the set of distributions to a singleton.

In every period an element $h_t \in H$ is observed: this pair consists of a realization of the stock return $w_t \in W$ and a realization of the labor efficiency unit shock $z_t \in Z$. At age t then the agent's information set consists of the history $h^t = (h_1, h_2, ..., h_t)$. Given that the horizon is finite the full state space will be H^T. The agent ranks consumption plans $c = \{c_t\}$ where consumption c_t depends on the history h^t. At any age $t = 1, 2, ..., T$ and given history h^t,

the agent's ordering of consumption plans is represented by a conditional utility function U_t defined recursively by:

$$U_t(c; h^t) = \min_{p \in \mathcal{P}_t(w^t)} E^p[u(c_t) + \beta E^{z_{t+1}} U_{t+1}(c; h^{t+1})] \tag{1}$$

where β and u are defined above. The set of probability measures $\mathcal{P}_t(w^t)$ models beliefs about the next realization of the stock return process w_{t+1} given history up to w_t. When this set is a non-singleton such beliefs reflect ambiguity and the minimization over p reflects ambiguity aversion.[2] The set of probability measures $\{\mathcal{P}\}$ is called *process of conditional one-step ahead beliefs* and together with $u(\cdot)$ and β constitute the primitives of the functional form.

2.2 Labor Income and Pensions

I use the indexed letter Y_t to denote income. During working life income is determined by an uncertain stream of labor earnings. Earnings can be expressed as the product of two components:

$$Y_t = G(t) z_t \tag{2}$$

where the function $G(t)$ is a deterministic function of age meant to capture the hump in life-cycle earnings that is observed in the data. The second term, z_t, is a stochastic component that follows an AR(1) process in logarithms:

$$ln(z_t) = \rho ln(z_{t-1}) + \varepsilon_t \tag{3}$$

where ε_t is an i.i.d. normal random variable.

In the retirement years agents receive a fixed pension benefit, so that

$$Y_t = Y_{ss}. \tag{4}$$

[2]The minimization is taken with respect to p only since the process for labor earnings is independent and is not ambiguous.

2.3 Financial Assets

Agents can use two different assets to carry out their investment plans. First there is a one period risk free bond with price q and return $R^f = \frac{1}{q}$. The second asset is a risky stock. Investors perceive the return to this asset ambiguous but the actual return R^s_{t+1} is generated by a single i.i.d. process that can take two values: $\mu \pm \delta$ with equal probability. Consequently $\mu - \frac{1}{q}$ is the average equity premium and δ is the standard deviation of the equity return. The adoption of this simple process for equity returns is needed to formulate the model of learning about ambiguous stock returns used here and described in the next section.

Trade in the two assets is subject to three frictions. First all households are prevented both from borrowing and from selling short stock. Denoting bond and stock-holdings with B_t and S_t respectively this implies:

$$B_t \geq 0 \tag{5}$$

$$S_t \geq 0. \tag{6}$$

Second, households who do participate in the stock market are subject to a minimum investment limit that I denote with $\underline{S}$, that is, the relevant constraint for them is

$$S_t \geq \underline{S}. \tag{7}$$

Third it is assumed that participation in the stock market requires payment of a fixed cost F_p in each period.

A further important assumption about the stock market is that in the model households that participate receive a signal about the ambiguous stock return

process for sure while households who do not participate receive it only probabilistically. The letter ξ will be used to denote the probability that a non stockholder can infer information about the return process from the observed realized return in any given year.

The minimum investment requirement and the differential flow of information to stock holders and non stock holders are non standard, hence require some comments. The fact that participants receive signals about the stock return process with greater probability than non participants can be justified based on ideas recently expressed in the "rational inattention" literature developed for example by Sims (2006) and applied to monetary theory by Mankiw and Reis (2006) and to consumer theory by Reis (2004). The founding principle behind this theory is the observation that even though information may be in principle free, still absorbing and processing it requires the allocation of resources to it so that agents may choose to disregard it. As in Mankiw and Reis (2006) the model presented here assumes exogenously a differential flow of information to different agents rather than deriving the result from an optimal information acquisition problem. [3] The advantage of stockholders though seems reasonable if one takes into account that stockholders may receive already processed information through their broker or other financial advisor or as a side product of activities required by stock-holding like compiling the relevant section of tax forms. More generally it is arguable that if an agent has only limited processing resources to allocate to her financial decisions she will follow more closely those assets she has in her portfolio. With respect to the minimum equity requirement

[3] Mankiw and Reis (2006) estimate the probabilities of information update. The complexity of the current model makes this impossible so that a baseline case plus sensitivity analysis will be carried out.

observe that as long as participating in the stock market gives an informational advantage towards resolving ambiguity, as assumed here, it can be optimal to pay the fixed cost even if the current "worst case"equity premium is negative. This could potentially make some agents pay the fixed participation cost but hold no stocks which would be contradictory. At an empirical level this choice can be justified for example by observing that mutual fund companies and brokerage houses often impose minimum investment limits. Also some work like Heaton and Lucas (2000) study stock portfolio allocation at the empirical level conditional on stock holding being above a threshold of 500 dollars to rule out occasional investors.

2.4 Learning

In the current section an informal description of the learning process is given while the mathematical details are left for an appendix. As it was said the stock return can be described by a two point i.i.d. process. Let us denote with w_t the stock return state, where $w_t = 1$ refers to the high stock return state and $w_t = 0$ denotes the low stock return state. It is assumed that the probability of a high and a low stock return are equal, that is, $pr(w_t = 1) = pr(w_t = 0) = 0.5$. Ambiguity arises because agents do not think they know the probability of a high stock return $pr(w_t = 1)$, but they think it can lie in some set. Starting from some set of priors on this probability, learning allows the size of this set to change over time in response to the information carried by the stock return realization. As it was said in the previous section the model assumes that stock market participants update the set of posteriors in each period, while non participants do so only with some probability.

The learning environment is described by two parameters.[4] A first parameter denoted with θ describes features of the environment that the agent can learn about and it is assumed that $\theta \in \Theta = [\bar{\lambda}, 1 - \bar{\lambda}]$ where $0 < \bar{\lambda} < \frac{1}{2}$. There are also features of the data generating process that the agent does not think can be learnt and this is reflected by a multiplicity of likelihoods. In practice the likelihood that a high stock return is observed given θ is $\ell(w_t = 1|\theta) = \theta + \lambda$ where $\lambda \in [-\bar{\lambda}, \bar{\lambda}]$. At the beginning of life agents believe that $\theta \in \mathcal{M}_0 \subset \Theta$ where $\mathcal{M}_0$ is the set of initial priors. Agents observe the sequence of stock returns. If the agent had the chance to see an infinite sequence of stock return realizations the set of posteriors $\mathcal{M}(w^t)$ would converge to the singleton $\theta = \frac{1}{2}$, that is, the true probability of a high stock return. Even in this case though ambiguity would not entirely vanish because of the multiplicity of likelihoods. With a finite horizon the set of posteriors tend to shrink but cannot converge to a singleton leaving a larger extent of ambiguity. A parameter that will be denoted α in what follows, regulates how fast the set of posteriors responds to new information.

As an illustration I report below in figure 1 the boundaries of the set of posteriors as a function of an agent's age under some special conditions that help highlighting the effects of learning. In particular it is assumed that the agent has updated the set of posteriors in every period and that the fraction of high stock returns is a half in any period. [5] The parameters of the learning model, that is, $\bar{\lambda}$, α and the initial set of priors $\mathcal{M}_0$ are chosen as in the baseline

[4] This framework of learning under multiple priors was set in Epstein and Schneider (2005). The reader is referred to that paper for a detailed description of the environment and the theoretical results. An appendix in this paper gives a short formal description of that environment.

[5] This is possible only for even periods, values for odd periods are interpolated.

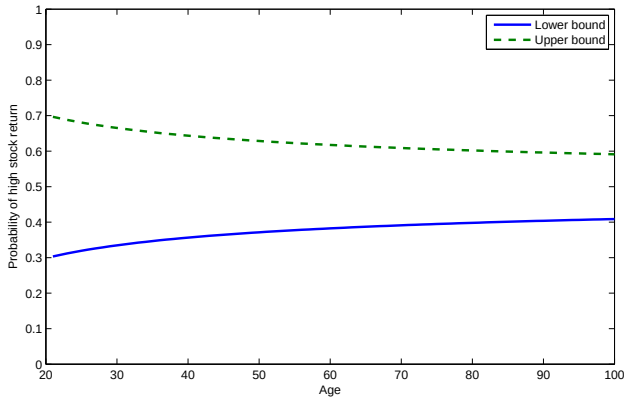

Figure 1: Posterior bounds as a function of the number of signals.

model in section 4.1.2. As it can be seen the set of posteriors shrinks over time and it does so at a decreasing pace: at the beginning of life the agent thinks that the probability of a high stock return can be anywhere between approximately 0.3 and 0.7, while after 80 signals she thinks this probability can be only between 0.4 and 0.6. Figure 2 reports the bounds of the range of expected equity returns corresponding to the bounds of the posteriors represented in figure 1. As it can be seen, the range of equity premia shrinks in response to the reduction in the posterior set. At the beginning of life the expected equity returns can be, in the agent's belief as low as 1.5 percent and as high as 13 percent; after the observation of 80 signals they can only range from about 4 percent to 10 percent. Notice that the upper dotted line represents the expected equity return corresponding to the largest probability of a high stock return in the posterior set and is also the max-min return for an ambiguity averse agent that has a short position in stocks. The lower dashed curve represents the expected equity return corresponding to the smallest probability of a high stock return in the

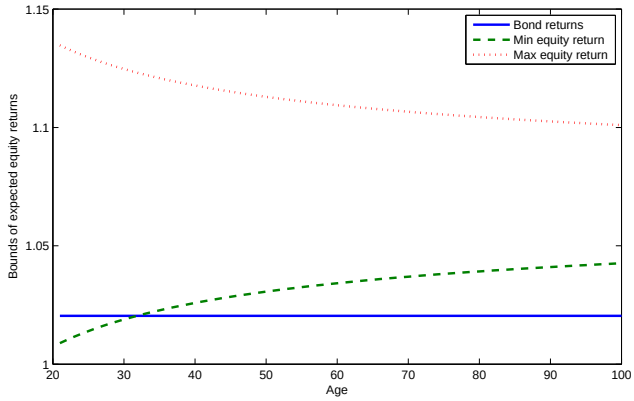

Figure 2: Bounds of expected equity returns as a function of the number of signals.

posterior set and is also the max-min return for an ambiguity averse agent that has a long position in stocks. Given that in the model short-selling is ruled out this latter curve represents the "worst case"equity premium in the baseline model; for comparison the straight continuous line reports the bond return, thus the difference between the two curves gives a glimpse at the evolution of the "worst case"equity premium over the life-cycle. For the purpose of illustration the curves in figures 1 and 2 are drawn for the very special case that at each time half of the past realizations of stock returns are high and half low, so in some sense they are also the average of a large number of simulated random draws. For more general random draws the posterior bounds and corresponding equity returns would show a sawtooth path with the tendency of the set of posteriors to shrink but with the possibility that it expands in response to a sequence of low stock returns.

Finally Epstein and Schneider (2005) proved that the set of posteriors can

be characterized as a function of two state variables only, that is, the fraction of prior stock return realizations that were high that we denote with ϕ_t and the number of prior observations on the stock return realization that we denote with n_t. This feature of the learning model used here is very convenient because it only adds two state variables to the dynamic programming problem, thus keeping it tractable even if somewhat burdensome.

2.5 The Optimization Problem

With the description of the model given above it is now possible to state the household's optimization problem in dynamic programming form. In order to make the description more readable I divide the section into two paragraphs, the first one describing the indirect utility of an agent if she chooses to participate in the stock market and the second one for an agent that chooses not to participate.

