Stock Market Participation and Portfolio Shares Over the Life-Cycle*

Francisco Gomes† Oksana Smirnova‡

October 18, 2021

Abstract:

We estimate the life-cycle profile of stock market participation and risky portfolio share. We address the classical identification problem by running the estimations in first differences, which allows us to estimate the age profiles without making any assumptions on time or cohort effects. We find that stock market participation is a hump-shaped function of age, increasing early in life and decreasing after age 60. The conditional risky share also decreases late in life, but it is a flat function of age before that. We also investigate the economic mechanisms driving this behavior. Our results provide empirical support for the importance of participation costs in explaining stock market participation, and for models where investors have decreasing relative risk aversion and where human capital is a close substitute for bonds, although not completely uncorrelated with stock returns. Finally, background risks are also likely to play a role, particularly late in life. We conclude by presenting a structural life-cycle model that closely replicates our empirical results.

JEL Classification: G5, G11, D14, D15.

Key Words: Life-cycle asset allocation, stock market participation, life-cycle models, human capital.

*We thank Joao Cocco, Evgenii Gorbatikov, Martin Holm, Anna Pavlova, Henri Servaes, Taisiya Sikorskaya, Paolo Surico, and seminar participants at LBS and at the 7th BI-SHOF Conference on Asset Pricing and Financial Econometrics for their helpful comments. We also thank Francesco Amodeo for research assistance, and the AQR Asset Management Institute at LBS for financial support.

†London Business School and CEPR, fgomes@london.edu

‡London Business School, osmirnova@london.edu

1 Introduction

Do younger households invest in more or less risky portfolios than older households, and what are the economic drivers of this behavior? The answers to these questions can provide important insights into theories of life-cycle portfolio choice. A simple life-cycle portfolio choice model with borrowing constraints and undiversifiable income risk predicts that households should decrease their risky investments as they approach retirement (Cocco, Gomes, and Maenhout, 2005), and this result forms the basis for the popular target-date funds offered in most defined-contribution pension plans.[1] However, several extensions of this baseline model can deliver a hump-shaped pattern, or even a risky share that increases with age (e.g. Cocco et al., 2005; Benzoni, Collin-Dufresne, and Goldstein, 2007; Campanale, Fugazza, and Gomes, 2015; Fagereng, Gottlieb, and Guiso, 2017; Bagliano, Fugazza, and Nicodano, 2017 or Catherine, 2019).[2]

Over the last two decades household finance researchers have been trying to estimate the age-profiles of the stock market participation and the share of wealth invested in risky assets (e.g. Poterba and Samwick, 1997; Ameriks and Zeldes, 2004; Fagereng et al., 2017, Catherine, 2019 and Parker, Schoar, and Cole, 2021). The challenge to this apparently straightforward task is the classic identification problem of separately identifying age, time and cohort effects.[3] In this paper we consider a simple approach to tackle this problem, by taking first differences of the data. In first differences cohort effects disappear, and we are left with age(-difference) and time(-difference) effects which we can estimate in a standard regression framework. This solution only has one drawback: we cannot recover the actual level of the conditional risky share or stock market participation at any age. However, if we limit our attention to the question of how these change as a function of age, this approach provides a very simple solution to the identification problem.

[1] Choi and Robertson (2020) survey households directly and find that "years left until retirement" is the most cited factor for the equity shares of stock market participants, reinforcing the importance of life-cycle considerations for household portfolio decisions.

[2] See Gomes (2020) for a detailed survey of this literature.

[3] Since "current year" = "birth year" + "age", the three variables are perfectly collinear and we can't identify them all simultaneously.

From a theoretical perspective, the average level of the risky share depends on a large range of factors that are independent of the particular theory being considered, such as return expectations, risk aversion, IQ, education/financial literacy or trust in the stock market.[4] On the other hand, the age profile is a more robust prediction of a given model, and can therefore be used to distinguish between alternative theoretical frameworks. Therefore our estimation in first differences is highly informative of the economic mechanisms driving the behavior of the risky share over the life-cycle.

Furthermore, since asset allocation decisions of households are determined by several often unobservable characteristics, such as the previously mentioned risk aversion, IQ, financial sophistication, return expectations, and others, this poses a problem for empirical work. The inclusion of time fixed effects in the regressions allows researchers to capture any time-variation in these variables that is common across households, but it does not control for cross-sectional heterogeneity. In the absence of consistent measures for all those variables, by taking first differences of the data, our results are more robust to this unobserved individual-level heterogeneity.[5,6]

Implementing our estimation approach requires household-level panel data on asset holdings, since we have to be able to follow individual portfolios over time. This rules out the Survey of Consumer Finances, and therefore we use the Panel Study for Income Dynamics (PSID). In our analysis we focus on stock market participation and the conditional risky share, i.e. the risky share of stockholders. We find that stock market participation is a hump-shaped function of age, increasing early in life, then flat around mid-life and finally decreasing as individuals approach retirement and again during retirement. For the conditional risky share we obtain an increasing profile before age 55 and flat after that when

[4]For empirical evidence on the impact of these on asset allocation decisions see, for example, Guiso, Haliassos, and Jappelli (2002), Calvet, Campbell, and Sodini (2007), Guiso, Sapienza, and Zingales (2008), Malmendier and Nagel (2011), Van Rooij, Lusardi, and Alessie (2011), Hurd, Van Rooij, and Winter (2011), Grinblatt, Keloharju, and Linnainmaa (2011), Dimmock, Kouwenberg, Mitchell, and Peijnenburg (2016), Black, Devereux, Lundborg, and Majlesi (2018).

[5]For this purpose, an alternative would be to include individual fixed effects in our regression. Calvet and Sodini (2014) control for this heterogeneity by using a sample of matched twins.

[6]Unobserved heterogeneity that varies over time in different ways for different households is still not captured, but the same concern is present with the alternative approaches.

considering total changes in the portfolio. However, when we consider active rebalancing decisions only, to control for the high-levels of household portfolio inertia previously documented in the literature, we find that the share of wealth invested in stocks is actually constant with age until ages 55-60, and decreasing after that.[7]

We contrast our results with those obtained under the alternative identification assumptions of no cohort effects or no time effects (as in Poterba and Samwick (1997) and Ameriks and Zeldes (2004)). Without cohort effects we would have estimated increasing profiles for both stock market participation profile and the conditional risky share. If we had ruled out time effects instead, we would conclude that stock market participation is strongly decreasing with age, while the conditional risky has an inverse hump-shape. Contrasting these results with ours highlights the importance of controlling for both time and cohort effects when estimating these life-cycle profiles.

In addition to estimating age profiles we also investigate the economic mechanisms driving risk taking behavior over the life cycle, by studying the impact of wealth and human capital on these decisions. More precisely, we consider both their direct effect on the risky share and the stock market participation decision, and how they influence the age profiles. The later evidence is informative of how risk taking behavior over the life cycle is shaped by the simultaneous evolution of those two particular variables as a function age, and the combined evidence allows us to disentangle the predictions implied by different asset allocation models.

We find evidence of positive wealth effects for both outcome variables. In fact, we find that, the increase in stock market participation early in life is almost fully explained by the increase in household wealth during this period. These positive wealth effects for participation provide support for theories emphasizing the role of stock market participation costs in explaining why a large fraction of households does not invest in equities (e.g. Vissing-Jørgensen, 2002; Haliassos and Michaelides, 2003; Gomes and Michaelides, 2005). The positive wealth effects in the conditional risky share are evidence for preferences with decreasing relative risk aversion (e.g. Constantinides, 1990; Campbell and Cochrane, 1999).[8]

[7]For the evidence on household portfolio inertia see the survey paper Gomes, Haliassos, and Ramadorai (forthcoming)

[8]Gomes and Michaelides (2003), Polkovnichenko (2007), Wachter and Yogo (2010) and Meeuwis (2020)

For modest changes in wealth the implied change in the conditional risky share is quite small, which helps to explain why previous literature has found mixed results when estimating wealth effects in the demand for risky assets.[9] However, both in the cross-section and over the life-cycle, we observe substantial changes in wealth which are then reflected in non-trivial implied differences in the portfolio allocation. The evidence for decreasing relative risk aversion can explain the increasing risky share early in life. This pattern can also reflect high background risk which leads young investors to hold more conservative portfolios, as shown by Viceira (2001) and Cocco et al. (2005).[10] Finally, a lower equity allocation early in life is also consistent with models with large return skewness (Fagereng et al., 2017) or transaction costs (Campanale et al., 2015).

The theoretical portfolio choice literature has identified the ratio of the present-value of future labor income to total wealth as an important determinant of the optimal asset allocation (e.g. Heaton and Lucas, 1996; Viceira, 2001; Cocco et al., 2005). In our paper we find that this ratio has a positive impact on both the risky share and stock market participation, consistent with models where labor income is a close substitute for bonds. Under these conditions the optimal equity allocation will decrease as agents approach retirement (see Viceira (2001) or Cocco et al. (2005)). Therefore this human capital channel, combined with decreasing relative risk aversion (and/or background risk) can explain the combined flat pattern of the risky share early in life, and decreasing late in life. In principle, increasing relative risk aversion and human capital being a close substitute for stocks could also deliver those results, with the two effects offsetting each other early in life, and increasing relative risk aversion dominating close to retirement. However, such a model would imply the opposite sign for the estimated coefficients on wealth and human capital.

---

solve life-cycle models of portfolio choice with decreasing relative risk aversion preferences.

[9] See, for example, Heaton and Lucas (1996), Campbell (2006), Brunnermeier and Nagel (2008), Wachter and Yogo (2010), Chiappori and Paiella (2011), Calvet and Sodini (2014), Bach, Calvet, and Sodini (2016), Fagereng, Guiso, Malacrino, and Pistaferri (2020) and Meeuwis (2020).

[10] Guiso, Jappelli, and Terlizzese (1996), Calvet and Sodini (2014), Bonaparte, Korniotis, and Kumar (2014), Knüpfer, Rantapuska, and Sarvimäki (2017), Fagereng et al. (2020) provide direct evidence for the impact of income risk on household portfolio decisions. Related, Cocco (2005), Yao and Zhang (2005), Chetty and Szeidl (2007), Chetty, Sándor, and Szeidl (2017) document the importance of risks resulting from mortgage commitments.

Despite being consistent with human capital as a close substitute for bonds, our evidence indicates that this substitutability is weaker than what is predicted by a model where income shocks are uncorrelated with stock returns. Our results thus provide support for a model with higher-order correlations as in Catherine (2019), or for a weaker version of the channel proposed by Benzoni et al. (2007). Consistent with previous literature (e.g. Brunnermeier and Nagel (2008) and Bonaparte et al. (2014)), we document significant stock market exit decisions, providing support for per-period stock market participation costs also playing an important role. Finally, the late-in-life decreases in both the conditional risky share and stock market participation gives empirical support for models with background risks during retirement (e.g. Hubbard, Skinner, and Zeldes, 1995; De Nardi, French, and Jones, 2010; Ameriks, Caplin, Laufer, and Van Nieuwerburgh, 2011; Yogo, 2016 or Koijen, Van Nieuwerburgh, and Yogo, 2016).[11]

Our paper is part of a large empirical literature studying stock market participation and the asset allocation decisions of stockholders.[12] In addition to the papers already cited in this introduction, our work is particularly related to previous studies that have tried to estimate the life-cycle profiles of stock market participation and the conditional risky share, namely Poterba and Samwick (1997), Ameriks and Zeldes (2004), Fagereng et al. (2017), Catherine (2019) and Parker et al. (2021). These papers differ with respect to the data that they use, but also in their approaches to address the identification problem. Poterba and Samwick (1997) use data from the U.S. Survey of Consumer Finances (SCF) and control for cohort effects. Ameriks and Zeldes (2004) report results both with SCF data and retirement wealth data from TIAA-CREF. They consider the two special cases of no cohort effects or no time effects. Fagereng et al. (2017) use Norwegian administrative data, and consider both cohort and time effects simultaneously, by making specific assumptions about those. In one specification they capture cohort effects using individuals' life-time return experiences (based on the work of Malmendier and Nagel (2011)), while in the other they impose the Deaton and Paxson (1994) restriction that time effects sum to zero once the variables have

[11]Guiso et al. (1996) empirically document the impact of health risks on household portfolio allocations.
[12]Detailed surveys of this literature can be found in Gomes (2020) and Gomes et al. (forthcoming).

been detrended. Catherine (2019) utilizes the SCF and the Deaton and Paxson (1994) methodology to capture time effects. Parker et al. (2021) use retirement wealth data only. Therefore, they do not consider stock market participation and instead focus on the life-cycle patterns of contribution rates to retirement accounts. With regards to portfolio allocation they find a hump-shape pattern with age, but also highlight the importance of institutional features, namely target-date funds as default options, in determining the asset allocations in these accounts.

Unfortunately, there is no consensus among the results obtained in the previous papers. When using cohort dummies Ameriks and Zeldes (2004) document increasing stock market participation until retirement and flat thereafter, and the same result is obtained by Catherine (2019). However, when using time dummies Ameriks and Zeldes (2004) estimate a flat profile almost until retirement and decreasing thereafter. Finally, Poterba and Samwick (1997) and Fagereng et al. (2017) estimate hump-shaped patterns, peaking around age 40 and age 60, respectively. Our results are therefore similar to these, particularly the ones in Fagereng et al. (2017). With respect to the conditional risky share our results are closest to the ones in Ameriks and Zeldes (2004) when considering time effects only. In this case they obtain a flat profile until around age 55 and decreasing after that.

Relative to these previous papers we use a different data set and a different time period, but crucially our results also rely on a different approach to the identification problem. By ignoring the levels of the risky share and stock market participation, and focusing exclusively on their age profiles, we can address the identification problem without having to make assumptions about cohort or time effects (as in Fagereng et al. (2017) and Catherine (2019)), or having to ignore one or the other (as in Ameriks and Zeldes (2004) and Poterba and Samwick (1997)). [13]

In the final section of the paper we present a structural life-cycle portfolio choice model which closely replicates our main empirical findings. The features included in the model are motivated by our previous evidence, and therefore include stock market participation costs,

[13] As previously-mentioned we repeat our estimations using the alternative identification approaches to highlight their importance for the results.

preferences with decreasing relative risk aversion, moderate correlation between labor income shocks and stock returns. In line with previous theoretical work on limited participation (e.g. Gomes and Michaelides, 2005), household preference heterogeneity also plays an important role in the model, but we also highlight the importance of including heterogeneity in financial literacy.

The remainder of the paper is structured as follows. Section 2 describes our estimation approach, while section 3 describes the data and presents some summary statistics. Sections 4 and 5 present results for the stock market participation decision and the conditional risky share, respectively. In section 6 we present the life-cycle portfolio choice model and Section 7 includes our concluding remarks.

