Experience-based Learning, Stock Market Participation and Portfolio Choice*

Richard Foltyn†

June 2020

Recent evidence suggests that lifetime experiences play an important role in determining households' investment choices. I incorporate these findings and the fact that household portfolios are underdiversified into an otherwise standard life-cycle model and examine to what extent they can help resolve long-standing puzzles in the literature regarding stock market participation and the fraction of financial wealth invested in risky assets. I show that experience-based learning about returns creates a positive correlation between a household's position in the wealth distribution and its optimism about future returns. The wealthy consequently increase their investment in risky assets, while participation is limited among poor households. I find that in a reasonably calibrated quantitative model, this mechanism is able to close approximately half of the gap between the participation rates observed in the data and the predictions from standard models. On the other hand, the average conditional risky share mostly remains unaffected.

Keywords: Stock-market participation, portfolio choice, learning, subjective beliefs, underdiversification

JEL classification: E21, E7, G40, G51

*I want to thank Tobias Broer, Jon de Quidt, Alexandre Kohlhas, Per Krusell, Kieran Larkin, Kurt Mitman, Arash Nekoei, Jonna Olsson, Kathrin Schlafmann, Kjetil Storesletten and Roine Vestman, as well as seminar participants at Stockholm University for valuable comments and discussions. The computations were enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) at HPC2N partially funded by the Swedish Research Council through grant agreement no. 2020/9-6.

†richard.foltyn@glasgow.ac.uk, University of Glasgow

1 Introduction

How do households split their financial wealth among different types of assets? From US data on portfolio allocations, it has long been known that wealthier households are more likely to participate in the stock market (Poterba and Samwick (1995), Haliassos and Bertaut (1995)), while the fraction invested in risky assets conditional on participation is almost flat across the wealth distribution. A similarly clear pattern emerges when examining portfolios over the life-cycle: participation is humped-shaped in age, with younger households being the least likely to invest in stocks despite their longer investment horizon. Recent evidence from Swedish and Norwegian administrative data (Bach, Calvet, and Sodini (2018), Fagereng et al. (2019)) confirms these patterns to be present in other countries and additionally documents vast return heterogeneity across the wealth distribution, which to a large extent stems from differences in exposure to risky assets.[1]

Standard models of portfolio choice, on the other hand, have largely been unsuccessful in replicating these empirical observations, in particular the limited participation (Campbell (2006), Guiso and Sodini (2013)). Early life-cycle models such as Cocco, Gomes, and Maenhout (2005) and Gomes and Michaelides (2005) more or less obtain the results established in Merton (1971), albeit in a more quantitative framework. These imply that young households, whose wealth primarily consists of non-tradable human capital (future earnings), choose to invest most if not all of their freely disposable financial wealth in risky stocks, while older households diversify their investment towards risk-free bonds. Since young households are, on average, poor in terms of financial wealth, this mechanism creates a risky share that is decreasing in wealth in the cross-section, contrary to what we observe in the data.

More recent attempts to reconcile household-finance models with empirical findings mostly work along two dimensions: the first approach abandons the almost ubiquitous assumption of a constant relative risk aversion and instead incorporates non-homothetic utility, as in Wachter and Yogo (2010).[2] A second strand of literature imposes particular assumptions on stochastic labor income to tilt the portfolio allocation of young or poor households away from the risky asset; papers in this group include Catherine (2019) and Chang, Hong, and Karabarbounis (2018). However, their mechanisms have

[1] Hubmer, Krusell, and Smith (2019) and Benhabib, Bisin, and Luo (2019), in turn, establish the central role of differences in portfolio returns as a major determinant of wealth inequality in quantitative models.

[2] Gomes and Michaelides (2003) introduce additive habits into a life-cycle model which can also generate a decreasing relative risk-aversion, but they find no improvement over a standard CRRA framework.

no effect on retired households and do not alter the extensive-margin participation decision.

In this paper, I explore an alternative explanation for the portfolio composition patterns observed in the data: based on evidence in Malmendier and Nagel (2011, 2016), the model presented here departs from rational expectations and instead lets households form beliefs about stock returns based on their history of previous realizations. Furthermore, in line with the data, I assume that households hold underdiversified portfolios, which results in idiosyncratic return histories. These two additions to an otherwise standard household-finance model give rise to positive sorting across beliefs and a household's position in the wealth distribution: rich households, to the extent that they are rich because of high returns in the past, are more optimistic and choose to invest a higher share in stocks than if they had known the true data-generating process. Poorer households who experienced low returns believe that risky returns will continue to be low, and hence shift their portfolios towards risk-free bonds, or exit the stock market altogether.

The findings from Malmendier and Nagel (2011, 2016) have recently been incorporated into models to investigate their implications for asset prices, trading volumes and return predictability (see Schraeder (2015), Nakov and Nuño (2015), Collin-Dufresne, Johannes, and Lochstoer (2016a, 2016b), Ehling, Graniero, and Heyerdahl-Larsen (2017), Malmendier, Pouzo, and Vanasco (2018), Nagel and Xu (2019)), and persistent investment slumps following severe recessions (Kozlowski, Veldkamp, and Venkateswaran 2015). These papers mainly either feature representative agents or representative cohorts in infinite-horizon settings and thus, they cannot speak to how the portfolio composition varies across the wealth distribution and over the life-cycle.

Conversely, I embed subjective belief heterogeneity into an incomplete-markets life-cycle model in which households face idiosyncratic earnings shocks. Additionally, returns on the risky asset are allowed to be imperfectly correlated within cohorts. Households know the true variance of risky returns but form beliefs about the mean excess return which they update following a rule similar to the one estimated in Malmendier and Nagel (2011, 2016).

I find that this mechanism can claim a partial success in reconciling portfolio choices with those observed in the data: compared to the standard model, the positive correlation between beliefs and a household's position in the wealth distribution induces limited participation even among middle-class households, something that cannot be achieved with participation costs of a reasonable magnitude. Subjective beliefs are thus able to close approximately half the gap between the participation rates observed in U.S. data and the predictions generated by a standard model. On the other hand, conditional on participation, the risky share in the cross-section mostly remains unchanged. While an individual household's optimal risky share responds to changes in beliefs in an intuitive way, these effects average out in the aggregate.

Unlike Chang, Hong, and Karabarbounis (2018) and Catherine (2019), which rely on certain properties of the stochastic earnings process to make the risky asset less desirable, the mechanism employed here continues to work for households once they retire. Furthermore, the model generates limited participation even when no participation costs are imposed. This is in contrast to the above papers and most other models where households face a positive expected excess return; in such a setting, absent any participation costs, agents choose to invest a (potentially small) positive fraction of their savings in the risky asset, irrespective of their risk aversion.

As a robustness test, I discuss an alternative model specification in which households update beliefs using Bayes' rule instead of the mechanism suggested in Malmendier and Nagel (2011, 2016). In this scenario, agents are perfectly rational (since Bayesian updating is the optimal strategy for updating beliefs as new information arrives), but suffer from limited information since they do not incorporate return realizations prior to their birth when forming beliefs. I show that the results are almost identical, even though in the Bayesian case, household expectations converge to the true excess return more quickly, thus muting the effect of subjective beliefs.

The remainder of the paper is organized as follows: In section 2, I discuss related papers in the literature. In section 3, I review the stylized facts on households' portfolio composition observed in U.S. data. Then, section 4 introduces a simple three-period model to illustrate the mechanism, while section 5 expands it to a quantitative life-cycle framework. I discuss the calibration in section 6 and present results for the benchmark model in section 7. Section 8 shows that these largely remain unchanged if agents use Bayes' rule to update beliefs, and section 9 demonstrates that the findings are robust to assuming more realistic levels of underdiversification. Section 10 concludes the paper.

2 Related literature

This paper relates to several strands of literature: first, to papers in household finance trying to explain limited stock market participation or the fraction of financial wealth invested in risky assets; second, to papers studying investors' financial decisions without imposing rational expectations; and third, to papers quantifying the importance of portfolio choice for wealth inequality. Additionally, it builds on a vast and rapidly growing empirical literature on beliefs and their implications for portfolio choice.

Household finance. Among the early seminal papers in the first group are Cocco, Gomes, and Maenhout (2005) and Gomes and Michaelides (2005).[3] As the former do not have any participation costs, in their paper household portfolio allocations do more or less reflect the results in Merton (1971): young (or poor) households invest all their financial wealth into stocks and rebalance their portfolio towards safe bonds as they age (or become more wealthy). Participation is universal for all households that choose to save a positive amount. Gomes and Michaelides (2005) extend this framework, incorporating preference heterogeneity and a fixed cost of entering the stock market. In their model, more risk-averse agents with higher precautionary savings are more likely hold stocks as they are those willing to pay the participation cost. Other papers introduce additional mechanisms to better match participation over the life-cycle: for example, Fagereng, Gottlieb, and Guiso (2017) include stock-market disaster risk, which combined with a per-period participation cost prompts older households to liquidate their stock holdings at more realistic rates.

A second group of papers proposes mechanisms that attempt to reconcile the model-generated conditional risky share with data. Catherine (2019) incorporates the earning process introduced in Guvenen, Ozkan, and Song (2014) and Guvenen et al. (2016) that exhibits cyclical skewness, i.e. the likelihood of large drops in earnings when risky returns are low, to match the portfolio composition over the life-cycle.[4] Households who are thus particularly exposed to this kind of earnings risk are more reluctant to hold

[3] In this discussion, I exclusively focus on papers with two liquid assets, a risk-free bond and a risky stock. The role of housing in households' portfolio decisions is investigated by, among others, Cocco (2004), Yao and Zhang (2004), and Vestman (2018) in a Swedish context. See Campanale, Fugazza, and Gomes (2015) for a model where the risky asset is illiquid.

[4] This is an application of an idea going back to Mankiw (1986), used in Krusell and Smith (1997) and Storesletten, Telmer, and Yaron (2007), among others, to generate higher risk premia.

a large fraction in risky assets.[5] On the other hand, Chang, Hong, and Karabarbounis (2018) use age-dependent unemployment risk with zero replacement rates and uncertainty about future earnings growth to achieve the same goal. However, once retired, households in these models revert to the same counterfactual investment choices as in the standard model, and neither mechanism affects participation.

Asset pricing and learning from experience. Another set of related papers investigates the implications of belief formation estimated in Malmendier and Nagel (2011, 2016) in an asset-pricing context. Schraeder (2015) and Malmendier, Pouzo, and Vanasco (2018) build representative-cohort models featuring CARA preferences to characterize price dynamics and household choices in the presence of non-Bayesian learning in an analytically tractable way. Collin-Dufresne, Johannes, and Lochstoer (2016b, 2016a) use a Bayesian-learning framework in which a representative agent (or two representative dynasties) are uncertain about the mean of stochastic consumption growth. In Collin-Dufresne, Johannes, and Lochstoer (2016a), the variance of the prior is reset whenever a new generation enters, thus mimicking the "learning from experience" mechanism identified in Malmendier and Nagel (2011, 2016). Collin-Dufresne, Johannes, and Lochstoer (2016a) find that this form of belief updating has a substantial effect on the the risk premium in their model. Ehling, Graniero, and Heyerdahl-Larsen (2017) introduce "learning from experience" into a Blanchard (1985)-type perpetual-youth model in which young households react more strongly to positive return shocks, thus decreasing the risk premium. Finally, in a business-cycle context, Kozlowski, Veldkamp, and Venkateswaran (2015) introduce belief updating into a representative-agent economy to generate long-lasting effects of severe recessions. In their paper, agents form beliefs about capital returns in a non-parametric fashion; new (extreme) observations can thus skew the estimated kernel density which agents use to forecast future returns for a prolonged period of time.

Portfolio choice and wealth inequality. A strand of literature that is not directly related to portfolio choice, but establishes its importance for wealth inequality, includes papers such as Hubmer, Krusell, and Smith (2019) and Benhabib, Bisin, and Luo (2019). Neither of these contains an explicit portfolio choice, but Hubmer, Krusell, and Smith (2019) incorporate exogenous heterogeneous returns into their model and find that

[5]One implication is that earnings-rich households have the lowest risky share as the earnings risk enters multiplicatively in that model.

these are a major driver of wealth inequality, in particular affecting the right tail of the wealth distribution. They calibrate a wealth-level-dependent return process to moments reported in Bach, Calvet, and Sodini (2018). However, it is important to stress that the increase in portfolio returns along the wealth gradient is primarily a result of household choices, and return heterogeneity conditional on asset class is only of secondary importance. To illustrate this, Bach, Calvet, and Sodini (2018) report that the lowest decile of the wealth distribution earns an annual excess return of 0.61% on their total financial wealth, whereas the corresponding value is above 4% for the wealthiest households. However, the return difference on *risky* financial assets is substantially smaller, since the corresponding excess returns are 5.9% for the lowest decile and around 7.7% for the richest households. Thus, the share of risky financial assets, which rises from 10% for the poorest to almost 60%, is the main driver of return heterogeneity, and highlights the importance of understanding the determinants of household portfolio choices. Kuhn, Schularick, and Steins (2019) provide empirical evidence for the importance of portfolio composition for aggregate wealth dynamics: Using decades of data from the SCF, they show that increases in house prices reduce wealth inequality as this mostly benefits the middle class who hold leveraged positions in housing. On the other hand, gains in the stock market have the opposite effect, i.e. increasing the wealth share held by the richest since middle-class households participate in stock markets to a lesser extent.

Empirical literature on beliefs and portfolio choice. There is a vast empirical literature on beliefs about asset returns and how these relate to household portfolios. One general conclusion is that beliefs in the population are very heterogeneous and belief updating can be classified into a few distinct types that show either extrapolative or mean-reverting behavior, or evolve as if returns were a random walk (Gaudecker and Wogrolly (2019), Heiss et al. (2019), Dominitz and Manski (2011)).[6] At an aggregate level, Greenwood and Shleifer (2014) examine the average expected returns from six different surveys and find them to be highly correlated with past returns, interpreting this as evidence for extrapolative beliefs. Similarly, Vissing-Jorgensen (2003) reports that investors are more optimistic about the stock market after high returns on their own portfolios; moreover, she documents heterogeneity in beliefs depending on investors' years of experience. Numerous papers find a positive relationship between higher expected returns and the probability of participating in the stock market (Arrondel,

[6]Giglio et al. (2019) argue that beliefs are very persistent and most of the dispersion is due to between-person variation that cannot be explained by observables.

Hector, and Tas (2014), Kézdi and Willis (2011), Hurd and Rohwedder (2012), Dominitz and Manski (2007)), even though the magnitudes are usually small. Similar findings have been established for the conditional risky share: while there is a statistically robust relationship, the effect of higher expected returns on risky shares is an order of magnitude lower than predictions from standard models (Giglio et al. (2019), Kézdi and Willis (2011), Ameriks et al. (2019)). Another group of papers does not directly observe beliefs, but relates return histories to individuals' investment decisions and finds evidence for reinforcement learning: Choi et al. (2009) report that higher past performance on retirement accounts is associated with increases in retirement savings, while Meyer and Pagel (2019) establish that the probability of reinvesting funds after quasi-random mutual fund closures is higher for investors who experienced gains. Kaustia and Knüpfer (2008) provide evidence that the likelihood of participating in IPOs decreases for investors who previously experienced poor performance. Malmendier and Nagel (2011) find that individuals who lived through times of high returns are more likely to participate in the stock market and invest a higher share in stocks. Lastly, Briggs et al. (2019) provide quasi-experimental evidence for the effect of wealth on stock market participation using data on Swedish lottery winners. They document that standard models with participation costs predict responses in the participation rate which are three to four times larger than those found in the data. On the other hand, they find that subjective beliefs elicited in a supplementary survey can account for about half of this gap. Additionally, splitting their sample into individuals who experienced high vs. low stock market returns prior to winning the lottery, they show that the participation responses are significantly larger within the first group.

3 Household portfolios in U.S. data

In this section, I document portfolio allocations in the Survey of Consumer Finances (SCF) using the 1998, 2001, 2004 and 2007 waves. I restrict the sample to include individuals aged 20 to 79. The analysis follows the one in Chang, Hong, and Karabarbounis (2018), but I additionally report portfolio composition across the wealth distribution, while they focus on the life-cycle. The variable definitions of aggregate asset categories (safe and risky financial assets) are identical to their definitions (see their appendix for a more detailed description of the SCF and the variables used).

I exclusively focus on how households invest their *financial wealth*, which does not include housing or actively managed businesses, but does include the value of businesses owned but not actively managed by households. For completeness, Table 5 and Table 6 in the appendix report detailed summary statistics for important components of households' balance sheets which also include asset and debt positions that are not part of financial wealth.

I will mostly be concerned with two statistics characterizing household portfolios across the wealth distribution and over the life-cycle:

1. The participation rate in risky assets, i.e. the fraction of households who hold any risky assets.
2. The conditional risky share, i.e. the amount of financial wealth invested in risky assets as a fraction of total financial wealth, conditional on a non-zero amount held in risky assets.

In the following graphs, I distinguish between the "gross share," which is defined as the risky share obtained when using *gross* safe assets, and the "net share" which controls for consumer debt. Using S and R to denote gross safe and risky financial assets, respectively, the gross risky share is defined as

$$\xi_{gross} = \frac{R}{R+S}$$

while the net risky share is

$$\xi_{net} = \frac{R}{R+S-B}$$

where B is the sum of credit card and consumer loans. The difference $S - B$ is thus a measure of net safe assets.

Figure 1 plots the portfolio composition along gross total wealth, which is the wealth measure I use when matching model moments to data. Two stylized facts emerge: the participation rate is almost monotonically increasing from zero to close to 100% along the wealth distribution, while the conditional risky share is more-or-less flat, except for the first wealth decile. The graphs show that for the most part, the gross and net risky shares are quite close, which is due to the fact that consumer debt is not very high on average. In the appendix in section A.2, I show that these plots are quite similar when using other wealth measures such as gross financial wealth or net worth.

In Figure 2, I report the evolution of the portfolio composition over the life-cycle, averaged over 5-year bins (ages 20–24, . . . , 75–79). The participation rate exhibits a hump-

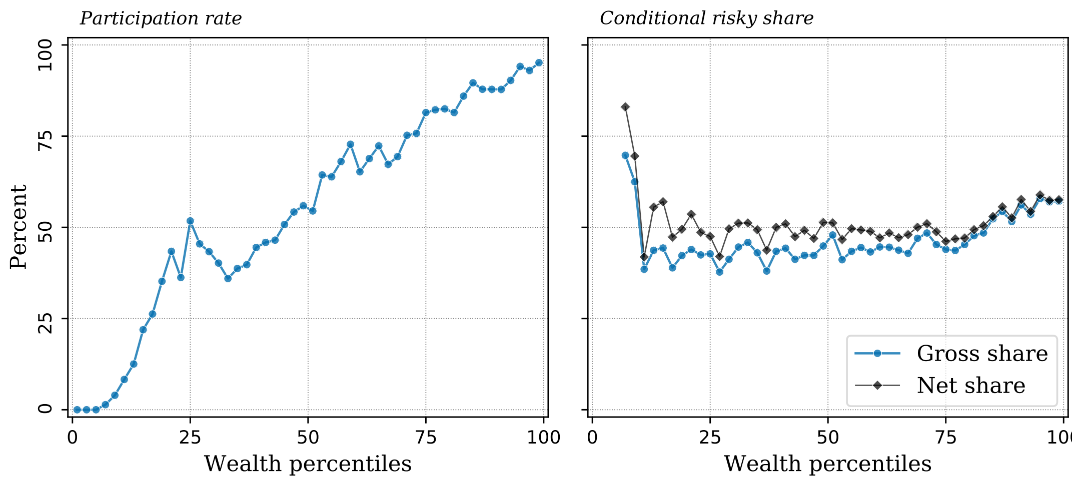


Figure 1: Portfolio composition along percentiles of gross total wealth. Each dot represents two percent of households. Data source: SCF 1998–2007.

shaped pattern: only 30% of young households invest in risky assets, whereas this value peaks at around 65% for ages close to retirement. After that, participation slopes down to approximately 40% for those aged 80.

On the other hand, the conditional risky share, shown in the right-hand panel of Figure 2, is almost flat across age. Similar to the previous graphs, it makes little difference whether the gross or net risky share is used.

The evidence presented above ignores that fact that the most important asset on many household's balance sheet is residential real estate (the homeownership rate is approximately 69.4% in this SCF sample). Since there is no housing in the model presented below, one potential concern is that financial wealth allocations differ systematically between home owners and renters because housing affects the investment decision of non-housing financial assets.

Figure 3 suggests that this is not the case: conditioning on wealth (in this case deciles of gross total wealth), homeowners and renters allocate their financial wealth similarly.[7] By construction, conditional on total gross wealth, the balance-sheet composition of homeowners and renters will look very different, so I present plots using alternative definitions of wealth in the appendix, section A.3. The findings remain unchanged for these wealth measures.

[7] All figures plotting portfolio choices disaggregated by home ownership show the gross risky share.

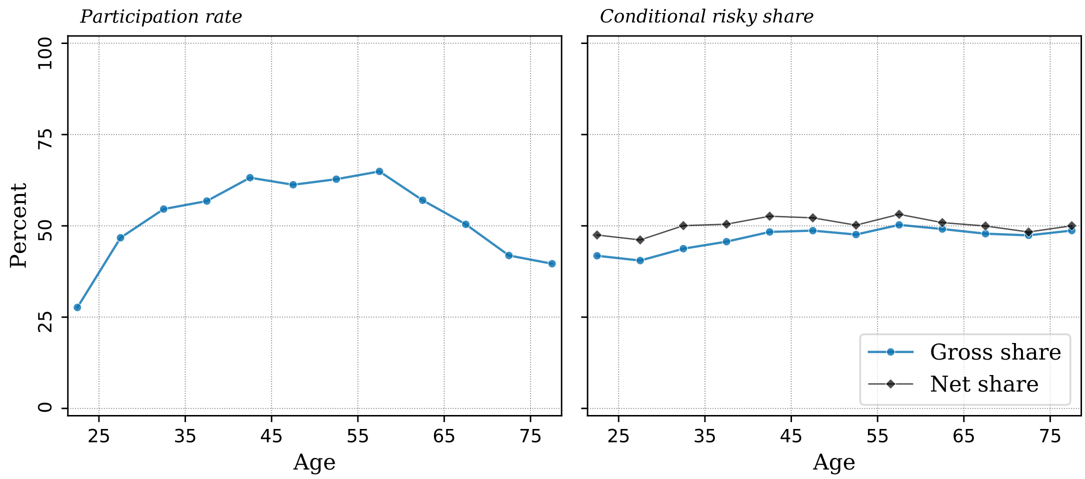


Figure 2: Portfolio composition over the life-cycle, 5-year averages. Data source: SCF 1998–2007.