Participation Indirect Utility The indirect utility of an agent if she decides to participate in the stock market is given by the following equation:

$$V_t^{part}(X_t, z_t, \phi_t, n_t) = \max_{c_t, B_{t+1}, S_{t+1}} \min_{p_t \in \mathcal{P}_t} \Big\{ u(c_t) + \\ + \beta E\Big[\pi_{t+1} V_{t+1}(X_{t+1}, z_{t+1}, \phi_{t+1}, n_{t+1}) + (1 - \pi_{t+1}) D(X_{t+1})\Big]\Big\} \quad (8)$$

subject to the budget constraint

$$c_t + qB_{t+1} + S_{t+1} \leq X_t + Y_t - F_p \quad (9)$$

the transition equation for financial resources

$$X_{t+1} = B_{t+1} + R(w_{t+1}) S_{t+1} \quad (10)$$

the transition equation for the fraction of time a high return was observed

$$\phi_{t+1}(w_{t+1}) = \frac{n_t\phi_t + w_{t+1}}{n_t + 1} \tag{11}$$

, the equation describing the number of past signals about the stock return process observed

$$n_{t+1} = n_t + 1 \tag{12}$$

the inequality constraints (5) and (7) and equations (2), (3) and (4) that define the nonfinancial income available to the agent from labor earnings or pensions. The agent's state variables are the amount of financial resources X_t, the labor earnings shock z_t, the fraction of past observations on the stock return that were high ϕ_t and the number of those signals observed n_t. The agent chooses the amounts of stocks, bonds and consumption that maximize his utility but since he has max-min preferences he minimizes these optimal values with respect to $\mathcal{P}_t$, the set of admissible beliefs. Given the way the learning process is modeled the set $\mathcal{P}_t$ is defined by $\{(\theta, \lambda)|\theta \in \mathcal{M}_t^{\alpha}\ , \lambda \in [-\bar{\lambda}, \bar{\lambda}]\}$. The argument of the function to be maximized that we find in curly braces is the sum of the utility of current consumption plus continuation utility which in turn is given with probability π_{t+1} — the probability of survival — by the continuation value function and with probability $1 - \pi_{t+1}$ by the utility from bequests function $D(X_{t+1})$. The expectation operator is taken with respect to the distributions p_t and the distribution of next period labor shock conditional on the current value z_t. Inequality (9) is a standard budget constraint: it states that the expenditures in consumption, bond and stock purchases must not exceed the sum of financial resources, plus the income from earnings or pensions minus the fixed participation cost. Equation (10) describes the evolution of financial

resources as the sum of one-period bonds plus stock times its gross return. This return can take a high value or low value depending on whether the state w_t takes a value of 1 or 0. Equation (12) shows that for an agent who decides to participate in the stock market the number of past signals on the stock return process increases by one between the current and the next period and finally equation (11) describes how, depending on whether the realized return is high or low, — w_{t+1} equal to 1 or 0 — the past fraction of high signals observed is updated.

Non Participation Indirect Utility The indirect utility of an agent who decides not to participate in the stock market is given by the following equation:

$$V_t^{nopart}(X_t, z_t, \phi_t, n_t) = \max_{c_t, B_{t+1}} \min_{p_t \in \mathcal{P}_t} \Big\{ u(c_t) + \\ \beta E\Big[\pi_{t+1} E V_{t+1}(X_{t+1}, z_{t+1}, \phi_{t+1}, n_{t+1}) + (1-\pi_{t+1}) D(X_{t+1})\Big]\Big\} \quad (13)$$

subject to the budget constraint

$$c_t + qB_{t+1} \leq X_t + Y_t, \quad (14)$$

the law of motion of financial resources

$$X_{t+1} = B_{t+1}, \quad (15)$$

the law of motion of the fraction of past high signals on the stock return process

$$\phi_{t+1} = \begin{cases} \phi_t & \text{with probability } 1-\xi \\ \frac{n_t \phi_t + w_{t+1}}{n_t + 1} & \text{with probability } \xi \end{cases}$$

and the law of motion of the past number of signals observed

$$n_{t+1} = \begin{cases} n_t & \text{with probability } 1-\xi \\ n_t + 1 & \text{with probability } \xi. \end{cases}$$

As it can be seen the problem has the same state variables as the one of an agent who chooses to participate and the maximization on the right hand side of equation (13) differs from the analogous equation (8) only in that the maximization is performed on consumption and bonds with the amount of stocks being zero by definition. Also notice that the minimization with respect to the distributions in the set $\mathcal{P}_t$ must take place even if the agent does not buy stocks. This is because the investor is still exposed to ambiguity through the probability that a signal about the stock market return process is observed. The budget constraint (14) simply states that for a non participant expenditures on consumption and bonds must not exceed income from labor or pensions plus financial resources and the law of motion (15) expresses that fact that for a non stockholder financial resources next period coincide with the amount of one period risk-free bonds purchased in the current period. The last two laws of motion reflect the probabilistic receipt of a signal about the stock return generating process by an agent who does not participate in the stock market. With probability ξ the agent receives the signal, hence the number of past observations received grows by one and the fraction of those that were high is updated based on the value of the shock w_t. With probability $1-\xi$ the investor does not observe a signal so that both the number of observations and the fraction of those that were high stay constant at their current value ϕ_t and n_t.

Finally the household's optimal value function will result by taking the maximum of the indirect utility from participating and from not participating in the stock market:

$$V_t(X_t, z_t, \phi_t, n_t) = max\Big\{V_t^{nopart}(X_t, z_t, \phi_t, n_t), V_t^{part}(X_t, z_t, \phi_t, n_t)\Big\}. \quad (16)$$

The problem has no analytical solution so that numerical methods are used

to examine its properties. The solution procedure consists of two parts: first decision rules are computed from the agent's dynamic programming problem; second these decision rules are used, together with random draws of the stochastic variables, to compute life-cycle profiles for 1000 agents. The simulation is repeated 30 times and the reported results are obtained by averaging over those repetitions. More details about the solution method are given in Appendix B.

3 Parameter Calibration

3.1 Preferences Parameters

Preferences are defined by the functional form and parameters of the period utility index and the function defining the utility of bequests plus the subjective discount factor. The utility index is chosen to be of the standard iso-elastic form: $u(c_t) = \frac{c_t^{1-\sigma}}{1-\sigma}$ and a baseline value of 2.5 is chosen for σ, the coefficient of relative risk aversion. A sensitivity analysis on this parameter will be performed, using values of 1.5 and 3.5 as well. These values are somewhat lower than those typically used in the life-cycle portfolio choice literature but more in line with the preferred values of macroeconomic studies. The utility of bequest function is defined as $D(X_{t+1}) = d\frac{(X_{t+1}/d)^{1-\sigma}}{1-\sigma}$, that is, I use the same functional form and curvature of the utility index. The additional parameter d which sets the strength of the bequest function is taken to be 2.5, one of the values used by Gomes and Michaelides (2005) who also use the same functional form. The subjective discount factor β is set equal to 0.95 a value commonly used in the macro and finance literature. The effective discount rate is determined also by the conditional survival probabilities which are taken from the male survival

probabilities available at the "Berkeley Mortality Database ".[6]

3.2 Learning Parameters

The process for learning is characterized by three quantities: the long run ambiguity, the initial ambiguity and the speed at which the agents are willing to get rid of ambiguity over the life-cycle by incorporating the new information contained in the sequence of realized stock returns. For this reason we need to specify three parameters to fully describe the features of learning in this ambiguous environment. The long run ambiguity, that is the component of ambiguity that the agent thinks he cannot get rid of even in the long run — i.e. asymptotically — is entirely fixed by the parameter $\bar{\lambda}$ whose value I fix at 0.01. Following Epstein and Schneider (2005) this value implies that in the long run the set of posteriors of the probability of high stock returns shrinks to $[0.49, 0.51]$ which implies a range of equity premia of 64 basis points. This number seems sufficiently small to leave substantial scope for learning in the model. The speed at which the agent is willing to get rid of ambiguity is governed by the parameter α and its value is taken to be 0.2 in the baseline case. Finally, once the speed of learning is fixed, I determine the initial extent of ambiguity perceived in the data by assuming that prior to entering the model agents observed a certain number of stock return realizations that follow exactly the data generating process, that is, are 50 percent high and 50 percent low. The number of such observations is fixed at 20 in the baseline case. It should be stressed that this statement is only technical in nature, that is, it only serves the purpose of fixing initial ambiguity and should not be taken literally as to imply that the agent observed

[6]The database is available at the website http://www.demog.berkeley.edu/ bmd/.

stock market realizations prior to entering the model.[7] Under the parameters chosen the initial set of posteriors of the probability of high stock returns is the interval $[0.30336, 0.69662]$; with such an interval the difference between the maximum and minimum expected equity premium implied by the agent's set of beliefs is 12.58 percentage points wide. Since there is no direct evidence on which one can base the choice of these two parameters a sensitivity analysis will be performed on both the speed of updating and the initial ambiguity. Also empirical evidence to support the claim that the extent of ambiguity over the life-cycle implied by this choice of parameters is reasonable will be discussed in the result section.

3.3 Labor Income and Pensions

The specification of the labor earnings process during working life requires fixing two sets of parameters. The first one refers to the function $G(t)$ which defines the deterministic hump-shaped component of earnings. This function is assumed to be a third degree polynomial in age and the coefficients are taken form the estimates by Cocco, Gomes and Maenhout (2005) for high school graduates. These estimates when aggregated over five year groups are also consistent with the ones of Hansen (1993) based on the whole population. The second one is the idiosyncratic component z_t which is assumed to follow an $AR(1)$ process with autocorrelation coefficient $\rho = 0.95$ and a standard deviation of the innovation $\sigma_\varepsilon = 0.158$, both values taken from Hubbard, Skinner and Zeldes (1994).

During retirement it is assumed that agents receive a fixed pension benefit

[7] In principle one could think that the agent got some information about the stock market process from family members, neighbors or other sources during childhood or teen-age but I don't want to stick to that interpretation.

equal to 68 percent of average lifetime earnings conditional on the last year of earnings. The replacement ratio implicit in this formula is chosen based on Gomes and Michaelides (2005).

3.4 Asset Returns and Fixed costs

The bond price is set at 0.98 which implies a risk free return of about 2 percent annually. The stock return process is modeled as a two point i.i.d. process with the expected value μ set at 7.5 percent percent annually and a standard deviation δ of 16 percent. The value of the risk free rate is close to but a little lower than what reported in Cecchetti et al. (2000) and a little higher than what other authors in the asset pricing literature have used.[8] The equity premium of 5.5 percent is a little below the values used in that literature that range from 5.75 in Cecchetti et al. (2000) to about 8 percent in Lettau (2003). As in other work on life-cycle portfolio allocation the use of a reduced equity premium may be thought to proxy for the existence of proportional transaction costs that the agent normally has to pay even after paying the fixed participation cost and that, if modeled explicitly, would add a non trivial extra burden on the numerical solution of a problem that is already quite demanding.[9]

Empirical work that tried to measure the magnitude of fixed stock market participation costs found values in the range of 50 to 200 dollars.[10] The cost in the model is then set so that when compared to model wages it is consistent with values at the lower end of that interval.

The minimum equity investment is set at about 4 percent of average annual

[8] See for example Mehra and Prescott (1985) or Lettau (2003).

[9] The choice to use a reduced equity premium can be found for example in Campbell et al. (2001), Cocco, Gomes and Michaelides (2005) and Gomes and Michaelides (2005).

[10] See Paiella (2001) and Vissing-Jørgensen (2002).

earnings in the economy. Assuming a plausible 35000 dollar average earnings this would be equivalent to 1400 dollars, a value in line with the minimum investment requirement at several large mutual fund companies. [11] As a robustness check the model was also solved assuming a smaller and larger minimum requirement of approximately 2.25 and 5.75 percent average annual earnings. Since the results are very similar to the baseline case they won't be reported.

Finally, to complete the description of the parameters concerning the assets in the economy we need to specify the probability with which agents that do not participate in the stock market get a signal about the process generating equity returns. Unfortunately here there is no empirical base for calibration hence I present results for a baseline value of 0.2 and present sensitivity analysis using values of 0.3 and 0.1 as well.

4 Results

In this section I report the results of the simulation of the model. The main focus throughout the section will be on average conditional allocations to stock and their patterns over the life-cycle and by wealth levels. As a check on the model also average participation rates and their pattern over the life-cycle will be reported. To economize on space though I will omit participation rates by wealth since those conform closely with what other models have found and with the data. The section is divided into three subsections. In the first one I report a benchmark case and, for comparisons, results of models that abstract from

[11] I performed a casual search of some large mutual fund companies' web-sites and found that they impose such requirements and that they range from 250 $ at American Funds to 3000 $ at Vanguard. The latter also provides brokerage services and imposes the same minimum investment on those.

learning and from both learning and ambiguity but that are otherwise similar to the complete model in the choice of parameters. In this section the intuition behind the results will also be described. In the second subsection I report the results of a sensitivity analysis on several parameters to check the robustness of the findings of the model. In the third subsection I report results that are obtained with sample draws of the stock return sequence to highlight the potential of the model to generate cohort effects. I also describe the implications of the different models presented here for the wealth distribution and examine briefly the results of models that make alternative assumptions about the difference of participants and non participants in absorbing the information contained in the return realizations into their posteriors.

4.1 A Benchmark Case

4.1.1 The Model without Learning

In this section I report the results that are obtained when learning is omitted from the model so that the extent of ambiguity that agents perceive in the stock return process is constant over time. Otherwise the model uses the same parameters as the baseline case with learning. The only other difference is that the per-period participation cost is set so as to obtain the same average participation rate as in the baseline case with learning. For exposition purposes I also report the results of a model that, keeping the same set of parameters, abstracts from ambiguity altogether and recalibrates the participation cost to get the same participation rate as the previous model.