2 Estimation Approach

We study both the stock market participation decision and the share of wealth invested in risky assets by stock market participants. Following the literature we refer to the latter as the conditional risky share, i.e. the risky share conditional on positive stock holdings. The discussion in this section focuses only on the conditional risky share, to avoid repetition.

2.1 Frictionless case

We first consider the outcome of a frictionless portfolio choice model, where investors can adjust their portfolios every period without cost.[14] The conditional risky share for individual i at time t (ω_{it}) is then represented in the following reduced form equation:

$$\omega_{it} = I(a_{it}) + I(t) + I(c_i) + \theta X_{it} + \Gamma F_i + \varepsilon_{it} \tag{1}$$

where the $I(a_{it})$, $I(t)$ and $I(c_i)$ are dummy variables for each age, time period and cohort, respectively.[15] The variable F_i is a vector of time-invariant individual characteristics, while

[14]Here we refer to a broad definition of frictions, including any constraints and/or biases that generate either partial or full inertia in portfolio adjustments.

[15]There are A age effects, T time effects and C cohort effects.

X_{it} is vector of time-varying characteristics which are not perfectly correlated across individuals.[16] Finally ε_{it} is a (regression) residual.

Estimating equation (1) faces the classical identification challenge that it is impossible to separately identify age, time and cohort effects, since these three variables are perfectly collinear:

$$t = c_i + a_{it} \tag{2}$$

Hence, under the classical identification challenge of age, time and cohort it is only possible to estimate A+T+C-3 effects in equation (1) (plus all other controls): two are missing because of the dummy variable trap and one is lost because age, time and cohort can always be represented as the sum of the other two.

Researchers have explored alternative methods for tackling this issue. These include imposing zero restrictions on either time or cohort effects[17] (Poterba and Samwick (1997) and Ameriks and Zeldes (2004)), or modelling them as functions of other variables (Fagereng et al. (2017) and Catherine (2019)). In our paper we take an alternative approach and estimate the equation in first differences, i.e. we replace equation (1) with

$$\Delta\omega_{it} = I(a_{it}) - I(a_{i,t-1}) + I(t) - I(t-1) + \theta\Delta X_{it} + u_{it} \tag{3}$$

where

$$u_{it} = \varepsilon_{it} - \varepsilon_{i,t-1} \tag{4}$$

Moving from equation (1) to equation (3) eliminates cohort effects ($I(c_i)$) and therefore resolves the identification problem. The drawback of this approach is that we cannot recover the actual level of the conditional risk share at a given age, which is why it was probably not previously considered. However, if we want to estimate how the conditional risky share changes with the age, and the economic channels that drive such behavior, this method provides a straightforward solution.

[16] Time-varying characteristics that are perfectly correlated across individuals (e.g. stock market returns) are captured by the time dummies.

[17] However, one would still need to

Another important advantage of our approach is that it is more robust to unobserved individual heterogeneity. Estimating equation (1) requires data on several individual characteristics which can be very hard to measure in a large cross-section of households, such as risk aversion, IQ, financial literacy, present-value of labor income, or return expectations. In equation (3) the F_i terms drop out, thus eliminating (unobserved) characteristics that are constant over the sample.[18] Individual characteristics that are time-varying and for which the time variation is driven by a common factor (e.g. realized returns which might drive changes in expected returns) are captured in the time fixed effects. That is the case both in equation (3) and in equation (1). However, due to the identification challenge, the latter is often estimated with restrictions on the time fixed effects or while ignoring them completely, in which case these characteristics are only imperfectly controlled for, or not at all.

For the rest of the paper we simplify the notation and re-write equation (3) as

$$\Delta\omega_{it} = \overline{I}(\Delta a_{it}) + \widetilde{I}(\Delta t) + \theta\Delta X_{it} + u_{it} \tag{5}$$

In this specification $\overline{I}(\Delta a_{it})$ refers to dummy variables that are equal to 1 when individual i has that specific age at time t. So they are identical to the dummy variables in equation (1), but they have a different interpretation, hence the new notation.[19] Estimating equation (5) still requires eliminating one of the age or time dummies, but this is simply a normalization. Although the estimated coefficients will depend on which dummy we choose to remove, the predicted average margin effects are unaffected by this choice.[20]

Finally, in our empirical implementation, the vector ΔX_{it} will include variables such as changes in household wealth (different measures), changes in the present-value of future labor income, and demographic variables.

[18] This is likely to be the case with IQ, education, or individual preferences, for example. These variables are expected to have very limited (if any) variation from adult age onward.

[19] Alternatively, we could write down the regression as it is directly implied by equation (3), where these variables would actually take three values, -1, 0 or 1. Naturally the estimated coefficients would be different, but the predicted average marginal effects are the same under both specifications.

[20] This is also the case if we use the previous approach, i.e. estimating equation (1) without cohort or time effects. We would have to remove one of the remaining dummies, and likewise the estimated coefficients would depend on the particular choice, but not the predicted margin effects.

2.2 Transaction costs and inertia

It has been largely documented that a significant fraction of households only rebalances their portfolios infrequently (see, for example, Choi, Laibson, Madrian, and Metrick (2002), Agnew, Balduzzi, and Sunden (2003), Ameriks and Zeldes (2004), Brunnermeier and Nagel (2008), Calvet, Campbell, and Sodini (2009), Bilias, Georgarakos, and Haliassos (2010) or Meeuwis (2020)).[21]

We address this concern by following the approach in Calvet et al. (2009). We first compute the active change in the risky share

$$\Delta\omega_{it}^{Active} = \omega_{it} - \omega_{it}^{P} \tag{6}$$

where ω_{it}^{P} is the passive risky share, i.e. the risky share that would be obtained in the absence of trading between $t-1$ and t,

$$\omega_{it}^{P} \equiv \frac{\omega_{i,t-1}R_t}{\omega_{i,t-1}R_t + (1-\omega_{i,t-1})R_t^f} \tag{7}$$

with R_t as the return on risky assets, and R_t^f the return on the riskless asset(s). We then use the active change in the risky share ($\Delta\omega_{it}^{Active}$) as the right-hand-side variable in our regressions, therefore replacing equation (5) with

$$\Delta\omega_{it}^{Active} = \overline{I}(\Delta a_{it}) + \widetilde{I}(\Delta t) + \theta\Delta X_{it} + u_{it} \tag{8}$$

2.3 Human capital

One important variable to include in X_{it} is the ratio of the present-value of future labor income (human capital) to current wealth (hereafter $PVYW$), as discussed in the theoretical work of Heaton and Lucas (1996), Viceira (2001) and Cocco et al. (2005), for example.

[21]For potential explanations of this behavior see, among others, Gabaix and Laibson (2001), Sims (2003), Alvarez, Guiso, and Lippi (2012), Abel, Eberly, and Panageas (2013), Campanale et al. (2015) or Pagel (2018).

2.3.1 Measuring changes in the ratio of human capital to financial wealth

The present-value of future labor income for individual i at age a_0 is given by

$$PVY_{i,a_0} = E_{a_0} \sum_{a=a_0+1}^{A} \frac{p_{i,a} Y_{i,a}}{(R_i^Y)^{a-a_0}} \tag{9}$$

where A is an arbitrary maximum age, $p_{i,a}$ is the conditional expected survival probability from age $a-1$ to age a, $Y_{i,a}$ is income at age a and R_i^Y is the discount rate for labor income.[22] From equation (9) the change in the ratio of the present-value of future labor income to current wealth over a two-year period, is given by[23]

$$\Delta PVYW_{i,a_0} = E_{a_0} \sum_{a=a_0+1}^{A} \frac{p_{i,a} Y_{i,a}}{(R_i^Y)^{a-a_0} W_{i,a_0}} - E_{a_0-2} \sum_{a=a_0-1}^{A} \frac{p_{i,a} Y_{i,a}}{(R_i^Y)^{a-a_0-2} W_{i,a_0-2}} \tag{10}$$

$$= \left(E_{a_0} \sum_{a=a_0+1}^{A} \frac{p_{i,a} Y_{i,a}}{(R_i^Y)^{a-a_0} W_{i,a_0}} - E_{a_0-2} \sum_{a=a_0+1}^{A} \frac{p_{i,a} Y_{i,a}}{(R_i^Y)^{a-a_0-2} W_{a_0-2}} \right) - E_{a_0-2} \left(\frac{p_{i,a_0} Y_{i,a_0}}{(R_i^Y)^2 W_{i,a_0-2}} + \frac{p_{i,a_0-1} Y_{i,a_0-1}}{R_i^Y W_{i,a_0-2}} \right) \tag{11}$$

These equations show that changes in $PVYW_{i,a_0}$ occur for four reasons. First, because of revisions in the expectation of income going forward: E_{a_0} replaces E_{a_0-2}. Second, because of changes in current wealth: W_{i,a_0} replaces W_{i,a_0-2} in the denominator. Third, because the income from future years is now discounted less: the exponent on returns in the denominator changes from $a-(a_0-2)$ to $a-a_0$. These three effects are captured in the first term in parenthesis in equation (11). Finally, $PVYW_{i,a_0}$ also changes because some of the labor income that was included in the measure of human capital at age a_{0-2}, has been earned over

[22]More generally, the discount rate for labor income would also be a function of age, so we would have

$$PVY_{i,a_0} = E_{a_0} \sum_{a=a_{0+1}}^{A} \frac{p_{i,a} Y_{i,a}}{\prod_{j=a_{0+1}}^{a} R_{i,j}^Y}$$

where $R_{i,j}^Y$ is the discount rate for labor income at age j.

[23]The frequency for our data is bi-annual, so we have to consider the change over a two-year period.

the last two years. This is captured in the last term of equation (11).[24] In our empirical analysis we consider specifications with $\Delta PVYW_{i,a_0}$and specifications with the two terms in equation (11) considered separately.

2.3.2 Empirical implementation

The two terms in equation (11) are not directly observable, so they have to be replace with empirical proxies. To construct these proxies we have to make some assumptions, which are discussed in this section.

Assumption 1: discount rates and the survival probabilities are non-stochastic.

Using assumption 1 the change in the ratio of the present-value of future labor income to financial wealth (equation (11)) becomes

$$\begin{aligned}\Delta PVYW_{i,a_0} &= \sum_{a=a_0+1}^{A}\left(\frac{p_{i,a}E_{a_0}Y_{i,a}}{(R_i^Y)^{a-a_0}W_{i,a_0}} - \frac{p_{i,a}E_{a_0-2}Y_{i,a}}{(R_i^Y)^{a-(a_0-2)}W_{i,a_{0-2}}}\right)\\ &\quad -\left(\frac{p_{i,a_0}E_{a_0-2}Y_{i,a_0}}{(R_i^Y)^2W_{i,a_{0-2}}} + \frac{p_{i,a_0-1}E_{a_0-2}Y_{i,a_{0-1}}}{R_i^Y W_{i,a_{0-2}}}\right) \qquad (12)\end{aligned}$$

We define the two terms in equation (12) as

$$\xi^1_{i,a_0} \equiv \sum_{a=a_0+1}^{A}\frac{p_{i,a}}{(R_i^Y)^{a-a_0}}\left(\frac{E_{a_0}Y_{i,a}}{W_{i,a_0}} - \frac{E_{a_0-2}Y_{i,a}}{(R_i^Y)^2W_{i,a_{0-2}}}\right) \qquad (13)$$

$$\xi^2_{i,a_0} \equiv \frac{p_{i,a_0}E_{a_0-2}Y_{i,a_0}}{(R_i^Y)^2W_{i,a_{0-2}}} + \frac{p_{i,a_0-1}E_{a_0-2}Y_{i,a_{0-1}}}{R_i^Y W_{i,a_{0-2}}} \qquad (14)$$

To compute these terms we need expectations of future labor income.[25]

Assumption 2: Log labor income is given by

$$y_{i,a} = \mu^y_a + v_{i,a} + \varepsilon^y_{i,a} \qquad (15)$$

$$v_{i,a} = v_{i,a-1} + u^y_{i,a} \qquad (16)$$

[24]The last two terms would be absent in an infinite-horizon setting, but appear in a life-cycle model.

[25]Here our approach is again beneficial. First we only require revisions in expectations as opposed to actual expectations. Second, if survival probabilities and discount rates are (approximately) constant over one period, then their exact values become less important.

From assumption 2 the expectation of future labor income is given by

$$\begin{aligned} E_{a_0}(Y_{i,a}) &= E_{a_0}(\exp(\mu_a^y + v_{i,a} + \varepsilon_{i,a}^y)) && (17)\\ &= \exp Ln[E_{a_0}(\exp(\mu_a^y + v_{i,a} + \varepsilon_{i,a}^y))] && (18)\\ &= \exp\left[E_{a_0}(\mu_a^y + v_{i,a} + \varepsilon_{i,a}^y) + \frac{1}{2}((a-a_0)\sigma_u^2 + \sigma_\varepsilon^2)\right] && (19)\end{aligned}$$

Since μ_a^y is a constant and

$$\begin{aligned} E_{a_0}(\varepsilon_{i,a}^y) &= 0 && (20)\\ E_{a_0}(v_{i,a}) &= v_{i,a_0} && (21)\end{aligned}$$

we are left with

$$E_{a_0}(Y_{i,a}) = \exp(\mu_a^y)\exp(v_{i,a_0})\exp\left(\frac{1}{2}(a-a_0)\sigma_u^2 + \sigma_\varepsilon^2\right) \tag{22}$$

Since we don't observe the permanent component of labor income ($v_{a_0}^i$) we first re-write this expectation as

$$E_{a_0}(Y_{i,a}) = \exp(\mu_a^y)\frac{Y_{i,a_0}}{\exp(\mu_{a_0}^y)\exp(\varepsilon_{i,a_0}^y)}\exp\left(\frac{1}{2}((a-a_0)\sigma_u^2 + \sigma_\varepsilon^2)\right) \tag{23}$$

and assume that the transitory shock equals its unconditional expectation ($\varepsilon_{i,a_0}^y = 0$) hence

$$E_{a_0}(Y_{i,a}) \simeq Y_{i,a_0}\frac{\exp(\mu_a^y)}{\exp(\mu_{a_0}^y)}\exp\left(\frac{1}{2}((a-a_0)\sigma_u^2 + \sigma_\varepsilon^2)\right) \tag{24}$$

Discount rates for future labor are typically small values (see Cocco et al. (2005) or Lustig, Van Nieuwerburgh, and Verdelhan (2013)), so we set R_a^Y equal to 1.02% in real terms, for all a. The survival probabilities are from the National Center for Health Statistics with $A = 100$. Estimates for σ_ε^2 , σ_u^2, and the deterministic income age profile (μ_a^y) are from Cocco et al. (2005). For each individual we use the values that match her education group.[26]

[26] As common in the literature we decrease the estimates of the volatility of the transitory shocks, σ_ε^2, by

3 Data

Since taking first differences is at the core of our empirical approach we require a panel dimension in our data, so we use the Panel Study of Income Dynamics (PSID) for our analysis. The survey starts in 1968, but the first collection of wealth data takes place in 1984 with a repeating the module each five years (i.e. in 1989 and 1994). Since 1997 the PSID became a biannual survey, and the wealth module was included in each wave from 1999, so we take this as the starting date for our analysis. The most recent available survey data is the 2019 wave.