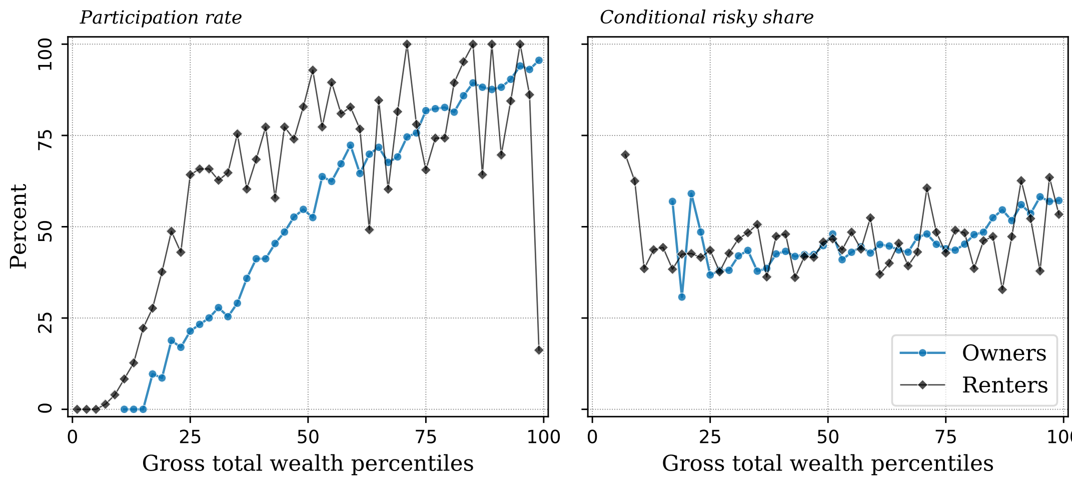


Figure 3: Portfolio composition along percentiles of gross total wealth by homeownership status. Wealth percentiles are computed for the *pooled* sample of owners and renters. Plots show averages conditional on wealth percentile and homeownership status. Data source: SCF 1998–2007.

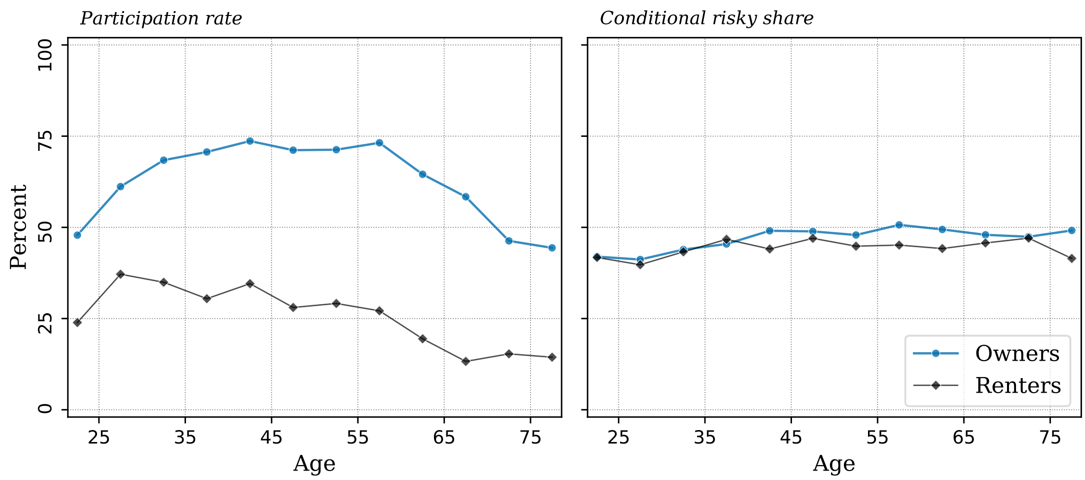


Figure 4: Portfolio composition over the life-cycle by homeownership status, 5-year averages. Data source: SCF 1998–2007.

The picture changes somewhat over the life-cycle, as shown in Figure 4. While the conditional risky share is almost identical for owners and renters, participation varies markedly between the two sub-populations. The reason is that renters are substantially poorer on average, which is not evident from Figure 3 since those moments condition on wealth. For example, in this SCF sample the median gross wealth among owners is $258,000 vs. $2,100 for renters, while the corresponding figures for net worth are $174,000 vs. $567, and $55,000 vs. $1,930 for gross financial wealth.

In Figure 5, I plot how the homeownership rate varies with wealth for several wealth aggregates, which illustrates a stark increase in homeownership along the wealth distribution, as expected. For example, while there are zero owners in the first decile of total gross wealth (left-hand panel), this fraction increases to almost 100% for the wealthiest. The picture is similar for various definitions of financial wealth (right-hand panel), where the homeownership rate is monotonically increasing in wealth except for the first decile.

To summarize the portfolio composition of homeowners vs. renters, the financial portfolios of these groups do not differ to any great extent conditional on wealth, and are particularly close when conditioning on financial wealth levels. Over the life-cycle, the conditional risky share is similar for both groups, while participation is uniformly shifted downwards for renters. This is driven by substantially lower wealth levels among renters and the fact that participation increases in wealth, as documented above.

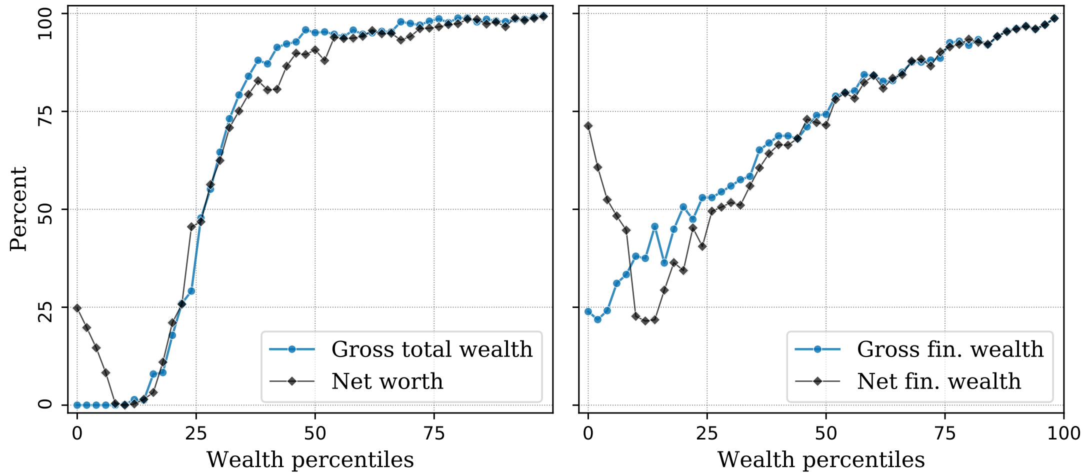


Figure 5: Share of homeowners along wealth percentiles. Left panel: total gross wealth and net worth. Right panel: Gross and net financial wealth. Data source: SCF 1998–2007.

In conclusion, the stylized facts highlighted in this section are as follows: participation substantially increases along the wealth distribution, from basically zero to 100%, while it is hump-shaped over the life-cycle, peaking around retirement. The share of financial wealth invested in risky assets, conditional on holding a positive amount of such assets, is mostly flat both along the wealth distribution and over the life-cycle.

In the remainder of the paper, I explore to what extent these patterns can be accounted for in a model with experience-based learning that gives rise to subjective beliefs about risky returns.

4 Simple three-period model

Before introducing the full quantitative model, I begin the exposition with a simplified three-period model which is nevertheless rich enough to illustrate the main mechanism. The terminal period is only needed so that households face a non-trivial decision problem in the second period, as opposed to simply consuming all resources, and can otherwise be ignored.

In the first period ($t = 1$), I assume that households indexed by i are ex ante identical. They decide on consumption c, total savings b and the share invested in the risky asset, denoted by ξ. Their problem can be written as

$$
\begin{aligned}
V_1\left(a_1\right) &= \max_{c_1, b_1, \xi_1} \left\{ u(c_1) + \beta \mathbf{E} V_2\left(a_2, \widehat{\mu}_{i2}\right) \right\} \\
\text{s.t.} \qquad a_1 &= c_1 + b_1, \qquad c_1 \geq 0,\ b_1 \geq 0 \\
a_2 &= \Big(\xi_1 R_{i2} + (1-\xi_1) R_f\Big) b_1, \qquad \xi_1 \in [0,1]
\end{aligned}
$$

where a_1 are beginning-of-period assets and $u(\bullet)$ is the standard power utility function with relative risk-aversion γ. Households can choose to invest either in a risk-free bond with gross return R_f, or in a risky asset with excess return

$$
R_{it+1} - R_f = \overline{\mu} + z_{it+1}, \qquad z_{it+1} \overset{\text{iid}}{\sim} \mathcal{N}\left(0, \sigma^2\right) \tag{1}
$$

Risky return realizations are i.i.d. in the cross-section and across time and thus, each household has a potentially different return realization R_{it+1}. Households are uncertain about the true value $\overline{\mu}$ and instead have household- and time-specific beliefs, denoted by $\widehat{\mu}_{it}$. However, the variance of risky returns σ^2 is known with certainty.

As households start out ex-ante identical with the same wealth level a_1 and the correct belief $\widehat{\mu}_{i1} = \overline{\mu}$, they choose the same portfolio allocation. Absent any labor income, the optimal choice is sufficiently well approximated by the standard formula $\xi_1 \approx \overline{\mu}/(\gamma\sigma^2)$.[8]

At the beginning of period two, after observing their risky return realizations R_{i2}, households update their beliefs according to

$$
\widehat{\mu}_{i2} = (1-\alpha)\widehat{\mu}_{i1} + \alpha\left(R_{i2} - R_f\right) \tag{2}
$$

where α is the weight put on the most recent return realization. I postpone the discussion on the exact way in which α is determined until later; for now it suffices to say that various updating strategies (experience-based learning as in Malmendier and Nagel (2011, 2016), Bayesian learning from experience, constant-gain learning) can be mapped into a corresponding value for α.

Equation (2) implies that starting in period two, there will be a dispersion of beliefs in the cross-section that closely mimics the risky return distribution. The optimization

[8] The chosen preference and return parameters ensure that the optimal risky share in the first period is interior, i.e. $\overline{\mu}/(\gamma\sigma^2) \in (0,1)$.

problem in the second period thus has an additional state variable $\widehat{\mu}_{i2}$ and can be stated as

$$
\begin{aligned}
V_2\left(a_2, \widehat{\mu}_{i2}\right) &= \max_{c_2, b_2, \xi_2} \left\{ u(c_2) + \beta \mathbf{E}_i u(c_3) \right\} \\
\text{s.t.} \qquad a_2 &= c_2 + b_2, \qquad c_2 \geq 0,\ b_2 \geq 0 \\
c_3 &= \left(\xi_2 R_{i3} + (1 - \xi_2) R_f \right) b_2, \qquad \xi_2 \in [0, 1]
\end{aligned}
$$

The problem is the same as in the first period, except that now households form subjective expectations about returns in $t = 3$, and will therefore choose different risky shares

$$
\xi_{i2} \approx \begin{cases} 0 & \text{if } \widehat{\mu}_{i2} \leq 0 \\ \widehat{\mu}_{i2} / (\gamma \sigma^2) & \text{if } 0 < \widehat{\mu}_{i2} < \gamma \sigma^2 \\ 1 & \text{else} \end{cases}
$$

Since there is no other heterogeneity, idiosyncratic return realizations do not only determine the beliefs in period two, but also a household's position in the wealth distribution. By construction, the wealthiest also end up being most overoptimistic about future returns, and consequently choose a higher risky share than in the fully rational model. This outcome is illustrated in Figure 6. By contrast, the optimal risky share with rational

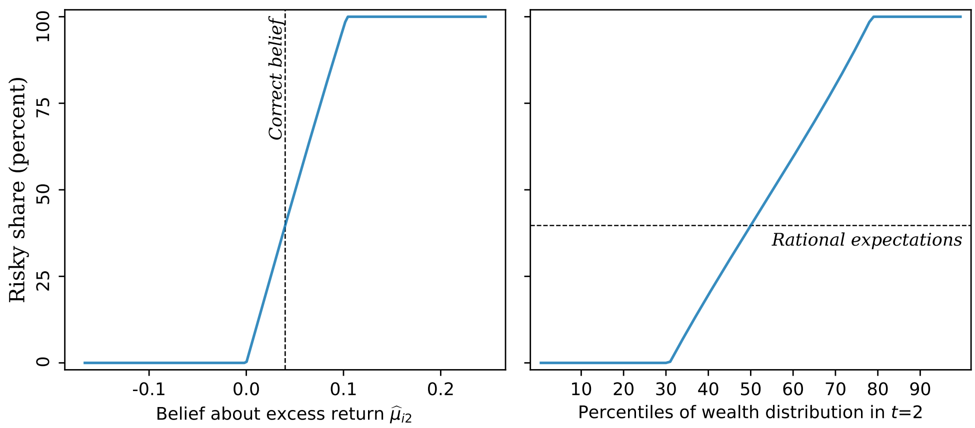


Figure 6: Risky shares across the belief and wealth distribution in the second period

expectations is illustrated by the dashed line, and is constant across the wealth distribution (in fact, it is the same as in the first period since the objective risky return distribution remains unchanged). Thus subjective belief heterogeneity creates an upward-sloping risky share in the cross-section and non-participation for low-wealth households. Since short-selling is not permitted, the poorest households (who at the same time are the most pessimistic about risky returns) choose not to hold any stocks, thus generating limited participation in the cross-section.

Naturally, this stylized setting vastly exaggerates the correlation between beliefs, wealth and portfolio choice. In reality, many factors other than asset returns affect a household's position in the wealth distribution, the most important being (uncertain) earnings and the life-cycle, which creates a hump-shaped age profile of asset holdings. To explore whether the mechanism outlined above carries over to a more quantitative framework which takes these aspects into account, I proceed to embed subjective beliefs into an otherwise standard life-cycle model.

5 Quantitative life-cycle model

Consider a discrete-time, partial-equilibrium life-cycle model along the lines of Cocco, Gomes, and Maenhout (2005), Gomes and Michaelides (2005) and many others in the household finance literature. I modify this framework along two dimensions, both of which are supported by empirical evidence: first, I impose that households form subjective beliefs about the risky asset's average return, and these beliefs are determined by each household's history of experienced risky returns. Second, I assume that households hold underdiversified portfolios with an idiosyncratic component, and they therefore differ in their return histories.

Besides their beliefs and return histories, households are heterogeneous in their wealth holdings, earnings capacity and age. They live up to a maximum of 89 years and face an age-dependent probability of death. Households supply labor inelastically, retire at an exogenously fixed age and receive a deterministic retirement income thereafter.

Markets are incomplete, so households cannot insure against earnings or survival risk. I additionally impose that households cannot borrow in either asset.

5.1 Subjective beliefs

Since subjective beliefs are the main non-standard building block of the model and work the same for both working-age and retired agents, I discuss these first.

As in the three-period example, households allocate their portfolio across a risk-free and a risky asset. Risky returns are assumed to be i.i.d. over time, and in the benchmark case I additionally impose that they are i.i.d. in the cross-section. While in the data individual returns on financial wealth are, of course, not independent of other investors' return realizations, Calvet, Campbell, and Sodini (2007) report that in Swedish register data, the idiosyncratic share of the variance of individual portfolio returns due to underdiversification is approximately 60%. Hence, there is something to be learned even from the i.i.d. assumption adopted in the benchmark case. I relax this assumption in section 9.

A household at age h chooses to invest in the risk-free asset with gross return R_f and a risky asset with excess return

$$r^e_{ih+1} \equiv R_{ih+1} - R_f = \overline{\mu} + z_{ih+1}\,, \qquad z_{it+1} \overset{\text{iid}}{\sim} \mathcal{N}\left(0, \sigma^2\right) \tag{3}$$

where the variance of risky returns σ^2 is constant and known.[9] After a return realization R_{ih} has been observed, agents update their beliefs according to

$$\widehat{\mu}_{ih} = (1 - \alpha_h)\widehat{\mu}_{ih-1} + \alpha_h \left(R_{ih} - R_f\right) \tag{4}$$

where $\widehat{\mu}_{ih-1}$ was the prevailing belief in the last period. The choice of learning mechanism determines how the weight α_h on the most recent observation is determined. I discuss three potential alternatives in turn.

Constant-gain learning. In this setting, $\alpha_h = \alpha$ is a time- and age-invariant parameter. This method is used by Nagel and Xu (2019) in a modified Bayesian framework in an infinite-horizon model. The findings in Malmendier and Nagel (2011, 2016), however, suggest that new information is assigned different weights depending on an individual's age, and are thus incompatible with a constant α.

Bayesian learning from experience (BLE). Bayes' rule is the optimal algorithm to update beliefs when new information arrives and hence, this approach is closer to

[9] For the remainder of the paper, I adopt the convention to subscript objects that only depend on age by h instead of t which is used to denote calendar time. In the benchmark model with i.i.d. returns in the cross-section and no time-varying aggregates, the household problem can be fully characterized in terms of household age h. Calendar time t will only play a role in section 9 when aggregate market returns are introduced into the model.

rational expectations.[10] However, agents only use information from their own experience to update their beliefs, thus discarding events prior to their birth. With a sufficiently long sequence of returns, agents' beliefs would converge to the true average excess return $\overline{\mu}$ and any subjective belief heterogeneity would consequently disappear in the long run. However, in a life-cycle setting, it is reasonable to assume that newborns do not factor in information that predates their own lifetime and instead impose Bayesian learning from experience.

Experience-based learning (EBL). A third belief formation method, proposed and estimated in Malmendier and Nagel (2011, 2016) and applied in an asset pricing context in Malmendier, Pouzo, and Vanasco (2018), postulates that the update weight is a function of age only.[11] Malmendier and Nagel (2011) show that an index of past returns, computed as

$$\overline{R}_{it}(\lambda) = \sum_{k=1}^{age_{it}-1} w\left(age_{it}, k; \lambda\right) R_{t-k} \tag{5}$$

(using their notation), with the weighting function given by

$$w(age_{it}, k; \lambda) = \frac{(age_{it} - k)^{\lambda}}{\sum_{j=1}^{age_{it}-1} (age_{it} - j)^{\lambda}} \tag{6}$$

is related to households' risk attitudes, stock market participation and the share invested in stocks conditional on participation. The functional form is sufficiently flexible to allow for past observations to receive more or less weight as one moves back in time, and collapses to the conventional estimate of the sample mean for $\lambda = 0$.

In the appendix, section B, I show how the weighting scheme in (5) and (6) can be transformed into a recursive formulation and stated in terms of the belief updating equation (4). Adapted to the notation used in my model, the update weight is given by

$$\alpha_h = \frac{(\underline{h} + h - 1)^{\lambda}}{\sum_{k=1}^{\underline{h}+h-1} (\underline{h} + h - k)^{\lambda}} \tag{7}$$

[10] I adopt the terminology used in Malmendier, Pouzo, and Vanasco (2018) who label the belief updating methods discussed here as "Bayesian learning from experience" and "experience-based learning."

[11] I combine the updating methods in Malmendier and Nagel (2011) and Malmendier and Nagel (2016) even though they use different functional forms. The latter paper assumes that people update their beliefs using a decreasing-gain learning algorithm where the gain is a function of age and birth year. However, as the authors show in the appendix to Malmendier and Nagel (2016), there is a direct mapping between the gain estimated in Malmendier and Nagel (2016) and the shape parameter λ of the weighting scheme in Malmendier and Nagel (2011), here restated in (5) and (6).

where $\underline{h}$ is the actual age corresponding to model age $h = 0$, i.e. $age = \underline{h} + h$.

Figure 7 shows the weights assigned to lagged observations implied by each of the three belief updating methods. The parameters are calibrated such that individuals at the age of 30 update their beliefs by approximately 2.5% (at a quarterly frequency) in all three cases. This is line with the estimates of λ in Malmendier and Nagel (2011) which at age 30 imply an update weight of about 10% at an annual frequency. The three mechanisms have starkly different implications for how past information is used (or discarded) to form beliefs: with constant-gain learning, lagged observations are weighted in a geometrically decreasing fashion, while with Bayesian learning from experience, all observations receive the same weight (since the returns are i.i.d.), even though this weight decreases as an individual collects longer time series of data. Lastly, the estimates in Malmendier and Nagel (2011, 2016) suggest that people aged 30 put a higher weight on the most recent observation (approximately 2.5%), while by the age of 70 this value drops to 1%. The effect of new information is thus muted for older individuals.

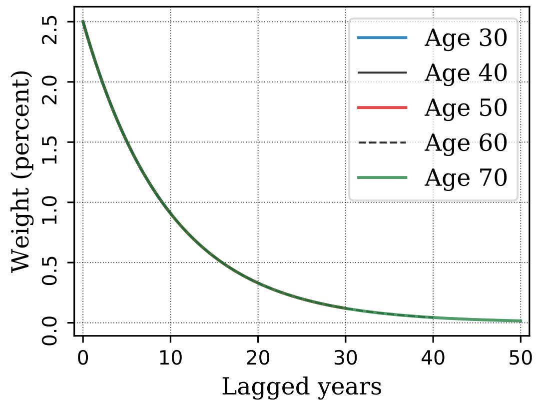
In this paper, I choose experience-based learning as the benchmark case for two reasons: first, Malmendier and Nagel (2011, 2016) provide evidence that historical observations weighted in this way have a significant impact on how people form beliefs and make financial decisions. Second, EBL with $\lambda = 0$ can replicate the same weights as BLE (for some initial prior), but $\lambda = 0$ is rejected by the findings in Malmendier and Nagel (2011, 2016).