The model with neither ambiguity nor learning can generate an average participation rate of 43.6 percent with a participation cost equal to about 11

percent of average annual earnings. Taking a rough estimate of the latter to be 35000 dollars implies a fixed cost of 3850 dollars per year. This cost is clearly huge and beside that the model generates an allocation to stocks for market participants of 99.7 percent. This result is not new and the intuition was well explained for example in Cocco, Gomes and Maenhout (2005): with stock returns uncorrelated with labor earnings the latter are a form of implicit holdings of a risk-free asset. Given this and the high equity premium households will want to invest almost entirely in risky stocks, hence the very high conditional stock share. On the other hand because of the large benefit of investing in the stock market, a large cost will be needed to deter agents from participating in it.

When ambiguity is added it is possible to obtain virtually the same participation rate — 43.7 percent — with a much smaller per period participation cost, that is, 0.7 percent of annual earnings. Again taking a reference value for average annual earnings of 35000 dollars this cost is equivalent to about 245 dollars. This number is reasonably small, although slightly larger than estimates provided by Paiella (2001) and Vissing-Jørgensen (2002) and ranging between 50 and 200 dollars. At the same time the model generates an allocation to stocks conditional on participation of about 58.2 percent a number close to the figures reported for example by Guiso et al. (2001). This result is obtained by assuming a "worst case"expected equity premium for a stock buyer of 0.9 percent. Since the relevant equity premium used by the agent in his optimal allocation problem is much lower this justifies the substantially smaller portfolio share of stocks. At the same time both the lower "worst case"equity premium, hence smaller perceived benefit of stock holding and the lower optimal share will

enable a reasonably small per period participation cost to generate moderate participation rates as in the data.

Even though the model with ambiguity seems to provide a straightforward explanation for the moderate stock market participation rate and conditional stock shares one should check that this is obtained with a reasonable amount of ambiguity and that the model is consistent with other features of investors' behavior. With respect to the first issue one useful comparison can be drawn with the amount of uncertainty about the expected equity premium from a survey of finance professors reported in Welch (2000). This seems particularly appropriate in light of the experimental findings of Fox and Tversky (1995) that it is precisely when an agent knows that there are experts that are more knowledgeable than her on a subject that ambiguity aversion arises more strongly. The author interviewed 226 academic professors in finance asking among else to report a measure of the central tendency and 95 percent confidence interval of the arithmetic 30-year average of their equity premium forecast. He found the average central tendency was about 7 percent with the average confidence interval ranging from 2.2 to about 13 percent. If one interprets the 95 percent confidence interval as a plausible representation of the multiplicity of stock return distributions entertained by those academic professors the "worst case"equity premium would be 2.2 percent versus the 0.9 percent used in the current experiment. Also more than a tenth of the respondents considered as a "pessimistic scenario"an equity premium of 0 percent or less. Hence the number used here seems quite reasonable or even conservative, especially if one considers that it is reasonable to assume that a lay person faces more uncertainty about the expected equity premium than an academic professor in finance. A second argument can be

developed along the lines of Jorion and Goetzmann (1999). The two authors claim that when we use the experience of the last century in the US to assess the equity premium we are actually conditioning on the experience of the most successful equity market in recent history. They compute return for other markets in the period 1921 to 1996 and find a median average return that is 3.5 percentage points below the one in the US with some countries without any mayor disruption in the functioning of the market — like Italy or New Zealand — at 4.5 percentage points below the US one. Something similar can be said for the US in previous times since according to Siegel (1992) the equity premium in the period 1802 to 1870 was only 1.5 percent, 6 and a half percentage points below the one in the period 1926 to 1990. To the extent that investors are aware of this, assuming that they think that the "worst case"equity premium is 4.5 percent less than the true one, as it is done here, seems reasonable. Finally, a third argument is based on the well known fact that the average of a volatile series cannot be pinpointed with high precision unless a very long draw of data is available. For example Cochrane (1997) reports that with a 50 year long sequence of US data running from 1947 to 1996 the 95 percent confidence interval for the average equity premium is $\pm$ 5 percent away from its 8 percent mean. Once again if we think about the 95 percent confidence interval as the set of plausible distributions, a "worst case"equity premium 4.5 percent below the true one as assumed here seems quite reasonable.

On the other test though the model fares quite worse. This can be seen with the help of Figure 3 where I report the life-cycle profiles of conditional allocations to stocks for both models considered in this section. The continuous line at the top represents the profile for the model without ambiguity and consistently with

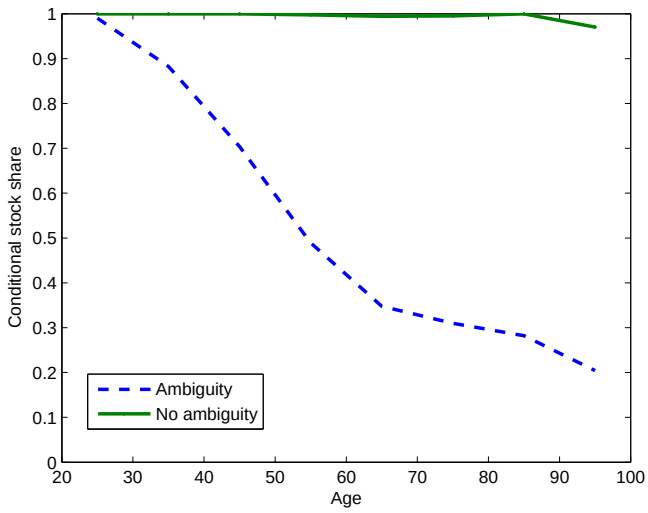

Figure 3: Life-Cycle profile of conditional stock shares, models without learning/ambiguity.

an overall average conditional share of 100 percent it is almost equal to that value at all ages. The dashed line represents instead the profile for the model with ambiguity. As it can be seen the conditional average allocation to stocks starts at 100 percent in the first decade of life and then monotonically declines to only 20 percent in the last decade of life. This profile is strongly at odd with the data where conditional shares tend to increase slightly early in the very first decade of life and then are virtually constant afterwards. Similar declining patterns have been found in other studies like Cocco, Gomes and Maenhout (2005) and Gomes and Michaelides (2005), especially during working life. Compared to those papers the decline in the model with ambiguity is quantitatively even stronger and persists all over the life-cycle. The intuition is that early in life the agent holds a large amount of relatively safe human

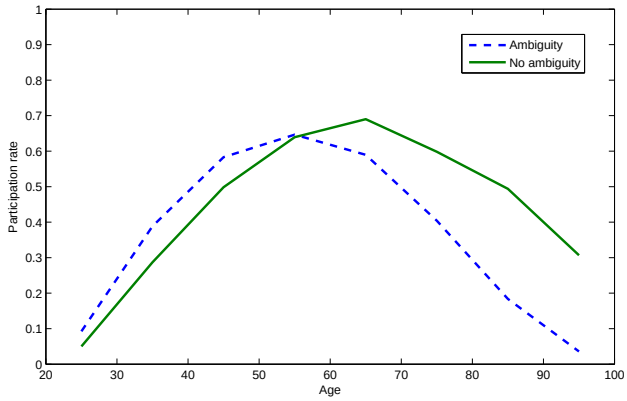

Figure 4: Life-Cycle profile of participation rates, models without learning/ambiguity.

capital and little financial wealth hence would like to invest the latter entirely in stocks to benefit from the equity premium. As it ages human capital gets progressively smaller and financial wealth accumulates inducing the agent to diversify towards safe bonds. After retirement the further path depends on whether wealth is drawn down more rapidly than the progressive reduction in the residual non financial wealth as the horizon shortens. While this intuition is common with the one in the models cited above, here the reduced perceived benefit of stock market participation associated with the "worst case"equity premium makes the decline in stock shares more substantial.

In Figure 4 I report the life-cycle profiles of participation rates for the models analyzed in this section. As it can be seen both the model without and the model with ambiguity generate hump-shaped life-cycle profiles of participation rates. The hump is quite pronounced in both models. In the model without ambiguity this is simply the consequence of the fact that in order to match the

average participation rate an implausibly large per-period participation cost is needed, hence participation rates tend to be very low at both ends of the life-cycle when the wealth stock has not yet formed or has been depleted. In the model with ambiguity the per-period participation cost is lower but the benefit of participation per unit invested is also lower because the relevant equity premium, corresponding to the "worst case" distribution of stock returns is small. As a consequence once again households will find it convenient to participate in the stock market only when the amount of wealth that can be invested in it is substantial, something that occurs only in the central part of the life-cycle.

Summarizing the assumption that the agent considers the stock return ambiguous and is averse to ambiguity helps generating moderate participation rates in the stock market and moderate portfolio allocations to stocks for participants but it does so at the cost of taking a step in the wrong direction as far as the life-cycle profile of conditional shares is concerned. Motivated by this finding in the next section I introduce learning in the model.

4.1.2 The Baseline Model with Learning

In this section I present the results of the model with learning in a multiple prior environment *à* la Epstein and Schneider (2005) described in the model section of this paper. The choice of parameters is the one described as the baseline case in the calibration section. Under that calibration the model produces an average participation rate of 43.1 percent and an average share invested in stocks for participants of 59.6 percent. Both figures are roughly consistent with the empirical evidence: the participation rate is a little below the most recent figures which are around 50 percent starting from the 1998 Survey of Consumer

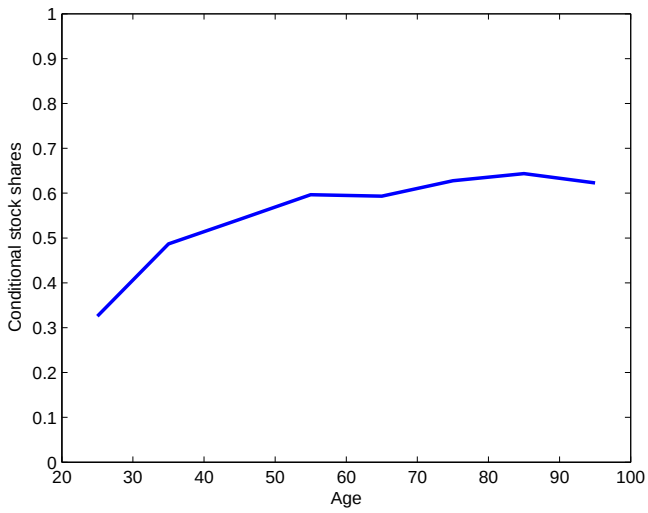

Figure 5: Life-cycle profile of conditional stock shares, baseline case.

Finances and somewhat above the ones for the preceding years — they were at 40.4 percent in 1995.[12] The conditional share was 59.4 percent in the 1998 SCF according to Guiso et al. (2001) and only slightly below that figure in the 1995 and 2001 SCF according to Heaton and Lucas (2000) and Gomes and Michaelides (2005) respectively. The life-cycle profile of conditional stock shares is reported in Figure 5. As it can be seen, when learning is allowed the strongly declining profile is overturned and substituted with an increasing profile that except for the first decade of life shows indeed a quite limited variation over the life-cycle. This brings the model predictions close to their empirical counterpart since in the data conditional stock shares tend to show little variation over age, with a small increase early in life. This is confirmed by Table 1 where I report the

[12]The sources for these figures are the studies by Bertaut and Starr-McCluer (2000) and Guiso el al. (2001).

Table 1: Conditional shares by age

	20-30	30-40	40-50	50-60	60-70	70+
Data	52.0	53.7	61.8	62.1	61.4	59.4
Ambiguity	98.9	88.2	70.4	48.9	34.8	30.5
Learning	32.6	48.7	54.1	59.7	59.3	63.2

data underlying Figure 5 together with data from the US presented in Guiso et al. (2001). In the model all wealth is financial, hence the reported shares are the shares of the risky asset in total financial wealth which coincides with total non-human wealth. For comparability the data from Guiso et al. (2001) refer to the same qualitative variable, that is, they represent the share of stocks — directly and indirectly held — in total financial wealth which excludes housing and business assets. As we can see by comparing the first and last row of the table the life-cycle profile generated by the model tracks the empirical one quite closely from the second decade to the end of life and only underestimates it somewhat in the first decade of life. For comparison the second row reports also the profile for the model with ambiguity but not learning presented in the previous subsection. This profile is by contrast strongly monotonically decreasing and quite far from the empirical one also in quantitative terms.