The timing of the data in the PSID is the following: all personal characteristics and wealth data are reported on the date of interview and relate to the survey year. Income reported in the survey year refers to the previous tax year. Wealth and income data are collected at the household level. We proxy that head of household/reference person is the main decision maker in the family unit, and therefore consider his/her age in our analysis.[27]

3.1 Sample Selection

We impose several requirements on the original data. First, as commonly done, we exclude the special SEO subsample, to keep a representative sample of households. This leaves us with 64,933 observations in total for 11 survey waves.[28] Since our methodology relies on taking first differences of the data, we drop "solo observations", i.e. those that are not part of a series of at least 2 consecutive observations. We also exclude data points for which there was a change in marital status of the reference person between t and t-1, and those which indicate new movers, to ensure that changes in assets are not due to the exclusion or addition of a family member. We impose a minimum level of liquid financial assets of $1,000, since for individuals with less than $1,000 the asset allocation decision is not particularly

1/2 to adjust for potential measurement error.

[27]Starting in 2017, the term "Reference Person" has replaced "Head of Household".

[28]The 1999 to 2019 waves do not include the Latino subsample, which is also typically excluded when considering earlier data.

relevant.[29] Likewise, we impose a minimum level of family income, which is set at $5,000, and exclude families who own a farm or a business.

Finally, since our goal is to estimate age profiles, we require a minimum number of observations for each age. Table 1 reports the number of observations at each age after applying the previous filters.[30] For the stock market participation regressions we exclude age changes 19 to 21 and 20 to 22 since we have few observations for those. The risky share sample is smaller since we also have to condition on stock market participation, and therefore we set the initial starting age change for this analysis at 24.[31] In both cases we consider a final age of 85, although with the caveat that the number of observations for the risky share is not very high for these observations.

After applying these filters we are left with 26,861 observations in our sample for the stock market participation regressions. For the risky share analysis, since we condition on stock market participants, and have a starting age of 24, the sample size is 5,094.[32]

3.2 Retirement wealth

The PSID does not include data on asset holdings in DC retirement accounts, such as those in 401(k) plans. Therefore, our analysis only considers non-retirement wealth.

With regards to stock market participation, we don't consider this to be a limitation. Even if we did have information on stock holdings in DC accounts, we would prefer to exclude it. Access to 401k plans is mostly determined by institutional features, namely whether the employer offers such a plan or not, and job choice is naturally dictated by many

[29]To calculate $\Delta PVYW$ and its terms, we further impose a minimum level of total wealth of $1,000 in both current and previous wave, as positive wealth levels are required for the variable construction.

[30]The interviews may take place in different times of the year, mostly between March and November. Hence, between two consecutive waves, the change in age of the individual can be 1, 2 or 3 years. Our empirical strategy assumes a uniform change in age for all agents, so we manually define auxiliary age as age in 1999 plus the step to the next wave (that is if in 1999 the person's age was 35, then in 2001 age it is 37, in 2003 age it is 39, etc)

[31]So the first observation corresponds to the age change from 24 to 26 for the risky share and from 21 to 23 for stock market participation.

[32]In some of our empirical specifications we include change in wealth and/or change in income as explanatory variables. For those we also have to impose the requirement that lagged wealth and lagged income are not missing, which eliminates a few additional observations.

other factors rather than the existing of a DC plan. Probably very few individuals had the option to choose between otherwise identical/similar jobs, but one with a DB plan and the other with a DC plan, particularly since similar jobs in similar industries tend to have the same pension plan structure (DB, DC or hybrid).[33]

For the risky share regressions, having access to retirement wealth data might be beneficial, but even here there are still concerns. Portfolio allocations in 401k plans are influenced by institutional features, namely the menu of investment choices, and the default options, making them less informative about the underlying economic mechanisms driving risk-taking over the life-cycle. Parker et al. (2021) show how the expansion of Target Date Funds, and their increasing importance as default options in 401k plans in the U.S., has significantly affected the portfolio allocation of investors.[34]

3.3 Variable construction

The risk-free asset in our analysis (B) is the cash category in the PSID, which includes money in checking or savings accounts, money market funds, certificates of deposit, government bonds and treasury bills (excluding employer-based pensions or I.R.A.s). Risky assets (S) include shares of stock in publicly held corporations, stock mutual funds, or investment trusts (not including stocks in employer-based pensions or I.R.A.s). We refer to the sum of these two as liquid financial assets (LFA). Stock market participation is defined as equal to 1 if the individual has positive value of S and 0 otherwise.

In our regressions we include four measures of wealth. Liquid financial assets (LFA), housing equity (HE) and two measures of total wealth. The first measure of total wealth (TW) is the sum of liquid financial assets and housing equity, while in the second definition (NTW) we also subtract uncollateralized debt. Housing equity is computed as the difference between house value and mortgage debt. House value is the present value of the home,

[33]This concern does not apply directly to IRAs. But they represent a much smaller fraction of total retirement savings, and individuals with a 401(k) have less need for an IRA, so selection concerns still apply.

[34]In theory, investors could potentially offset these "distortions" by adjusting their portfolio of non-retirement assets, but several households only hold stocks through their retirement account, and for others retirement wealth represents most of their total financial wealth.

including the value of the lot if applicable, while mortgage debt is the total value currently owed on mortgages. Uncollaterized debt, used in our second definition of total wealth, is the sum of credit card debt, student loan debt, medical debt, legal debt, family loan debt.

Total family income is the sum total of taxable, transfer, and social security income of the head, spouse/partner, and other family units. We also use the data on age, employment, marital status, and number of children in the household, and deflate all nominal variables to 2017 using the consumer CPI series from the St. Louis Federal Reserve. Log changes in financial assets, home equity and total wealth are direct computations from the original data, while the ratio of the present-value of future labor income (human capital) to current wealth ($\Delta PVYW_{i,a_0}$) is computed as outlined in section 2.3, where current wealth is total wealth or financial wealth, depending on the empirical specification. Finally, we winsorize all explanatory variables in the different regressions at the 1st and 99th percentiles.

3.4 Summary statistics

Table 2 provides initial summary statistics for our data. Panel (a) reports values for the sample used in the participation analysis (henceforth, full sample), while the values in Panel (b) refer to the sample of stock market participants (henceforth, participation sample), which is used in the conditional risky share regressions.[35] Note that all growth rates and changes are over a *two-year period*, since that is the frequency of our data. We deflate income and wealth/asset data to 2017 dollars.

In the full sample the average age is 49, while for the sample of stock market participants the value is higher (54). The reference person gender in both samples is skewed towards males and roughly 70% of the households in the sample are married couples. 38% (31%) of the full (participation) sample households have children. Family income is smaller in the full sample compared to the participation sample, both in means ($101,557 versus $140,849) and in medians ($85,186 versus $113,927). 77% of the full sample own home, while in the participation sample the share of home owners is 88%. Unconditional average home equity

[35] In the sample, 28% of households are stock market participants, with an average (median) risky share of 65% (73%).

value is also higher in the participation sample (\$122,835 versus \$231,731). The levels of financial wealth and total wealth are higher in the participation sample as well.

In the full sample the average 2-year growth rate of financial assets is 77%, but this reflects a very skewed distribution, with the median being a more modest 23%, corresponding to about 10% per year. For the participation sample those numbers are even smaller, 9% and 7%, respectively for mean and medians.[36] For housing equity the distribution of growth rates is even more skewed, with a mean of 94% (39%) and a median of 11% (5%) over 2-years (in annual terms). The values are again smaller for the participation sample: 63% (27.7%) and 8% (3.9%) for the 2-year (annual) mean and median, respectively. Interestingly, the growth rate of total wealth is much less skewed than the other two, suggesting that, in a given year most individuals either have a high (low) growth rate of financial wealth or a high (low) growth rate of home equity, but rarely both at the same time. The median growth rate of total wealth for the full (participation) sample is 16% (9%), corresponding to 7.7% (4.4%) per-year.

Figure 1 plots wealth as a function of age, for the three wealth categories, financial wealth, home equity and total wealth.[37] Panel (a) reports the mean, while panel (b) reports the median. All three measures of wealth increase with age until retirement. Housing equity accumulation is more pronounced early in life, while financial assets increase more rapidly from age 50 onwards, as households increase their retirement savings. After retirement, financial assets continue to increase, while housing equity falls slightly.

Figure 2 plots the ratio of human capital to total wealth ($\Delta PVYW_{i,a_0}$) as a function of age, as well as its two components, ξ^1_{i,a_0} and ξ^2_{i,a_0}. Panel (a) reports means and Panel (b) reports medians. As previously discussed, the term ξ^1_{i,a_0} captures revisions in the expectation of the ratio going forward, and the lower discount rate applying to future income. This term is always negative and is initially very large, reflecting the rapid increase in wealth at this stage of the life-cycle, but then gradually converges to zero. The second term ξ^2_{i,a_0} captures the reduction in the ratio because of the two years of income that have been earned. Since

[36] These correspond to 4.4% and 3.4% in annual terms.

[37] Naturally these are just descriptive statistics, since we are not controlling for time and cohort effects.

both income and wealth are always positive this term is also positive by construction. It is also particular high early in life, when total wealth is relatively low and hence the ratio is much larger. The difference between these two terms gives us $\Delta PVYW_{i,a_0}$.

4 Stock market participation

We first study the determinants of stock market participation, while in the next section we consider the portfolio allocation of stockholders. We define stock market participation as equal to 1 if the individual has a positive value of risky financial assets, and equal to 0 otherwise. The variable change in participation can then take three values: 1 (stock market entry), 0 (no change of previous status), or -1 (stock market exit), but we also study entry and exit decisions separately.

4.1 Descriptive statistics

We first present and discuss descriptive statistics which provide a background for our empirical analysis, and motivate our approach.

4.1.1 Stock market participation over the time

Table 3 presents summary statistics for participation over time. Our sample starts in 1999, and since we estimate the data in first differences the first observations in our regressions will be for 2001. Column 3 reports the stock market participation rate in each wave of the PSID. The average participation rate during the time span of the 10 waves from 2001 to 2019 is around 28%. This value is comparable to the one obtained in other sample periods and other datasets for direct participation.[38,39]

From 2001 onward the raw series has a clear downward trend. Direct stock market

[38]When including wealth held in retirement accounts, such as those in IRA or 401k plans, stock market participation in the US population is closer to 50%. The PSID however does not include this data.

[39]See Mankiw and Zeldes (1991), Haliassos and Bertaut (1995), Guiso et al. (2002), Guiso et al. (2008), Christelis, Georgarakos, and Haliassos (2013), Badarinza, Campbell, and Ramadorai (2016) for evidence on stock market participation rates across countries.

participation increased during the period of the "tech bubble", and then gradually returned to previous levels. The increasing pattern is visible when comparing the 1999 and 2001 values, while the rest of our sample captures the subsequent decreasing trend. For comparison, Panel a) of Figure 3 plots average stock market participation from our data and from the Survey of Consumer Finances (SCF). Average participation is higher in the SCF, since it over-samples wealthy individuals, but the same pattern is visible in both series. These time effects are an example of the patterns that we want to control for, when estimating age profiles.

Column 4 of Table 3 reports stock market participation in the previous wave, and column 5 reports the difference between columns 3 and 4, i.e. the change in participation. It is important to note that the values in column 4 are not merely the values of the previous row in column 3. For example, average participation in 2001 is 39.71%, but lagged participation in 2003 is 36.08%. This occurs because these are not necessarily the same individuals in each year. Consider, for example, an individual who is present in two consecutive waves, 2003 and 2005. Those two observations will enter the participation statistics for each of those years, but for change in participation we only have one observation, corresponding to the change between 2005 and 2003 (so there will be no observation for the changes from 2001 to 2003 or from 2005 to 2007). This distinction will become even more important particularly later on, as we study the conditional risky share. Even if we had started with a balanced panel of households, the risky share sample would be an unbalanced panel due to the dynamics of stock market entry and exit. Therefore, comparing averages, either across time or across age, is not the same as actually computing the differences at the individual level and then taking the average.

The difference in results is also visible by comparing Panel (b) of Figure 3, which plots the change in participation over time (so column 5 of Table 3) with its panel a). When we measure changes in participation by conditioning on the same individuals, thus controlling for cohort effects, we find some subtle differences relative to the previous results. There are indeed visible decreases in participation in 2003 and 2011, and a clear increase in 2001, but for the rest of the sample period participation has remained fairly constant.

4.1.2 Age profiles

Figure 4 plots average stock market participation over the life-cycle based on simply averages across the full sample.[40] From this we would conclude that stock market participation increases quite significantly with age, almost doubling over the life-cycle, and that the relationship is very close to linear. In Figure 5 we report the same calculation but for each separate PSID wave in our sample. We observe an overall increasing pattern for every cross-section, but these are no longer simple linear fits, and in some cases the relationship is no longer monotonic.

The results in Figure 5 are based on cross-sectional data, so for each series we don't have to worry about time effects, but conclusions can still be influenced by cohort effects. The data is not a representation of the same individuals over time, it captures individuals for the same age at a different time, depending on their cohort. Each individual observation therefore reflects both age, time and the individual's cohort. By taking first differences we eliminate the cohort effects, and then by taking averages across time we average out the time effects. This is what we report in Figure 6 next, and also addresses the concern highlighted in Table 3, about taking differences of means in an unbalanced panel.

In Figure 6 we construct the life-cycle profile of stock market participation by first computing changes in participation at the individual level and then averaging these by age. The life-cycle profile is generated by computing the cumulative changes over age, thus making it comparable to the results reported in Figures 4 and 5.[41] The only difference between the two figures is the scale, since the line in 6 starts from zero. As previously discussed, by taking first differences we are only able to report how participation changes with age, not its absolute value at any given age.

Figure 6 reveals a hump-shaped pattern for stock market participation. The fraction of individuals investing in stocks increases early in life (until the early 40s), remains approximately constant at mid-life (until the late 50s), and eventually decreases, both as individuals

[40] We present the results in 5-year age groups to facilitate the visual interpretation, but all estimations throughout the paper consider each year-age as the unit of analysis, which are then aggregated to 5 years.