In the next section, I describe how subjective beliefs are incorporated into the household's optimization problem.

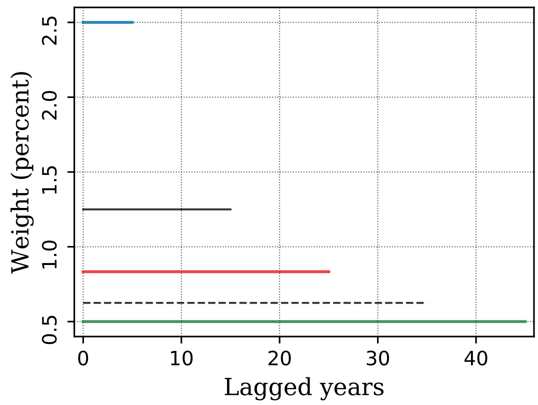
5.2 Retired households

Households exogenously retire at age H_r and continue to live up to a maximum age H, facing a stochastic survival risk along the way. Once retired, they receive retirement benefits that are proportional to their last pre-retirement labor income, which is summarized by the state p and remains unchanged through the retirement.

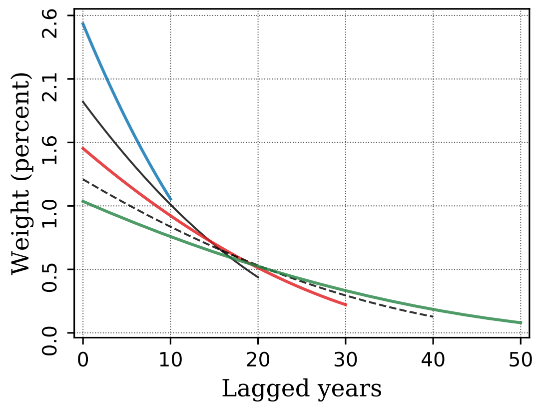
A retired household's state is represented by the tuple $\boldsymbol{x} = (h, a, p, \widehat{\mu}_i, j)$, where h denotes age, a is beginning-of-period cash-at-hand (i.e. any assets plus retirement benefits), $\widehat{\mu}_i$ is the current belief about excess returns and j indexes a (fixed) preference type. The



(a) Constant-gain learning



(b) Bayesian learning from experience (BLE)



(c) Experience-based learning (EBL)

Figure 7: Weights assigned to past observations for constant-gain learning, Bayesian learning from experience and experience-based learning.

household optimally chooses consumption c, total savings b and the share invested in the risky asset ξ in order to maximize

$$V_{jh}^{r}\left(a,p,\widehat{\mu}_{i}\right)=\max_{c,b,\xi}\left\{c^{1-\psi}+\beta_{j}\left[\pi_{h}^{s}\mathbf{E}_{i}\left[\left(V_{jh+1}^{r}\right)^{1-\gamma}\right]+(1-\pi_{h}^{s})\mathbf{E}_{i}\left[\left(V_{j}^{b}\right)^{1-\gamma}\right]\right]^{\frac{1-\psi}{1-\gamma}}\right\}^{\frac{1}{1-\psi}}$$

with continuation values

$$\begin{aligned}V_{jh+1}^{r}&\equiv V_{jh+1}^{r}\left(a',p,\widehat{\mu}_{i}'\right)\\ V_{j}^{b}&\equiv V_{j}^{b}\left(a_{b}'\right)\end{aligned}$$

subject to the budget constraint

$$a=c+b+\kappa\cdot\mathbf{1}_{\{\xi>0\}} \tag{8}$$

and the usual non-negativity constraints $c\geq 0$, $b\geq 0$. Moreover, households are not allowed to take short positions in either asset, restricting the risky share to $\xi\in[0,1]$. I assume that the preferences are of the Epstein-Zin-Weil form (Epstein (1988) and Epstein and Zin (1989), Weil (1990)) such that the EIS ψ^{-1} is not restricted to be the inverse of the relative risk aversion γ. Decoupling the EIS from the risk aversion is helpful when trying to match the wealth distribution (and hence savings).

To this end, households are also permitted to be heterogeneous in terms of preferences. These differences are assumed to be randomly drawn at birth and remain unchanged throughout their lifetime. A household's preference type, indexed by j, then determines its discount factor β_j and the bequest utility weight ϕ_j, which is discussed below. I assume that there are two types, with type 1 being the impatient household which also has a lower bequest motive.

In the benchmark calibration, I impose a fixed per-period participation cost κ when households choose to invest a positive amount in the risky asset, but additionally report results for $\kappa = 0$.

Given optimal choices and next-periods shock realizations, the ex-post return on a household's portfolio is

$$R_p'=\xi(R_i'-R_f)+R_f$$

and hence next-period cash-at-hand is

$$a' = R'_p b + \rho_{ss} p$$

with ρ_{ss} being the replacement rate relative to the earnings level p just prior to retirement.

Households survive with an age-dependent survival probability π^s_h with $\pi^s_H = 0$ in the terminal period. Upon death, they derive utility from "warm-glow" bequests as in De Nardi (2004), summarized by the function[12]

$$V^b_j(a_b) = \phi_j^{1/(1-\psi)} a_b \tag{9}$$

The parameter ϕ_j scales the utility derived from bequests relative to current-period consumption and the continuation value conditional on survival. While it is common in the macroeconomic literature to include a non-homothetic luxury-good component in the bequest motive such that it predominantly affects wealthy households, this induces a relative risk aversion that is increasing in wealth, which is undesirable in a context studying portfolio choice.

Subjective beliefs enter the household's problem via the state variable $\widehat{\mu}_i$, which I subscript by i to make explicit the link between household i's beliefs about mean excess returns and its subjective expectations, denoted by $\mathbf{E}_i[\dots]$. In the benchmark model, I assume that households are naive with respect to any future belief updating, i.e. they do not anticipate that they will revise their beliefs about risky returns as they observe realizations in the future. Therefore, their perceived law-of-motion for $\widehat{\mu}_{ih}$ is $\widehat{\mu}_{ih+1} = \widehat{\mu}_{ih}$. Alternatively, households could be sophisticated and factor in that for a given realization of R'_i tomorrow, they will update their beliefs according to (4) with the weight on the last observations given by (7). This difference arises only in terms of how expectations are formed; when simulating households, their beliefs are always updated according to (4). I report results for an economy with sophisticated households in the appendix, since for the benchmark model, the differences are modest.

[12]This simple linear function can be obtained from a more general functional form for the case of EZW preferences suggested in Bommier, Harenberg, and Le Grand (2017) by setting all their additive terms to zero (in particular the bequest shifter). With power utility, when $\gamma = \psi$, the bequest utility implied by (9) is the familiar $v^b_j(a_b) = \phi_j a_b^{1-\gamma}/(1-\gamma)$.

5.3 Working-age households

Working-age households are assumed to supply labor inelastically and therefore, their optimization problem is almost identical to the one discussed above. An agent with a state vector given by $\boldsymbol{x} = (h, a, p, \widehat{\mu}_i, j)$ for $0 \leq h < H_r$ solves

$$V_{jh}\left(a, p, \widehat{\mu}_i\right) = \max_{c,b,\xi} \left\{ c^{1-\psi} + \beta \left[\pi_h^s \mathbf{E}_i \left[\left(V_{jh+1}\right)^{1-\gamma} \right] + (1-\pi_h^s) \mathbf{E}_i \left[\left(V_j^b\right)^{1-\gamma} \right] \right]^{\frac{1-\psi}{1-\gamma}} \right\}^{\frac{1}{1-\psi}}$$

with continuation values

$$V_{jh+1} \equiv \begin{cases} V_{jh+1}\left(a', p', \widehat{\mu}_i'\right) & \text{if } h < (H_r - 1) \\ V_{jh+1}^r\left(a', p, \widehat{\mu}_i'\right) & \text{else} \end{cases}$$

and

$$V_j^b \equiv V_j^b\left(a_b'\right)$$

The constraints are identical to those of the retired household. However, conditional on survival, next-period cash-at-hand is now given by

$$a' = R_p' b + y'$$

where y' are stochastic earnings. I adopt the standard convention of modeling these as

$$\log y_{ih+1} = \log \omega_{h+1} + \log p_{ih+1} + \log \epsilon_{ih+1} \tag{10}$$

where ω_h is a hump-shaped, deterministic age-profile that reflects the average earnings growth over the life-cycle. Idiosyncratic uncertainty enters via the persistent labor component p_{ih}, which follows an AR(1) in logs,

$$\log p_{ih+1} = \rho_p \log p_{ih} + \nu_{ih+1}, \qquad \nu_{ih+1} \overset{\text{iid}}{\sim} \mathcal{N}\left(-\frac{1}{2}\sigma_\nu^2, \sigma_\nu^2\right)$$

as well as the purely transitory shock ϵ_{ih}, distributed as

$$\log \epsilon_{ih+1} \overset{\text{iid}}{\sim} \mathcal{N}\left(-\frac{1}{2}\sigma_\epsilon^2, \sigma_\epsilon^2\right).$$

I normalize the average wage level to unity and omit it from earnings altogether.

Belief updating and subjective expectations are identical to the retired household's problem. The only additional assumption needed is the distribution of beliefs of newborns with $h = 0$, which I discuss in the calibration section.

6 Calibration

6.1 Externally calibrated parameters

6.1.1 Demographics

Households enter the economy at the age of 20 and live for a maximum of 70 years up to a maximum age of 89. They exogenously retire after the initial 45 years. The age-dependent survival probabilities are taken from the U.S. period life tables (Arias and Xu 2018) for the year 2015, using the values for the white male sub-population. For the quarterly calibration, I assume that the survival probabilities are constant within a year and given by $\pi_h^{s,qtr} = \left(\pi_h^{s,annual}\right)^{1/4}$. The implied life expectancy at birth is 80.4 years (i.e. 60.4 model years).

6.1.2 Labor income

For the deterministic age profile of earnings $(\omega_h)_{h=0}^{H}$ in (10), I use the estimates for the high-school-educated sub-sample reported in Cocco, Gomes, and Maenhout (2005). This profile also includes the replacement rate ρ_{ss} for retirement benefits, which Cocco, Gomes, and Maenhout (2005) estimate to be approximately 68% for this group.

I take the parameters (ρ_p, σ_v^2) that govern the persistent earnings component as well as the variance of the transitory shocks σ_ϵ^2 from Krueger, Mitman, and Perri (2016) which they estimate from the PSID.

These parameter values are listed in Table 1. I normalize the (cohort-size-weighted) deterministic earnings profile for working-age households to one, such that the average earnings of working-age households are always unity.

	Description	Value	Source
Demographics			
$\underline{h}$	Initial age (in years)	20	
H	Maximum attainable age (in years)	89	
H_r	Retirement age (in years)	65	
π_h^s	Survival probabilities	–	US life tables (2015)
Earnings			
ρ_p	Auto-correlation of persistent earnings	0.9695	Krueger, Mitman, and Perri (2016)
σ_v	Std. dev. of persistent earnings shock	0.1960	Krueger, Mitman, and Perri (2016)
σ_ϵ	Std. dev. of transitory earnings shock	0.2284	Krueger, Mitman, and Perri (2016)
ω_h	Age-dependent earnings profile	–	Cocco, Gomes, and Maenhout (2005)
ρ_{ss}	Retirement income replacement rate	0.6828	Cocco, Gomes, and Maenhout (2005)

Table 1: Demographic and earnings parameters (annual)

	Description	Value	Source
Returns			
R_f	Gross risk-free return	1.02	Cocco, Gomes, and Maenhout (2005)
$\overline{\mu}$	Risk premium	0.04	Cocco, Gomes, and Maenhout (2005)
σ	Volatility of risky returns	0.16	Cocco, Gomes, and Maenhout (2005)
Subjective beliefs			
λ	Belief updating weight	1.5	Malmendier and Nagel (2011)
–	Cross-sectional mean of initial beliefs	0.04	–
–	Cross-sectional std. dev. of initial beliefs	0.09	–

Table 2: Risky return and belief formation parameters (annual)

6.1.3 Beliefs and returns

I adopt the standard values for the risk-free interest rate, the risk premium and the volatility of risky returns used in the household-finance literature. These are reported in Table 2.

Turning to beliefs, Malmendier and Nagel (2011) estimate the shape parameter λ for the experience-based learning model in (5), which governs the weighting of past realizations, for several outcome variables including stock market participation and the share invested in stocks conditional on participation. Their point estimates are in the range of approximately 1.3–1.9 at an annual frequency, depending on the exact specification. I choose λ to be 1.5 for an annual model, which implies that a 30-year-old individual assigns a weight of about 10% to the most recent observation. For the quarterly model, I therefore choose $\lambda = 2.0$ so that households aged 30 update their beliefs by $\approx 2.5\%$ as new information arrives.

The initial distribution of beliefs used to simulate the economy is determined as follows:

I assume that households observe risky returns for five years (20 periods in the quarterly model) before becoming economically active and making investment choices. Their initial belief distribution five years prior to entering the economy is identical to the risky return distribution (as they only have this single observation), which they then update according to the algorithm described above. This implies that when agents become economically active at the age of 20, the cross-sectional distribution of their beliefs about excess returns has a standard deviation of approximately 0.09 at an annual or 0.037 at a quarterly frequency.[13]

Since excess returns are Gaussian and the updating rule prescribes a deterministic (age-dependent) weight to be applied to new observations, this allows me to exactly characterize the distribution of beliefs at each age. As the initial distribution is a linear combination of Gaussian realizations, it is itself Gaussian at age $h = 0$. Applying the belief updating rule in (4), the cross-sectional distribution at any age $h + 1$ is therefore again Gaussian with a mean that evolves according to

$$\mathbf{E}\widehat{\mu}_{ih} = (1 - \alpha_h)\mathbf{E}\widehat{\mu}_{ih-1} + \alpha_h \mathbf{E}\left[R_{ih} - R_f \right] = (1 - \alpha_h)\mathbf{E}\widehat{\mu}_{ih} + \alpha_h \overline{\mu} \tag{11}$$

where the expectation is taken over the distribution of households aged h. The law of motion for the cross-sectional variance of within-cohort beliefs is

$$\begin{aligned} \mathrm{Var}\left(\widehat{\mu}_{ih} \right) &= (1 - \alpha_h)^2 \, \mathrm{Var}\left(\widehat{\mu}_{ih-1} \right) + \alpha_h^2 \, \mathrm{Var}\left(R_{ih} - R_f \right) \\ &= (1 - \alpha_h)^2 \, \mathrm{Var}\left(\widehat{\mu}_{ih-1} \right) + \alpha_h^2 \sigma^2 \end{aligned} \tag{12}$$

The evolution of beliefs in the cross-section is plotted in Figure 8. As equations (11) and (12) suggest, if newborns' beliefs are centered around the true excess return, beliefs will on average remain unbiased at any age. Furthermore, the within-cohort variance of beliefs collapses as a cohort ages, albeit at a slower pace than if agents had updated their beliefs optimally, which I discuss in section 8 in the context of Bayesian learning from experience.

[13] These differences stem from the fact that λ depends on the period length, and households have four times as many observations in the quarterly model.

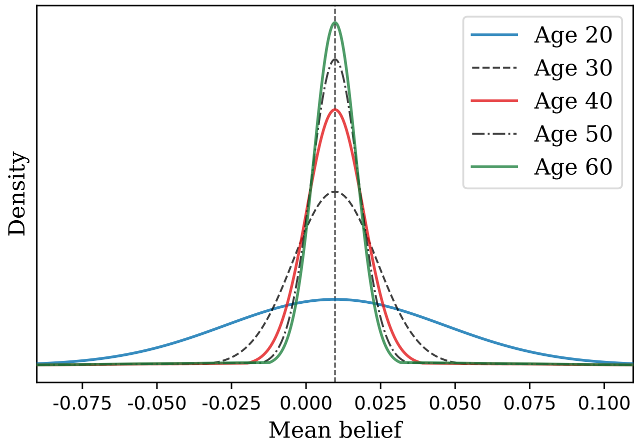


Figure 8: Evolution of beliefs in the cross-section with experience-based learning as in Malmendier and Nagel (2011, 2016).

6.2 Parameters determined by moment matching

I use the remaining preference parameters to approximately match the wealth distribution relative to average earnings as obtained from the SCF.[14] To this end, during the simulation newborn households draw initial wealth levels that correspond to the wealth distribution observed in the SCF for ages 20–25. This, however, turns out not to matter much in this class of models as in the presence of an upward-sloping earnings profile young households optimally consume their assets, thus leveling any differences in initial endowments.[15]

The relative risk aversion γ is assumed to be 5 for both household types, as a value of this magnitude is commonly used in the recent literature (Chang, Hong, and Karabarbounis (2018), Catherine (2019), Vestman (2018)).[16] The remaining parameters $\{\beta_j\}_j$, $\{\phi_j\}_j$, the elasticity of intertemporal substitution ψ^{-1} and the fraction of impatient households are set to values reported in Table 3.

A few comments are in order: while the discount factor heterogeneity seems large compared to calibrations in other contexts such as Krusell and Smith (1998), who were

[14] As in Chang, Hong, and Karabarbounis (2018) who use the same SCF sample, the average earnings in the data are assumed to be $40,000 per year.

[15] If permitted to do so, young households would in fact prefer to borrow against their future income to smooth lifetime consumption.

[16] Setting a higher value such as $\gamma = 10$, which is also a commonly used value, makes matching the wealth distribution almost impossible as it induces substantial precautionary savings in the presence of highly persistent earnings risk. Thus the "poor" households end up being substantially richer than in the data.

	Description	Value	Source
γ	Relative risk aversion	5	Standard
β_j	Discount factor	0.88, 1.06	–
ψ^{-1}	Elasticity of intertemporal substitution	0.35	–
ϕ_j	Weight on bequest utility	0.45, 10.0	–
–	Fraction of impatient households	0.85	–

Table 3: Preference parameters (annual)

the first to introduce β-heterogeneity into a quantitative model, these magnitudes are required to generate a plausible level of wealth inequality in this setting. The reason is that in life-cycle models which already have a great deal of background risk, heterogeneity in the discount factors is less powerful (see Hendricks (2007), and Foltyn and Olsson (2019) for a setting in which discount factor heterogeneity results from differences in life expectancy). In contrast, in Krusell and Smith (1998), households only face a low-persistence unemployment risk. Furthermore, without intergenerational wealth transfers, as in the present model, any wealth built up by a cohort vanishes as that generation exits the economy, thus preventing the accumulation of substantial wealth which would be the case in infinite-horizon models such as Krusell and Smith (1998).

Turning to the heterogeneity in the weight ϕ_j that households assign to their "warm-glow" bequest utility, the differences seem large in the context of CRRA utility. However, with EZW preferences, these values interact with the EIS and γ, so the effective difference for this parametrization is much smaller.

Figure 9 compares the resulting wealth moments from the model to those obtained from the SCF. The match is reasonably good, even though the model struggles to generate the wealth concentration at the very top. In the tenth decile, which is not shown as it makes the differences in the lower parts of the distribution hard to read, the model can generate only about 60% of wealth holdings observed in the data.

7 Results

Before reporting the results for the benchmark calibration with per-period participation costs, I first discuss the model *without* participation costs to illustrate to what extent

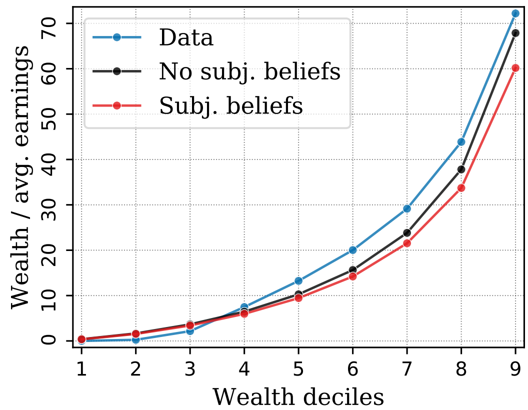


Figure 9: Wealth distribution in SCF vs. model by wealth decile. Each dot shows the mean wealth conditional on being in a given decile, normalized by average quarterly earnings. Data source: SCF 1998–2007.

subjective beliefs can generate limited participation on their own.[17]

7.1 No participation costs

Starting with the model without participation costs, Figure 10 shows the average simulated household portfolio allocations in the cross-section and Figure 11 plots the corresponding graphs along the life-cycle. While neither model comes close to the data moments, these figures illustrate the mechanism and the improvements due to subjective beliefs.

First, subjective beliefs generate limited participation in the risky asset *without* imposing any participation costs. In contrast, as is well-known, the participation rate conditional on saving a positive amount is 100% in the "standard" model. Introducing subjective beliefs pushes this value down to about 70% in the first wealth decile, and participation is upward-sloping along the wealth distribution. The conditional risky share is slightly tilted since poor households no longer choose to invest all their savings in stocks, while the wealthy increase their stock holdings. The model with subjective beliefs performs somewhat better than the rational-expectations variant, lowering the conditional risky

[17]In the appendix, section C additionally shows the effect of introducing subjective beliefs using an alternative calibration along the lines of Cocco, Gomes, and Maenhout (2005), which has become a benchmark paper in the household-finance literature. I argue that this parametrization generates wealth levels that are highly counterfactual and is thus not well-suited to investigate portfolio choices along the wealth distribution.