In order to develop the intuition to explain this result we can look at Figures 6 and 7. The first of the two figures represents the "worst case"expected equity return for a buyer as a function of the number of observations of the realized stock return assuming that these were 50 percent of the time high and 50 percent of the time low. In the same graph also reported by the flat continuous line is the return on the risk-free bond. As it can be seen the "worst case"equity

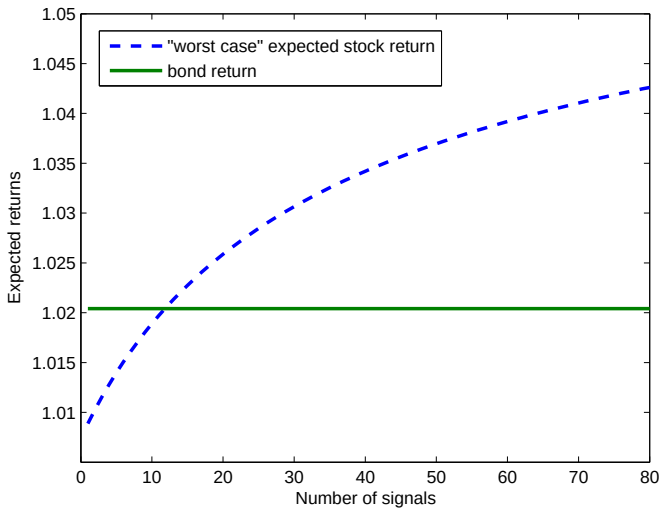

Figure 6: Worst case expected equity premium.

premium moves from -1 percent for an agent with no prior observations of the realized returns to a little more than 2 percent for an agent that received the signal for almost all his life. This follows from the fact that as more and more signals reflecting the underlying process generating the stock returns are received, the set of possible distributions shrinks, hence the "worst case"expected equity premium increases. This is reflected in the decision rules presented in Figure 7. As an example the figure reports decision rules at ages 30, 40, 50, 60 and 70 for an agent that participated in the stock market in every period and observed prior to each of those dates 50 percent of the times a high return realization and 50 percent of the times a low return realization. As one may see those decision rules are higher and higher as the agent gets older, hence had the time to observe more signals. Notice that the work of Cocco, Gomes and Maeanhout (2005) pointed out that in a conventional

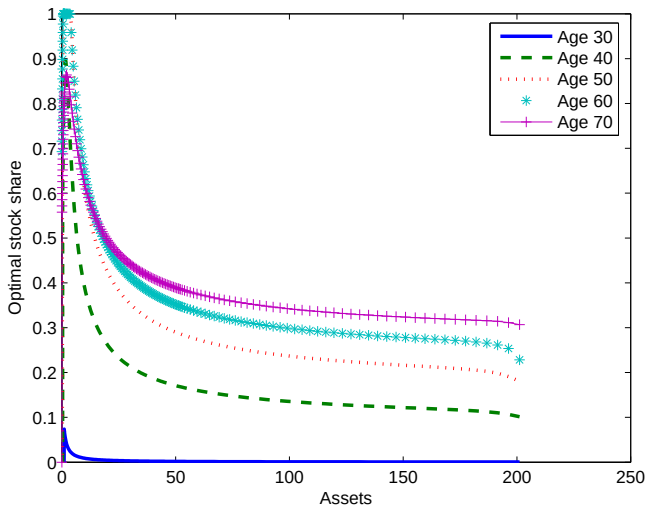

Figure 7: Stock portfolio decision rules.

framework stock decision rules should move inwards as the agent ages to reflect the progressive reduction in relatively safe human wealth. While this phenomenon still exists in the current model, the shrinking of the posterior sets determines a counteracting force that overturns the age-decision rules relationship. Summarizing, along the life-cycle two forces would push the household towards a smaller exposition to stock market risk: one is the above mentioned reduction in the holdings of safe human wealth and the second is the progressive accumulation of wealth to finance retirement consumption — corresponding in the graph to a movement outwards along a given decision rule. On the other hand learning adds a force in the opposite direction given by the reduction in the posterior set and the ensuing improvement in the "worst case"equity premium. Overall as Figure 5 shows this turns the life-cycle profile of conditional stock shares into one that exhibits a mildly increasing pattern.

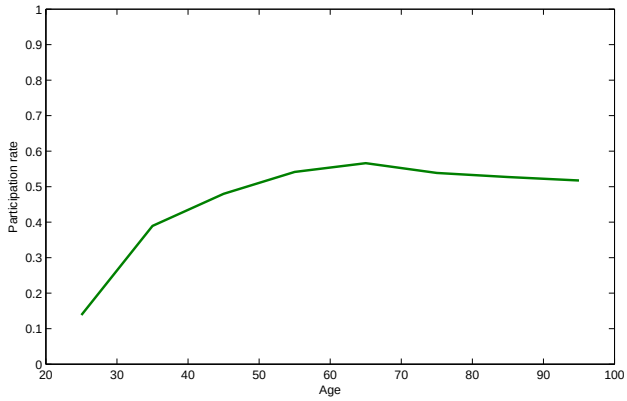

Figure 8: Life-cycle profile of participation rates, baseline case.

The life-cycle profile of participation rates is reported in Figure 8. As it can be seen that profile is increasing in the first part of life and up to retirement age; after that it is modestly declining. The initial steep increase is justified by the need of accumulating wealth to pay the per-period fixed cost and the fact that early in life the set of posteriors for the stock return distribution is still rather large leaving room for a substantial increase of the "worst case"equity premium. During retirement, while wealth is depleted possibly driving households out of the stock market, the set of beliefs tends to further shrink improving the "worst case"equity premium and making households more willing to pay the cost even in the face of smaller sizes of the investment. The mild decrease in participation rates is then the balance of these two opposing forces. As it can be seen in Table 2 the model is qualitatively consistent with the data. Quantitatively the fit is good as far as the increasing part of the profile during working life is concerned. [13] Afterwards though, while participation rates are strongly declining in the

[13] In judging this statement observe that the average participation rate in the model is about 43 percent, that is, an average of participation rates in the last few editions of the SCF, while

Table 2: Participation rates by age

	20-30	30-40	40-50	50-60	60-70	70+
Data	34.3	51.8	58.3	61.4	47.1	32.4
Ambiguity	9.25	38.9	58.3	64.6	58.9	32.5
Learning	13.8	38.9	47.9	54.1	56.6	53.2

data, they are only mildly so in the model. Despite the apparent inconsistency one should be warned that for participation rates there is a debate over whether life-cycle profiles are actually hump-shaped or increasing. For example Ameriks and Zeldes (2004) suggest that older households belong to cohorts that were less likely to participate in the stock market at all ages compared to younger cohorts, so that the decline in participation rates that we observe in cross-sectional data may not indeed reflect a tendency for older households to shift out of the stock market.

Finally we want once more to comment on the extent of ambiguity implied by the parameters chosen in this calibration of the model. We can see from figure 6 that starting from about 20 observations the "worst case"equity premium is about 1 percent and increases to a bit more than 2 percent after 70 observations. In light of the discussion in the previous section these numbers seem reasonable. In particular, they amount to say that after about 20 signals the agent's uncertainty about the equity premium is only 1 percentage point larger than the average for academic professors in finance reported by Welch (2000) and that after a life of observations it becomes indeed smaller than that. As far as agents with fewer than 10 years of observations on the stock return process

in the 1998 issue of the survey it had climbed up to 49 percent.

are concerned the current parametrization entails a "worst case" equity premium that is negative. This can be justified in light of the results of field studies reported in Lusardi and Mitchell (2007). The two authors describe findings of several surveys about financial literacy in the general population. Perhaps surprisingly to the academician, when asked about the historical performance of stocks versus fixed income investments a third or more of the respondents did not know that the former yielded a higher return than the latter. If we interpret this finding in terms of multiple priors this suggests that those households deem possible stock return distributions that entail a negative equity premium. [14] Also as it was pointed out in section 4.1.1 even in the sample of finance professors surveyed in Welch (2000), about ten percent admitted a negative "worst case" equity premium over a thirty-year horizon.

4.1.3 Conditional Shares by Wealth

In this section I examine the results of the model with and without learning along another dimension, that is, the allocation to stocks for participants along the wealth distribution. This is done in Table 3. The table reports in the first row the empirical conditional stock shares by wealth quartiles and the top 5 percent of the distribution. The source for these data is Guiso et al. (2001) based on the 1998 edition of the Survey of Consumer Finances; the definition of stocks in

[14] Strictly speaking the question reported in Lusardi and Mitchell refers to knowledge of the historical record of the equity premium rather than long term forecasts so to make the statement reported in the text one needs to add that expectations are formed based on past observations. This assumption is consistent with the learning model used here. Moreover at the empirical level it is reasonable that the long term forecast of the equity premium reflects the long term historically recorded one. To support this just observe that the central tendency for the 30-year forecast of finance professors reported in Welch (2000) is about 7 percent, very close to historical records.

Table 3: Conditional shares by wealth percentiles

	Quartile I	Quartile II	Quartile III	Quartile IV	Top 5%
Data	40.7	45.0	49.0	60.4	64.0
Ambiguity	99.9	99.7	75.4	41.3	32.7
Learning	57.0	56.7	61.3	60.3	58.5

their data includes both directly and indirectly held equity. In the second and third rows the corresponding figures from the model with ambiguity but not learning and for the baseline model with learning in ambiguous environment are reported. As we can see in the data the share of stocks conditional on participation is increasing over the whole wealth distribution. On the other hand the model without learning generates a very strongly negative relationship between wealth and conditional stock shares. As we can see the share allocated to the risky asset declines from virtually 100 percent to only 32.7 when moving from the bottom quartile to the top 1 percent of the wealth distribution. When learning is added to the model this declining pattern disappears and the share of assets invested in risky equity remains more or less constant over the whole distribution. Although this change is insufficient to match the data still it represents an important step in the right direction.

To understand better where this result comes from in Table 4 we report the patterns of conditional stock shares by wealth levels but conditioning on age. This is done both for the model with ambiguity only and for the model with learning in a multiple prior framework. To keep the table at a manageable size the last two decades of life are not reported but their properties are similar to those of the nearby decades of life. What we can see from the table is that in

Table 4: Conditional shares by wealth percentiles and age

	Quartile I	Quartile II	Quartile III	Quartile IV	Top 5%
Ambiguity					
20-30	n.a	n.a	n.a	98.9	94.9
30-40	n.a	99.9	99.6	81.7	58.0
40-50	99.9	98.1	78.0	52.7	39.6
50-60	98.9	67.9	49.1	37.5	30.7
60-70	71.8	45.5	35.6	30.2	26.3
70-80	n.a	37.1	33.4	29.5	26.7
Learning					
20-30	99.9	98.8	37.2	35.9	39.5
30-40	65.9	51.3	50.5	49.1	47.0
40-50	61.8	57.6	55.2	52.2	47.8
50-60	67.5	63.5	60.2	55.8	50.8
60-70	66.8	62.2	59.0	56.1	51.7
70-80	68.2	65.5	63.6	60.5	57.2

the model with ambiguity only, if we condition on age, the patterns of equity shares by wealth are strongly monotonically decreasing: for example in the 40 to 50 age group the share declines from almost 100 percent to only about 40 percent and in the 60 to 70 age group it declines from about 72 percent to 26 percent.[15]

If we then look at the bottom panel where the portfolio allocation in the model with learning is portrayed we see that while conditional on age the pattern of stock shares by wealth is still declining, it is much less so than in the model with ambiguity only. For example in the 40 to 50 year of age group, when moving

[15]The entries in the table marked with "n.a."correspond to age-wealth cells where the participation rate in the model is 0 so that a conditional equity share cannot be computed.

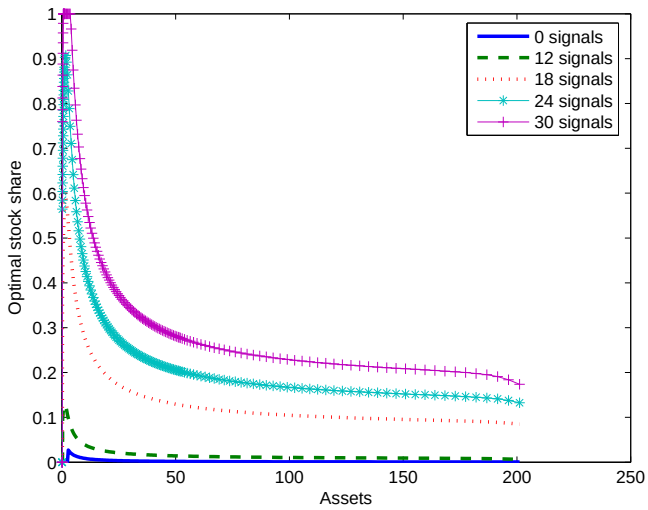

Figure 9: Stock portfolio decision rules.

from the bottom quartile to the top 5 percent of the distribution, the decline in the conditional stock share is only from 61.8 percent to 47.8 percent and in the 60 to 70 year group it is from 66.8 percent to 51.7 percent. The ability of the model with learning to reduce the slope of the conditional share-wealth relationship can be understood by looking at Figure 9. This figure is similar to Figure 7 but it focuses on a household of a given age, more precisely, 50 years old and plots the optimal portfolio allocation as a function of the past number of signals about the stock return process observed prior to that age. As in the previous figure the decision rules reported here are plotted assuming a past fraction of high signals observed equal to 50 percent. As we can see these decision rules are declining in wealth as in more conventional models and decision rules corresponding to more observations are higher. This reflects the reduction in size of the set of posterior distributions that the agent deems possible as more signals that reflect

the underlying data generating process accrue. In a more conventional model or in one with ambiguity only, for a given age the conditional stock share could only decline with the amount of asset holdings, reflecting movements to the right along a given curve. In a model with learning though a wealthier agent can potentially invest a larger fraction of her wealth in stocks than a poorer one. This is because a wealthier agent will have faced a better past history of labor shocks, hence may have started to participate in the stock market earlier and for this reason have observed more signals of the underlying return process. This corresponds to movements towards higher decision rules. Although this mechanism can potentially generate a positive relationship between wealth and stock shares for market participants conditional on age, what we see from Table 4 is that, at least with the set of parameters used here, quantitatively the model can take a step in the right direction by reducing the magnitude of the negative slope of that relationship but not overturn its qualitative pattern. The explanation for this can be found by observing that average wealth in model units is about 9. Given the skewness of the wealth distribution most agents will be close to the vertical axis, where the decision rules are very steeply sloping downward. On the other hand for given wealth we can see in figure 9 that the increase in the stock share with the number of signals is not huge: for example the difference between an agent that has observed 24 signals and an agent that has observed 30 signals is only about 10 percentage points. Clearly for the mechanism proposed here to make the relationship between wealth and stock shares conditional on age positive, large differences in the number of observed signals between wealthy and poor agents are needed, but these are constrained by the need to generate reasonable participation rates which forces the model

to produce most entry in a narrow period of the life-cycle as it can be seen for example from figure 8.