[41] As before, we further average these to 5-year age groups to facilitate the graphical exposition.

approach retirement and during retirement itself. Comparing the results in Figures 4 and 5 we see that the increasing profile early in life is very similar, but from then on the two diverge significantly. The previous results suggest that stock market participation continues to increase with age after the early 40s, and even accelerates slightly. This is in sharp contrast with the evidence presented in Figure 6. The comparison between Figure 4 and Figure 6 illustrates the intuition behind our identification approach and its importance.[42] In the next section we consider a more formal regression analysis where we will be able to include additional explanatory variables, which will help us understand better the underlying economic mechanisms.

4.1.3 Stock market entry and exit

In the final subsection of the descriptive statistics we decompose the dynamics of stock market participation in entry and exit decisions separately. We consider two definitions of entry and exit shares. The first definition scales both entry and exit by the total population in that age group:

$$Entry^1_{at} = \frac{\text{number of new stock market participants in age group } a \text{ at time } t}{\text{total population in age group } a \text{ at time } t-1} \tag{25}$$

$$Exit^1_{at} = \frac{\text{number of new stock market nonparticipants in age group } a \text{ at time } t}{\text{total population in age group } a \text{ at time } t-1} \tag{26}$$

The advantage of these two measures is that, since they use the same denominator, they are directly comparable and therefore can be combined to compute the net change in stock market participation at each time. However, following up on the previous literature, we also compute entry and exit shares relative to the fraction of non-participants and participants,

[42] We discuss this more formally later in the paper, by presenting estimation results associated with different identification schemes.

respectively:[43]

$$
\begin{aligned}
Entry^2_{at} &= \frac{\text{number of new stock market participants in age group } a \text{ at time } t}{\text{total stock market nonparticipants in age group } a \text{ at time } t-1} && (27)\\
Exit^2_{at} &= \frac{\text{number of new stock market nonparticipants in age group } a \text{ at time } t}{\text{total stock market participants in age group } a \text{ at time } t} && (28)
\end{aligned}
$$

Figure 7 plots entry and exit shares using the two alternative definitions. Panels (a) and (b) report results using the first classification (equations (25) and (26)), respectively over time and as a function of age, while Panel (c) considers the alternative definition (equations (27) and (28)). When expressed as a fraction of the total population (Panels (a) and (b)), entry and exit shares have fairly similar values. Average entry and average exit over the sample are almost identical, and equal to 8.4%. The former is roughly flat with age, while the exit share shows a mildly increasing pattern.

Turning to Panel (c), where exit and entry shares are computed relative to their corresponding initial populations (participants and non-participants, respectively), we find that the exit rate is much higher than the entry rate. As a function of age the exit share now shows a pronounced decreasing pattern, while the entry share is mildly increasing.

4.2 Regression analysis

4.2.1 Specifications

We first estimate the age-profile of the stock market participation decision using the approach outlined in section 2. More precisely, we consider equation (5), but with the left-hand-side variable being the change in stock market participation (ΔP_{it})

$$\Delta P_{it} = \overline{I}(\Delta a_{it}) + \widetilde{I}(\Delta t) + \theta \Delta X_{it} + u_{it} \tag{29}$$

The right-hand-side variables will vary across specifications, as described below. Since participation is a binary variable, the left-hand-side variable can only take values of 1, 0 or -1.

[43]These are the measures reported in Brunnermeier and Nagel (2008) and Fagereng et al. (2017).

In our main analysis we estimate equation (29) using a linear regression model to facilitate the interpretation of the coefficients. Later in the paper we estimate separate probit models for entry and exit decisions.

In the first specification the vector ΔX_{it} includes only the age dummies and two control variables, change in homeownership status and change in the number of children in the household. In the second specification we study wealth effects by including changes in total wealth, and in specifications three and four we further examine the effect of changes in the ratio of the present-value of human capital to total wealth.[44] The estimation results are shown in Table 4. We do not include the coefficients on the age dummies as otherwise the table would become too large. Instead we report the implied age-profiles in Figure 8. The standard errors in all regressions are cluster-robust where cluster is a household identifier.

4.2.2 The age profile of stock market participation

The first specification confirms, in a formal regression setting, the descriptive results presented in Figure 6. As shown in Figure 8, participation is a hump-shaped function of age, increasing until the 40s, then remaining flat at mid-life, and finally decreasing from age 55 onward, as individuals approach retirement and further still during retirement itself.

The increase in participation in the early stage of the life cycle is about 5 percentage points, which is economically significant when compared with the average participation rate of 28.44%. The decrease late in life is even larger, such that the participation rate from the mid 70s onward is lower than for our starting age group (mid 20s). Our results are consistent with the findings of Poterba and Samwick (1997) and Fagereng et al. (2017) who also estimate a hump-shaped life-cycle profile for stock market participation. Catherine (2019) also documents an increasing pattern early in life but constant thereafter, and this is also the result obtained by Ameriks and Zeldes (2004) when using cohort dummies in their estimation. By contrast, when using time dummies, Ameriks and Zeldes (2004) estimate a decreasing profile late in life, same as in our paper, but their results show a flat stock market

[44]Here we consider the first definition of total wealth discussed in the Data section (TW). Results with the alternative definition (NTW) are reported in the appendix, and they are qualitatively identical and quantitatively very similar.

participation at young ages.

The increasing pattern early in life indicates that stock market entry costs are likely important in explaining participation decisions, while the exits late in life suggest that per-period costs are probably also needed to match the empirical evidence. The latter result is also consistent with models where background risks are particularly important late in life, namely health shocks and/or medical expenditure risk, as in Hubbard et al. (1995), De Nardi et al. (2010), Ameriks et al. (2011), Yogo (2016) or Koijen et al. (2016). These risks decrease the optimal risky share, as shown empirically by Guiso et al. (1996), making it less likely that households would be willing to pay a per-period cost to remain as stockholders.

4.2.3 The roles of wealth and human capital

Specifications 2, 3 and 4 reveal that (total) wealth is an important determinant of stock market participation. The coefficients on log changes in wealth are always positive and statistically significant, with p-values less than 0.000. The estimated coefficients in the more complete specifications (3 and 4) imply that a 10% increase in total wealth over 2-years leads to an increase in the probability of stock market participation over the same period of 1.1 percentage points.[45] These effects are economically important when compared with an average participation rate of 28.44% in the sample.

The importance of these wealth effects is highlighted in Figure 8. The age-pattern of stock market participation changes quite significantly from specification 1 to specification 2, where it becomes essentially flat over the life cycle. This indicates that, conditional on having the same change in wealth, a 50-year old household is just as likely to become a new stock market participant (or to exit the stock market) as a 30-year old, for example. In other words, the increase in participation revealed in specification 1, can be fully explained by the observed changes in wealth at this stage of the life-cycle. These large wealth effects provide support for participation costs being a key determinant of stock market participation decisions, consistent with results in Vissing-Jørgensen (2002), Haliassos and Michaelides (2003)

[45] The dependent variable can take the values of 1, 0 or -1, so the coefficient captures the increase (decrease) in the probability of stock market entry (exit).

and Gomes and Michaelides (2005). Moreover, the hump-shape pattern in specification 1, and the large documented number of stock market exits, are consistent with per-period participation costs also playing an important role.[46]

In specification 3 we introduce the change in the ratio of human capital to total wealth ($\Delta PVYW_{i,a_0}$) as an additional explanatory variable. The estimated coefficient is positive and highly statistically significant. Individuals who have experienced an increase in this ratio are more likely to become stock market participants. This is consistent with life-cycle models with uninsurable labor income risk, where labor income is a close substitute for bonds. Since we are already controlling for changes in wealth, this ratio affects the participation decision indirectly through its impact on the optimal equity allocation.[47] As individuals age this ratio typically falls, and therefore their optimal equity allocation decreases, reducing the incentive to participate in the stock market. In specification 4 we decompose $\Delta PVYW_{i,a_0}$ in the two components, ξ^1_{i,a_0} and ξ^2_{i,a_0}. The coefficient on ξ^2_{i,a_0} is not statistically significant but the coefficient on ξ^1_{i,a_0} is significant and positive, again consistent with the notion that labor income is a closer substitute for bonds than for stocks.

The role of human capital is also visible in Figure 8, by comparing the age patterns resulting from specification 3 with the other two.[48] Under this specification we obtain a strongly decreasing age profile. This implies that, conditional on the same change in wealth, and the same change in the ratio of human capital to total wealth, older individuals are less likely to participate in the stock market. This is consistent with a model with stock market entry costs, since older individuals expect to have less years over which to amortize that cost. It is also consistent with a combination of per-period participation costs and increased background risks at retirement.

Comparing the age profile from specification 1 with the other two, we conclude that the increase in participation early in life is largely explained by the increase in financial wealth over this period and, to a lower extent, also by the reduction in the present-value of

[46]Entry costs alone can generate stock market nonparticipation, but not stock market exit decisions.

[47]For that reason we also postpone the discussion of the economic magnitude of the estimates to the section on the conditional risky share.

[48]Specification 4 produces a very similar pattern hence it is excluded from the figure to facilitate the comparisons.

human capital. The decrease in participation from age 60 onwards is largely driven by the reduction in human capital, as can be seen from specification 3. In fact, from age 70 onwards, this effect dominates the wealth channel which would predict a constant participation rate (specification 2).

4.3 Robustness of wealth effects: non-linear effects and decomposing total wealth

In Table 5 we estimate separate regressions for each wealth quartile, to explore potential heterogeneity in the effects by wealth levels. We find that the coefficient on total wealth is highly statistically significant in all regressions. Interestingly the point estimate increases quite noticeably across wealth quartiles, again consistent with the importance of participation costs. For a poor household a 10% increase in wealth, for example, might not be enough to push her over the participation threshold. On the other hand, for a rich household, such an increase represents a much more significant change in the dollar value of her portfolio. Likewise, the coefficients on the human capital variable also increase with wealth, and in this case the coefficient for the first quartile is not even significant. The interpretation is similar. Individuals with low wealth are sufficiently below the participation threshold, such that changes in the optimal portfolio are often not enough to justify paying the fixed costs.

In Table 6, we consider financial assets instead of total wealth, and also explore the role of home equity (specification 6). Interestingly, when we decompose changes in total wealth into changes in financial assets and changes in housing equity, only the former is statistically significant, which is why we do not include the home equity variable in specifications 7 and 8. The conclusions are similar to the ones obtained in specifications 2 to 4. The coefficient on (log differences in) financial assets is highly significant (both economically and statistically), with the point estimate even slightly higher than the one previously obtained for (log differences in) total wealth. The coefficient on (log differences in) the ratio of human capital variable to financial wealth is also statistically significant and positive, as in the previous results.

4.4 Assessing the role of time and cohort effects

Our methodology allowed us to identify the life-cycle profile of stock market participation without having to make specific assumptions on time and cohort effects. In this section we compare our results with those obtained under the two extreme identification assumptions of setting either time effects or cohort effects equal to zero (as in Poterba and Samwick (1997) and Ameriks and Zeldes (2004)).

More precisely we estimate equation (1) for the participation decision,

$$P_{it} = I(a_{it}) + I(t) + I(c_i) + \theta X_{it} + \Gamma F_i + \varepsilon_{it} \tag{30}$$

with either $I(t) = 0$ for all t, or $I(c_i) = 0$ for all i. Under both of these assumptions the age effects become identified, so we don't need to estimate the equation in first differences.

The age profiles of stock market participation implied by these two specifications are shown in Figure 9, where we can see that the results are very different under the two identification schemes. The age pattern of participation is decreasing after age 31 if we rule out time effects, while when we ignore cohort effects it becomes strongly increasing. Both of these are in contrast with our baseline results (Figure 8) which identify a hump-shape pattern, highlighting the importance of properly controlling for both cohort and time effects when estimating the age profile of stock market participation.

4.5 Probits for entry and exit

Participation is a binary variable, so its first difference can only take three values. In our main analysis we estimated linear regressions to facilitate the interpretation of the coefficients. Now we re-examine our conclusions in the context of probit models. The results are shown in Tables 7 and 8 for entry decisions and 9 and 10 for exit decisions. Here we consider the more standard classification of entry and exit shares, given by equations (27) and (28).[49]

The conclusions are very similar to the ones previously obtained. Consistent with models

[49] Results with the alternative definitions are reported in the Appendix 1 and deliver the same conclusions.

of participation costs, and with the results in Brunnermeier and Nagel (2008), changes in wealth, and in particular changes in financial wealth are positively (negatively) associated with stock market entry (exit). Changes in the ratio of human capital to total wealth are also positive (negatively) associated with stock market entry (exit), consistent with models where labor income is a close substitute for the riskless asset.

5 Conditional Risky share

Having established the evolution of the participation decision over the life-cycle, we now turn our attention to the portfolio decision of stock market participants. The object of interest is the conditional risky share, i.e. the risky share conditional on stock market participation.

5.1 Descriptive statistics

Figure 10 plots the average cross-sectional conditional risky share across time (Panel (a)) and across age groups (Panel (b)). Time variation in the conditional risky share is fairly moderate, and much smaller than previously reported for the participation rate (Figure 3). In Panel (b) we find that, in the absence of any controls, the conditional risky share is a positive function of age until retirement, and essentially flat after retirement. The increasing pattern before retirement is economically large, with a rise of 20 percentage points in the average allocation to risky assets over this period. However, as previously discussed, the data in Figure 10 combines age, time and cohort effects.

Repeating the same approach as for participation, we isolate age effects by first taking first differences of the conditional risky share for each individual at each date, and then averaging across all observations for the same age. The result, reported in Figure 11, reveals an increasing profile until around age 55 and flat after that. As in Figure 10 we observe an increasing pattern early in life, but now it peaks before retirement. Furthermore, the magnitude of this increase is about one third smaller in Figure 11: 13 percentage points versus 20 in the previous results.

5.2 Estimation results

We now consider a formal regression framework, by estimating equation (5) for the conditional risky share. With regards to the explanatory variables included in the vector ΔX_{it} we consider the same specifications as for the participation decision. So, the first regression includes the age dummies and two controls, change in home-ownership and change in number of children. In the second regression we add log changes in total wealth and the third and fourth we further include the human capital variables. The estimated coefficients are reported in Table 11 while Figure 12 shows the corresponding age effects. In the same way as for stock market participation, we also consider specifications with financial wealth instead of total wealth, and those results are reported in 12.

Finally, as discussed in Section 2.2, a significant fraction of households only rebalances their portfolios infrequently. Therefore we repeat our estimations following the approach in Calvet, Campbell and Sodini (2009) and considering only active rebalancing decisions. More precisely, we replace changes in the risky share with active changes in the risky share, computed from equations (6) and (7). In our implementation of these equations the return on risky asset (R_t) is the value-weighted return on equities from CRSP, while the return on the riskless asset (R_f) is the real return on the 90-Day T-Bill. These results are reported in Tables 13 and 14 and Figure 13.