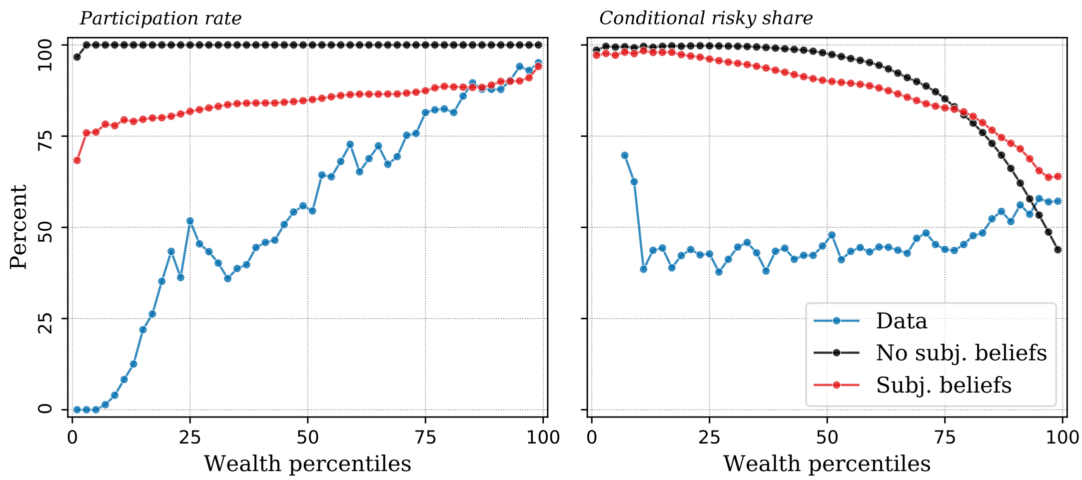


Figure 10: Portfolio composition along the wealth distribution. Calibration *without* participation costs.

share for households with fewer assets, and increasing it in the right-hand tail of the distribution.

Over the life-cycle, Figure 11 shows that the differences between the fully-rational and the subjective-beliefs models are again predominantly on the extensive margin, as younger households are less likely to enter the stock market. Compared to the rational-expectations model where almost all households choose to hold some risky assets, this figure drops to 65% for those aged 20–25.

7.2 Benchmark model with participation costs

Since subjective beliefs alone do not bring participation down to the levels observed in the data, a natural follow-up question is to what extent fixed participation costs improve the model fit, and how much the effects differ from a fully-rational model in that scenario. I therefore set κ, the fixed per-period participation cost that enters the household budget constraint in (8), to $\kappa = 0.01$, i.e. to 1% of average per-period earnings (this amounts to around \$500 annually).

Figure 12 illustrates that participation along the wealth distribution in the model with subjective beliefs improves considerably, and more so than in the standard model with the same participation costs. In the latter setting, the fixed cost has no effect beyond the second wealth decile, while it pushes participation down towards its empirical

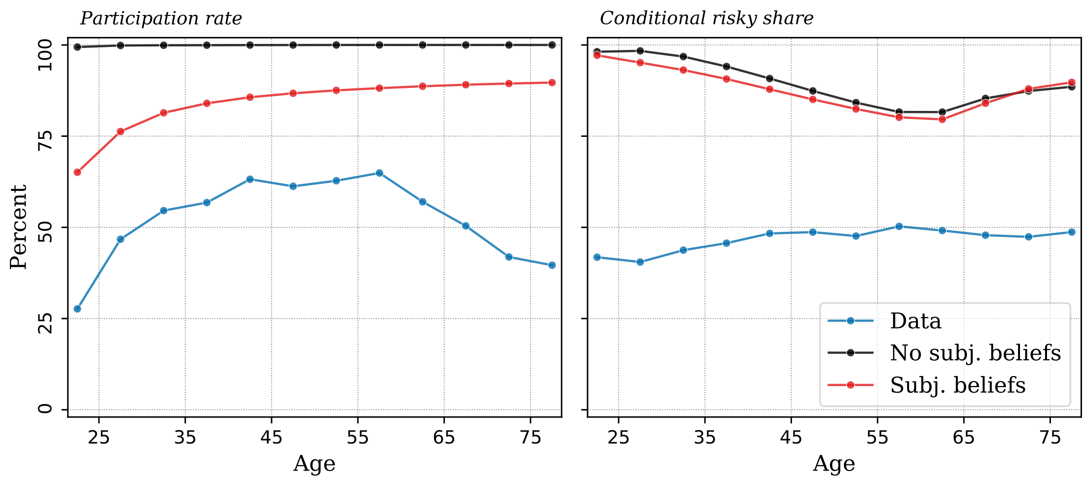


Figure 11: Portfolio composition over the life-cycle. Calibration *without* participation costs.

counterpart in the presence of subjective beliefs. The findings for the fully-rational model are in line with previous literature on participation in the presence of fixed costs: modest fixed participation costs only affect poor households with little financial wealth, but they cannot rationalize limited participation among middle-class households. For example, Vissing-Jorgensen (2003) reports that when attempting to match the participation rates in each wealth decile using heterogeneous participation costs, she finds that the required median cost ranges from $650–$1,450 (in 2003 USD), depending on the PSID wave. While not being directly comparable, Briggs et al. (2019) estimate that a one-time entry cost of more than $30,000 is needed in an otherwise standard life-cycle model to rationalize the (low) participation response among lottery winners. In comparison, in the presence of subjective beliefs, households who experienced low returns in the past expect low or even negative returns in the future, and thus small or even no participation costs are sufficient to exclude them from the stock market.

Whereas introducing participation costs on top of subjective beliefs substantially improves the fit with data compared to a standard model, both models produce qualitatively similar predictions when it comes to the conditional risky share. Between the 20th and 70th percentiles of the wealth distribution, the risky share is almost identical for the rational-expectations and subjective-belief models, while in the latter setting it is somewhat higher for the wealthiest households since these are on average more optimistic about stock returns.

A similar picture emerges over the life-cycle, shown in Figure 13: both the fully-rational

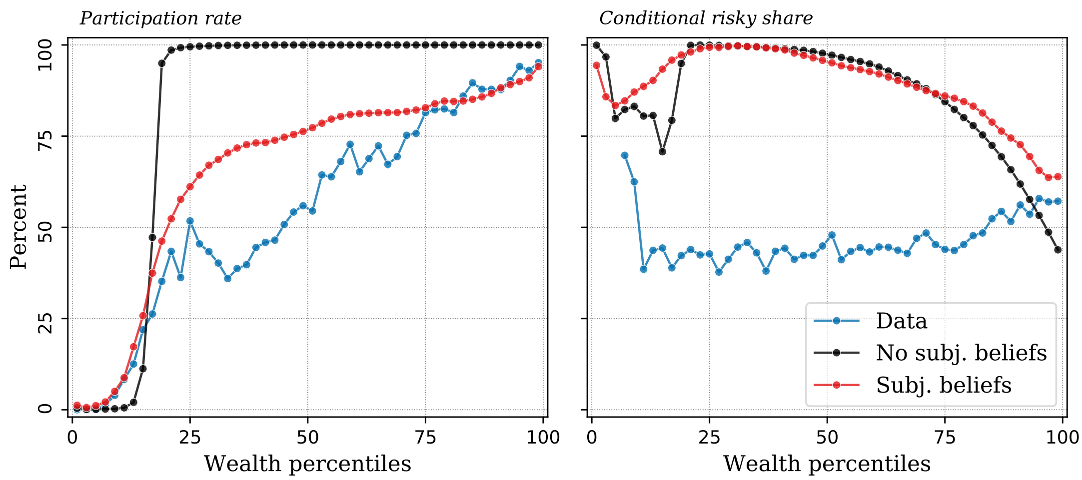


Figure 12: Portfolio composition along the wealth distribution. Benchmark calibration *with* participation costs.

and the subjective-beliefs model see improvements in the participation rates along the age dimension, but the latter model is considerably closer to the data. However, the conditional risky share is again very similar in both scenarios.

7.3 Mechanism: positive sorting over beliefs and wealth

The model with subjective beliefs helps explain the limited stock market participation because it induces positive sorting across beliefs and a household's position in the wealth distribution. Consider a poorer household: one reason why a household has less wealth in this model is that it experienced repeated low returns on its investment in the past. These low return realizations at the same time induced the household to revise its beliefs about average stock returns downward, and it consequently chooses to decrease the fraction of financial wealth invested in the risky asset, or stays out of the stock market altogether. The opposite holds for a wealthy household.

In the benchmark model, earnings uncertainty and life-cycle savings behavior also determine a household's position in the wealth distribution in addition to past returns, unlike in the stylized example in section 4. There might very well be *richer* households who are *more pessimistic* about risky returns because these households are wealthy due to high earnings, or due to being close to retirement when the life-cycle asset profile peaks, i.e. for reasons unrelated to their realized stock market returns. However, *on*

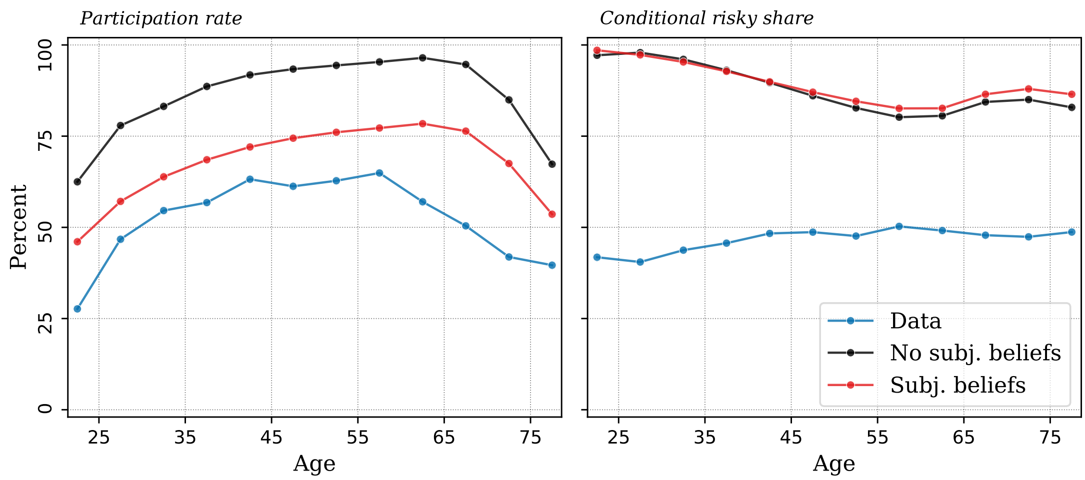


Figure 13: Portfolio composition over the life-cycle. Benchmark calibration *with* participation costs.

average, wealthier households are more optimistic about the stock market than poorer ones.

To make this point, Figure 14 plots the CDF over wealth for ages 20,. . . , 80 for selected percentiles of the subjective belief distribution.[18] At the age of 20, the sorting mechanism is not yet operational since newborns' initial wealth levels and beliefs are assumed to be uncorrelated. As households grow older, their position in the wealth distribution diverges conditional on their belief. By the age of 80, households that are in approximately the 75th percentile of the belief distribution are substantially wealthier than those less optimistic about excess returns: compared to a household with median beliefs, their CDF is uniformly shifted to the right.

While subjective beliefs due to learning from experience induce a positive correlation between beliefs and wealth in the population, for an individual household the shape of the risky share policy function is qualitatively the same as in the standard model: as shown in Figure 15, households diversify away from the risky asset as their cash-at-hand increases, thereby creating the usual downward-sloping optimal risky share. However, beliefs shift a household's optimal risky share up or down, with more optimistic households choosing higher risky shares, while pessimistic households (here illustrated by the 25th percentile of the belief distribution) choose to not invest in stocks at all. Even though beliefs have large effects on individual behavior, these approximately average

[18]Given that newborn households start with subjective beliefs which are centered on the true excess return in the cross-section, the median household shown in Figure 14 has the correct belief at any age.

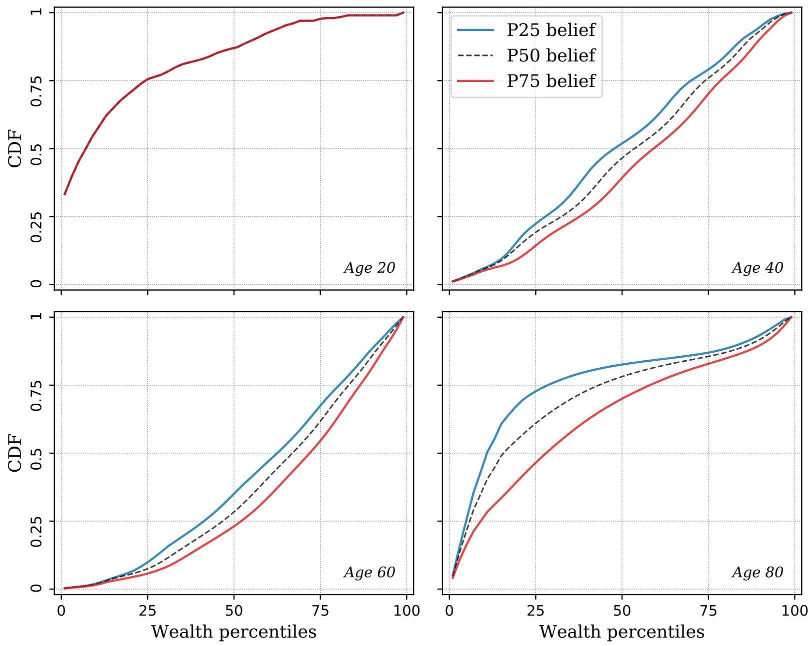


Figure 14: CDF over wealth for selected ages and percentiles of the subjective belief distribution. Each line shows the distribution of wealth conditional on age and beliefs. Wealth percentiles on the x-axis are computed for the overall population.

out in the aggregate, yielding an average conditional risky share that is not too different from the standard model.

I end this section with a more technical discussion on why the model with subjective beliefs creates limited participation even without imposing any participation costs. To simplify the exposition, I consider a one-period portfolio-choice problem with CRRA preferences and no participation costs, but the same reasoning carries over to the full model. Household i in this case maximizes

$$\max_{\xi} \ \mathbf{E}_i \left[\frac{(aR^p_{ih+1} + y_{ih+1})^{1-\gamma}}{1-\gamma} \right]$$
$$\text{s.t.} \quad R^p_{ih+1} = \xi \left(R_{ih+1} - R_f \right) + R_f \, , \qquad \xi \in [0,1]$$

where $a > 0$ is cash-at-hand, y_{ih+1} is some (stochastic) realization of non-financial income tomorrow and R^p_{ih+1} is tomorrow's return on the household's portfolio. Denoting the Lagrange multipliers on the constraints $\xi \geq 0$ and $\xi \leq 1$ by λ_0 and λ_1, respectively, the

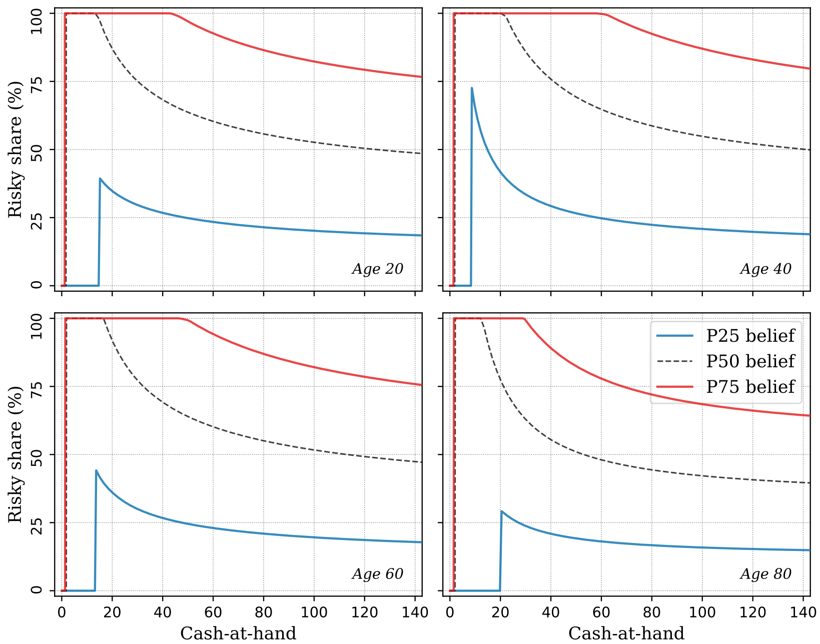


Figure 15: Household policy function for the optimal risky share. The red line shows the risky share adopted by an overly optimistic household (located at the 75th percentile of the belief distribution), while the dashed black line represents the choices of a household with median beliefs, and the blue line those of a pessimistic household at the 25th percentile.

first-order condition for the optimal risky share ξ is given by

$$\mathbf{E}_i\left[\left(aR^p_{ih+1}+y_{ih+1}\right)^{-\gamma}\left(R_{ih+1}-R_f\right)\right]+\lambda_0-\lambda_1=0$$

where $\lambda_0 \geq 0, \lambda_1 \geq 0$ and the usual complementary-slackness conditions apply.

Consider a household that chooses to not participate, i.e. $\xi = 0$ which implies that $\lambda_0 > 0$ and $\lambda_1 = 0$: assuming that y_{ih+1} is independent of R_{ih+1}, as is the case in this paper as well as in many others in the household-finance literature, the expectation can be split to obtain

$$\underbrace{\mathbf{E}_i\left[\left(aR_f+y_{ih+1}\right)^{-\gamma}\right]}_{\text{expected MU}}\times\underbrace{\mathbf{E}_i\left[R_{ih+1}-R_f\right]}_{\text{risk premium}}+\lambda_0=0 \tag{13}$$

where the first term represents the expected marginal utility when saving everything in the risk-free asset, and the second term is the expected excess return (or risk premium). Expected marginal utility is unambiguously positive, and a rational expectations model usually imposes that $\mathbf{E}_i\left[R_{ih+1}-R_f\right] > 0$ in line with historical data on stock-market performance. In that case, the condition in (13) cannot hold for $\xi = 0$ as all terms on the l.h.s. are positive, and non-participation is therefore sub-optimal. On the other hand, in a model with subjective beliefs and learning from experience, households who had repeated low return realizations might, in fact, expect $\mathbf{E}_i\left[R_{ih+1}-R_f\right] < 0$, and thus (13) is satisfied for $\xi = 0$ and some $\lambda_0 > 0$. These households will choose not to hold the risky asset even in the absence of participation costs. As discussed above, due to the positive sorting across beliefs and wealth, such households will one average be poorer. This gradient of the perceived risk premium along the distribution is illustrated in Figure 16 for the benchmark model with participation costs.

8 Bayesian learning from experience (BLE)

The benchmark calibration uses the belief updating method from Malmendier and Nagel (2011, 2016). A natural alternative is to assume that agents incorporate new data in an optimal way using Bayes' rule, thus efficiently weighting all past observations, but ignoring information prior to becoming economically active themselves. The only difference to a standard rational-expectations model is thus the restriction that households only learn from their *own* experience.

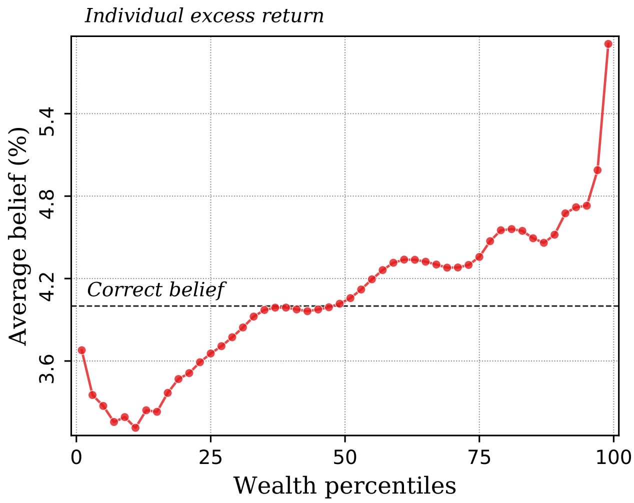


Figure 16: Average beliefs about individual excess returns along the wealth distribution. Each dot represents the average beliefs of households conditional on being in a given wealth percentile.

8.1 Belief updating

In the Bayesian framework, households at any age h have a prior belief about mean excess returns, which I assume to be Gaussian and given by

$$\mu \sim \mathcal{N}\left(\widehat{\mu}_{ih}, \tau_h \sigma^2\right) \tag{14}$$

The prior is centered around $\widehat{\mu}_{ih}$, and the uncertainty associated with the belief is expressed as a scaling factor τ_h applied to the known variance of risky returns σ^2. Within a cohort, there will be dispersion in households' mean belief $\widehat{\mu}_{ih}$ due to different realized return histories, while the variance $\tau_h \sigma^2$ is identical for all agents of the same age and is thus not indexed by i. This is a consequence of assuming that all newborns are assigned the same uncertainty about their beliefs τ_0 and update their beliefs each period, irrespective of whether they invest in stocks or not.

Newborns draw the initial beliefs about excess returns from a Gaussian distribution,

$$\widehat{\mu}_{i0} \overset{\text{iid}}{\sim} \mathcal{N}\left(\overline{\mu}, \sigma_0^2\right) \tag{15}$$

which is centered around the true excess return. The variance σ_0^2 is set to the same value as in the benchmark model. Thus, both economies start off with the same distribution of mean beliefs.

As new return realizations are observed, households update their beliefs as follows: both the prior and the excess return were assumed to be Gaussian to obtain the usual analytically tractable case in which the posterior distribution is Gaussian and the updating rule is given by

$$\begin{aligned}\widehat{\mu}_{ih} &= \left[\frac{\left(\tau_h\sigma^2\right)^{-1}}{\left(\sigma^2\right)^{-1}+\left(\tau_h\sigma^2\right)^{-1}}\right]\widehat{\mu}_{ih-1}+\left[\frac{\left(\sigma^2\right)^{-1}}{\left(\sigma^2\right)^{-1}+\left(\tau_h\sigma^2\right)^{-1}}\right]\left(R_{ih}-R_f\right)\\ &= \frac{1}{1+\tau_h}\widehat{\mu}_{ih-1}+\frac{\tau_h}{1+\tau_h}\left(R_{ih}-R_f\right)\end{aligned}$$

Bayesian learning from experience therefore leaves the household problem virtually unchanged, except that the update weight α_h in (4) is now defined as

$$\alpha_h = \frac{\tau_h}{1+\tau_h}$$

The variance of the posterior distribution is given by the standard formula

$$\tau_{h+1}\sigma^2 = \frac{1}{\left(\tau_h\sigma^2\right)^{-1}+\left(\sigma^2\right)^{-1}}$$

which implies that the scaling factor evolves according to the non-linear difference equation

$$\tau_{h+1}^{-1} = \tau_h^{-1}+1$$

Since $\lim_{h\to\infty}\tau_h = 0$, for infinitely-lived households their beliefs eventually collapse into a degenerate distribution around the true expected excess return, thus converging to the rational expectations equilibrium.