Summarizing, the fact that the model is able to generate a relationship between conditional risky investment and wealth that is slightly increasing is in part the result of a less negatively sloped relationship conditional on age and in part the consequence of the reduction of conditional equity shares of young households, who hold little wealth, compared to older and wealthier ones.

4.2 Sensitivity Analysis

Summarizing the results of the two previous subsections we can say that a reasonable amount of uncertainty about the stock return process and ambiguity aversion enable an otherwise conventional life-cycle portfolio choice model to predict moderate participation rates and moderate conditional stock shares with low risk aversion and reasonably small participation costs. On the other hand they generate conditional stock shares that are strongly decreasing in both age and wealth moving a step in the wrong direction compared to a conventional model. Adding learning in this multiple prior environment, while preserving the ability of the model to generate moderate average participation rates and conditional shares, also enables it to produce a life-cycle profile of stock allocation for participants that is slightly increasing but with little variation as in the data. It also makes conditional stock investment roughly constant in wealth thus taking a step in the right direction towards matching the empirical evidence.

In choosing the parameters to calibrate the learning process I insured that the resulting amount of ambiguity perceived by agents over their life-cycle was reasonable. However in a dynamic context of learning it is not only the average

extent of ambiguity that matters but also the way it unfolds over time. Since there is no hard evidence based on which one can fix the parameters that determine this property the only thing that can be done is to perform a sensitivity analysis. The present section is devoted to this task. As a side output this section will also help further understanding the mechanisms at work in the model. Beside the parameters related to learning the current section will perform a sensitivity analysis also on the coefficient of relative risk aversion.

Initial Ambiguity. In this first sensitivity analysis I change the size of the set of distributions of the stock return that the household includes in its beliefs at the beginning of life. I consider two cases: one with more ambiguity where the initial set of priors for the probability of a high stock return is the interval $[0.276, 0.724]$ and one with less initial ambiguity where that set is the interval $[0.323, 0.677]$. These imply an initial "worst case"equity premium of about -2 percent and -0.5 percent respectively. The remaining parameters are left unchanged from the baseline case. The model with less initial ambiguity generates an average participation rate of 50.5 percent and an average share for participants of 64.6 percent. The model with higher initial ambiguity produces an average participation rate of 36.7 percent and a conditional allocation to stocks of 56.1 percent. As we can see the numbers are close to the data although the participation rate in the case of more initial ambiguity is somewhat below. The life-cycle profiles of conditional stock shares are reported in Figure 10. The figure also reports the same profile for the baseline case. As we can see the profile in the case of less initial ambiguity lies above and is flatter than the baseline case, while the life-cycle profile in the case of higher initial ambiguity lies below

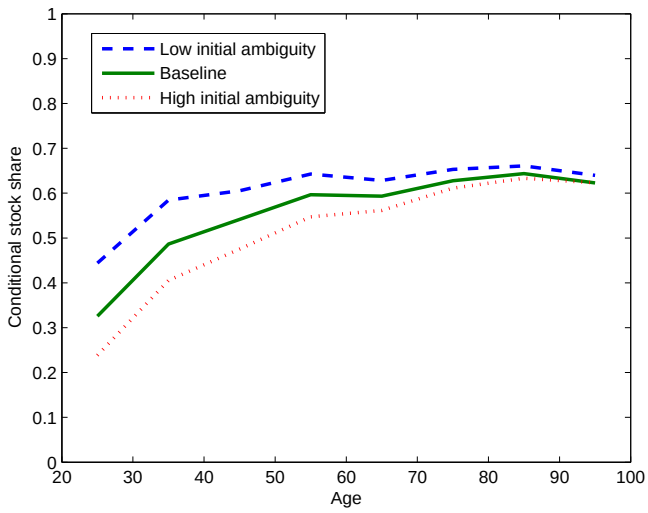

Figure 10: Life-cycle profile of conditional stock shares; changing initial ambiguity.

and is steeper. As a result they converge later in life.

To understand this result observe that less initial ambiguity means a smaller set of beliefs at the beginning of life, hence the "worst case"expected equity premium starts out larger. This means that the benefits of stock holding are perceived to be larger. As a result more households will participate from the very beginning and they will invest larger shares of their wealth in the stock market. Also, since the evolution of the "worst case"equity premium tends to follow an asymptotic path, the initial differences in participation rates and conditional stock shares tend to vanish over the life-cycle. The life-cycle profiles of participation rates are reported in figure 11. Qualitatively they are similar to the ones in the baseline case, that is, they are increasing up to the 60 to 70 year old group and then flatten out or decrease mildly. Consistent with

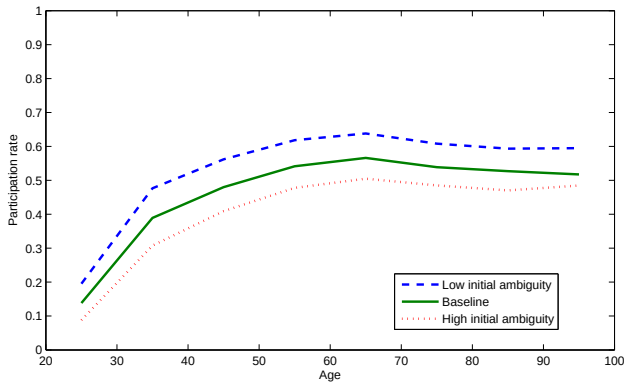

Figure 11: Life-cycle profile of stock market participation rates; changing initial ambiguity.

the overall average participation rate, the profile is higher for the case when initial ambiguity is smaller represented by the dashed line in the figure and lower for the case when initial ambiguity is larger, represented by the dotted line. The intuition is straightforward: higher initial ambiguity implies a smaller "worst case" equity premium, hence smaller benefits from participation so that households will start to participate later in life and do that to a lesser degree all over the life-cycle.

"Speed" of Learning. In this paragraph the sensitivity analysis on the parameter α is performed. Two values are taken, that is, $\alpha = 0.1051$ and $\alpha = 0.2991$ and the remaining parameter that controls the properties of the evolution of the posterior set is changed so that the set itself at the beginning of life is unchanged. The wording "speed" of learning refers to the fact that α determines how strict is the statistical test based on which stock return distributions are discarded or kept in the set of beliefs in the face of new signals:

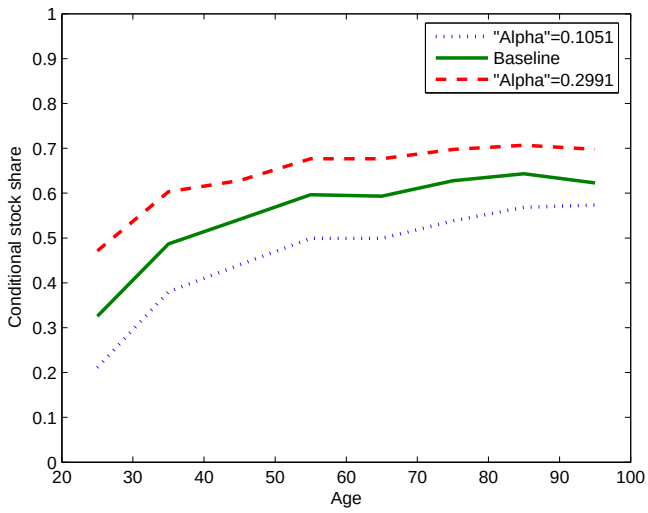

Figure 12: Life-cycle profile of conditional stock shares; changing the "speed" of learning.

a higher α implies a stricter test hence a quicker adaptation of beliefs to new signals. In the case of $\alpha = 0.2991$ the average participation rate is 46.9 percent and the average equity share for participants is 68.6 percent, while in the case of $\alpha = 0.1051$ the average participation rate is 30.8 percent and stock holding households invest on average 51.1 percent of their wealth in the risky asset. The life-cycle profiles of conditional stock shares are reported in Figure 12, where as usual, the continuous line reports the baseline case for comparison. Consistently with the overall average, the line representing the conditional share for the higher value of α lies above and the one for the lower value of α lies below the one of the baseline case. Also the profile for $\alpha = 0.2991$ shows an increase from 50 percent to a little bit more than 60 percent between the first and the second decade of life and then little variation thereafter, while the profile for

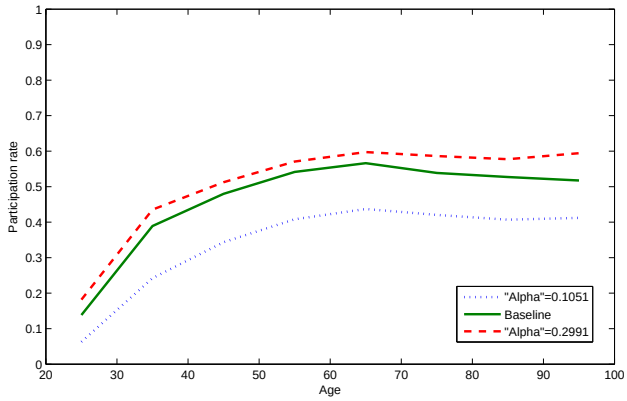

Figure 13: Life-cycle profile of stock market participation rates; changing the "speed" of learning.

$\alpha = 0.1051$ follows a steeper pattern especially in the first four decades of life. To understand this result observe that upon entry in the workforce in both cases agents face the same amount of ambiguity. After that in the model with larger α they update their beliefs more quickly which, on average, translates into better "worst case" expected equity premia. This is reflected both in the higher average participation rate and higher conditional stock share. Also since with larger α the household gets rid of ambiguity more quickly, apart from the first decade of life the life cycle profile of conditional stock shares shows a much weaker slope. The life-cycle profile of participation rates is reported in figure 13. As it can be seen a higher "speed" of learning leads to a higher profile of participation rates, given that the "worst case" equity premium will be higher. As far as the shape of the profiles is concerned they are qualitatively similar to the ones of the baseline case.

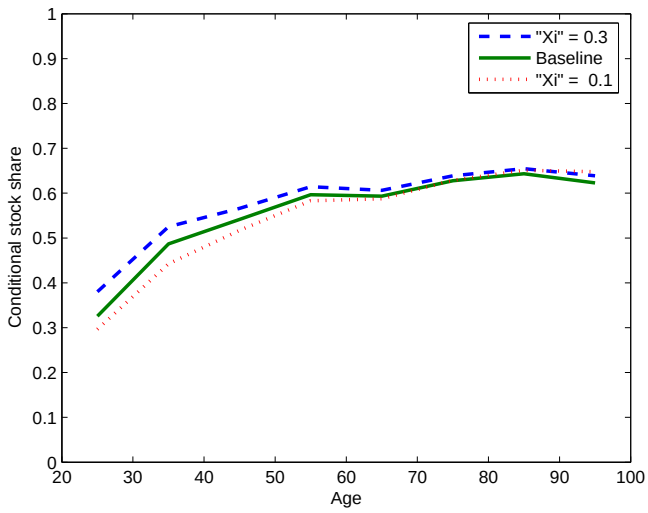

Figure 14: Life-cycle profile of conditional stock shares, changing the informational advantage of participants.