5.2.1 Age profiles

Figure 12 plots the age profile of the conditional risky share implied by the two specifications for the left-hand-side variable, namely total changes and active changes only. When considering total changes we obtain an initially increasing and then flat profile, but when we focus on active changes only the conditional risky share is constant until age 55-60 and then decreases as agents approach retirement and again during the retirement period. The difference in results indicates that the increase in the conditional risky share early in life is actually the result of portfolio inertia combined with a positive realized risk premium. Household portfolios are shifting towards equities not because of active rebalancing decisions, but simply because the value of their stock holdings is increasing faster than the value of their bond holdings.[50]

The evidence for a flat or even increasing risky share early in life is consistent with models with decreasing relative risk aversion (as in Gomes and Michaelides (2003), Polkovnichenko (2007) or Wachter and Yogo (2010)), background risks, and where labor income is close substitute for bonds. Background risks and decreasing relative risk aversion (DRRA) generate the flat or increasing pattern early in life. This pattern can also arise in the presence of liquidity demands (as in Campanale et al. (2015)) and/or large skewness in risky asset returns (as in Fagereng et al. (2017)). Finally, the substitutability of human capital and bonds delivers the decreasing profile as individuals approach retirement and during retirement. The presence of DRRA preferences explains why the reduction in the risky share late in life is much weaker than predicted by models with constant relative risk aversion.

The previous evidence could also be consistent with labor income being a close substitute for stocks and increasing relative risk aversion. The former would deliver the flat/increasing profile for young individuals, with the latter being the dominating effect late in life. However, as discussed below, regressions specifications 2 through 8, with the wealth and human capital variables, help us to disentangle between these different economic mechanisms. The coefficients on wealth and human capital in the conditional risky share regressions are consistent with decreasing relative risk aversion and human capital being a close substitute for bonds, respectively, and this was already the case in the stock market participation regressions.

Models with CRRA preferences and labor income uncorrelated with stock returns imply a pronounced decreasing risky share with age. Preferences with DRRA and/or background risks can deliver the flat/increasing pattern early in life, as discussed. But the mild decrease in the risky share late in life suggests that the substitutability between labor income and bonds, although positive, is not as strong as predicted by the model where labor income shocks are completely uncorrelated with stock returns. In principle this could be evidence for strong DRRA, but our regressions indicate that this channel is fairly moderate. Therefore,

[50] It is, of course, theoretically possible that households do not rebalance their portfolios to offset this effect, because this shift in their portfolio mix matches their optimal age profile.

our combined results suggest that there is some correlation between labor income and stock returns, but not too high.[51] This could arise from the cointegration channel proposed by Benzoni et al. (2007), but weaker than under their calibration, such that human capital is still a closer substitute for bonds than for stocks. Alternatively, it can result from the higher-order correlations identified in Catherine (2019).

Our results are closest to the ones obtained by Ameriks and Zeldes (2004) when considering time effects. They find a mildly decreasing risky share late in life, and a flat profile before that. For the pre-retirement period Catherine (2019) estimates a mildly increasing conditional risky share until retirement, while Ameriks and Zeldes (2004) estimate a strong increasing pattern when controlling for cohort effects. Neither of these estimations delivers a decreasing risky share late in life, and the results for earlier ages are also at odds with our evidence, particularly when considering active rebalancing decisions. Finally, Fagereng et al. (2017) estimate that the conditional risky asset share is a decreasing function of age.[52]

5.2.2 The role of wealth

The estimated coefficients of log changes in wealth are almost identical in the regressions for total changes in risky share and in the regressions for active changes only, respectively the ones reported in Tables 11 and 12 and the ones in Tables 13 and 14. Therefore, in our discussion, we only refer to the active share results to avoid repetition. The coefficients on log differences in total wealth (specifications 2 to 4) and log differences in financial assets (specifications 5 to 8) are highly significant and their coefficients are also very similar across the different regressions. Qualitatively these results are essentially identical to those obtained for the stock market participation regressions, providing additional supporting evidence for decreasing relative risk aversion: wealthier investors hold riskier portfolios.

The economic magnitude of the wealth effects is fairly moderate. For example, the

[51] Consistent with this, as discussed next, we find that the coefficient on human capital is not significant in some regression specifications.

[52] Poterba and Samwick (1997) don't estimate a conditional risky share but they obtain very similar hump-shaped age-profiles for both stock market participation and the unconditional risky share. These results therefore suggest that the conditional risky share is approximately constant with age.

estimated coefficient of 0.0711 in specification 2 of Table 13, implies that that a 10% growth in total wealth leads an increase in the conditional risky share by 0.711 percentage points. On the other hand, when we compare the age profile from specification 1 with specification 2 (in Figure 13), we find more significant differences.[53] This apparent contrast can be understood from the substantial cross-sectional differences in the growth rate of wealth. As shown in Table 2 Panel (b), the 25th and 75th percentiles of the growth rate of total wealth are -18% and 41% respectively. In other words, although moderate changes in wealth have a small impact on the risky share, in the data we often observe very large fluctuations in wealth, and hence significant portfolio reallocations as a result. This contrast might help explain why previous literature has found conflicting results when testing for wealth effects in portfolio demands.[54] These effects will be hard to detect when considering modest changes in wealth.

Finally, and just as we found in the participation decision regression, the coefficient on log changes in home equity is not statistically significant (specification 6 in Table 12), and the point estimate is also very small.

5.2.3 The role of human capital

The coefficient on changes in the ratio of human capital to total wealth in specification 3 is positive, consistent with models where labor income is a close substitute for bonds.[55] When decomposing this variable in its two components, ξ^1_{i,a_0} and ξ^2_{i,a_0}, we find that only the estimated coefficient on the first term is statistically significant, just as in the stock market participation regressions.

The point estimate for the PVYW coefficient in specification 3 in Table 13 implies that a 1 standard deviation increase in this ratio leads to an increase in the risky share by 1.58 percentage points. This significant but moderate role of human capital is also visible in

[53]As before, we omit Specifications 4 to 8 in Figure 13 to facilitate the visual comparison.

[54]Heaton and Lucas (2000), Campbell (2006), Wachter and Yogo (2010), Calvet and Sodini (2014), Bach et al. (2016), Meeuwis (2020) and Fagereng et al. (2020) find evidence for increasing risky shares as a function of wealth, although the estimated effects are relatively small in several cases. On the other hand, the estimates in Brunnermeier and Nagel (2008) and Chiappori and Paiella (2011) find that the wealth effects in the risky share of households are not statistically significant.

[55]As before, the conclusions are identical if we consider total changes in risky share or active changes only, and therefore we only discuss the results for the second case.

Figure 13 when comparing the results from specifications 2 and 3. This again suggests that, although human capital is a closer substitute for bonds than stocks, this substitutability is not as high as in a model where labor income and stock returns are assumed to be uncorrelated. Along these lines, the estimate for the PVYW variables in specification 7 in Table 14, although also positive, is not statistically significant.

5.3 Non-linear wealth effects

If households have a strong demand for liquidity, for example due to transaction costs or precautionary savings, this could generate the positive correlation between wealth and conditional risky share, as these precautionary savings and liquid assets should be primarily invested in safe assets. For less wealthy households those savings represent a higher fraction of their total wealth, hence the impact on their optimal risky share would be larger. Consequently, when households become wealthier (poorer), we would observe an increase (decrease) in their conditional risky share.

We explore this possibility, and more generally analyze whether the previous results change with wealth, by estimating specifications two and three by total wealth quartiles. The regression results are reported in Table 15. Interestingly, the wealth coefficient is very similar across all eight specifications. This is consistent with the DRRA interpretation and suggests that the liquidity and precautionary savings channels, if present, are quantitatively small. One possible explanation for the later implication is that, most households with a higher demand for bonds due to these two channels are not even stock market participants. Given to their low demand for stocks they do not have an incentive to pay the participation costs.

5.4 Assessing the role of time and cohort effects

Repeating our exercise for stock market participation, we now compare the results for the conditional risky share with those obtained under the identification assumptions of setting either time effects or cohort effects equal to zero. More precisely, we estimate regression (1) with either $I(t) = 0$ for all t, or $I(c_i) = 0$ for all i.

The age profiles of the conditional risky share implied by these two specifications are shown in Figure 14. We find that the results are significantly different under the two identification schemes, same as for the stock market participation regressions. When controlling for cohort effects only we conclude that the conditional risky is an inverse hump-shape function of age. On the other hand if we only include time effects, we estimate an increasing profile. Both of these are qualitatively very different from our baseline results (specification 1 in Figure 13), again re-emphasizing the importance of an estimation approach that controls for time, cohort and age simultaneously.

6 Life-Cycle Model

In this section we present a relatively simple life-cycle model which captures most of important economic mechanisms identified in our empirical analysis. There are certainly additional features that we could have included, or that could have replaced some of the ones that we have chosen, while still delivering similar results. We discuss these options when presenting the set-up of the model, but the objective here is to present a framework with a minimum set of features that can closely replicate our empirical findings, while preserving the important economic channels. Making the model more complicated along some dimensions could potentially bring its results closer to our empirical estimates, and we discuss these later.

Finally, we abstract from multiple sources of risk that are particularly relevant late in life, such as longevity risk (e.g. Horneff, Maurer, Mitchell, and Stamos (2009), Cocco and Gomes (2012) and Yogo (2016)), medical expenditure risk (e.g. De Nardi et al. (2010) and Ameriks et al. (2011)), and health risks (e.g. Yogo (2016)).[56] For this reason we focus on the asset allocation and stock market participation decisions before retirement only.

[56]Having a realistic representation of these risks would be challenging given that we consider a utility function with an exogenous subsistence level, as described below.

6.1 Model Set-up

6.1.1 Labor income process and retirement income

Households have a stochastic finite horizon, divided into two periods: working-life and retirement. Before retirement they earn labor income subject to undiversifiable shocks, and after retirement they receive a fixed pension.

The estimated age profile for the risky share suggests that background risks are likely to be important early in life.[57] Therefore the labor income process before retirement extends the standard permanent and transitory combination, by considering a separate unemployment state.[58] Specifically, households suffer an unemployment spell with probability π^u, in which case they receive a replacement income (Y^u), while with probability $1-\pi^u$ their labor income is given by

$$Ln(Y_t) = f(t,\theta) + p_t + u_t \tag{31}$$

$$p_t = p_{t-1} + z_t,\ \ z_t \sim N(0,\sigma_z^2) \tag{32}$$

$$u_t \sim N(0,\sigma_u^2) \tag{33}$$

where $f(t,\theta)$ is a deterministic function of age and other household characteristics (θ), namely education. For convenience, we refer to the current level of permanent income as

$$Y_t^P = \exp(p_t)\exp(f(t,\theta)) \tag{34}$$

Income in a year with an unemployment spell is specified as a deterministic function of the current level of permanent income

$$Y^u = \lambda^u Y_t^P \tag{35}$$

Households retire at age K, and retirement income is a deterministic function of perma-

[57] A similar result could be obtained by replacing this additional background risk with negative return skewness instead, as in Fagereng et al. (2017).

[58] For an alternative formulation to capture time-varying downside risk see Catherine (2019) and Shen (2019), which considered the income process estimated by Guvenen, Ozkan, and Song (2014).

nent income in the last year of working life:

$$Y_t = \lambda^R Y_K^P, \; t > K \tag{36}$$

where λ^R is the retirement replacement ratio.

6.1.2 Preferences

Households have Epstein-Zin utility functions (Epstein and Zin, 1989) defined over the consumption of a single non-durable good (C_t), but with a consumption floor ($\overline{C}_t$):[59]

$$V_t = \{(1-\beta)(C_t - \overline{C}_t)^{1-1/\psi} + \beta E_t \left[\pi_t^s [V_{t+1}^{1-\gamma}]\right]^{\frac{1-1/\psi}{1-\gamma}}\}^{\frac{1}{1-1/\psi}} \tag{37}$$

where ψ is the elasticity of intertemporal substitution, β is the subjective discount factor and π_t^s is the conditional survival probability from age t to age $t+1$.[60] The parameter γ determines atemporal risk aversion, and would be equal to the coefficient of relative risk aversion if we had $\overline{C}_t = 0$.

The consumption floor ($\overline{C}_t$) generates decreasing relative risk aversion in the model, consistent with our empirical evidence.[61] In line with the habit formation interpretation, we scale the consumption floor by the permanent level of income

$$\overline{C}_t = \overline{C} Y_t^P \tag{38}$$

This formulation guarantees that $\overline{C}_t$ increases with age, as consumption also rises.

[59] Since our focus is on the pre-retirement period we do not consider a bequest motive to simplify the model and reduce the number of free parameters.

[60] The conditional survival probability is equal to zero at a pre-determined maximum age, which we calibrate to 100.

[61] In the models of Gomes and Michaelides (2003) and Polkovnichenko (2007) decreasing relative risk aversion arises from habit formation preferences. Wachter and Yogo (2010) obtain decreasing relative risk aversion from non-homothetic preferences over multiple goods, a basic good and a luxury good. The consumption floor can also represent consumption commitments, as in the work of Chetty and Szeidl (2007).

6.1.3 Financial Assets, Participation Costs and Budget Constraint

Households can invest in a riskless one-period bond, and in an aggregate stock market index. The return on the riskless bond (R_f) is constant, while the return on the stock market follows a normal distribution, and is potentially correlated with the labor income shocks as discussed in the calibration section. Investing in the stock market requires the payment of both a first-time entry cost (F^0) and a per-period cost (F^1). As standard in the literature (e.g. Gomes and Michaelides (2005)), these costs are expressed as a fraction of permanent income. This is done both for tractability and because they are partially capturing an opportunity cost of time.

Letting W_t and α_t denote, respectively, wealth and the risky share at time t, the household's budget constraint is

$$W_{t+1} = (\alpha_t R_{t+1} + (1-\alpha_t)R_f)(W_t - C_t - F^0 Y_t^P I^{firstentry} - F^1 Y_t^P I^{\alpha>0}) + Y_{t+1} \tag{39}$$

where $I^{firstentry}$ is a dummy variable that is equal to 1 if the household is participating in the stock market for the first time, and $I^{\alpha>0}$ is a dummy variable that is equal to 1 if the household has positive stock holdings this year.

6.1.4 Household heterogeneity

Following Gomes and Michaelides (2005), we consider two groups of households with heterogeneous savings motives. This is important for matching limited participation with moderate participation costs. One group of households has a low discount factor and high retirement replacement ratio, and as such they will have low wealth accumulation over the life cycle. In the presence of stock market participation costs, most of those households will not have an incentive to become stockholders. By contrast, the other households have a high discount factor and a lower retirement replacement ratio, and consequently most of them will be regular stock market participants.[62]

[62] More generally we could consider many forms of heterogeneity in preferences and income profiles. This a simple two point-distribution that captures high savers and low savers, but several other combinations

In addition, we also consider heterogeneity in financial literacy building on the work of Lusardi and Mitchell (2011) and Lusardi and Mitchell (2014). Some households have high financial literacy and therefore their stock market participation costs are very low, while for others those costs are more significant. This heterogeneity will be important for matching our empirical evidence on stock market participation.