As illustrated in Figure 7, the weights assigned to past observations under the assumption of BLE differ from the setup in Malmendier and Nagel (2011, 2016). With BLE, the weight assigned to the most recent observation decreases more rapidly as households age, which implies that the *cross-sectional* distribution of mean beliefs clusters more tightly around the true excess return for older cohorts. This is illustrated in Figure 17 for selected ages: while the distributions are identical by construction at age 20 when newborns enter the economy, with Bayesian learning the cross-sectional dispersion collapses somewhat faster around the true expected excess return (which is $\approx 0.97\%$ in the quarterly calibration) as a cohort grows older. This, in turn, dampens any effect of subjective beliefs on the portfolio composition in the cross-section, moving the model closer to the rational-expectations framework.

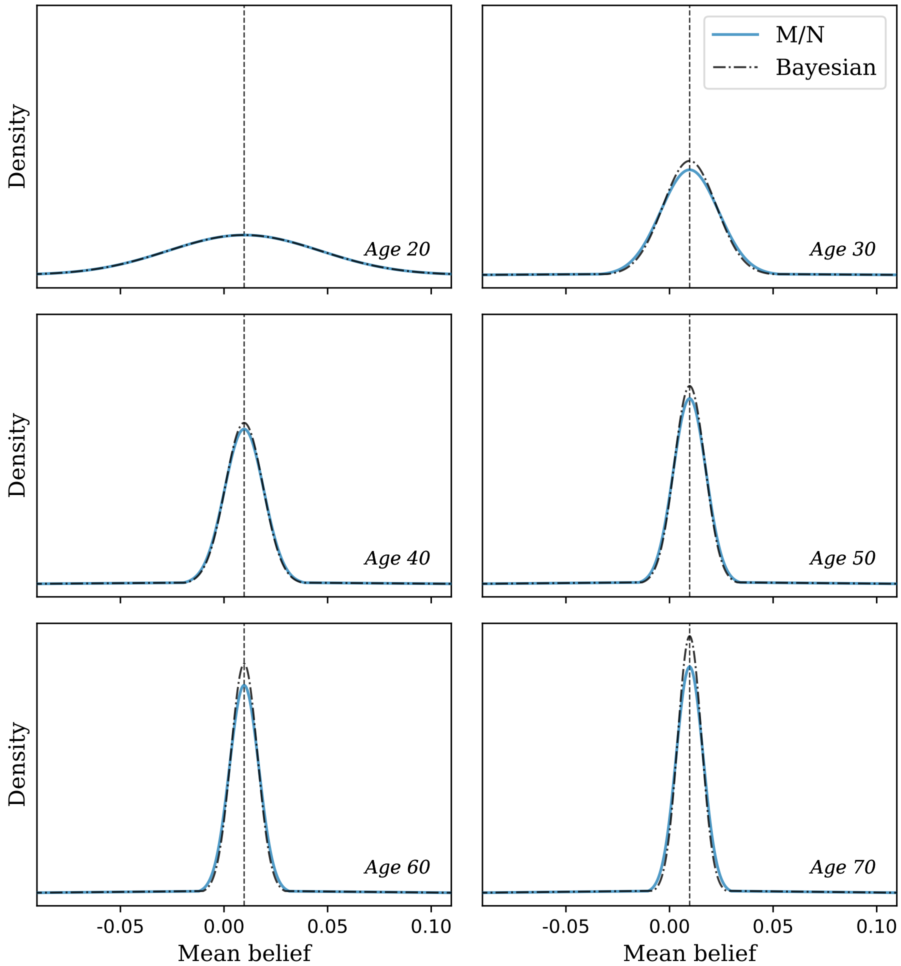


Figure 17: Cross-sectional distribution of mean beliefs at selected ages for Bayesian learning from experience and experience-based learning as in Malmendier and Nagel (2011, 2016).

8.2 Predictive distribution

Besides the faster convergence towards the true expected excess return, there is an additional conceptual difference as compared to the benchmark model: households factor in the uncertainty associated with their mean belief when evaluating the riskiness of their portfolio. For a household i of age h, the subjective expected risky return conditional on some belief μ is given by

$$\mathbf{E}_i\left[R_{ih+1} - R_f \,\middle|\, \mu\right] = \mathbf{E}_i\left[\mu + z_{ih+1} \,\middle|\, \mu\right] = \mu$$

and therefore, unconditionally, applying the law of iterated expectations, one obtains the same expression as in the Malmendier and Nagel (2011, 2016) setting, i.e.

$$\mathbf{E}_i\left[R_{ih+1} - R_f\right] = \mathbf{E}_i\left[\mathbf{E}_i\left[R_{ih+1} - R_f \,\middle|\, \mu\right]\right] = \mathbf{E}_i\mu = \widehat{\mu}_{ih}$$

This no longer holds for the variance, however. With Bayesian learning, the conditional variance is

$$\mathrm{Var}_i\left(R_{ih+1} - R_f \,\middle|\, \mu\right) = \mathrm{Var}_i\left(\mu + z_{ih+1} \,\middle|\, \mu\right) = \sigma^2$$

while the unconditional variance, obtained from the law of total variance, evaluates to

$$\begin{aligned}\mathrm{Var}_i\left(R_{ih+1} - R_f\right) &= \mathbf{E}_i\left[\mathrm{Var}_i\left(R_{ih+1} - R_f \,\middle|\, \mu\right)\right] + \mathrm{Var}_i\left(\mathbf{E}_i\left[R_{ih+1} - R_f \,\middle|\, \mu\right]\right) \\ &= \sigma^2 + \tau_h\sigma^2\end{aligned}$$

The difference as compared to experience-based learning emerges because in the latter case, the beliefs are not uncertain from a household's perspective and thus, they do not amplify the perceived variance of risky returns. Conversely, the Bayesian updater acts as if the risky return process was effectively defined as

$$R_{ih+1} - R_f = \widehat{\mu}_{ih} + \sqrt{1+\tau_h} \cdot z_{ih+1}\,, \qquad z_{ih+1} \overset{\text{iid}}{\sim} \mathcal{N}\left(0, \sigma^2\right)$$

instead of (3). In principle, this additional perceived variance of risky returns stemming from the uncertainty of beliefs could be used to generate a lower participation or risky shares, in particular for young households with a higher τ_h. However, as I discuss in the next section, the effect turns out to be negligible.

8.3 Calibration

I use the same calibration as for the benchmark model. In particular, as mentioned above, I set the variance of initial mean beliefs σ_0^2 in (15) to the same value as in the benchmark listed in Table 2.

The Bayesian learning model has one additional parameter τ_0 which controls newborn households' uncertainty about their beliefs. I set this parameter such that households at the age of 25 apply the same weight to the most recent realization as under experience-based learning, which is approximately 3.04% at a quarterly frequency. The initial scaling factor required to obtain the same value with Bayesian learning is $\tau_0 = 0.085$. Thus, newborns perceive the effective return variance to be only 8.5% higher than in the benchmark model, which is not sufficient to substantially affect their portfolio choices.

8.4 Results for the economy with Bayesian updating

The portfolio composition for the economy populated with Bayesian updaters is shown in Figure 18 and Figure 19 together with the benchmark model. Qualitatively, there are hardly any differences as compared to the benchmark model, neither across the wealth distribution, nor along the life-cycle. This result is hardly surprising given the preceding discussion, as the cross-sectional distributions of beliefs are close in both economies, and the perceived additional variance of risky returns due to belief uncertainty is modest. The results for the model *without* fixed participation costs (not shown) exhibit the same pattern, i.e. there are almost no differences between the two updating rules.

9 Correlated returns in the cross-section

In this section, I relax the assumption that returns are i.i.d. in the cross-section. An alternative model of household i's risky return is to assume a common component that arises because households invest a fraction of their portfolio into a market index. Denote by r_{mt}^e the excess return of such an index fund. Individual excess returns r_{it+1}^e in $t+1$

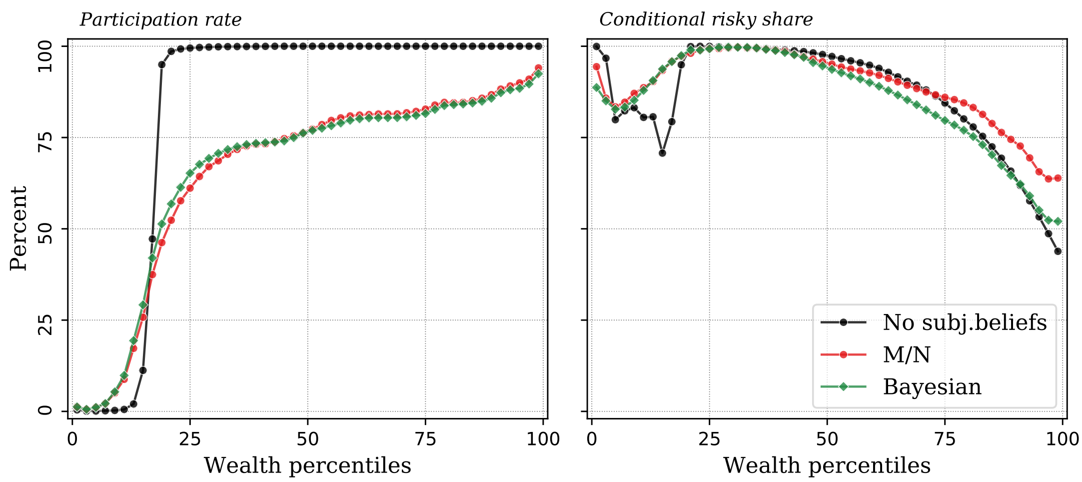


Figure 18: Portfolio composition along the wealth distribution. Bayesian learning from experience vs. benchmark.

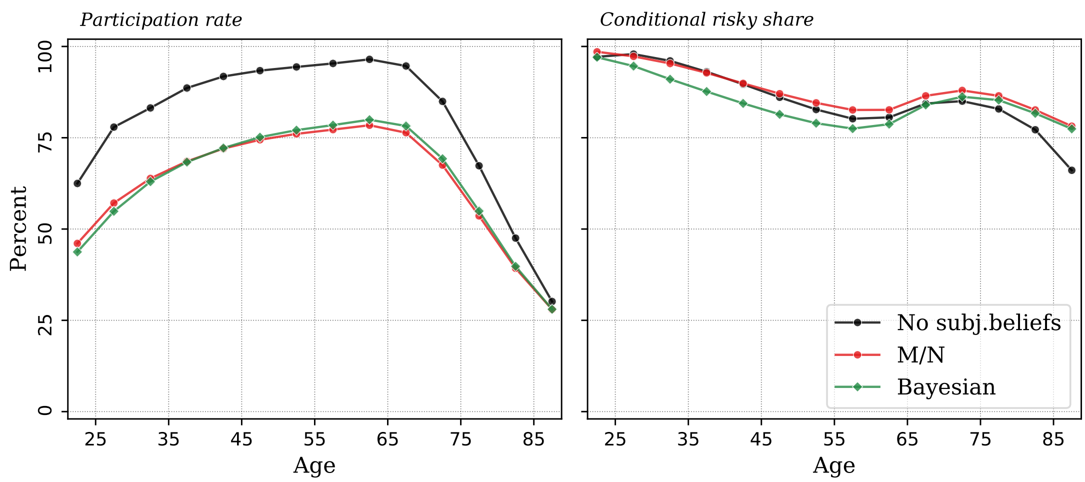


Figure 19: Portfolio composition over the life-cycle. Bayesian learning from experience vs. benchmark.

are then given by[19]

$$r^e_{it+1} = R_{it+1} - R_f = \beta_m r^e_{mt+1} + u_{it+1}$$

with

$$r^e_{mt+1} \overset{\text{iid}}{\sim} \mathcal{N}\left(\overline{\mu}^m, \sigma^2_m\right) \qquad u_{it+1} \overset{\text{iid}}{\sim} \mathcal{N}\left(\overline{\mu}^u, \sigma^2_u\right) \qquad \text{Corr}\left(r^e_{mt+1}, u_{it+1}\right) = 0$$

Here u_{it+1} is an uncorrelated idiosyncratic component that arises due to underdiversification. I assume that a household's portfolio choice is between investing in the risk-free asset and this composite risky asset, but households cannot choose the fraction of risky assets invested in the market index. The (objective) risk premium and the variance of this individual excess return are thus given by

$$\mathbf{E}\left[r^e_{it+1}\right] = \overline{\mu}^u + \beta_m \overline{\mu}^m \qquad \text{Var}\left(r^e_{it+1}\right) = \beta^2_m \sigma^2_m + \sigma^2_u$$

Furthermore, the time-t correlation between any two households' gross returns is

$$\text{Corr}\left(R_{it}, R_{jt}\right) = \frac{\beta^2_m \sigma^2_m}{\beta^2_m \sigma^2_m + \sigma^2_u}$$

I use the empirical counterparts of the above moments reported in Calvet, Campbell, and Sodini (2007) for Swedish data to pin down the additional parameters β_m, σ^2_m and σ^2_u, as described in the calibration section below.

Implementing this model poses two additional challenges: first, households can be uncertain about both $\overline{\mu}^m$ and $\overline{\mu}^u$ and thus have to form beliefs about both, which introduces an additional continuous state variable into the household problem. Second, the economy no longer has a time-invariant ergodic distribution over the household's state variables, but depends on the sequence of aggregate realizations $(r^e_{mt})_t$ which go back to the oldest cohort's date of birth.

In addition, this setting allows me to distinguish between public and private information. In the benchmark model it was assumed that households always observe the (potentially counterfactual) realization of their idiosyncratic return, regardless of whether they chose to invest in the risky asset or not. In this section, I instead assume that households always see the excess return on the market index r^e_{mt} which is public information. However, they observe the purely idiosyncratic return component u_{it} only if they choose to participate.

[19] Unlike in the previous sections where there was no need to distinguish between a household's age h and calendar time t, the notation in this section is adapted to reflect the fact that investors of different ages now receive the same market returns r^e_{mt} at time t.

This introduces a role for active learning, as households who pay the participation cost to invest in the stock market at the same time pay to acquire new information about $\overline{\mu}^u$ and can thus update their belief $\widehat{\mu}^u_{it}$.

9.1 Modified household problem

The household problem mostly remains the same as in the benchmark model of section 5, so I discuss the changes for the retired household only. Assuming that households are uncertain about the means of both the market and idiosyncratic returns, the expanded state vector is now given by $\boldsymbol{x} = \left(h, a, p, \widehat{\mu}^m_i, \widehat{\mu}^u_i, j\right)$ where $\widehat{\mu}^m_i$ is the belief about the market-return risk premium $\overline{\mu}^m$ and $\widehat{\mu}^u_i$ is the belief about $\overline{\mu}^u$. Retired households maximize

$$V^r_{jh}\left(a, p, \widehat{\mu}^m_i, \widehat{\mu}^u_i\right) = \max_{c,b,\xi} \left\{ c^{1-\psi} + \beta_j \left[\pi^s_h \mathbf{E}_i \left[\left(V^r_{jh+1}\left(a', p, (\widehat{\mu}^m_i)', (\widehat{\mu}^u_i)'\right) \right)^{1-\gamma} \right] + (1-\pi^s_h)\mathbf{E}_i \left[\left(V^b_j\left(a'_b\right) \right)^{1-\gamma} \right] \right]^{\frac{1-\psi}{1-\gamma}} \right\}^{\frac{1}{1-\psi}}$$

subject to the same constraints as before. The belief about the market excess return is updated according to

$$(\widehat{\mu}^m_i)' = \left(1-\alpha_{h+1}\right)\widehat{\mu}^m_i + \alpha_{h+1} r^e_{mt+1}$$

each period. On the other hand, $\widehat{\mu}^u_i$ is updated only if the household chooses a non-zero risky share ξ, i.e.

$$(\widehat{\mu}^u_i)' = \begin{cases} (1-\alpha_{h+1})\widehat{\mu}^u_i + \alpha_{h+1} u_{it+1} & \text{if } \xi > 0 \\ \widehat{\mu}^u_i & \text{else} \end{cases}$$

9.2 Calibration

Calvet, Campbell, and Sodini (2007) report that in Swedish register data on household portfolios $\beta_m \approx 0.87$, which they obtain by regressing individual portfolio returns onto the MSCI world index. Additionally, the average idiosyncratic variance share in the

	Description	Value	Source
Risky return			
$\overline{\mu}$	Risk premium	0.04	Cocco, Gomes, and Maenhout (2005)
σ	Volatility of risky return	0.16	Cocco, Gomes, and Maenhout (2005)
β_m	Market-β	0.87	Calvet, Campbell, and Sodini (2007)
–	Share of idiosyncratic variance	60%	Calvet, Campbell, and Sodini (2007)
Market return			
$\overline{\mu}^m$	Risk premium of market return	0.04	Cocco, Gomes, and Maenhout (2005)
σ_m	Volatility of market return	0.16	Cocco, Gomes, and Maenhout (2005)
Individual return			
$\overline{\mu}^u$	Risk premium of idiosyncratic return	$(1-\beta_m)\overline{\mu}$	–
σ_u	Volatility of idiosyncratic return	0.1705	–

Table 4: Parameters for the model with cross-sectionally correlated returns (annual). All remaining parameters are unchanged from the benchmark calibration.

Swedish data is reported to be

$$\frac{\sigma_u^2}{\beta_m^2\sigma_m^2+\sigma_u^2} \approx 60\% \;, \tag{16}$$

which implies that the cross-sectional correlation of returns is approximately 35%.[20]

To make the results as comparable to the benchmark model as possible, I choose the same overall risk premium and variance as reported in Table 2, i.e.

$$\mathbf{E}_t\left[r^e_{it+1}\right] = \overline{\mu} \qquad\qquad \mathrm{Var}_t\left(r^e_{it+1}\right) = \sigma^2 \tag{17}$$

I assume that $\overline{\mu}^m = \overline{\mu}$ and $\sigma_m = \sigma$ since the values for $(\overline{\mu}, \sigma)$ imposed here are used in the household-finance literature to capture the moments of a broad market index in the first place. This implies that $\overline{\mu}^u = (1-\beta_m)\overline{\mu}$ needs to be imposed for (17) to hold. Finally, given $\beta_m = 0.87$, equation (16) pins the remaining parameter to be $\sigma_u = 17.05\%$ annually. This calibration is summarized in Table 4.

While one could alternatively assume that individual portfolios have different return moments than the market index, the above assumptions ensure that in a partial-equilibrium setting, a rational-expectations investor is indifferent to which share of the risky portfolio is invested in the market index, as the individual return moments are unaffected.

[20] Calvet, Campbell, and Sodini (2007) point out that both β_m and the idiosyncratic variance share are quite heterogeneous in the population. Given the already large state space of the model presented in this section, I ignore this additional source of heterogeneity.

Lastly, in the model with market returns, the question arises which initial belief newborn investors should be assigned. Here I exactly follow the approach in Malmendier and Nagel (2011, 2016), who postulate that an individual of age $age_{it} = \underline{h} + h_{it}$ will form beliefs based on the returns observed in $t-1, \ldots t - \underline{h} - h_{it}$ going back to his or her year of birth. Thus, every newborn at time t has the same belief about market returns, as they all experienced the same history. As for the newborns' distribution of $\widehat{\mu}^{u}_{it}$, I adopt the same approach as in the benchmark model to generate an initial distribution that is on average unbiased and has the cross-sectional standard deviation previously reported in Table 2.

9.3 Results

To interpret the results presented below, it is worthwhile to understand how these are generated. Since the economy now has an aggregate state given by the market return history going back to the birth date of the oldest cohort, every statistic used to characterize the portfolio composition in the cross-section or along the life-cycle depends on a particular realization of this market return time series. I address this issue by "bootstrapping" $N = 250$ such market return time series and computing the implied portfolio composition moments for each particular history.

Figure 20 shows the distribution of portfolio allocations along the wealth distribution from this bootstrapping exercise. Each dot represents the average of the 250 simulated economies for a given wealth bin, while the darker shaded areas represent the interquartile range, and the lighter shaded areas indicate the range which brackets 95% of the simulated economies.

This graph should be compared to the benchmark economy without aggregate market returns that was previously shown in Figure 12. Clearly the average of all simulated economies aligns well with both the participation rates and the conditional shares in that figure and thus, the findings from the benchmark model with i.i.d. returns continue to hold. Figure 21 shows the portfolio composition over the life-cycle and should be contrasted with Figure 13 for the benchmark model. Again, the mean over all simulations is quite close the the portfolio composition of the benchmark economy.

Both graphs show that the bootstrapped moments of the model without subjective beliefs (shown in black) exhibit substantially less variation. In that economy, market

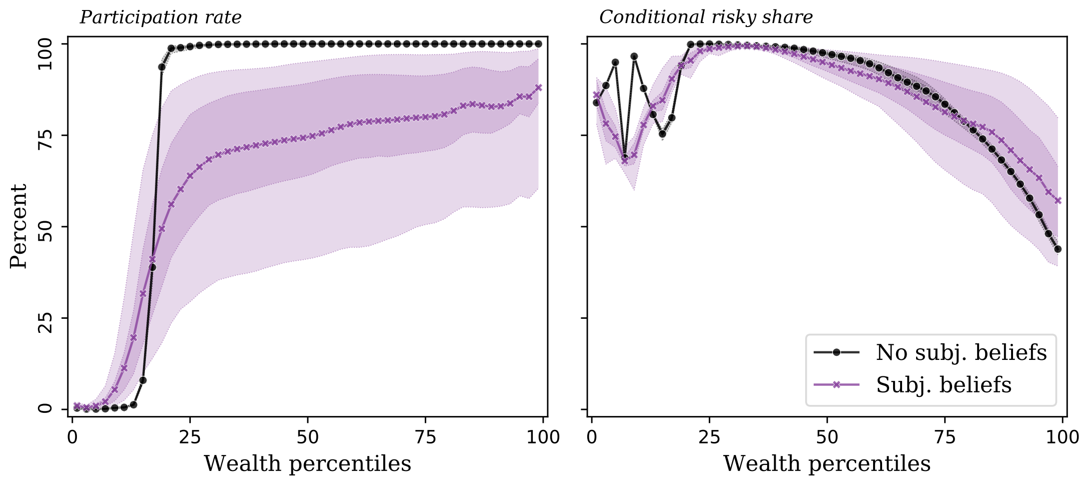


Figure 20: Portfolio composition along the wealth distribution. Model *with* participation costs and aggregate market returns. Dots indicate bin averages over simulated economies. Dark shaded areas show the interquartile range of simulated moments, while light shaded areas represent the range containing 95% of simulations.