Informational Advantage of Participants. In this section I consider a further sensitivity analysis in which I change the probability that an agent that does not participate in the stock market receives the signal about the stock return process. In general, for a given "worst case"expected equity premium, the higher that probability the smaller is the relative benefit of participation compared to non participation. I consider two cases that move in opposite directions starting from the baseline case. When the probability ξ that a non-stockholding household receives a signal is set at the higher value of 0.3 the average participation rate is 46 percent and the average conditional share is 62.6 percent, while when ξ is equal to 0.1 the average participation rate is 39.8 percent and the average conditional share is 56.8 percent. The life-cycle profiles of conditional stock shares are reported in figure 14. As we can see the

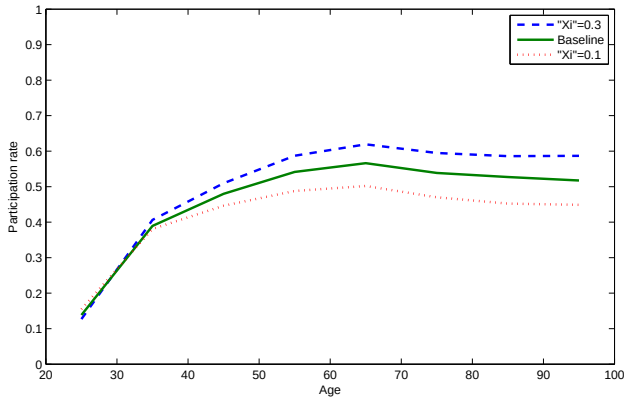

Figure 15: Life-cycle profile of stock market participation rates, changing the informational advantage of participants.

two profiles are not very different from each other and from the baseline case. The life-cycle profile for $\xi = 0.1$, that is, a lower probability of receiving a signal is only a little steeper since it starts about 7 percentage points below the one for $\xi = 0.3$ and then converges to it at the end of life. The intuition for these results is the following: an increase in ξ implies that a larger proportion of agents that do not find it optimal to participate in the stock market will receive the signal about the return process anyway. On average this will imply that over time their set of beliefs will shrink, hence the "worst case"expected equity premium improves. This may induce them to participate in the market earlier than they would have otherwise done. As a result the average percentage of participants in the population increases. At the same time, as agents start participating earlier, hence receive signals about the return process systematically from a younger age, this will lead to earlier reduction of ambiguity and the observed increase in the conditional share as well. The life-cycle profiles of participation rates are

reported in figure 15. The upper dashed line represents the profile for the case where non participants have a higher probability of receiving the signal while the lower dotted line represents the case when they have a lower probability. The profiles are increasing up to the age group 60 to 70 and then they flatten out as in the baseline case. A higher probability of receiving the signal will tend to improve the "worst case"equity premium for non participants which explains why more agents will tend in the end to hold stocks, hence the higher life-cycle profile. At the beginning of life the two profiles overlap and then diverge as agents get older. This reflects the cumulative effect of a permanently higher probability of receiving the signal for non stockholders in one case compared to the other.

Risk Aversion. In this paragraph I explore the implications of changing risk aversion in the current model and consider two alternative values, that is, 1.5 and 3.5. In the low risk aversion case the average participation rate is 40.9 percent and the average conditional share is 74.3 percent; in the high risk aversion case the average participation rate is 41.9 percent and the average conditional share is 45.3 percent. The results concerning the conditional share are quite standard since it is well known that an increase in the risk aversion coefficient will cause the agent to reduce its exposition to the risky asset. As far as participation rates are concerned, earlier work like Gomes and Michaelides (2005) pointed out that an increase in risk aversion has two effects working in opposite directions. On the one hand the reduction in the conditional share would make agents less willing to pay the fixed participation cost thus reducing participation. On the other hand though, given the functional form chosen for the utility

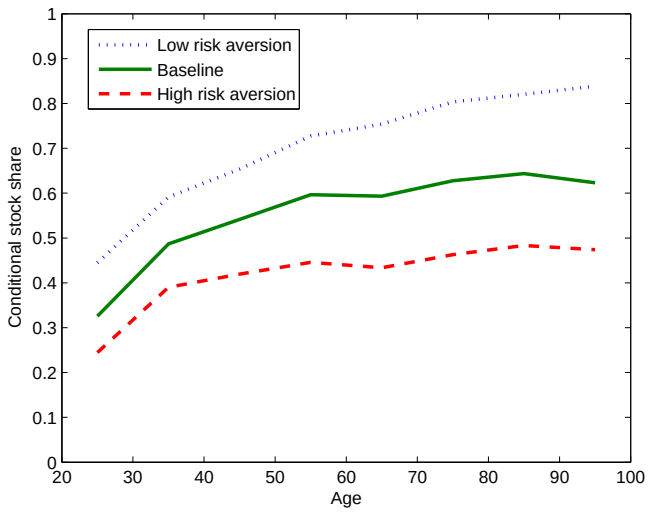

Figure 16: Life-Cycle profile of conditional stock shares; changing risk aversion.

function a higher risk aversion is associated with more prudence inducing higher precautionary savings. This makes the agent more willing to pay the fixed cost, hence increase participation. In their model the overall effect is to increase the average participation rate while in

the current model with ambiguity and learning the two forces approximately balance out leaving the average participation rate almost unaffected by the coefficient of risk aversion, at least for the range of values considered. This reflects the fact that the unitary benefit of investing in equity, which is driven by the "worst case" equity premium is smaller hence the change of the overall benefit is dampened. Results concerning the life-cycle profiles of conditional shares are reported in Figure 16 where the middle continuous line again represents the baseline case and is reported for comparison. Qualitatively the patterns do not change with risk aversion: conditional shares are increasing and more than in

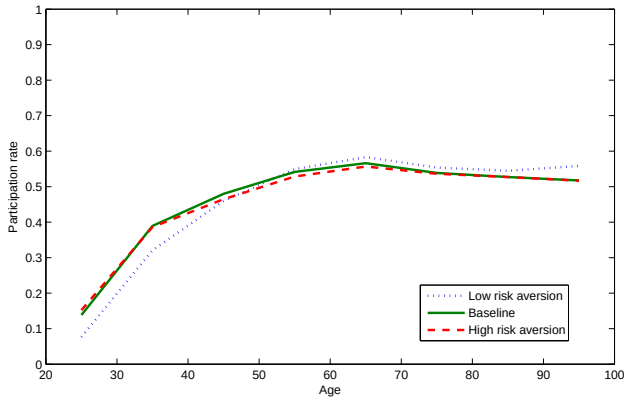

Figure 17: Life-Cycle profile of stock market participation rates; changing risk aversion.

the baseline case when the coefficient of relative risk aversion is 1.5 while they are virtually constant, except for the first decade of life when risk aversion is increased to 3.5. The life-cycle profiles of stock market participation rates for the three different values of the coefficient of relative risk aversion are reported in figure 17. Consistent with the fact that the average participation rate in the overall population is very similar in the three cases the three lines in the figure are very close to each other and actually intersect. All the profiles are increasing up to the 60 to 70 year of age group and then flatten out or decrease mildly.

Sensitivity Analysis of Conditional Shares by Wealth. In this paragraph I summarize the results of the sensitivity analysis performed above but with respect to the pattern of conditional stock holdings as a function of wealth. In order to economize on space the tables with the conditional allocations to stock by wealth and age groups are not reported. [16] The results of this anal-

[16] These tables are available from the author upon request.

ysis are reported in Table 5. A look at the table confirms the result of the baseline case that the patterns of conditional portfolio allocations by wealth do not exhibit the strongly declining shape of the model with ambiguity only or of a conventional model with sufficiently large risk aversion to generate reasonable average shares. Within this general observation there are some differences among the cases considered. For example some of the patterns are a little more increasing although still less than in the data. One is the case of $\alpha = 0.2991$ where the conditional allocation to stocks moves from 59.6 percent in the bottom quartile of the wealth distribution to 71.3 percent in the third quartile to stay roughly constant thereafter. If we go back to Figure 12 we see that this case is also one where the life-cycle profile is flatter so that what we observe across wealth levels is least driven by young poor agents holding lower conditional shares. The main reason then is that conditional on age the relationship between wealth and portfolio share of stocks is less declining than in the baseline case. This can be explained by the fact that when α is higher learning occurs faster, hence the benefits in term of a smaller sets of posteriors for those who get the signals are larger. Wealthier agents in general start to participate earlier, get more signals hence benefit more from the faster reduction in ambiguity. A similar argument in the opposite direction explains why the case $\alpha = 0.1051$ instead exhibits an overall declining pattern of conditional stock shares over the wealth distribution. The other case that shows a more pronounced increase in conditional stock shares over the wealth distribution is the one with $\xi = 0.1$ where the conditional share moves from 39.2 percent in the bottom quartile to 61.4 percent in the top one and then only modestly declines to 59.7 percent in the top 5 percentiles of the distribution. The basic principle to interpret

Table 5: Conditional shares by wealth percentiles

	Quartile I	Quartile II	Quartile III	Quartile IV	Top 5%
Baseline	57.0	56.7	61.3	60.4	58.5
HIA	58.3	52.4	56.6	57.3	56.6
LIA	61.1	62.9	67.3	64.1	60.5
$\alpha = 0.1051$	61.1	50.8	51.9	50.2	48.1
$\alpha = 0.2991$	59.6	66.7	71.3	69.7	68.0
$\xi = 0.1$	39.2	49.0	59.6	61.4	59.7
$\xi = 0.3$	70.3	62.8	63.3	60.9	58.6
$\sigma = 1.5$	91.0	67.6	73.5	77.6	78.8
$\sigma = 3.5$	42.9	47.1	47.6	42.6	40.2

this result is the same as the one mentioned in the previous case: wealthier agents through the observation of signals can reduce the size of the belief set, hence increase the "worst case"equity premium better than poorer agents. Here the reason is that given the lower probability of observing such signals for non participants the difference between wealthier agent who participate more often and poorer agents who participate less frequently becomes larger. Once again moving the parameter in the other direction — $\xi = 0.3$ — leads to one of the most declining patterns.

4.3 Other Issues

In this section three more sets of results will be presented that further help understanding the model and its potential in explaining households' financial behavior. One concerns the effects that alternative histories of stock return realizations have on "worst case"equity premia and households financial choices,

the second the impact of different assumptions about ambiguity and learning on the wealth distribution in the model. Finally the third is a brief analysis of an alternative assumption about how signals are received and incorporated in the belief set by participants and non participants.

4.3.1 Results of Sample Simulations

In order to study the main features of the model, in the previous sections this was simulated for many different draws of the sequence of return shocks, and results were obtained by averaging out the different simulations. In this section we take two of those simulations and show the potential of the current setup in explaining other features of households' financial choices. For the purpose of illustration two draws that display somewhat "extreme"patterns of stock return realizations are chosen. The treatment given here is short, leaving for future research the exploration of the full potential of the model along the directions outlined here. Figures 18 and 20 report three variables: the bottom dashed line represents the fraction of past stock returns that were high for the particular cohort in the simulation considered. The top dotted line represents the "worst case"expected equity return that corresponds to that particular history of shocks, while the continuous line represents the constant bond return and is reported for comparison. The lines represent gross returns so that, for example, the number 1.1 on the vertical axis when referring to returns means a 10 percent net return. Figures 19 and 21 report the corresponding portfolio allocation decisions, that is, the life-cycle participation rates and conditional stock shares for that cohort. The simulations are performed using the parameters of the baseline model. As it can be seen in Figure 18 that cohort experiences first an

initial period of low returns, then from age 25 to 45 a long period of high stock returns, followed by about 10 years of low returns. Finally these are followed by high stock returns until death. These patterns in the realized sequence of stock returns are mirrored in the sequence of expected "worst case"equity returns: after starting out below the bond return this variable monotonically increases up until age 45, then in response to the decade of low stock returns it declines somewhat and finally it increases again starting from age 55. Figure 19 shows that conditional stock shares, represented by the dashed line increase from 50 percent in the youngest age group to about 80 percent in the 40 to 50 year of age group, then, in response to the decade of bad stock returns this share declines to 70 percent in the 50 to 60 year of age group and finally increases over the rest of the life-cycle. Participation rates increase monotonically over the whole life-cycle due to the important effect of wealth accumulation during working life, however, around age 50 the period of low stock returns flattens the profile.[17]

A similar reasoning can be applied to figures 20 and 21. Looking at figure 20 we see that in this simulation an initial period of 5 years of high returns is followed by a period of low returns up to age 30 or slightly more and a long period of slightly less than average returns up to age 55. Later a period of moderately high returns follows until death. In response to this pattern the "worst case"expected equity return first climbs up but then starting at age 30

[17]Both life-cycle profiles are substantially higher than what is observed in the data. Looking at figure 18 it is clear that this depends on the fact that the sequence of shocks represented in this graph implies a substantially higher average stock return than the historical experience, reaching about 10 percent in the last 40 years of life. This in turn has a strong effect in generating very high "worst case"equity premia.

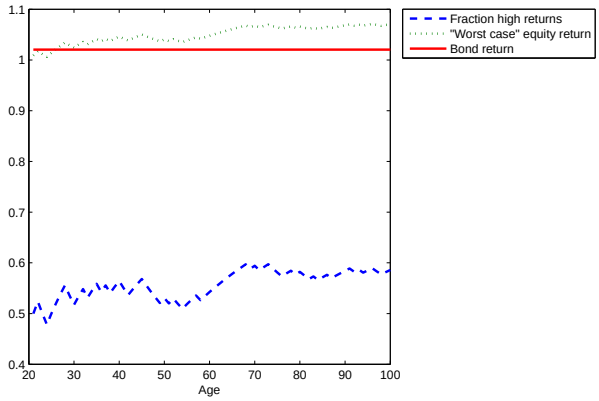

Figure 18: Fraction of past high returns and "worst case"equity premium.