6.2 Empirical counterparts to the model predictions

Our previous estimates are based on PSID data which does not include retirement wealth. Therefore, we now discuss how we adjust our empirical results to make them more directly comparable with the predictions of the model. Crucially, as shown below, the different adjustments that we consider imply very small changes to our previous results.

First, we consider our estimates of the cumulative changes in stock participation. Here we use data from the Survey of Consumer Finances (SCF) to scale our empirical estimates. More precisely, we first use the SCF to compute the ratio of total stock market participation to direct stock market participation (i.e. excluding retirement accounts), and obtain the value of 1.45. The series "Empirical 1" in Figure 15 is then obtained by applying this factor uniformly to our previous estimates.

One potential concern with the previous adjustment is that it assumes that participation in DC accounts follows the same pattern as direct stock market participation, which is unlikely. For example, in the former, we should expect a much slower decreasing pattern in participation late in life. Therefore, we also consider a second adjustment where we assume that only direct stock market participation decreases late in life, and adjust the scaling factor for the previous ages accordingly. The corresponding profile is the series "Empirical 2" in Figure 15, and can see that the two life-cycle profiles are actually very similar.

Next, we consider the cumulative changes in the risky share. Our first scenario takes the conservative assumption that the behavior of the risky share with age is identical in both the tax-deferred retirement accounts (TDAs) which we do not observe, and the liquid taxable

could deliver the same result.

accounts (TAs) that are included in our data. Under this assumption, there is no adjustment required. Therefore, the series "Empirical 1" in Figure 16 , is identical to the one reported before.

In our second scenario we take into account that, in the later years of our sample, a significant fraction of retirement wealth is being invested in Target-Date Funds (TDF) which have a strong age-profile for their risky share. Therefore, denoting by θ the fraction of wealth in tax-deferred accounts that is invested in Target-Date Funds (TDF), the adjusted risky portfolio share becomes:

$$\alpha^{adjusted} = \frac{\alpha^{TA}W^{TA} + \alpha^{TA}W^{TDA}(1-\theta) + \alpha^{TDF}W^{TDA}\theta}{W^{TA} + W^{TDA}} \tag{40}$$

where W^{TA} and W^{TDA} are, respectively, total wealth in taxable and tax-deferred accounts, and α^{TDF} is the risky share in the Target-Date Fund.

To implement this adjustment we use data from the SCF to compute the ratio of W^{TDA} to W^{TA} at each age, we obtain the fraction of TDA wealth invested in TDFs (θ) from the Vanguard surveys, "How America Saves" (Vanguard, 2009, 2020), and we take the portfolio allocation of the target date fund (α^{TDF}) from Vanguard's website. The implied cumulative change in the risky share is the series "Empirical 2" in Figure 16 below.

6.3 Calibration

6.3.1 Income process and preferences

As standard in the literature we set the starting age to 20, the retirement age to 65 and the maximum age to 100. The deterministic income profile and the income retirement replacement ratio are taken from Cocco et al. (2005). For the variances of the incomes shocks we consider the values in Brown, Fang, and Gomes (2015) for their more recent sample period ($1991-2011$), since it matches more closely with the sample period in our empirical analysis.[63]

[63] Their estimate for the standard deviation of transitory income shocks is 28.1%. We consider 20% to take into account for potential measurement error.

The probability of unemployment (π^u) represents the probability of suffering an unemployment spell throughout the year, and therefore the corresponding income includes both unemployment subsidy received during the unemployment spell and labor income earning during the rest of that year. Estimates for these values are taken from Brown et al. (2015), giving us $\pi^u = 0.14$ and $\lambda^u = 0.7$.

As previously discussed we have two equal-sized group of households with different discount factors and retirement replacement ratios. The more patient households have $\beta = 0.99$ and $\lambda^R = 0.68$, while the others have $\beta = 0.9$, and $\lambda^R = 0.9$. For both groups we set the parameter γ equal to 6 and the EIS (ψ) equal to 0.5. The retirement replacement ratios reflect the range reported in Cocco et al. (2005), while the preference parameters are fairly standard and in line with those estimated in Calvet, Campbell, Gomes, and Sodini (2019).[64]

We set the consumption floor parameter ($\overline{C}$) to 0.5, corresponding to 50% of current permanent labor income. This delivers only a moderate increase in risk aversion relative to the power utility case, and a moderate decreasing relative risk aversion, consistent with our empirical estimates. All households receive an initial wealth endowment equal to 50% of their age-20 income.

6.3.2 Returns and correlation with income

We set the real riskless rate to 1.5% and the equity premium to 4%. The correlation between stock market returns and permanent and transitory labor income shocks, are both set to 0.15. Empirical estimates of these correlations are essentially equal to zero (see Campbell, Cocco, Gomes, and Maenhout (2001) and Davis and Willen (2013)). These values are meant to capture, in a reduced-form the higher-order correlations estimated in Catherine (2019) and the low-frequency correlations suggested by Benzoni et al. (2007). Nevertheless, the values that we consider are not very high because our empirical results indicate that human capital is a closer substitute for bonds than for stocks.[65]

[64]The low discount rate value for the second group can also be interpreted as a reduced-form for hyperbolic discounting (see Laibson, 1997).

[65]In addition, models with high correlation between labor income and stock returns imply a strong (counter-factual) increase in the risky share at retirement, when the correlation effect suddenly disappears. Even though we focus here on the pre-retirement period, we prefer not to consider a calibration with that

6.3.3 Participation costs and financial literacy

We consider heterogeneity in financial literacy, captured by heterogeneity in the stock market participation costs. We have three groups of households with different levels of financial literacy. The more financially educated group represents 40% of the population, while the others capture 30% each. For simplicity we assume that financial literacy is uncorrelated with preferences, i.e. the two sources of ex-ante households heterogeneity are independent.

For the group of households with high financial literacy we set $F^0 = F^1 = 0$, for simplicity. A second group has lower financial literacy and faces an entry cost (F^0) of 3% of annual income and a per-period cost (F^1) of 0.5%. Finally, there is a third group which faces very high participation costs: $F^0 = 25\%$ and $F^1 = 2.5\%$. These very high values should be interpreted as a reduced form for other concerns that make individuals extremely unwilling to invest in stocks.[66] We could exclude this group from our model and still match overall participation rates, as done in Gomes and Michaelides (2005) for example. However, under such formulation it would be hard to prevent most households from temporarily becoming stockholders around retirement, and that would be inconsistent with our empirical estimates (see Figure 8). Modelling directly some of those other features (e.g. low trust in the stock market, pessimistic beliefs or first-order risk aversion) would replace the need to include this group with these high costs of participation.

6.4 Results

6.4.1 Baseline model

We first solve and simulate the model for the different groups of agents. We then compute the model-implied counterparts to our empirical life-cycle profiles of the risky share and stock market participation, and compare the two (model-implied and data) in Figure 15 and Figure 16, respectively. In addition to our baseline results (model version: "Baseline"), we

implication.

[66]For example, lack of trust in financial markets (Guiso et al., 2008), pessimism about expected returns, or preferences with first-order risk aversion (see, for example, Chapman and Polkovnichenko (2009), Campanale (2011), Peijnenburg (2018) or Pagel (2018)).

also plot results for a case without ex-ante household heterogeneity (model version: "NH"), and for a case where we also set the consumption floor to zero (model version: "NH_$\overline{C}$0").[67]

Figure 15 shows that the baseline model (series "Baseline") matches very well the life-cycle pattern of stock market participation. In the model stock market entry takes place slightly earlier than in the data and there is an additional increase just before retirement, when household wealth accumulation is at its maximum, but for all other ages the fit is extremely good. By comparison, the model with no ex-ante heterogeneity (series "NH") fails quite dramatically. Unless we include a group of households with a low savings motive, eventually all households will decide to participate in the stock market, as shown by Gomes and Michaelides (2005).[68] Extremely high participation is also obtained when, in addition to excluding ex-ante heterogeneity, we also remove the consumption floor (series "NH_$\overline{C}$0"), except that now all households optimally become stockholders from a very young age. Therefore, by age 23 the participation rate is already almost 100% and there is no further cumulative increase reported in Figure 15.

In Figure 16 we plot the cumulative change in the risky share. The baseline model (series "Baseline") delivers a moderate decreasing pattern over the life-cycle, in contrast to the steep decreasing profiles of the standard model without heterogeneity and with constant relative risk aversion (series "NH_$\overline{C}$0"). The cumulative decrease in risky share over working life is very similar in the baseline model and in the data, particular if we consider the series "Empirical 2", although in the former the reduction in the equity share takes place slightly earlier than in the data. Including decreasing relative risk aversion plays an important role in bringing the theoretical predictions close to the data. Without the consumption floor the optimal risky share decreases quite significantly with age (series "NH_$\overline{C}$0"), while if we keep the consumption floor and remove ex-ante heterogeneity instead, the results are very similar (series "NH" and series "Baseline"). Although this last set of results (series "NH") matches

[67] In the two cases without ex-ante heterogeneity we set the common discount factor to 0.96, the common retirement replacement ratio to 0.8 and the unique level of financial literacy to that of the "medium group" ($F^0 = 3\%$ and $F^1 = 0.5\%$). The results are not particularly sensitive to modest deviations around these values.

[68] Unless we assume very high costs of participation. In this version without ex-ante heterogeneity every households faces the same (moderate) costs $F^0 = 3\%$ and $F^1 = 0.5\%$.

the cumulative change in risky share slightly better, it fails quite dramatically on the stock market participation decision, as previously shown in Figure 15.

6.4.2 Additional comparative statics

To further understand the role of the different features of the model, we consider two additional restricted versions: no ex-ante heterogeneity in financial literacy (model version: "NHFL") and no ex-ante heterogeneity in preferences and retirement replacement ratio (model version: "NHPR"). In addition, we also report results with a higher value of the correlation between stock returns and income shocks, namely a 0.2 correlation with both transitory and permanent income innovations (model version: "HighCorr"). The results are shown in Figures 17 and 18, respectively for cumulative changes in stock market participation and cumulative changes in the risky share.

Similarly to the results obtained when excluding all ex-ante heterogeneity (series "NH"), the results for considering homogeneous financial literacy or homogeneous savings motives (preferences and retirement replacement ratio), both deliver very high stock market participation (Figure 17). Furthermore, compared with the baseline model, the implied change in the risky share (Figure 18) is either even a steeper function of age (model version: "NPHR"), or essentially the same ("model version: "NHFL"). Increasing the correlation between stock returns and labor income shocks improves the fit of the model with regards to the cumulative changes in the risky share, since labor income becomes a weaker substitute for bonds. However, as shown in Figure 17, since the optimal risky share of young households is now lower, they are also less likely to pay the participation cost and become stockholders. Therefore, under this level of correlation, stock market participation increases gradually with age, in contrast to our empirical results.

7 Conclusion

We estimate the life-cycle profiles of stock market participation and the conditional risky share. We find that the former is a hump-shaped function of age, while the conditional risky share is flat early on, and decreases late in life. We address the classical identification problem of separately disentangling time, age and cohort effects by running the estimations in first differences. Although we cannot recover the levels of the risky share and stock market participation at each age, we can identify the age profiles without making assumptions on time or cohort effects. While the levels of are likely influenced by several (potentially unobserved) individual characteristics, the age profiles are more direct implications of different theories.

By estimating these profiles, and showing how they are related to changes in wealth and changes in human capital, we can understand which life-cycle models are consistent with our findings. In particular, we find empirical support for stock market participation costs, background risks, decreasing relative risk aversion and human capital that has some correlation with stock returns but remains a closer substitute for bonds than for stocks. We confirm these implications in a structural life-cycle model of portfolio choice.

Crucially we also show that our empirical results are qualitatively different from those obtained if we impose the common exclusion restrictions on time or cohort effects. For both stock market participation and the conditional risky share, results without time effects or without cohort effects give rise to very different age profiles, and neither of them replicates our baseline results.

References

Abel, A. B., Eberly, J. C., and Panageas, S. (2013). Optimal inattention to the stock market with information costs and transactions costs. *Econometrica*, *81*(4), 1455–1481.

Agnew, J., Balduzzi, P., and Sunden, A. (2003). Portfolio choice and trading in a large 401 (k) plan. *American Economic Review*, *93*(1), 193–215.

Alvarez, F., Guiso, L., and Lippi, F. (2012). Durable consumption and asset management with transaction and observation costs. *American Economic Review*, *102*(5), 2272–2300.

Ameriks, J., Caplin, A., Laufer, S., and Van Nieuwerburgh, S. (2011). The joy of giving or assisted living? using strategic surveys to separate public care aversion from bequest motives. *The Journal of Finance*, *66*(2), 519–561.

Ameriks, J., and Zeldes, S. P. (2004). How do household portfolio shares vary with age. *working paper, Columbia University*.

Bach, L., Calvet, L. E., and Sodini, P. (2016). Rich pickings? risk, return, and skill in the portfolios of the wealthy. *CEPR Discussion Paper No. DP11734*.

Badarinza, C., Campbell, J. Y., and Ramadorai, T. (2016). International comparative household finance. *Annual Review of Economics*, *8*, 111–144.

Bagliano, F. C., Fugazza, C., and Nicodano, G. (2017). A life-cycle model with unemployment traps. *Working Paper*.

Benzoni, L., Collin-Dufresne, P., and Goldstein, R. S. (2007). Portfolio choice over the life-cycle when the stock and labor markets are cointegrated. *The Journal of Finance*, *62*(5), 2123–2167.

Bilias, Y., Georgarakos, D., and Haliassos, M. (2010). Portfolio inertia and stock market fluctuations. *Journal of Money, Credit and Banking*, *42*(4), 715–742.

Black, S. E., Devereux, P. J., Lundborg, P., and Majlesi, K. (2018). Learning to take risks? the effect of education on risk-taking in financial markets. *Review of Finance*, *22*(3), 951–975.

Bonaparte, Y., Korniotis, G. M., and Kumar, A. (2014). Income hedging and portfolio decisions. *Journal of Financial Economics*, *113*(2), 300–324.

Brown, J. R., Fang, C., and Gomes, F. (2015). Risk and returns to education over time.

Brunnermeier, M. K., and Nagel, S. (2008). Do wealth fluctuations generate time-varying risk aversion? micro-evidence on individuals. *American Economic Review*, *98*(3), 713–36.

Calvet, L. E., Campbell, J. Y., Gomes, F. J., and Sodini, P. (2019). The cross-section of household preferences. *Working Paper*.