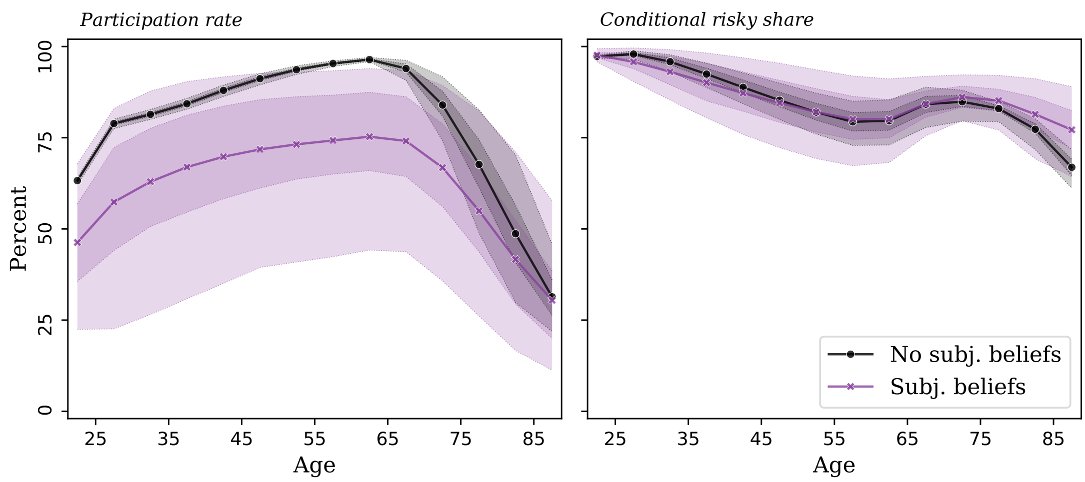


Figure 21: Portfolio composition over the life-cycle. Model *with* participation costs and aggregate market returns. Dots indicate bin averages over simulated economies. Dark shaded areas show the interquartile range of simulated moments, while light shaded areas represent the range containing 95% of simulations.

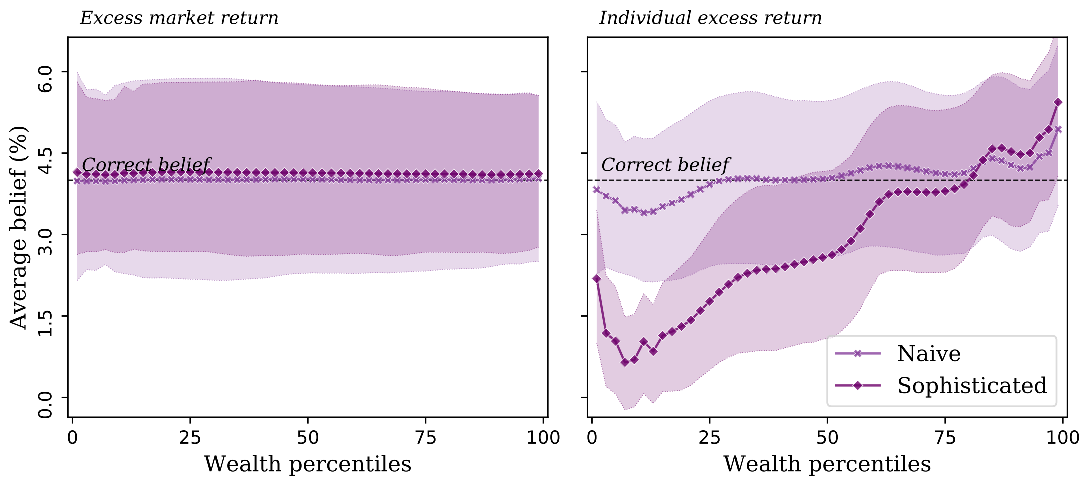


Figure 22: Beliefs about the excess market return r^e_{mt} and the individual excess return r^e_{it} for naive vs. sophisticated households in the model *with* participation costs. Dots indicate bin averages over simulated economies. Shaded areas show the interquartile range of simulated moments.

return realizations (which are assumed to be i.i.d. over time) have almost no persistent effects other than on wealth accumulation, while in the model with subjective beliefs, they affect cohort-specific beliefs about market returns for a prolonged period of time.

In Figure 22, I report the analogue of Figure 16 for the economy with aggregate market returns. The beliefs about individual mean excess returns shown in Figure 22 are computed as $\widehat{\mu}_{it} = \beta_m \widehat{\mu}^m_{it} + \widehat{\mu}^u_{it}$. The graph illustrates that just like in the benchmark model, there is a positive sorting across beliefs about *individual* returns and wealth (right-hand panel); however, on average no such sorting exists with respect to beliefs about *market* returns (left-hand panel). This is due to the fact that conditional on calendar year and age, all households hold the same beliefs about the market return irrespective of their wealth.

The right-hand panel of Figure 22 also highlights that in the setting discussed in this section, the differences between economies populated by naive vs. sophisticated agents are more pronounced. Conversely, in the benchmark model, where households update their beliefs even if they do not participate, naive and sophisticated households hold almost identical beliefs, as shown in Figure 39 in the appendix.

If belief updating is conditional on participation, and participation is costly due to per-period participation costs, the beliefs held by naive and sophisticated agents diverge.

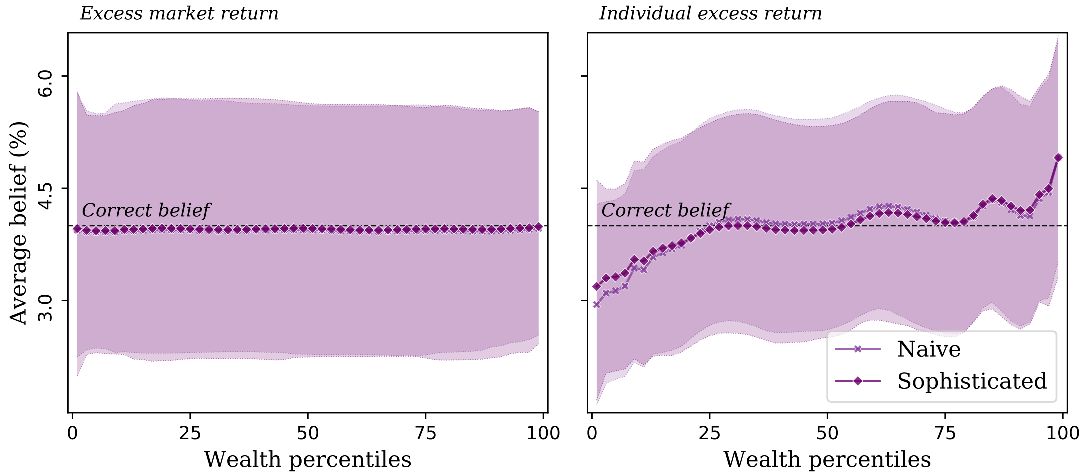


Figure 23: Beliefs about the excess market return r^e_{mt} and the individual excess return r^e_{it} for naive vs. sophisticated households in the model *without* participation costs. Dots indicate bin averages over simulated economies. Shaded areas show the interquartile range of simulated moments.

The reason is the following: sophisticated households are in a sense more risk-averse because they anticipate that in case of a low return realization tomorrow, they will adjust their beliefs about excess returns downward. Thus, they expect to earn a lower return on the investment choices they make tomorrow. Given that the EIS is assumed to be lower than one, they anticipate to *increase* their savings tomorrow when a bad shock hits, thereby lowering consumption. Low returns tomorrow consequently coincide with higher marginal utility tomorrow, which tilts the risky share chosen *today* downward (this is evident from the policy functions shown in Figure 40 in the appendix for the benchmark model). In the presence of participation costs, households who would choose a lower risky share conditional on participating have a higher incentive not to hold the risky asset at all. This effect is even more pronounced for households who are already more pessimistic about (idiosyncratic) returns, and since they choose not to participate in this case, they never update their pessimistic beliefs about idiosyncratic returns. Via this mechanism, sophisticated households are more likely to end up in non-participation as an absorbing state. Figure 23 shows that costly participation is a key for this mechanism to be operational: without it, there is almost no difference between the cross-sectional beliefs in an economy with naive vs. one with sophisticated agents.

Lastly, Figure 24 and Figure 25 contrast the portfolio composition across the wealth distribution and over the life-cycle for naive vs. sophisticated agents. These graphs

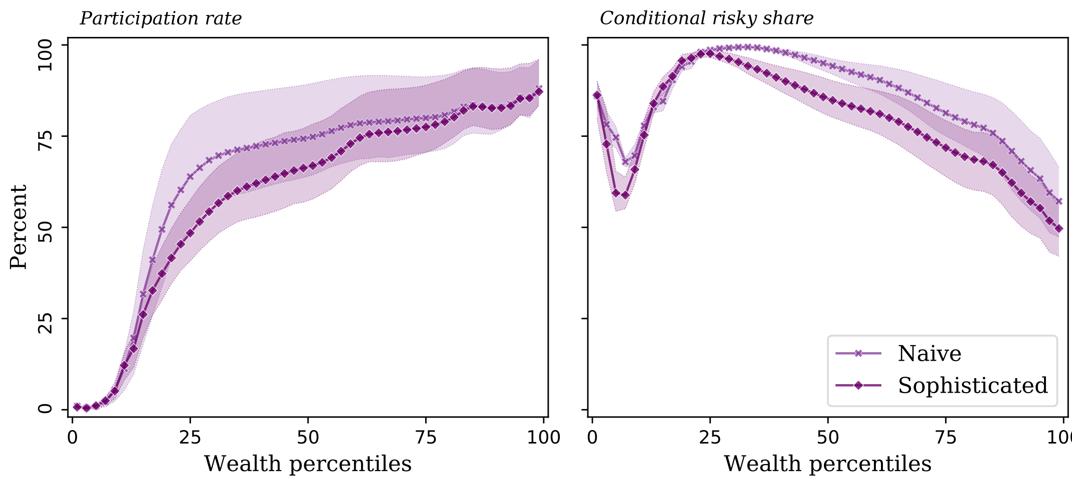


Figure 24: Portfolio composition along the wealth distribution. Naive vs. sophisticated households. Model *with* participation costs and aggregate market returns. Dots indicate bin averages over simulated economies. Shaded areas show the interquartile range of simulated moments.

should be compared to those for the benchmark economy shown in Figure 37 and Figure 38 in the appendix. For the reasons discussed above, the differences in the present model compared to the benchmark are now more substantial, also among middle-class and younger households.

10 Summary and conclusion

In this paper, I propose a mechanism to help explain the empirically observed heterogeneity of households' financial portfolios along the wealth distribution, in particular the almost monotonically increasing stock market participation rate. To this end, I incorporate empirical evidence on how households form beliefs based on past realizations of asset returns into an otherwise standard household-finance model. Unlike in a fully rational model, households overweight past observations of returns on their idiosyncratic portfolios when forming beliefs about excess returns, thus generating a belief dispersion in the cross-section that does not fully disappear over time – in this setting, households do not necessarily learn the true excess return in their lifetime.

Combined with underdiversification and thus idiosyncratic return histories, experience-based learning of this kind introduces heterogeneity in households' participation and portfolio allocation choices in a way that moves the model's predictions closer to the

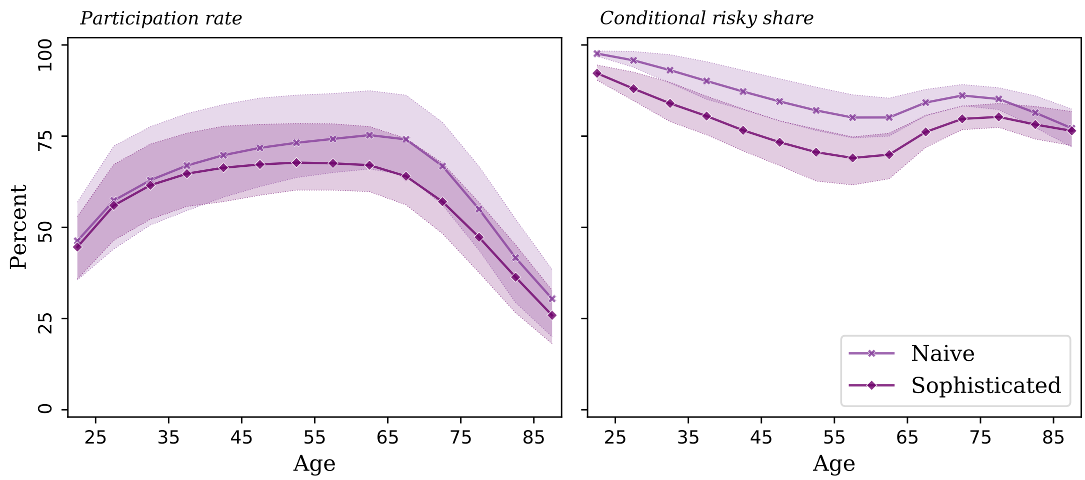


Figure 25: Portfolio composition over the life-cycle. Naive vs. sophisticated households. Model *with* participation costs and aggregate market returns. Dots indicate bin averages over simulated economies. Shaded areas show the interquartile range of simulated moments.

data. As above-average return realizations simultaneously move a household upwards in the wealth distribution and improve the household's outlook on future returns, this mechanism induces positive sorting in the joint distribution of wealth and beliefs: on average, the most wealthy households end up being the most optimistic and, consequently, they increase the fraction held in risky assets compared to a fully-rational agent. The opposite holds for households at the other end of the wealth distribution, at least to the extent that they are poor due to low return realizations in the past, and the participation rate within this group drops as compared to a standard model.

It is a quantitative question of how much of this intuitively appealing mechanism survives in a full life-cycle model where additional factors affect a household's wealth but not its beliefs about stock returns. With a highly-persistent earnings uncertainty and a hump-shaped wealth profile over the life-cycle, asset returns are neither the only nor the most dominant driver of a household's position in the wealth distribution. The wealthy in this framework are rich because they were lucky in how their investment paid off, or because they had repeated high earnings draws – the latter, however, have no effect on their beliefs about excess returns, and hence on their portfolio choice. Additionally, in a life-cycle framework, the newborns enter the economy without a history of past returns correlated with their position in the wealth distribution, further muting the effect of subjective beliefs.

The preceding sections have shown that subjective beliefs and experience-based learning are able to generate limited participation even in such a setting. Adding a reasonably small participation cost further improves the model fit with the data, above and beyond what a standard model with the same participation cost can achieve, in particular when it comes to the limited participation observed among middle-class households. This indicates that the mechanism presented here can potentially interact with other building blocks to move the model prediction even closer to the data. Exploring such extensions is left to future research.

References

Ameriks, John, Gábor Kézdi, Minjoon Lee, and Matthew D. Shapiro. 2019. Heterogeneity in Expectations, Risk Tolerance, and Household Stock Shares: The Attenuation Puzzle. *Journal of Business & Economic Statistics* 0 (0): 1–27.

Arias, Elizabeth, and Jianquan Xu. 2018. United States life tables, 2015. *National vital statistics reports,* Working Paper Series, 67 (2).

Arrondel, Luc, Calvo-Pardo Hector, and Derya Tas. 2014. *Subjective return expectations, information and stock market participation: evidence from France.* Discussion Paper Series In Economics And Econometrics. University of Southampton, Economics Division, School of Social Sciences.

Bach, Laurent, Laurent E. Calvet, and Paolo Sodini. 2018. *Rich Pickings? Risk, Return, and Skill in the Portfolios of the Wealthy.* Technical report.

Benhabib, Jess, Alberto Bisin, and Mi Luo. 2019. Wealth Distribution and Social Mobility in the US: A Quantitative Approach. *American Economic Review* 109 (5): 1623–1647.

Blanchard, Olivier J. 1985. Debt, Deficits, and Finite Horizons. *Journal of Political Economy* 93 (2): 223–247.

Bommier, Antoine, Daniel Harenberg, and François Le Grand. 2017. Recursive Preferences, the Value of Life, and Household Finance.

Briggs, Joseph S., David Cesarini, Erik Lindqvist, and Robert Östling. 2019. *Windfall Gains and Stock Market Participation.* Technical report. National Bureau of Economic Research, Inc.

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

Campanale, Claudio, Carolina Fugazza, and Francisco Gomes. 2015. Life-cycle portfolio choice with liquid and illiquid financial assets. *Journal of Monetary Economics* 71:67–83.

Campbell, John Y. 2006. Household Finance. *The Journal of Finance* 61 (4): 1553–1604.

Catherine, Sylvain. 2019. *Countercyclical Labor Income Risk and Portfolio Choices over the Life-Cycle.*

Chang, Yongsung, Jay H. Hong, and Marios Karabarbounis. 2018. Labor Market Uncertainty and Portfolio Choice Puzzles. *American Economic Journal: Macroeconomics* 10 (2): 222–262.

Choi, James J., David Laibson, Brigitte C. Madrian, and Andrew Metrick. 2009. Reinforcement Learning and Savings Behavior. *Journal of Finance* 64 (6): 2515–2534.

Cocco, João F. 2004. Portfolio Choice in the Presence of Housing. *The Review of Financial Studies* 18 (2): 535–567.

Cocco, João F., Francisco J. Gomes, and Pascal J. Maenhout. 2005. Consumption and Portfolio Choice over the Life Cycle. *Review of Financial Studies* 18 (2): 491–533.

Collin-Dufresne, Pierre, Michael Johannes, and Lars A. Lochstoer. 2016a. Asset Pricing When 'This Time Is Different'. *The Review of Financial Studies* 30 (2): 505–535.

———. 2016b. Parameter Learning in General Equilibrium: The Asset Pricing Implications. *American Economic Review* 106 (3): 664–698.

De Nardi, Mariacristina. 2004. Wealth Inequality and Intergenerational Links. *Review of Economic Studies* 71 (3): 743–768.

Dominitz, Jeff, and Charles F. Manski. 2007. Expected Equity Returns and Portfolio Choice: Evidence from the Health and Retirement Study. *Journal of the European Economic Association* 5 (2-3): 369–379.

———. 2011. Measuring and interpreting expectations of equity returns. *Journal of Applied Econometrics* 26 (3): 352–370.

Druedahl, Jeppe, and Thomas Høgholm Jørgensen. 2017. A general endogenous grid method for multi-dimensional models with non-convexities and constraints. *Journal of Economic Dynamics and Control* 74:87–107.

Ehling, Paul, Alessandro Graniero, and Christian Heyerdahl-Larsen. 2017. Asset Prices and Portfolio Choice with Learning from Experience. *The Review of Economic Studies* 85 (3): 1752–1780.

Epstein, Larry G. 1988. Risk aversion and asset prices. *Journal of Monetary Economics* 22 (2): 179–192.

Epstein, Larry G., and Stanley E. Zin. 1989. Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. *Econometrica* 57 (4): 937–969.

Fagereng, Andreas, Charles Gottlieb, and Luigi Guiso. 2017. Asset Market Participation and Portfolio Choice over the Life-Cycle. *Journal of Finance* 72 (2): 705–750.

Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri. 2019. *Heterogeneity and Persistence in Returns to Wealth.* Technical report. National Bureau of Economic Research, Inc.

Foltyn, Richard, and Jonna Olsson. 2019. *Subjective life expectancies, time preference heterogeneity and wealth inequality.*

Gaudecker, Hans-Martin von, and Axel Wogrolly. 2019. The dynamics of households' stock market beliefs.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus. 2019. *Five Facts About Beliefs and Portfolios.* Working Paper, Working Paper Series 25744. National Bureau of Economic Research.

Gomes, Francisco, and Alexander Michaelides. 2003. Portfolio choice with internal habit formation: a life-cycle model with uninsurable labor income risk. Finance and the Macroeconomy, *Review of Economic Dynamics* 6 (4): 729–766.

———. 2005. Optimal Life-Cycle Asset Allocation: Understanding the Empirical Evidence. *Journal of Finance* 60 (2): 869–904.

Greenwood, Robin, and Andrei Shleifer. 2014. Expectations of Returns and Expected Returns. *The Review of Financial Studies* 27 (3): 714–746.

Guiso, Luigi, and Paolo Sodini. 2013. Household Finance: An Emerging Field. Chap. 21, edited by George M. Constantinides, Milton Harris, and Rene M. Stulz, 2:1397–1532. Handbook of the Economics of Finance. Elsevier.

Guvenen, Fatih, Fatih Karahan, Serdar Ozkan, and Jae Song. 2016. *What Do Data on Millions of U.S. Workers Reveal about Life-Cycle Earnings Risk?* Technical report.

Guvenen, Fatih, Serdar Ozkan, and Jae Song. 2014. The Nature of Countercyclical Income Risk. *Journal of Political Economy* 122 (3): 621–660.

Haliassos, Michael, and Carol C. Bertaut. 1995. Why do so Few Hold Stocks? *The Economic Journal* 105 (432): 1110–1129.

Heiss, Florian, Michael Hurd, Maarten van Rooij, Tobias Rossmann, and Joachim K Winter. 2019. Dynamics and heterogeneity of subjective stock market expectations.

Hendricks, Lutz. 2007. How important is discount rate heterogeneity for wealth inequality? *Journal of Economic Dynamics and Control* 31 (9): 3042–3068.

Hubmer, Joachim, Per Krusell, and Anthony A. Jr. Smith. 2019. *Sources of U.S. Wealth Inequality: Past, Present, and Future.* Technical report.