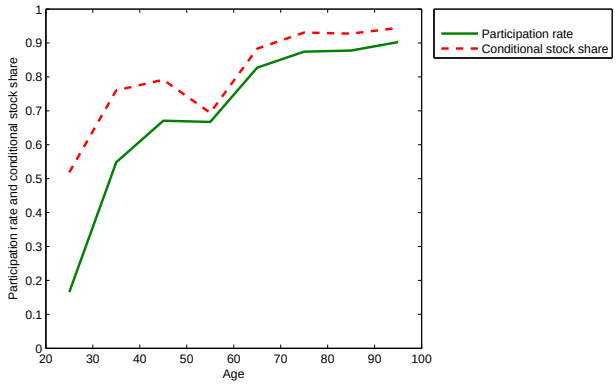

Figure 19: Participation rate and conditional stock share.

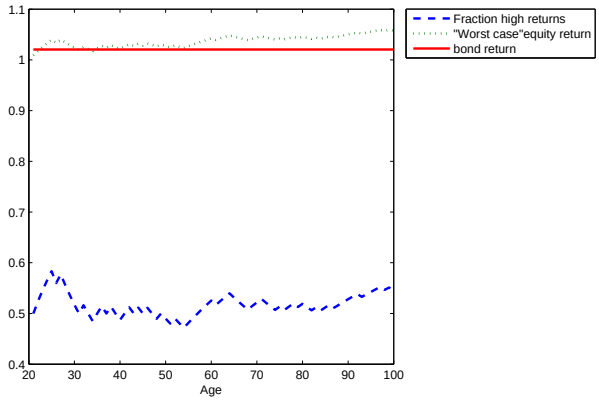

Figure 20: Fraction of past high returns and "worst case"equity premium.

it falls to values only slightly above the bond return; it is only after age 55 that it systematically rises to generate a sizeable premium over the bond. Mirroring these patterns, figure 21 shows that in the first decade of life, the average conditional share in stocks is about 45 percent, after ten years of low stock returns it goes down to 25 percent and then when the periods of"average"or above "average"returns sets in and the "worst case"equity premium starts increasing the profile itself becomes strongly monotonically increasing. As for the participation profile it is monotonically increasing except between the third and fourth decade of working life. Comparing figure 19 and 21 and interpreting them as the simulation for two different cohorts facing a different history of stock returns, we see how their choices may be quite different. For example the 30 to 40 year of age group holds 25 percent of its portfolio in stocks in the second simulation but 75 percent in the first one, reflecting the fact that in the latter there is an initial string of high stock returns while the reverse occurs in the former.

These two sample simulations highlight two interesting points about the

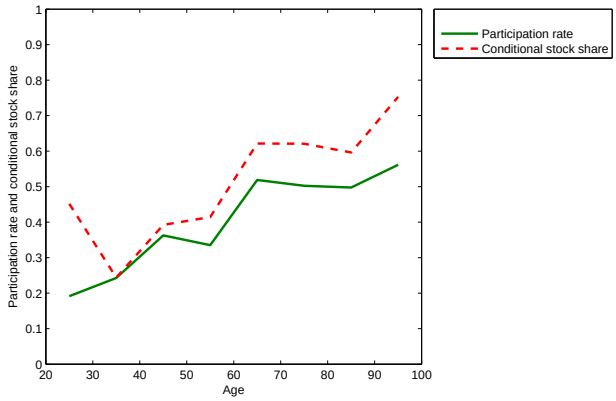

Figure 21: Participation rate and conditional stock share.

learning model. First even with an i.i.d. stock return process the model can generate time-varying return expectations that depend on the experience of a given cohort. Second, as a consequence, it introduces cohort effects in household financial decisions in the sense that choices can be different for the same age group belonging to cohorts with different stock market experiences. While it is not the purpose of this paper to further pursue this line of research it is worth noticing that recent empirical work has confirmed that these effects can be found in the data. In particular Malmendier and Nagel (2009), using the Reuters/Michigan Survey of Consumers found that individuals of different ages differ substantially in their inflation expectations with those differences being related to their experienced inflation histories. These differences can reach several percentage points in times with recent inflation experiences that are far from the historical average. The fact that investors' expectations about stock market returns are higher after a period of high stock returns, consistent with a story of learning from own experience, and that this is stronger for investors

with shorter investment periods is documented in Vissing-Jørgensen (2003) who also documents that actual stockholding choices do depend on these beliefs.

4.3.2 Wealth Distribution

In this subsection the implications for the wealth distribution of the models presented in the paper are analyzed. This is done with the help of table 6 that reports the share of wealth of the bottom 40 percent and the top 20, 10 and 5 percent households in the data, in the no learning models, both with and without ambiguity aversion and in the learning model with the baseline choice of parameters. [18] The first thing that can be noticed is that the models generate a wealth concentration that is largely inferior to the one in the data: they overestimate the share of total wealth held by the bottom 40 percent of the population and underestimate the one of the top 20, 10 and 5 percent. This is no surprise since the current model features a simple earnings process that itself generates little earnings inequality and it does not possess those ingredients like intergenerational transmission of wealth and entrepreneurship that have been recognized as important in order to generate sizeable estates at the top of the distribution. When it comes to the comparison among models we see that they generate roughly the same amount of wealth concentration. For example, comparing the last two rows of the table we see that the model with ambiguity but no learning generates a smaller share of wealth for the bottom 40 percentiles of the distribution but at the same time also a smaller share of wealth for the top 10 and top 5 percentiles when compared with the baseline model with learning. The differences though are in the order of one percentage

[18] The data are author's calculations based on the 1998 issue of the Survey of Consumer Finances.

Table 6: The distribution of wealth

	Wealth shares			
Percentiles	0-40	80-100	90-100	95-100
Data	1.35	79.5	66.1	53.5
No ambiguity	4.2	64.5	42.7	26.7
Ambiguity	4.7	63.7	41.8	26.2
Learning	5.7	62.2	42.2	26.7

point or less. In principle the model with learning generates a more positively sloped relation between wealth and the share of stocks suggesting that the higher average return faced by richer agents should lead to higher wealth concentration, however, as we saw in section 4.1.3 this is to a large extent the result of lower shares of stocks in the portfolio of young households who are on average poorer while conditional on age it is still true the stock shares are declining in wealth. This is especially important for the issue of wealth concentration because the top percentiles of the wealth distribution are made by agents around retirement for whom this age-conditional negative relationship holds.

4.3.3 Alternative Signal Structure

In this subsection I describe the different results that the model produces under an alternative signal structure. I will consider a case when agents learn in every period but stock market participants learn faster in the sense that they have a higher value of α than non participants. A full analysis of the model under this assumption would require a lengthy discussion so what is done here is to present just one case where parameters are chosen so that average participation rates and conditional shares of stocks are in the ballpark of the empirical evidence

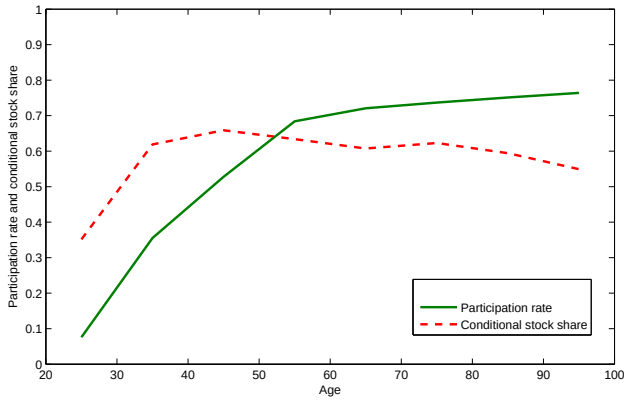

Figure 22: Participation rate and conditional stock share.

and of the models considered in the previous sections. Results for the life-cycle profiles are reported in figure 22 and those for the wealth-conditional share relationship are reported in table 7. The model parameters are obtained by taking the baseline calibration of section 4.1.2 and changing the values of α to 0.0552 for investors who decide to participate in the stock market and 0.02 for those who decide not to do so. Notice that the value of α for participants is below the one of the baseline calibration. This is needed because now agents are learning in every period so that keeping the baseline value of α and assigning a lower one to non participants would make the model over-predict both the participation rate and the conditional stock share. Under the parameters chosen here the average participation rate is 49.7 percent and the average share invested in stocks by participants is 61.8 percent. Looking at figure 22 we see that the life-cycle profile of participation rates is strongly increasing up to the 50 to 60 year old group and then is moderately so thereafter. The life-cycle profile of conditional shares is increasing between the first two decades of life and then it is

Table 7: Conditional shares by wealth percentiles

	Quartile I	Quartile II	Quartile III	Quartile IV	Top 5%
Shares	81.5	70.1	64.9	62.7	63.5

virtually constant thereafter. These profiles are qualitatively very similar to the ones of the previous sections where the informational advantage of participants came through a higher probability of incorporating the stock return realization into the belief set. As far as conditional shares of stocks by wealth levels are concerned we can see from table 7 that they decline from 81.5 percent in the bottom quartile of the wealth distribution to 64.9 percent in the third one and then they remain constant for higher percentiles. Although slightly declining this profile still represents an important improvement compared to the model with ambiguity but not learning confirming the main qualitative results of the model with learning.

Further signal structures can be imagined, for example one where agents incorporate the information about the stock return in their beliefs in every period and in the same way whether they participate in the stock market or not. Unreported simulations though show that this assumption leads to counterfactual results. To understand why observe that in the other cases considered so far, stock market participation gives two expected benefits, one coming from the "worst case"equity premium, the other from faster learning. If learning occurs equally, independently of participation, this latter benefit vanishes and participation will occur only if the perceived benefit from the equity premium is larger than the per-period participation cost. This benefit can be roughly thought of as the product of the "worst case"equity premium times the amount invested

in stock. Young agents hold very little wealth which limits the amount they can invest in stock, so a large "worst case"equity premium would be needed to induce them to pay the fixed cost and participate. This leads the model to counterfactual implications. First with a relatively high "worst case"equity premium and a very high human-to-financial wealth ratio conditional shares for the youngest age group will tend to be high creating an initial downward stretch in the life-cycle profile. Second with an initially high "worst case"equity premium and further learning the overall average of the portfolio stock share for participants will be well above the empirical evidence. Adding ex-ante heterogeneity in the αs does not seem to help resolving the tension. This leads to the conclusion that some advantage in the learning process for participants is an important feature in explaining the empirical evidence. Once this advantage is allowed the model is robust to the particular form chosen, as this subsection shows.

5 Summary and Conclusions

In the present paper I have introduced ambiguity and learning in an ambiguous environment in an otherwise basic model of life-cycle portfolio allocation. It was shown that a plausible amount of ambiguity can rationalize moderate conditional stock shares and moderate participation rates, as observed in the data, by resorting to relatively small fixed costs of participating in the stock market. Ambiguity alone though did not prove adequate to represent household behavior since it generated patterns of conditional stock shares that were counter-factually declining in both age and wealth. When learning is introduced the model, while still delivering moderate average participation rates and con-

ditional stock shares for a wide range of parameters it also moves towards a better fit with the data along two dimensions. First it generates life-cycle profiles of conditional stock shares that are mildly increasing in age. Second, for some set of parameters it is also able to generate patterns of conditional stock holdings over the wealth distribution that are mildly increasing. This second result though is less robust to changes in key parameters. Also it is obtained in part by the reduction in conditional shares for young agents that are on average poorer, thus leaving the wealth-share relationship conditional on age still declining, albeit to a lesser degree than more conventional models or the model with ambiguity but without learning. As such in this dimension the current theory could be complemented by alternative models, like the non homothetic utility model presented in Wachter and Yogo (2008) to match the empirical evidence more closely.

Appendix

A The Learning Model:A Formal Description

This appendix gives a minimal mathematical description of the learning model informally outlined in section 2.4. The investor in the model perceives stock returns as ambiguous. In particular he thinks that they are generated by the same memoryless mechanism in each period and that even if there are features of the data generating process that can be learned others are not.[19] Mathematically learning is represented by a tuple $(\Theta, \mathcal{M}_0, \mathcal{L}, \alpha)$ where Θ is a parameter space whose elements θ represent features of the data generating process that the agents think are learnable. The set $\mathcal{M}_0$ is the set of priors on Θ and $\mathcal{L}$ is a set of likelihood functions whose multiplicity reflects the existence of poorly understood factors driving the returns. Finally α is a parameter that governs the reevaluation process through which posteriors are constructed based on the past observed returns. The set of posteriors is constructed based on a likelihood ratio test and will be defined as:

$$\mathcal{M}_t^{\alpha}(w^t) = \{\mu_t(w^t; \mu_0 \in \mathcal{M}_0, \ell^t \in \mathcal{L}^t)| \\ \int \prod_{j=1}^{t} \ell_j(w_j|\theta) d\mu_0(\theta) \geq \alpha \max_{\tilde{\mu}_0 \in \mathcal{M}_0, \tilde{\ell}^t \in \mathcal{L}^t} \int \prod_{j=1}^{t} \tilde{\ell}_j(w_j|\theta) d\tilde{\mu}_0(\theta)\}. \quad (17)$$

In this specific context Θ is assumed to be a one-dimensional set with elements $\theta \in [\bar{\lambda}, 1-\bar{\lambda}]$ where $\bar{\lambda} < \frac{1}{2}$. The set of likelihoods is defined by $\ell(1|\theta) = \theta + \lambda$ for some $\lambda \in [-\bar{\lambda}, \bar{\lambda}]$ and $\ell(1|\theta)$ is the probability of observing a high stock return given the value of θ. The set of priors $\mathcal{M}_0$ is given by all the Dirac measures on

[19] For a complete treatment of the subject the reader is referred to the original paper by Epstein and Schneider (2005).