Calvet, L. E., Campbell, J. Y., and Sodini, P. (2007). Down or out: Assessing the welfare costs of household investment mistakes. *Journal of Political Economy*, *115*(5), 707–747.

Calvet, L. E., Campbell, J. Y., and Sodini, P. (2009). Measuring the financial sophistication of households. *American Economic Review*, *99*(2), 393–98.

Calvet, L. E., and Sodini, P. (2014). Twin picks: Disentangling the determinants of risk-taking in household portfolios. *The Journal of Finance*, *69*(2), 867–906.

Campanale, C. (2011). Learning, ambiguity and life-cycle portfolio allocation. *Review of Economic Dynamics*, *14*(2), 339–367.

Campanale, C., Fugazza, C., and Gomes, F. (2015). Life-cycle portfolio choice with liquid and illiquid financial assets. *Journal of Monetary Economics*, *71*, 67–83.

Campbell, J. Y. (2006). Household finance. *The journal of finance*, *61*(4), 1553–1604.

Campbell, J. Y., Cocco, J. F., Gomes, F. J., and Maenhout, P. J. (2001). Investing retirement wealth: A life-cycle model. In *Risk aspects of investment-based social security reform* (pp. 439–482). University of Chicago Press.

Campbell, J. Y., and Cochrane, J. H. (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of political Economy*, *107*(2), 205–251.

Catherine, S. (2019). Countercyclical labor income risk and portfolio choices over the life-cycle. *HEC Paris Research Paper No. FIN-2016-1147*.

Chapman, D. A., and Polkovnichenko, V. (2009). First-order risk aversion, heterogeneity, and asset market outcomes. *The Journal of Finance*, *64*(4), 1863–1887.

Chetty, R., Sándor, L., and Szeidl, A. (2017). The effect of housing on portfolio choice. *The Journal of Finance*, *72*(3), 1171–1212.

Chetty, R., and Szeidl, A. (2007). Consumption commitments and risk preferences. *The Quarterly Journal of Economics*, *122*(2), 831–877.

Chiappori, P.-A., and Paiella, M. (2011). Relative risk aversion is constant: Evidence from panel data. *Journal of the European Economic Association*, *9*(6), 1021–1052.

Choi, J. J., Laibson, D., Madrian, B. C., and Metrick, A. (2002). Defined contribution pensions: Plan rules, participant choices, and the path of least resistance. *Tax policy and the economy*, *16*, 67–113.

Choi, J. J., and Robertson, A. Z. (2020). What matters to individual investors? evidence from the horse's mouth. *The Journal of Finance*, *75*(4), 1965–2020.

Christelis, D., Georgarakos, D., and Haliassos, M. (2013). Differences in portfolios across countries: Economic environment versus household characteristics. *Review of Economics and Statistics*, *95*(1), 220–236.

Cocco, J. F. (2005). Portfolio choice in the presence of housing. *The Review of Financial Studies*, *18*(2), 535–567.

Cocco, J. F., and Gomes, F. J. (2012). Longevity risk, retirement savings, and financial innovation. *Journal of Financial Economics*, *103*(3), 507–529.

Cocco, J. F., Gomes, F. J., and Maenhout, P. J. (2005). Consumption and portfolio choice over the life cycle. *The Review of Financial Studies*, *18*(2), 491–533.

Constantinides, G. M. (1990). Habit formation: A resolution of the equity premium puzzle. *Journal of political Economy*, *98*(3), 519–543.

Davis, S. J., and Willen, P. S. (2013). Occupation-level income shocks and asset returns: Their covariance and implications for portfolio choice. *Working Paper*.

Deaton, A., and Paxson, C. (1994). Intertemporal choice and inequality. *Journal of political economy*, *102*(3), 437–467.

De Nardi, M., French, E., and Jones, J. B. (2010). Why do the elderly save? the role of medical expenses. *Journal of political economy*, *118*(1), 39–75.

Dimmock, S. G., Kouwenberg, R., Mitchell, O. S., and Peijnenburg, K. (2016). Ambiguity aversion and household portfolio choice puzzles: Empirical evidence. *Journal of Financial Economics*, *119*(3), 559–577.

Epstein, L. G., and Zin, S. E. V. (1989). Risk aversion, and the temporal behavior of consumption and asset returns: a theoretical framework. *V Econometrica*, *57*(4), 937.

Fagereng, A., Gottlieb, C., and Guiso, L. (2017). Asset market participation and portfolio choice over the life-cycle. *The Journal of Finance*, *72*(2), 705–750.

Fagereng, A., Guiso, L., Malacrino, D., and Pistaferri, L. (2020). Heterogeneity and persistence in returns to wealth. *Econometrica*, *88*(1), 115–170.

Gabaix, X., and Laibson, D. (2001). The 6d bias and the equity-premium puzzle. *NBER macroeconomics annual*, *16*, 257–312.

Gomes, F. (2020). Portfolio choice over the life cycle: A survey. *Annual Review of Financial Economics*, *12*, 277–304.

Gomes, F., Haliassos, M., and Ramadorai, T. (forthcoming). Household finance. *Journal of Economic Literature*.

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

Gomes, F., and Michaelides, A. (2005). Optimal life-cycle asset allocation: Understanding the empirical evidence. *The Journal of Finance*, *60*(2), 869–904.

Grinblatt, M., Keloharju, M., and Linnainmaa, J. (2011). Iq and stock market participation. *The Journal of Finance*, *66*(6), 2121–2164.

Guiso, L., Haliassos, M., and Jappelli, T. (2002). *Household portfolios.* MIT press.

Guiso, L., Jappelli, T., and Terlizzese, D. (1996). Income risk, borrowing constraints, and portfolio choice. *The American Economic Review*, 158–172.

Guiso, L., Sapienza, P., and Zingales, L. (2008). Trusting the stock market. *the Journal of Finance*, *63*(6), 2557–2600.

Guvenen, F., Ozkan, S., and Song, J. (2014). The nature of countercyclical income risk. *Journal of Political Economy*, *122*(3), 621–660.

Haliassos, M., and Bertaut, C. C. (1995). Why do so few hold stocks? *the economic Journal*, *105*(432), 1110–1129.

Haliassos, M., and Michaelides, A. (2003). Portfolio choice and liquidity constraints. *International Economic Review*, *44*(1), 143–177.

Heaton, J., and Lucas, D. (2000). Portfolio choice and asset prices: The importance of entrepreneurial risk. *The journal of finance*, *55*(3), 1163–1198.

Heaton, J., and Lucas, D. J. (1996). Evaluating the effects of incomplete markets on risk sharing and asset pricing. *Journal of political Economy*, *104*(3), 443–487.

Horneff, W. J., Maurer, R. H., Mitchell, O. S., and Stamos, M. Z. (2009). Asset allocation and location over the life cycle with investment-linked survival-contingent payouts. *Journal of Banking & Finance*, *33*(9), 1688–1699.

Hubbard, R. G., Skinner, J., and Zeldes, S. P. (1995). Precautionary saving and social insurance. *Journal of political Economy*, *103*(2), 360–399.

Hurd, M., Van Rooij, M., and Winter, J. (2011). Stock market expectations of dutch households. *Journal of Applied Econometrics*, *26*(3), 416–436.

Knüpfer, S., Rantapuska, E., and Sarvimäki, M. (2017). Formative experiences and portfolio choice: Evidence from the finnish great depression. *The Journal of Finance*, *72*(1), 133–166.

Koijen, R. S., Van Nieuwerburgh, S., and Yogo, M. (2016). Health and mortality delta: Assessing the welfare cost of household insurance choice. *The Journal of Finance*, *71*(2), 957–1010.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics*, *112*(2), 443–478.

Lusardi, A., and Mitchell, O. S. (2011). Financial literacy around the world: an overview. *National Bureau of Economic Research Working Paper Series*(w17107).

Lusardi, A., and Mitchell, O. S. (2014). The economic importance of financial literacy: Theory and evidence. *Journal of economic literature*, *52*(1), 5–44.

Lustig, H., Van Nieuwerburgh, S., and Verdelhan, A. (2013). The wealth-consumption ratio. *The Review of Asset Pricing Studies*, *3*(1), 38–94.

Malmendier, U., and Nagel, S. (2011). Depression babies: Do macroeconomic experiences affect risk taking? *The quarterly journal of economics*, *126*(1), 373–416.

Mankiw, N. G., and Zeldes, S. P. (1991). The consumption of stockholders and nonstockholders. *Journal of financial Economics*, *29*(1), 97–112.

Meeuwis, M. (2020). Wealth fluctuations and risk preferences: Evidence from us investor portfolios. *Available at SSRN 3653324*.

Pagel, M. (2018). A news-utility theory for inattention and delegation in portfolio choice. *Econometrica*, *86*(2), 491–522.

Parker, J. A., Schoar, A., and Cole, A. (2021). Household portfolios and retirement saving over the life cycle. *Working Paper*.

Peijnenburg, K. (2018). Life-cycle asset allocation with ambiguity aversion and learning. *Journal of Financial and Quantitative Analysis*, *53*(5), 1963–1994.

Polkovnichenko, V. (2007). Life-cycle portfolio choice with additive habit formation preferences and uninsurable labor income risk. *The Review of Financial Studies*, *20*(1), 83–124.

Poterba, J. M., and Samwick, A. (1997). Household portfolio allocation over the life cycle..

Shen, J. (2019). Countercyclical risks and portfolio choice over the life cycle: Evidence and theory. *Working Paper*.

Sims, C. A. (2003). Implications of rational inattention. *Journal of monetary Economics*, *50*(3), 665–690.

Vanguard. (2009). How america saves 2009, a report on vanguard 2008 defined contribution plan data.

Vanguard. (2020). How america saves 2020, a report on vanguard 2019 defined contribution plan data.

Van Rooij, M., Lusardi, A., and Alessie, R. (2011). Financial literacy and stock market participation. *Journal of Financial Economics*, *101*(2), 449–472.

Viceira, L. M. (2001). Optimal portfolio choice for long-horizon investors with nontradable labor income. *The Journal of Finance*, *56*(2), 433–470.

Vissing-Jørgensen, A. (2002). Limited asset market participation and the elasticity of intertemporal substitution. *Journal of political Economy*, *110*(4), 825–853.

Wachter, J. A., and Yogo, M. (2010). Why do household portfolio shares rise in wealth? *The Review of Financial Studies*, *23*(11), 3929–3965.

Yao, R., and Zhang, H. H. (2005). Optimal consumption and portfolio choices with risky housing and borrowing constraints. *The Review of Financial Studies*, *18*(1), 197–239.

Yogo, M. (2016). Portfolio choice in retirement: Health risk and the demand for annuities, housing, and risky assets. *Journal of Monetary Economics*, *80*, 17–34.

Table 1: Number of observations in the sample

Table 1 reports the number of observations at each age in the PSID data after applying our sample restrictions. Columns 1 to 4 show number of observations for change in participation for each age change starting with 19 to 21 and ending at 88 to 90 for the full sample. Columns 5 to 8 show the number of observations for change in risky share for the stock market participants sample, which additionally conditions on stock market participation in both current and previous waves.

All				Stock market participants			
ΔAge	N	ΔAge	N	ΔAge	N	ΔAge	N
19-21	42	54-56	519	19-21	2	54-56	117
20-22	88	55-57	496	20-22	6	55-57	112
21-23	115	56-58	525	21-23	4	56-58	114
22-24	233	57-59	490	22-24	22	57-59	110
23-25	322	58-60	480	23-25	21	58-60	100
24-26	420	59-61	454	**24-26**	53	59-61	88
25-27	496	60-62	460	25-27	54	60-62	87
26-28	528	61-63	413	26-28	76	61-63	98
27-29	586	62-64	424	27-29	68	62-64	102
28-30	577	63-65	397	28-30	77	63-65	94
29-31	606	64-66	376	29-31	75	64-66	90
30-32	601	65-67	344	30-32	84	65-67	92
31-33	632	66-68	346	31-33	92	66-68	86
32-34	612	67-69	328	32-34	101	67-69	83
33-35	619	68-70	308	33-35	94	68-70	81
34-36	612	69-71	281	34-36	91	69-71	78
35-37	609	70-72	276	35-37	89	70-72	76
36-38	577	71-73	270	36-38	96	71-73	73
37-39	570	72-74	224	37-39	86	72-74	68
38-40	563	73-75	242	38-40	98	73-75	69
39-41	566	74-76	228	39-41	90	74-76	63
40-42	569	75-77	216	40-42	91	75-77	61
41-43	550	76-78	194	41-43	90	76-78	56
42-44	530	77-79	225	42-44	89	77-79	62
43-45	532	78-80	177	43-45	87	78-80	62
44-46	551	79-81	185	44-46	92	79-81	52
45-47	547	80-82	156	45-47	94	80-82	57
46-48	525	81-83	174	46-48	110	81-83	51
47-49	553	82-84	136	47-49	102	82-84	44
48-50	519	83-85	143	48-50	105	83-85	41
49-51	526	84-86	103	49-51	106	84-86	36
50-52	548	85-87	116	50-52	116	85-87	23
51-53	521	86-88	80	51-53	97	86-88	26
52-54	542	87-89	89	52-54	116	87-89	16
53-55	517	88-90	64	53-55	108	88-90	14
Total			27443	Total			5264

Table 2: Summary statistics

Table 2 reports summary statistics for the variables in our sample. Panel (a) reports values for the sample used in participation analysis, while the values in Panel (b) refer to the sample of stock market participants, which is used in the conditional risky share regressions. Row 2 in Panels (a) and (b) describes age structure in the samples. Rows 3 to 7 provide summary of other individual characteristics. Row 8 summarizes household income. Rows 9, 10 and 11 (LFA, HE and TW correspondingly) provide descriptive statistics for wealth variables in corresponding sample. Row 12 (partic and ω) in Panels (a) and (b) summarizes participation and risky share, respectively. For both panels, Rows 13 (ΔlogLFA), 14 (ΔlogHE) and 15 (ΔlogTW) show growth rates for wealth variables in corresponding sample. Rows 16 (ΔPVYW), 17 (ξ^1) and 18 (ξ^2) summarize changes in the ratio of human capital to financial wealth explained in section 2.3. All growth rates and changes are over a two-year period since that is the frequency of our data. Income and wealth/asset data is deflated to 2017 dollars