Hurd, Michael D, and Susann Rohwedder. 2012. *Stock Price Expectations and Stock Trading.* Working Paper, Working Paper Series 17973. National Bureau of Economic Research.

Iskhakov, Fedor, Thomas H. Jørgensen, John Rust, and Bertel Schjerning. 2017. The endogenous grid method for discrete-continuous dynamic choice models with (or without) taste shocks. *Quantitative Economics* 8 (2): 317–365.

Kaustia, Markku, and Samuli Knüpfer. 2008. Do Investors Overweight Personal Experience? Evidence from IPO Subscriptions. *Journal of Finance* 63 (6): 2679–2702.

Kézdi, Gábor, and Robert J Willis. 2011. *Household Stock Market Beliefs and Learning.* Working Paper, Working Paper Series 17614. National Bureau of Economic Research.

Kozlowski, Julian, Laura Veldkamp, and Venky Venkateswaran. 2015. *The Tail that Wags the Economy: Beliefs and Persistent Stagnation.* Working Paper, Working Paper Series 21719. National Bureau of Economic Research.

Krueger, Dirk, Kurt Mitman, and Fabrizzio Perri. 2016. Macroeconomics and Household Heterogeneity, edited by John B. Taylor and Harald Uhlig, 2:843–921. Handbook of Macroeconomics. Elsevier.

Krusell, Per, and Anthony A. Smith. 1997. Income And Wealth Heterogeneity, Portfolio Choice, And Equilibrium Asset Returns. *Macroeconomic Dynamics* 1 (02): 387–422.

———. 1998. Income and Wealth Heterogeneity in the Macroeconomy. *Journal of Political Economy* 106 (5): 867–896.

Kuhn, Moritz, Moritz Schularick, and Ulrike I Steins. 2019. *Income and Wealth Inequality in America.*

Malmendier, Ulrike, and Stefan Nagel. 2011. Depression Babies: Do Macroeconomic Experiences Affect Risk Taking? *The Quarterly Journal of Economics* 126 (1): 373–416.

———. 2016. Learning from Inflation Experiences. *The Quarterly Journal of Economics* 131 (1): 53–87.

Malmendier, Ulrike, Demian Pouzo, and Victoria Vanasco. 2018. *Investor Experiences and Financial Market Dynamics.* NBER Working Papers 24697. National Bureau of Economic Research, Inc.

Mankiw, N.Gregory. 1986. The equity premium and the concentration of aggregate shocks. *Journal of Financial Economics* 17 (1): 211–219.

Merton, Robert C. 1971. Optimum Consumption and Portfolio Rules in a Continuous-Time Model. *Journal of Economic Theory* 3 (4): 373–413.

Meyer, Steffen, and Michaela Pagel. 2019. *Fully Closed: Individual Responses to Realized Gains and Losses.* Working Paper, Working Paper Series 25542. National Bureau of Economic Research.

Nagel, Stefan, and Zhengyang Xu. 2019. Asset Pricing with Fading Memory.

Nakov, Anton, and Galo Nuño. 2015. Learning from experience in the stock market. *Journal of Economic Dynamics and Control* 52:224–239.

Poterba, James M., and Andrew A. Samwick. 1995. Stock Ownership Patterns, Stock Market Fluctuations, and Consumption. *Brookings Papers on Economic Activity* 1995 (2).

Schraeder, Stefanie. 2015. Information Processing and Non-Bayesian Learning in Financial Markets. *Review of Finance* 20 (2): 823–853.

Storesletten, Kjetil, Chris Telmer, and Amir Yaron. 2007. Asset Pricing with Idiosyncratic Risk and Overlapping Generations. *Review of Economic Dynamics* 10 (4): 519–548.

Vestman, Roine. 2018. Limited Stock Market Participation Among Renters and Homeowners. *Review of Financial Studies* 32 (4): 1494–1535.

Vissing-Jorgensen, Annette. 2003. Perspectives on Behavioral Finance: Does "Irrationality" Disappear with Wealth? Evidence from Expectations and Actions. *NBER Macroeconomics Annual* 18:139–194.

Wachter, Jessica A., and Motohiro Yogo. 2010. Why Do Household Portfolio Shares Rise in Wealth? *The Review of Financial Studies* 23 (11): 3929–3965.

Weil, Philippe. 1990. Nonexpected Utility in Macroeconomics. *The Quarterly Journal of Economics* 105 (1): 29–42.

Yao, Rui, and Harold H. Zhang. 2004. Optimal Consumption and Portfolio Choices with Risky Housing and Borrowing Constraints. *The Review of Financial Studies* 18 (1): 197–239.

A Additional evidence from the Survey of Consumer Finances

A.1 Detailed household balance sheets

	Avg. (in $)	Median (in $)	Participation (in %)
Checking accounts	4,769	1,211	82.1
Savings accounts	10,243	105	55.0
Life insurance	9,131	0	26.0
Mutual funds (safe)	7,454	0	5.8
Retirement accounts (safe)	31,084	0	43.0
Total safe assets	98,638	14,149	92.9
Stocks	41,262	0	20.0
Mutual funds (risky)	20,749	0	14.2
Retirement accounts	44,957	0	46.2
Total risky assets	128,682	1,514	54.8
Total financial assets	227,320	24,631	93.1
Consumer debt	3,000	0	47.4
Education loans	2,633	0	13.7
Home ownership	–	–	69.4
Housing wealth (owners)	344,354	181,109	100.0
Mortgages (owners)	106,355	59,057	70.2
Net housing wealth (owners)	237,999	101,346	100.0
Actively managed businesses	89,669	0	11.3
Passively-held businesses	10,439	0	1.4
Total gross wealth	555,657	162,606	94.8
Net worth	477,458	93,742	–

Table 5: Summary statistics for disaggregated households' balance sheets. Housing wealth and mortgage statistics are reported for the sub-sample of homeowners. Data source: SCF 1998–2007.

	Avg. (in $)	Median (in $)	Participation (in %)
Checking accounts	5,122	1,333	88.2
Savings accounts	11,002	263	59.1
Life insurance	9,807	0	27.9
Mutual funds (safe)	8,006	0	6.2
Retirement accounts (safe)	33,376	0	46.2
Total safe assets	105,940	17,697	99.8
Stocks	44,316	0	21.5
Mutual funds (risky)	22,285	0	15.2
Retirement accounts	48,272	0	49.6
Total risky assets	138,208	3,356	58.8
Total financial assets	244,148	31,254	100.0
Consumer debt	3,196	0	49.8
Education loans	2,781	0	14.1
Home ownership	–	–	72.8
Housing wealth (owners)	350,836	182,850	100.0
Mortgages (owners)	108,341	61,060	70.7
Net housing wealth (owners)	242,495	103,491	100.0
Actively managed businesses	96,255	0	12.0
Passively-held businesses	11,212	0	1.5
Total gross wealth	595,420	186,239	100.0
Net worth	511,933	111,811	–

Table 6: Summary statistics for disaggregated households' balance sheets for sub-sample with *positive* financial wealth. Housing wealth and mortgage statistics are reported for the sub-sample of homeowners. Data source: SCF 1998–2007.

A.2 Portfolio composition across the wealth distribution

In this section, I present additional evidence on the composition of household portfolios observed in the SCF. In Figure 26, I plot net total wealth (net worth) on the x-axis, which includes all wealth components (including housing) after deducting all household liabilities (including mortgages). Figure 27 shows the portfolio allocation along percentiles of gross financial wealth, while Figure 28 reports portfolio choices over the distribution of financial wealth net credit-card debt and consumer loans. Last, Figure 29 plots those moments after additionally subtracting education loans from safe financial assets.

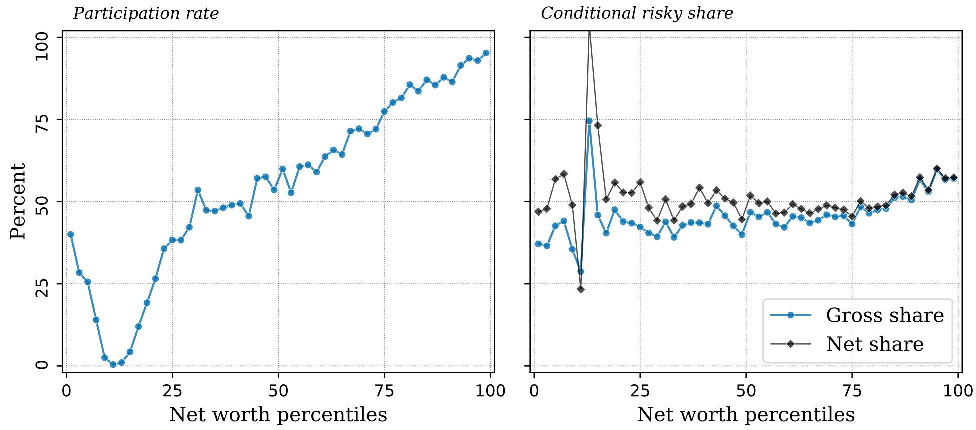


Figure 26: Portfolio composition along percentiles of *net worth*. Each dot represents two percent of households. Data source: SCF 1998–2007.

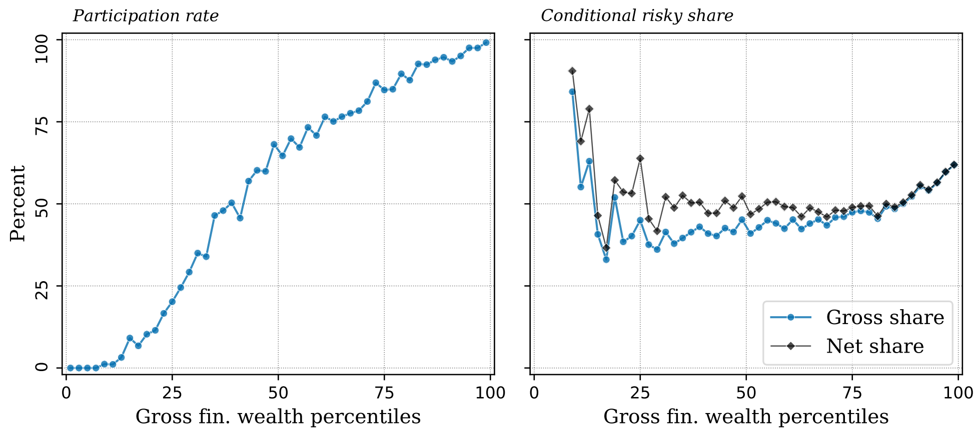


Figure 27: Portfolio composition along percentiles of *gross financial wealth*. Each dot represents two percent of households. Data source: SCF 1998–2007.

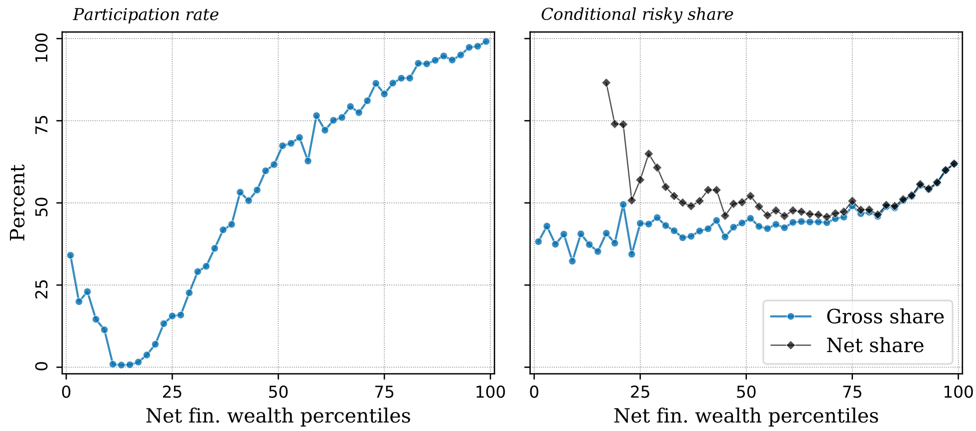


Figure 28: Portfolio composition along percentiles of *financial wealth net of consumer debt*. Each dot represents two percent of households. Data source: SCF 1998–2007.

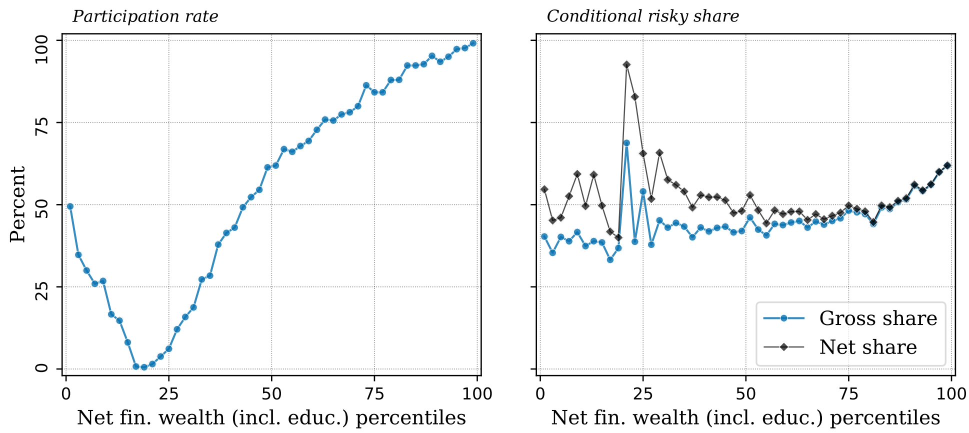


Figure 29: Portfolio composition along percentiles of *financial wealth net of consumer debt and education loans*. Each dot represents two percent of households. Data source: SCF 1998–2007.

A.3 Disaggregation by homeownership status

This section contains additional graphs disaggregating household portfolio allocations by homeownership status. In Figure 30, I show that financial portfolios are comparable conditioning on net worth. Figure 31 reports portfolio choices by gross financial wealth, while Figure 32 and Figure 33 illustrate that the differences are even smaller when plotting against the deciles of net financial wealth.

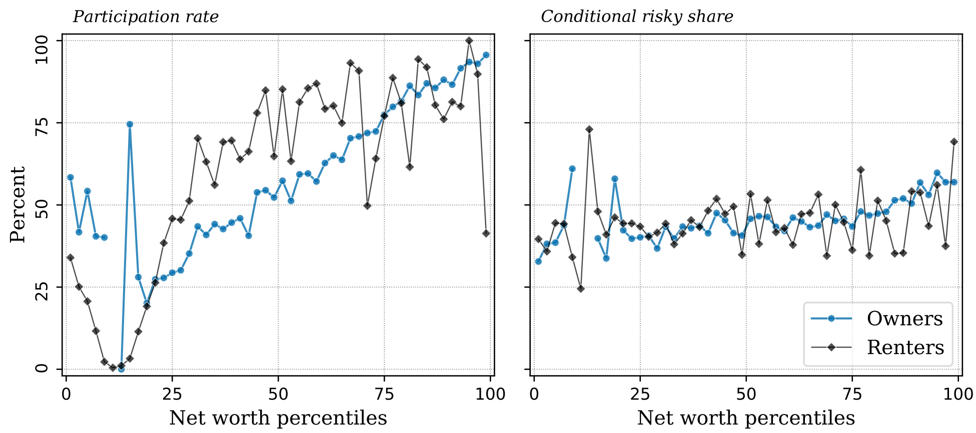


Figure 30: Portfolio composition along percentiles of *net worth* by homeownership status. Wealth percentiles are computed for the *pooled* sample of owners and renters. Plots show averages conditional on wealth percentile and homeownership status. Data source: SCF 1998–2007.

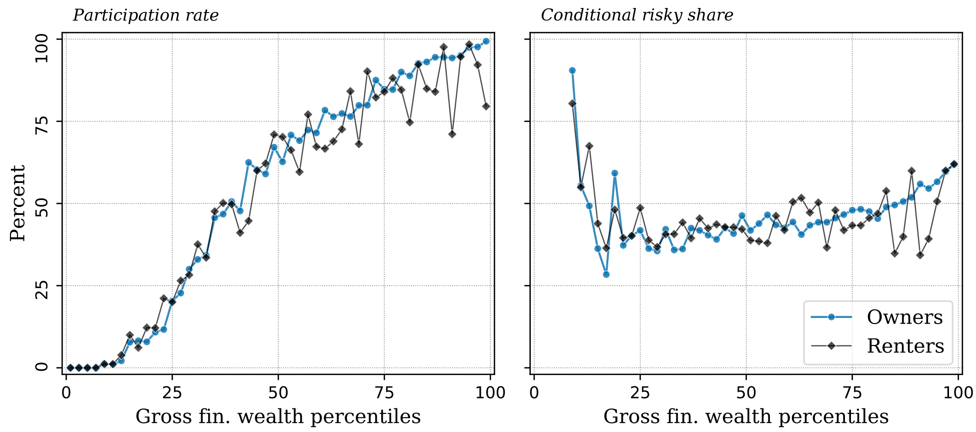


Figure 31: Portfolio composition along percentiles of *gross financial wealth* by homeownership status. Wealth percentiles are computed for the *pooled* sample of owners and renters. Plots show averages conditional on wealth percentile and homeownership status. Data source: SCF 1998–2007.

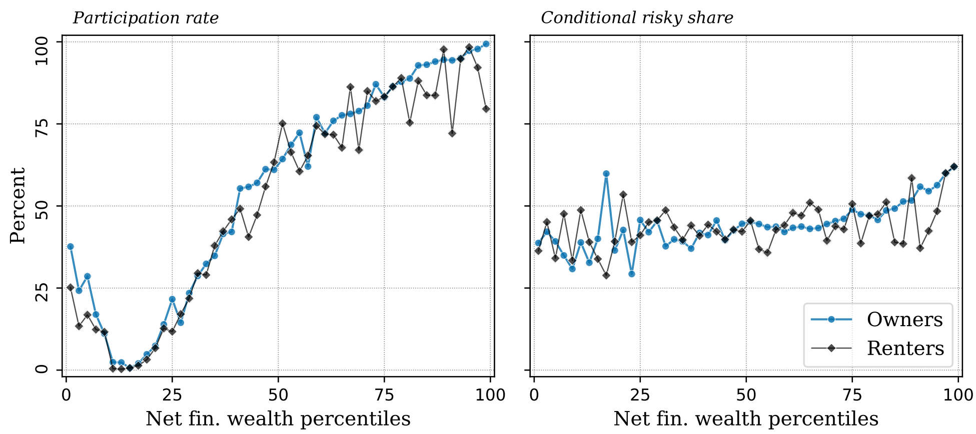


Figure 32: Portfolio composition along percentiles of *financial wealth net of consumer debt* by homeownership status. Wealth percentiles are computed for the *pooled* sample of owners and renters. Plots show averages conditional on wealth percentile and homeownership status. Data source: SCF 1998–2007.

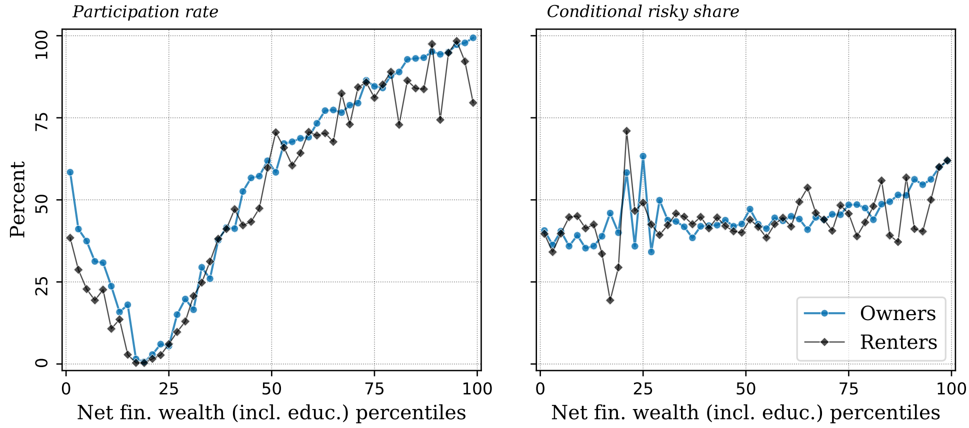


Figure 33: Portfolio composition along percentiles of *financial wealth net of consumer debt and education loans* by homeownership status. Wealth percentiles are computed for the *pooled* sample of owners and renters. Plots show averages conditional on wealth percentile and homeownership status. Data source: SCF 1998–2007.

B Recursive formulation of experience-based learning

In their paper, Malmendier and Nagel (2011) postulate that the beliefs about returns in period t are a weighted average of past realizations prior to t, going back as far as a person's year of birth. The index of historical returns proposed by Malmendier and Nagel (2011) is given by equations (5) and (6) in the main text.

Denote the new excess return observation at the beginning of period t by $r^e_{it} = R_t - R_f$. In this section, I ignore whether R_t has an idiosyncratic return component and drop any household-specific index i. The historical return index in Malmendier and Nagel (2011), adjusted to the notation used in the present paper and the fact that age $h = 0$ in the model corresponds to $\underline{h}$ actual years, can then be written as

$$\widehat{\mu}_{it} = \frac{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda} r^e_{t-k+1}}{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}}$$

where $\widehat{\mu}_{it}$ is the belief about mean excess returns by agent i of age h_{it} in period t after observing the risky realization at the beginning of period t. Note that, as in Malmendier and Nagel (2011), the above expression is not defined for $\underline{h} + h_{it} \leq 1$, i.e. they assume that the first return realization taken into account is the one that occurred at age one.