Θ. Finally α is a constant that determines how the set of posteriors responds to new information: were it equal to zero the set of posteriors $\mathcal{M}_t$ would be equal to $\mathcal{M}_0$ for all t and no updating would occur. A higher value of α implies a more stringent test so that a wider set of distributions is discarded from the set of possible posteriors which then changes more quickly in response to new observations. A value of $\bar{\lambda} > 0$ is needed for returns to be ambiguous signals.

It can be proved that under the simple specification used here the set of posteriors depends on the sample only through the fraction of high stock returns ϕ_t observed before t. More specifically it will obey the following law:

$$\mathcal{M}_t^{\alpha}(w^t) = \left\{\theta \in \Theta : g(\theta; \phi_t) \geq \max_{\tilde{\theta} \in \Theta} g(\tilde{\theta}; \phi_t) + \frac{log(\alpha)}{t}\right\} \tag{18}$$

where $g(\theta; \phi_t) = \phi_t \log(\theta + \bar{\lambda}) + (1 - \phi_t) \log(1 - \theta + \bar{\lambda})$. This specification is very convenient for the current problem since it only adds two state variables to the agent's optimal dynamic program — the number of signals observed and the fraction of those that were high stock returns — allowing it to retain numerical tractability.

B Numerical Solution Method

In this section I will briefly describe the numerical methods used to solve the model presented in the paper. The procedure requires two steps, that is, first solving the agent's dynamic programming problem and second simulating the model by using the decision rules obtained in the first step and draws of the realizations of the stock return and the individual histories of earnings and mortality shocks. Because of the minimization with respect to the set of beliefs the dynamic programming problem turns out to be more demanding than in

a standard problem. In practice one has to compute the set of posteriors $\mathcal{M}_t$ and then choose the pairs $\theta \in \mathcal{M}_t$ and $\lambda \in [-\bar{\lambda}, \bar{\lambda}]$ that minimize maxima with respect to the distributions in the set of admissible beliefs. The computation of such set then requires the addition of two state variables that were labeled ϕ_t and n_t in the text and that represent the total number of stock return realizations observed as signals of the underlying process and the fraction of those that were high. The assumption that short selling the stock was exogenously ruled out though, allows the model to retain tractability since under that assumption the minimizing distribution will be the one that minimizes the probability of a high stock return. The state space is discretized along the asset dimension using a grid of 165 points that is finer close to the origin and coarser away from it. The process for the labor earnings shock is also discretized by using 7 points and approximated with the method in Tauchen (1986). As far as ϕ_t and n_t are concerned observe that in principle those are discrete variables. However the number of values they can take over a 80 year period — one like the lifespan — is very high. For this reason the value function was computed only on a subset of 11 points in each of those dimensions and interpolation was used elsewhere. Functions were approximated via cubic splines along all the dimensions that required interpolation. The maximization with respect to stocks and bonds was performed using Brent's method: the method consists of bracketing the maximum with a triple, fitting a parabola through it and use it to eliminate the lowest point in the initial triple, then iterating until no further increase in the value of the function to optimize can be obtained. [20] The method is repeated

[20] See Brent (1973) for the theory and description of the method and Press et al. (1992) for the actual algorithm.

along each dimension — that is, each of the two assets — in turn by exploiting the relation $\max_{x,y} f(x,y) = max_x\{max_y f(x,y)\}$.

Once the decision rules are obtained the model is simulated. Multi-linear interpolation of the decision rules is used whenever the participation decision is the same at the 8 grid points defining a cube in the state space. Otherwise the optimal decision is recalculated using the value functions obtained in the previous step. This procedure is more accurate than simply applying multi-linear interpolation in all cases, however it is also more time consuming. Also the computation of statistics by wealth levels is also somewhat time consuming. For this reason the simulation is done by considering a cohort of 1000 agents and repeating it 30 times. Results are obtained by averaging over the 30 repetitions. To insure robustness for some set of parameters the simulations were also repeated by doubling both the number of households and the number of repetitions. Results were always very similar to the ones obtained with fewer agents.

References

[1] Ameriks, John and Zeldes, Stephen. 2004. "How Do Household Portfolio Shares Vary with Age." Working Paper.

[2] Benzoni, Luca, Pierre Collin-Dufresne and Robert S. Goldstein Robert. 2007. "Portfolio Choice over the Life-Cycle in the Presence of 'Trickle-down' Labor Income ". Working paper.

[3] Bertaut Carol C. and Starr-McCluer Martha. 2000. "Household Portfolios in the United States". *Federal Reserve Board of Governors*, *Finance and Economics Discussion Paper*: 26 (April).

[4] Brent R. 1973. "Algorithms for Minimization without Derivatives". Englewood Cliffs, NJ: Prentice-Hall.

[5] Cagetti Marco, Hansen Lars P., Sargent Thomas, Williams Noah. 2002. "Robustness and Pricing with Uncertain Growth". *Review of Financial Studies*. 15 (2). pp. 363-404.

[6] Calvet Laurent E., Campbell John Y. and Sodini Paolo. 2007. "Down or Out: Assessing the Welfare Costs of Household Investment Mistakes". *Journal of Political Economy*. 115 (5). pp. 707-747.

[7] Campbell John Y., Cocco Joao, Gomes Francisco and Maenhout Pascal. 1999. "Investing Retirement Wealth ? A Life-cycle Model". Working paper, Harvard University.

[8] Cao Henry H., Tang Tan and Zhang Harold H. 2005. "Model Uncertainty, Limited Asset Market Participation, and Asset Prices". *Review of Financial Studies*. 18 (4). pp. 1219-1251.

[9] Cecchetti Stephen G., Pok-Sang Lam and Nelson C. Mark. 2000. "Asset Pricing with Distorted Beliefs: Are Equity Returns too Good to be True ?". *American Economic Review*. 90. pp. 787-805.

[10] Cocco Joao; Gomes Francisco and Maenhout Pascal. 2005. "Portfolio Choice over the Life-Cycle". *Review of Financial Studies*, 18(2): 491-533.

[11] Cochrane John H. 1997. "Where Is the Market Going: Uncertain Facts and Novel Theories". Working paper 6207, NBER, (Cambridge, MA).

[12] Curcuru Stephanie, Heaton John, Lucas Deborah and Moore Damien. 2004. "Heterogeneity and Portfolio Choice: Theory and Evidence". Working Paper.

[13] Epstein, Larry G. and Schneider Martin. 2003. "Recursive Multiple Prior". *Journal of Economic Theory*. 113. pp. 1-31.

[14] Epstein, Larry G. and Schneider Martin, 2005, "Learning under Ambiguity". Working Paper.

[15] Epstein Larry G. and Wang Tan. 1994. "Intertemporal Asset Pricing under Knightian Uncertainty". *Econometrica*. 62 (3). pp. 283-322.

[16] Fox Craig R. and Tversky Amos. 1995. "Ambiguity Aversion and Comparative Ignorance". *Quarterly Journal of Economics*. 100 (3). pp. 585-603.

[17] Garlappi Lorenzo, Uppal Raman and Wang Tan. 2006. "Portfolio Selection wtih Parameter Uncertainty: A Multi-Prior Approach". *Review of Financial Studies*. 20. pp. 41-81.

[18] Gilboa Itzach and David Schmeidler. 1989. "Maxmin Expected Utility with Non-unique Priors". *Journal of Mathematical Economics*. 18. pp 141-153.

[19] Gomes, Francisco and Michaelides Alexander. 2005. "Optimal Life-Cycle Asset Allocation: Understanding the Empirical Evidence". *The Journal of Finance*. 60(2): 869-904.

[20] Guiso L., Haliassos M. and Jappelli T. eds. 2001. "Household Portfolios". Cambridge and London: MIT press.

[21] Haliassos, Michael and Michaelides Alexander. 2003. "Portfolio Choice and Liquidity Constraints". *International Economic Review*. 44(1): 143-177.

[22] Hansen Gary D. 1993. "The Cyclical and Secular Behaviour of the Labour Input: Comparing Efficiency Units and Hours Worked". *Journal of Applied Econometrics*. 8(1): 71-80.

[23] Heaton John and Lucas Deborah. 1997. "Market Frictions, Savings Behavior, and Portfolio Choice", *Macroeconomic Dynamics*. 1: 76-101.

[24] Heaton John and Lucas Deborah. 2000. "Portfolio Choice in The Presence Of Background Risk", *Economic Journal*. 110: 1-26.

[25] Hubbard Glenn, Skinner Jonathan and Zeldes Stephen. 1994. "The Importance of Precautionary Motives in Explaining Individual and Aggregate Saving". *Carnegie-Rochester Conference Series in Public Policy*. 40: 59-126.

[26] Jorion Philippe and Goetzmann William N. 1999. "Global Stock Markets in the Twentieth Century ". *Journal of Finance*. 54 (3). pp. 953-980.

[27] Ju Nengjiu and Miao Jianjun. 2007. "Ambiguity, Learning and Asset Returns". Working Paper.

[28] Klibanoff Peter, Marinacci Massimo and Mukerji Sujoy. 2006. "Recursive Smooth Ambiguity Preferences". Working Paper.

[29] Leippold Markus, Trojani Fabio and Vanini Paolo. 2007. "Learning and Asset Prices under Ambiguous Information". *Review of Financial Studies.* 20 (5). pp. 1-33.

[30] Lettau Martin. "Inspecting the Mechanism: Closed-form Solutions for Asset Prices in Real Business Cycle Models". *The Economic Journal.* 113 (July). pp. 550-575.

[31] Lusardi Annamaria and Mitchell Olivia S. 2007. "Financial Literacy and Retirement Preparedness: Evidence and Impications for Financial Education". *Business Economics.* January. pp. 35-44.

[32] Lynch Anthony W., and Sinan Tan. 2004. "Labor Income Dynamics at Business Cycle Frequencies: Implications for Portfolio Choice ". Working paper 11010, NBER (Cambridge, MA).

[33] Maenhout Pascal. 2004. "Robust Portfolio Rules and Asset Pricing ". *Review of Financial Studies.* 17 (4). pp. 951-983.

[34] Malmendier Ulrike, and Nagel Stefan. 2009. "Learning from Inflation Experience". Working Paper.

[35] Mankiw Gregory N., and Reis Ricardo. 2007. "Sticky Information in General Equilibrium". *Journal of the European Economic Association.* 5 (2-3). pp. 603-613.

[36] Mehra Rajnish and Prescott Edward C. 1985. "The Equity premium: A Puzzle". *Journal of Monetary Economics.* 15. pp. 145-161.

[37] Paiella Monica. 2001. "Limited Financial Market Participation: A Transaction-Cost Based Explanation". The Institute for Fiscal Studies. WP 01/06.

[38] Press William H., Flannery Brian P., Teukolsky Saul A., and Vetterling William T. 1992. "Numerical Recipes in Fortran 77: The Art of Scientific Computing". Cambridge University Press, (Cambridge, MA).

[39] Reis Ricardo. 2006. "Inattentive Consumers". *Journal of Monetary Economics*. 53 (8). pp. 1761-1800.

[40] Siegel Jeremy J. 1992. "The Equity Premium: Stock and Bond Returns since 1802". *Financial Analyst Journal*. 48 (1). pp. 28-38.

[41] Sims Christopher A. 2006. "Rational Inattention: A Research Agenda". working paper.

[42] Tauchen George. 1986. "Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions". *Economics Letters*, 20. pp. 177-181.

[43] Vissing-Jørgensen Annette. 2002. "Towards an explanation of household portfolio choice heterogeneity: Nonfinancial income and participation cost structures. ". Working paper 8884, NBER, (Cambridge, MA).

[44] Vissing-Jørgensen A. 2003. "Perspectives on Behavioral Finance: Does "Irrationality" Disappears with Wealth? Evidence from Expectations and Actions ", in Mark Gertler and Kenneth Rogoff eds. NBER Macroeconomics Annual, MIT Press, Cambridge, MA.

[45] Wachter Jessica A. and Yogo Motohiro. 2007. "Why Do Household Portfolio Shares Rise in Wealth". Working Paper.

[46] Welch Ivo. 2000."Views of Financial Economists on the Equity Premium and on Professional Controversies". *The Journal of Business*. 73 (4). pp.501-537.