The recursive updating weight can be derived as follows:

$$\begin{aligned}
\widehat{\mu}_{it} &= \frac{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda} r^e_{t-k+1}}{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}} \\
&= \frac{(\underline{h} + h_{it} - 1)^{\lambda}}{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}} r^e_t + \frac{\sum_{k=2}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda} r^e_{t-k+1}}{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}} \\
&= \frac{(\underline{h} + h_{it} - 1)^{\lambda}}{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}} r^e_t + \frac{\sum_{k=1}^{\underline{h}+(h_{it}-1)-1} (\underline{h} + (h_{it} - 1) - k)^{\lambda} r^e_{t-(k-1)+1}}{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}} \\
&= \frac{(\underline{h} + h_{it} - 1)^{\lambda}}{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}} r^e_t \\
&\quad + \frac{\sum_{k=1}^{\underline{h}+(h_{it}-1)-1} (\underline{h} + (h_{it} - 1) - k)^{\lambda}}{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}} \frac{\sum_{k=1}^{\underline{h}+(h_{it}-1)-1} (\underline{h} + (h_{it} - 1) - k)^{\lambda} r^e_{t-(k-1)+1}}{\sum_{k=1}^{\underline{h}+(h_{it}-1)-1} (\underline{h} + (h_{it} - 1) - k)^{\lambda}} \\
&= \left[\frac{(\underline{h} + h_{it} - 1)^{\lambda}}{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}} \right] r^e_t + \left[\frac{\sum_{k=1}^{\underline{h}+(h_{it}-1)-1} (\underline{h} + (h_{it} - 1) - k)^{\lambda}}{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}} \right] \widehat{\mu}_{it-1} \\
&= \left[\frac{(\underline{h} + h_{it} - 1)^{\lambda}}{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}} \right] r^e_t + \left[\frac{\sum_{k=2}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}}{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}} \right] \widehat{\mu}_{it-1}
\end{aligned}$$

Furthermore, since

$$1-\frac{(\underline{h}+h_{it}-1)^{\lambda}}{\sum_{k=1}^{\underline{h}+h_{it}-1}(\underline{h}+h_{it}-k)^{\lambda}}=\frac{\sum_{k=2}^{\underline{h}+h_{it}-1}(\underline{h}+h_{it}-k)^{\lambda}}{\sum_{k=1}^{\underline{h}+h_{it}-1}(\underline{h}+h_{it}-k)^{\lambda}}$$

the recursive updating equation can be written as

$$\widehat{\mu}_{it}=(1-\alpha_{it})\widehat{\mu}_{it-1}+\alpha_{it}r_t^e$$

with the weight on the most recent observation given by

$$\alpha_{it}=\frac{(\underline{h}+h_{it}-1)^{\lambda}}{\sum_{k=1}^{\underline{h}+h_{it}-1}(\underline{h}+h_{it}-k)^{\lambda}}$$

C Alternative calibration similar to Cocco et al. (2005)

This section illustrates how the model with subjective beliefs performs using a more "standard" calibration. To this end, to the extent possible, I adopt the parameters used in Cocco, Gomes, and Maenhout (2005), which have become somewhat of a benchmark model in the household-finance literature, and contrast the resulting portfolio allocation across the wealth distribution and over the life-cycle.[21] Unlike in the rest of the paper, I therefore set the risk-aversion to $\gamma = 10$, the elasticity of intertemporal substitution to $1/\gamma$, the annual discount factor to $\beta = 0.96$, and turn off the bequest motive. As in Cocco, Gomes, and Maenhout (2005), the participation cost κ is set to zero and there is no preference heterogeneity, unlike in the main text. These parameters are summarized in Table 7.

	Description	Value	Source
β	Discount factor (annualized)	0.96	Cocco, Gomes, and Maenhout (2005)
γ	Risk aversion	10	Cocco, Gomes, and Maenhout (2005)
ψ^{-1}	Elasticity of intertemporal substitution	0.1	Cocco, Gomes, and Maenhout (2005)
ϕ	Weight on bequest utility	0	Cocco, Gomes, and Maenhout (2005)
κ	Participation cost	0	Cocco, Gomes, and Maenhout (2005)

Table 7: Parameters used for alternative calibration, taken from Cocco, Gomes, and Maenhout (2005) where applicable. The remaining parameters are unchanged from the benchmark calibration in the main text.

The resulting portfolio composition across the wealth distribution is shown in Figure 34, and Figure 35 plots averaged household choices over the life-cycle.

Whereas participation rates along the wealth distribution look similar to the results reported for the model without participation costs in the main text, the conditional risky share in this calibration is considerably lower. The reason is that this "standard" calibration performs poorly in terms of matching the wealth distribution, generating households that hold substantially more assets than in the data. Since the optimal risky share is decreasing in wealth, this generates a lower conditional risky share in the cross-section.

To illustrate this, I plot gross total wealth as reported in the SCF for each wealth decile (approximately $117,000 for the 5th decile), expressed as multiples of average quarterly earnings, and contrast these values with their model counterparts.[22] Evidently, both mod-

[21] My model implementation does not have any heterogeneous income profiles for three education groups as in Cocco, Gomes, and Maenhout (2005). Additionally, their stochastic earnings process features a unit root, while the one used here follows an AR(1), albeit with a very high persistence. The results for the "standard" model presented here are thus not identical to those in Cocco, Gomes, and Maenhout (2005).

[22] I omit the highest decile as this makes the differences for the lower deciles hard to read. Both models do even worse when it comes to matching the 10th decile, as there is no mechanism present to generate the thick right tail of the wealth distribution.

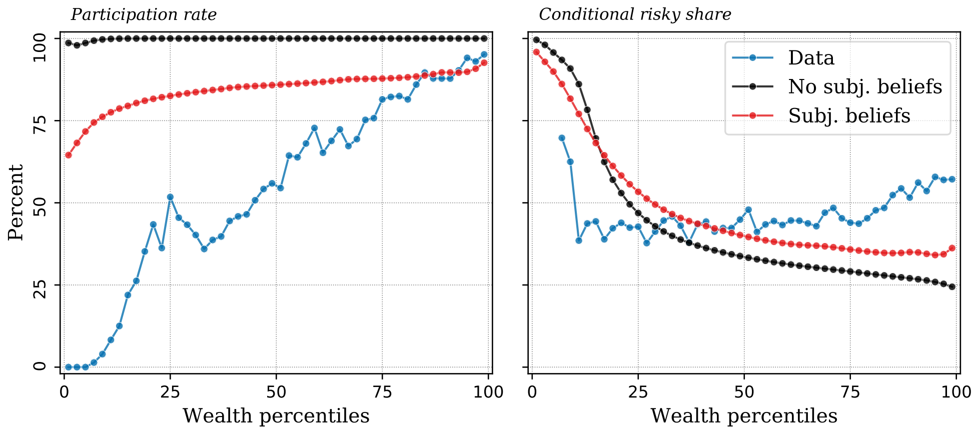


Figure 34: Portfolio composition along the wealth distribution. Parameters as reported in Table 7.

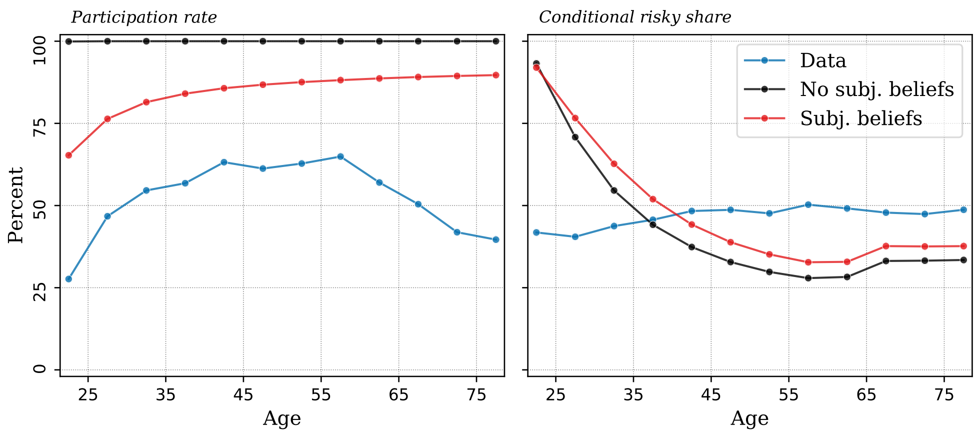


Figure 35: Portfolio composition over the life-cycle. Parameters as reported in Table 7.

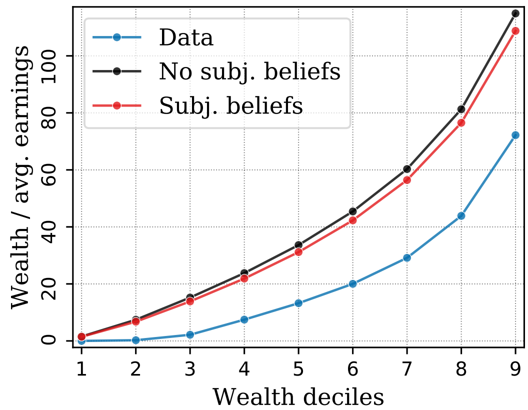


Figure 36: Wealth distribution obtained using "standard" calibration. Each dot shows the mean wealth conditional on being in a given decile, normalized by average quarterly earnings. Parameters as reported in Table 7. Data source: SCF 1998–2007.

els miss the empirical wealth distribution by a wide margin. For example, median wealth is approximately three times higher in the model than in the data.

D Results for the benchmark model with sophisticated households

This section reports the results for the benchmark calibration, but assuming that households are sophisticated and thus anticipate that they will update their beliefs in the future as new return realizations are observed. I focus on the case with per-period participation costs since the differences between the naive and the sophisticated models are very similar in both cases. Figure 37 shows the cross-sectional composition generated by the model with sophisticated agents, and how it compares to the benchmark calibration with naive agents. Figure 38 repeats the exercise for portfolio allocations over the life-cycle. Figure 39 shows that the beliefs do not differ to any considerable extent between naive and sophisticated households which is due to the assumption that investors update their beliefs independently of their participation decision. Figure 40 illustrates that while the optimal risky share chosen by naive vs. sophisticated households can differ substantially, these differences predominantly arise for high cash-at-hand levels and therefore have small effects on the portfolio composition in the cross-section.

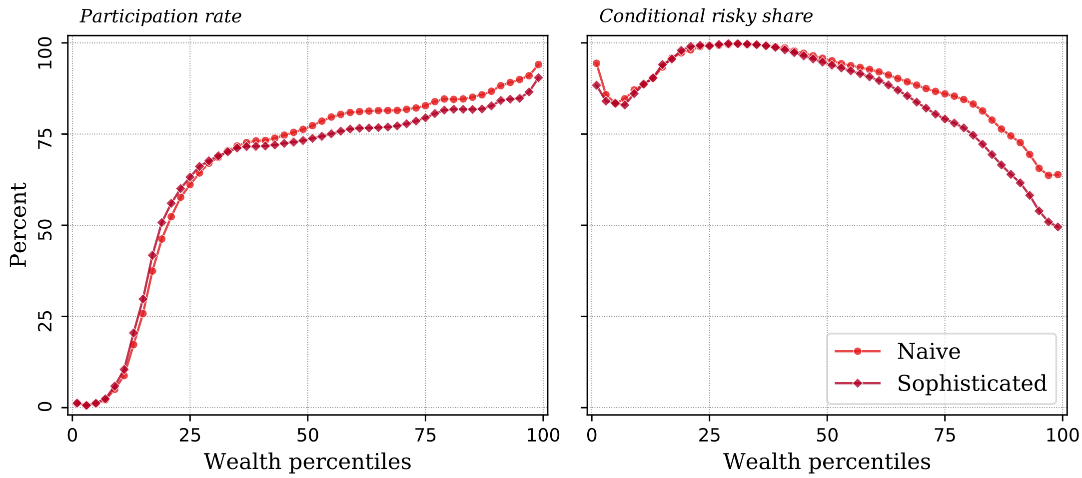


Figure 37: Portfolio composition over the life-cycle. Benchmark calibration *with* participation costs, naive vs. sophisticated households with subjective beliefs.

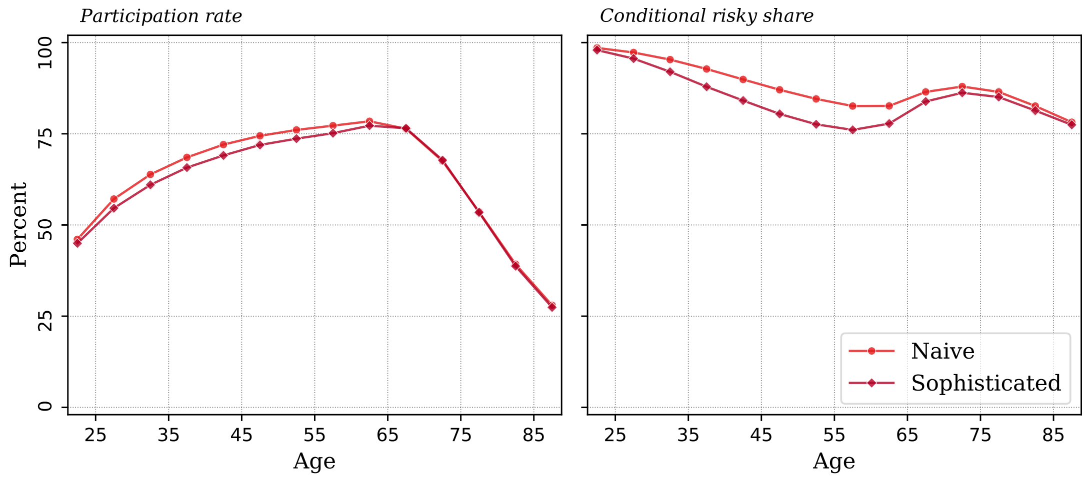


Figure 38: Portfolio composition over the life-cycle. Benchmark calibration *with* participation costs, naive vs. sophisticated households with subjective beliefs.

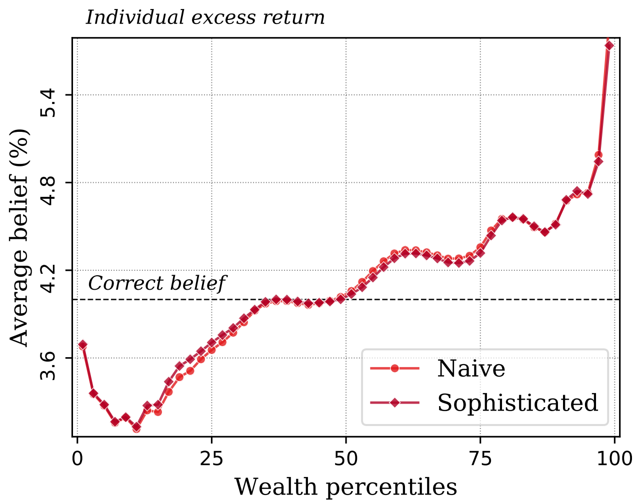
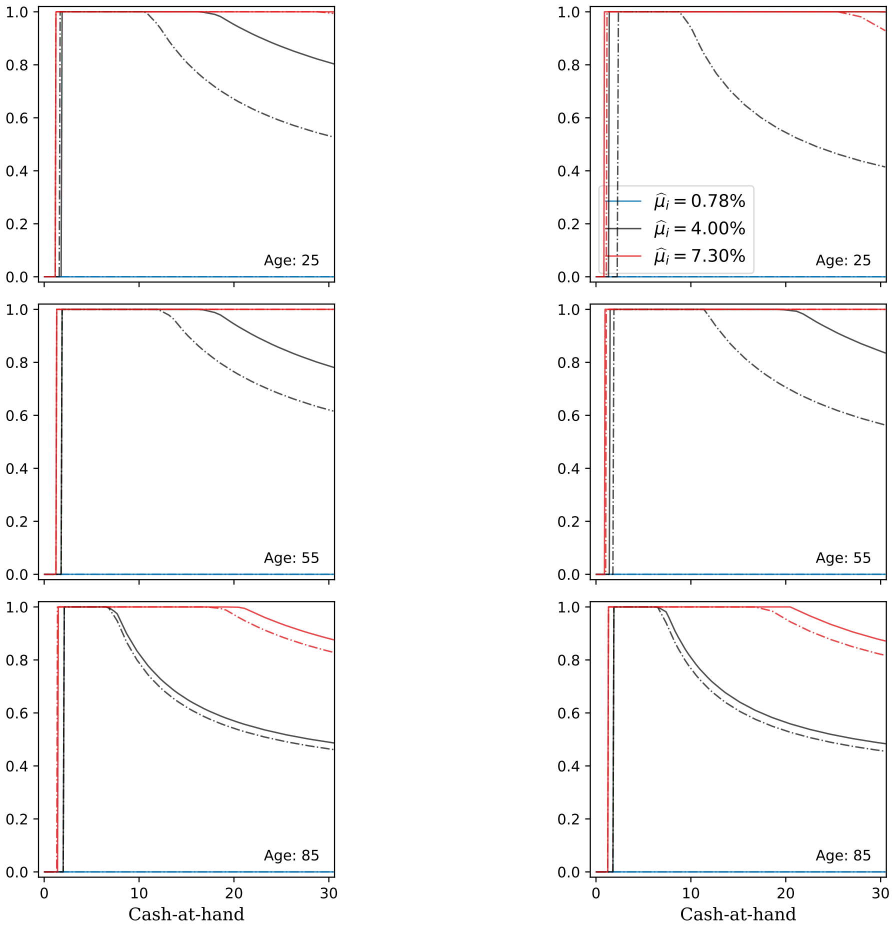


Figure 39: Average beliefs about excess returns, naive vs. sophisticated households.



(a) Preference type 1

(b) Preference type 2

Figure 40: Optimal risky share for naive (solid line) vs. sophisticated (dashed line) households, shown for a household with the median persistent labor productivity. Impatient households (type 1) in the left-hand column, patient households (type 2) to the right. Colors denote different beliefs about excess returns.

E Numerical solution

E.1 Household problem

Solving for optimal choices. I solve the household problem using a hybrid endogenous grid point method (EGM). Plain EGM was originally proposed to solve the consumption-savings problem without a portfolio choice, but can be extended to portfolio-choice problems in a straightforward way by adding a root-finding step at each exogenous savings level.

Root-finding needs to be performed on the first-order condition with respect to ξ, the optimal risky share. For the case of EZW preferences, this portfolio Euler equation (P-EE) is given by

$$\pi_h^s \mathbf{E}_i \left[(V')^{-\gamma} V_1' \left(R' - R_f\right) \right] + (1 - \pi_h^s) \mathbf{E}_i \left[((V^b)')^{-\gamma} (V_1^b)' \left(R' - R_f\right) \right] = 0$$

for the working-age household, where the continuation values are given by

$$\begin{aligned} V' &\equiv V_{jh+1}\left(a', p', \widehat{\mu}_i'\right) \\ (V^b)' &\equiv V_j^b\left(a_b'\right) \end{aligned}$$

and their derivatives with respect to to the first argument (i.e. cash-at-hand) are denoted as

$$\begin{aligned} V_1' &\equiv \frac{\partial V_{jh+1}'\left(a', p', \widehat{\mu}_i'\right)}{\partial a'} \\ (V_1^b)' &\equiv \frac{\partial V_j^b\left(a_b'\right)}{\partial a_b'} \end{aligned}$$

Using the envelope condition, the P-EE can be stated as

$$\pi_h^s \mathbf{E}_i \left[(V')^{\psi-\gamma} (c')^{-\psi} \left(R' - R_f\right) \right] + (1 - \pi_h^s) \phi_j^{\frac{1-\gamma}{1-\psi}} \mathbf{E}_i \left[\left(a_b'\right)^{-\gamma} \left(R' - R_f\right) \right] = 0 \tag{18}$$

As in plain EGM, I solve for the optimal solution conditioning on an exogenously imposed savings level $b > 0$ (the portfolio choice is indeterminate for $b = 0$, so the boundary case can be ignored). Then (18) is an implicit function of ξ via the continuation values, their derivatives and tomorrow's consumption c', which depend on the optimal choice of ξ via its effect on a' and a_b'. It turns out that (18) is in general a monotonically decreasing function of ξ, a fact that can be used to easily identify corner solutions as well as interior portfolio choices. Once the optimal portfolio choice has been determined, one can compute the return on the total portfolio $R_{p,t+1}$ which enters the consumption-savings Euler equation. The remaining steps are the same as in the standard EGM procedure.

In the presence of fixed participation costs, EGM has the problem that the first-order conditions are only necessary, but no longer sufficient. Intuitively, this is the case because while the above hybrid method can be used to find an interior optimal solution *conditional on paying the participation cost*, not saving in the risky asset and thus not incurring the participation cost might still yield a higher utility. This potentially creates downward jumps in the consumption policy function which plain EGM cannot deal with. I implement the approach suggested in Iskhakov et al. (2017) and Druedahl and Jørgensen (2017) to address this issue (the variant without taste shocks).

Exogenous grids. I discretize cash-at-hand and the exogenous savings grid in the usual way, putting more points in the region where policy and value functions have more curvature.

I approximate mean beliefs on an age-specific grid of 23 points as follows: I compute the percentiles (P0, P1, P2.5, P7.5, ..., P92.5, P97.5, P99, P100) of the cross-sectional belief distribution for each age, and place grid points at the expected values conditional on falling into the regions bracketed by these percentiles. Thus, the grid points are automatically placed in regions where the beliefs for each cohort are concentrated.

The persistent earnings component is approximated using the Rouwenhorst procedure with seven grid points. Transitory earning shock realizations are discretized to five possible realizations.

E.2 Simulation

Following an approach commonly used in the literature (see, for example, Gomes and Michaelides (2005)), I simulate the distribution of a "representative" cohort over its entire life-cycle instead of rolling forward a cross-section of a limited number of households.

The initial distribution of newborns is computed as follows:

1. Their initial wealth is assigned according to the distribution of wealth in the SCF for the age group 20–25, discretized onto 500 bins.
2. The initial distribution over persistent labor productivity states is the ergodic distribution implied by the Markov process obtained from the Rouwenhorst procedure.
3. Beliefs are drawn from the normal distribution that is assumed for newborns, discretized onto the 23 grid points as outlined above.

Using the households policy rules and the exogenous transition processes, I then build a (sparse) transition matrix that transforms the distribution of households over exogenous and endogenous states at age t into the corresponding distribution at age $t + 1$. This process is repeated for all ages $t = 0, \ldots, T - 1$. Since there are no endogenous intergenerational linkages (such as bequests between a specific parent and descendant household), the stationary distribution over all cohorts can then be obtained as a cohort-size-weighted average of these age-specific distributions.