Stats	N	Mean	SD	p25	p50	p75	Max
age	26861	49.42	15.7	36	48	61	85
male	26861	0.82	0.39	1	1	1	1
married	26861	0.67	0.47	0	1	1	1
has kids	26861	0.38	0.49	0	0	1	1
# of kids	26861	0.71	1.07	0	0	1	9
home owner	26861	0.77	0.42	1	1	1	1
income	26861	101557.3	70962.8	52343.8	85186.8	129887.6	408856.0
LFA	26861	84174.8	212592.2	4136.5	12547.1	54650.1	1418440.0
HE	26861	122835.2	168312.2	0	67114.1	169385.2	934903.0
TW	26861	209675.1	332561.8	22233.7	95139.6	245714.3	2073684.0
partic	26861	0.28	0.45	0	0	1	1
ΔlogLFA	26860	0.77	2.43	-0.43	0.23	1.16	9.93
ΔlogHE	19606	0.94	2.87	-0.09	0.11	0.49	12.00
ΔlogTW	26078	0.49	1.72	-0.19	0.16	0.67	8.93
ΔPVYW	24188	-24.28	193.73	-12.89	-1.06	2.17	717.76
ξ^1	24188	-16.28	183.81	-9.22	-0.23	3.47	736.11
ξ^2	24188	7.89	16.98	0.66	1.79	5.64	102.77

(a) Full sample

Stats	N	Mean	SD	p25	p50	p75	Max
age	5094	54.02	15.7	41	54	66	85
male	5094	0.85	0.36	1	1	1	1
married	5094	0.73	0.44	0	1	1	1
has kids	5094	0.31	0.46	0	0	1	1
# of kids	5094	0.55	0.94	0	0	1	6
home owner	5094	0.88	0.32	1	1	1	1
income	5094	140849.9	111551.4	68337.1	113927.7	175334.5	697494.3
LFA	5094	314323.6	538573.9	38857.1	119197.0	331428.6	3401662.0
HE	5094	231731.9	261916.4	60942	160065.4	305010.9	1464714.0
TW	5094	550553.3	713769.0	140000.0	320000.0	651965.5	4446743.0
risky share	5094	0.65	0.30	0	0.73	0.92	1
ΔlogLFA	5094	0.09	0.99	-0.40	0.07	0.55	3.27
ΔlogHE	4366	0.62	2.37	-0.07	0.08	0.35	12.20
ΔlogTW	5041	0.13	0.67	-0.18	0.09	0.41	2.42
ΔPVYW	5027	-8.34	48.52	-4.18	-0.47	0.75	143.68
ξ^1	5027	-6.13	44.79	-2.73	-0.06	1.32	147.41
ξ^2	5027	2.13	4.28	0.30	0.76	1.94	30.05

(b) Stock market participants

Table 3: Participation in stock market over time

Table 3 reports means for participation, past participation and change in participation over the PSID waves, in %. Stock market participation equals to 1 if the individual has positive value of risky financial assets and equals to 0 otherwise. In the PSID, risky financial assets are shares of stock in publicly held corporations, stock mutual funds, or investment trusts (not including stocks in employer-based pensions or I.R.A.s). Column 2 shows number of observations from each wave. Column 3 reports mean participation in the sample by wave. Column 4 reports stock market participation in the previous wave for the same wave subsample as in Column 3. Column 5 reports the change in participation, calculated as the difference between Columns 3 and 4.

year	N	Participation_t	$\text{Participation}_{t-1}$	ΔParticipation
1999	2838	37.03%		
2001	2402	39.38%	35.51%	3.87%
2003	2444	34.98%	36.33%	-1.35%
2005	2536	32.65%	32.29%	0.35%
2007	2610	31.07%	30.27%	0.80%
2009	2713	29.45%	29.45%	0.00%
2011	2680	24.85%	28.84%	-3.99%
2013	2760	24.24%	23.73%	0.51%
2015	2629	22.48%	22.67%	-0.19%
2017	3036	20.92%	19.96%	0.96%
2019	3051	19.47%	20.35%	-0.88%
Total	29699	28.44%	27.55%	-0.02%

Table 4: Regression results for change in participation (with total wealth)

Table 4 reports results of OLS regressions of change in participation on wealth, human capital variables and controls. Specification 1 in Column 2 includes only two control variables, change in home ownership status and change in the number of children in the household. Column 3 reports results for specification 2 which include changes total wealth. Columns 4 and 5 show Specifications 3 and 4 with indicators for the ratio of the present-value of human capital to total wealth. All Specifications include change in age (starting with 21 to 23 age change and to 83 to 85 age change) and year dummies. The coefficients for wealth variables are the 2-year growth rates.

	(1)	(2)	(3)	(4)
owner	0.00948	-0.0118	-0.0164*	-0.0178*
	(1.11)	(-1.36)	(-1.68)	(-1.82)
kids	-0.00875	-0.0113	-0.0108	-0.0108
	(-1.29)	(-1.63)	(-1.46)	(-1.46)
ΔlogTW		0.0310***	0.110***	0.111***
		(20.49)	(22.89)	(22.92)
ΔPVYW			0.000174***	
			(9.08)	
ξ^1				0.000192***
				(8.97)
ξ^2				0.000141
				(1.14)
N	26861	26078	24188	24188
adj. R^2	0.002	0.018	0.044	0.044
F	1.801	7.491	9.635	9.592

cluster (id) robust t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Regression results for change in participation (by total wealth quartiles)

Table 5 reports results of OLS regressions of change in participation on different wealth, human capital variables and controls. Columns 2 to 5 report results for quartile regressions for specification 2 (with total wealth). Columns 6 to 9 show estimation results for quartile regressions in specification 3 (with total wealth and the ratio of the present-value of human capital to total wealth). All Specifications include change in home ownership status and change in the number of children in the household as controls. Regression equations also include change in age (starting with 21 to 23 age change and to 83 to 85 age change) and year dummies. The coefficients for wealth variables are the 2-year growth rates.

	(2)				(3)			
	q1	q2	q3	q4	q1	q2	q3	q4
owner	-0.00374	-0.0464***	-0.119***	-0.145***	0.0164	-0.0413**	-0.0794**	-0.0935**
	(-0.30)	(-2.60)	(-3.91)	(-3.82)	(1.12)	(-2.05)	(-2.42)	(-2.50)
kids	-0.00210	-0.0131	-0.0290*	0.00704	0.00407	-0.0120	-0.0273*	0.0118
	(-0.21)	(-1.02)	(-1.90)	(0.35)	(0.33)	(-0.92)	(-1.79)	(0.59)
ΔlogTW	0.0121***	0.0402***	0.0802***	0.152***	0.0487***	0.0925***	0.153***	0.234***
	(9.68)	(10.11)	(11.53)	(13.60)	(7.21)	(10.50)	(13.83)	(19.42)
ΔPVYW					-0.0000112	0.000169***	0.000541***	0.00104***
					(-0.55)	(3.39)	(6.37)	(4.55)
N	6104	6590	6680	6704	4526	6374	6610	6678
adj. R^2	0.014	0.024	0.030	0.072	0.037	0.033	0.043	0.098
F	2.284	2.653	3.255	4.076	2.284	2.939	4.079	7.118

cluster (id) robust t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Regression results for change in participation (with liquid financial assets)

Table 6 reports results of OLS regressions of change in participation on liquid financial assets, human capital variables and controls. Column 2 reports results for specification 5 which include changes in liquid financial assets. Specification 6 in Column 3 includes liquid financial assets and home equity. Columns 4 and 5 show Specifications 7 and 8 with liquid financial assets and indicators for the ratio of the present-value of human capital to liquid financial assets. All Specifications include change in ownership and change in kids as well as change in age (starting with 21 to 23 age change and to 83 to 85 age change) and year dummies. The coefficients for wealth indicators are the 2-year growth rates.

	(5)	(6)	(7)	(8)
owner	0.00569	0.0186	0.0166*	0.0164*
	(0.69)	(0.63)	(1.76)	(1.74)
kids	-0.0155**	-0.0164**	-0.00592	-0.00634
	(-2.34)	(-1.98)	(-0.79)	(-0.84)
ΔlogLFA	0.0415***	0.0526***	0.152***	0.152***
	(32.04)	(28.66)	(39.74)	(39.74)
ΔlogHE		-0.00223		
		(-0.83)		
ΔPVYW			0.000101***	
			(9.63)	
ξ^1				0.000107***
				(9.54)
ξ^2				0.00000227
				(0.04)
N	26860	19606	22687	22687
adj. R^2	0.062	0.079	0.164	0.164
F	16.09	13.31	28.23	27.88

cluster robust t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Regression results for entries (with total wealth)

Table 7 reports results of Probit regressions of stock market entries (as share of total stock market nonparticipants in previous wave) on wealth, human capital variables and controls. The numbers reflect marginal effects on the probability to entry, evaluated at the sample means of the explanatory variables. Specification 1 in Column 2 includes only two control variables, change in home ownership status and change in the number of children in the household. Column 3 reports results for Specification 2 which include changes total wealth. Columns 4 and 5 show Specifications 3 and 4 with indicators for the ratio of the present-value of human capital to total wealth. All Specifications include change in age (starting with 21 to 23 age change and to 83 to 85 age change) and year dummies. The coefficients for wealth variables are the 2-year growth rates.

	(1)	(2)	(3)	(4)
owner	-0.00231	-0.00902	-0.0326***	-0.0192*
	(-0.29)	(-1.08)	(-3.04)	(-1.80)
kids	-0.0111*	-0.0123*	-0.0124*	-0.0121*
	(-1.72)	(-1.86)	(-1.76)	(-1.75)
ΔlogTW		0.00672***	0.0672***	0.0653***
		(6.54)	(19.09)	(18.46)
ΔPVYW			0.000127***	
			(7.86)	
ξ^1				-0.0000437
				(-1.58)
ξ^2				-0.00284***
				(-9.99)
N	19461	18777	16892	16892
pseudo R^2	0.022	0.024	0.059	0.073
chi2	282.6	317.1	639.9	649.5

cluster (id) robust t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Regression results for entries (with liquid financial assets)

Table 8 reports results of Probit regressions of stock market entries (as share of total stock market nonparticipants in previous wave) on liquid financial assets, human capital variables and controls. The numbers reflect marginal effects on the probability to entry, evaluated at the sample means of the explanatory variables. Column 2 reports results for Specification 5 which include changes in liquid financial assets. Specification 6 in Column 3 includes liquid financial assets and home equity. Columns 4 and 5 show Specifications 7 and 8 with liquid financial assets and indicators for the ratio of the present-value of human capital to liquid financial assets. All Specifications include change in ownership and change in kids as well as change in age (starting with 21 to 23 age change and to 83 to 85 age change) and year dummies. The coefficients for wealth indicators are the 2-year growth rates.

	(5)	(6)	(7)	(8)
owner	-0.00402	-0.0504**	-0.00279	-0.00129
	(-0.50)	(-2.10)	(-0.31)	(-0.16)
kids	-0.0129**	-0.0180**	-0.00256	0.00579
	(-2.01)	(-2.19)	(-0.37)	(0.90)
ΔlogLFA	0.0107***	0.0150***	0.0856***	0.0731***
	(14.92)	(15.10)	(31.03)	(24.81)
ΔlogHE		0.00219		
		(1.01)		
ΔPVYW			0.0000758***	
			(10.12)	
ξ^1				-0.0000709***
				(-4.80)
ξ^2				-0.00203***
				(-16.98)
N	19460	13375	15364	15364
pseudo R^2	0.034	0.041	0.154	0.192
chi2	494.1	446.3	1243.0	1384.5

cluster robust t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Regression results for exits (with total wealth)

Table 9 reports results of Probit regressions of stock market exits (as share of total stock market participants in previous wave) on wealth, human capital variables and controls. The numbers reflect marginal effects on the probability to entry, evaluated at the sample means of the explanatory variables. Specification 1 in Column 2 includes only two control variables, change in home ownership status and change in the number of children in the household. Column 3 reports results for Specification 2 which include changes total wealth. Columns 4 and 5 show Specifications 3 and 4 with indicators for the ratio of the present-value of human capital to total wealth. All Specifications include change in age (starting with 21 to 23 age change and to 83 to 85 age change) and year dummies. The coefficients for wealth variables are the 2-year growth rates.

	(1)	(2)	(3)	(4)
owner	0.0298	0.0449*	0.0400	0.0240
	(1.30)	(1.83)	(1.59)	(0.95)
kids	0.0298*	0.0333*	0.0315*	0.0322*
	(1.79)	(1.93)	(1.83)	(1.87)
ΔlogTW		-0.116***	-0.153***	-0.155***
		(-11.62)	(-16.48)	(-16.65)
ΔPVYW			-0.000241***	
			(-3.85)	
ξ^1				-0.0000218
				(-0.32)
ξ^2				0.00768***
				(7.19)
N	7400	7301	7248	7248
pseudo R^2	0.019	0.054	0.063	0.072
chi2	166.2	289.9	435.8	484.8

cluster (id) robust t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 10: Regression results for exits (with liquid financial assets)

Table 10 reports results of Probit regressions of stock market exits (as share of total stock market participants in previous wave) on liquid financial assets, human capital variables and controls. The numbers reflect marginal effects on the probability to entry, evaluated at the sample means of the explanatory variables. Column 2 reports results for Specification 5 which include changes in liquid financial assets. Specification 6 in Column 3 includes liquid financial assets and home equity. Columns 4 and 5 show Specifications 7 and 8 with liquid financial assets and indicators for the ratio of the present-value of human capital to liquid financial assets. All Specifications include change in ownership and change in kids as well as change in age (starting with 21 to 23 age change and to 83 to 85 age change) and year dummies. The coefficients for wealth indicators are the 2-year growth rates.

	(5)	(6)	(7)	(8)
owner	-0.0100	-0.174**	-0.00367	-0.00971
	(-0.41)	(-2.35)	(-0.15)	(-0.39)
kids	0.0253	0.0169	0.0245	0.0198
	(1.46)	(0.92)	(1.41)	(1.12)
ΔlogLFA	-0.163***	-0.173***	-0.163***	-0.169***
	(-29.20)	(-28.67)	(-24.53)	(-25.71)
ΔlogHE		0.0103		
		(1.61)		
ΔPVYW			0.0000425	
			(1.09)	
ξ^1				0.000242***
				(5.71)
ξ^2				0.00838***
				(11.10)
N	7400	6231	7279	7279
pseudo R^2	0.167	0.186	0.172	0.217
chi2	919.0	884.9	969.3	1133.1

cluster robust t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 11: Regression results for change in risky share (with total wealth)

Table 11 reports results of OLS regressions of change in conditional risky share on total wealth, human capital variables and controls. Sample includes only those, who are stock market participants in current and previous waves. Specification 1 in Column 2 includes only two control variables, change in home ownership status and change in the number of children in the household. Column 3 reports results for specification 2 which include changes total wealth. Columns 4 and 5 show Specifications 3 and 4 with indicators for the ratio of the present-value of human capital to total wealth. All Specifications include change in age (starting with 24 to 26 age change and to 83 to 85 age change) and year dummies. The coefficients for wealth variables are the 2-year growth rates.

	(1)	(2)	(3)	(4)
owner	0.0102	-0.000626	0.00829	0.00721
	(0.55)	(-0.03)	(0.44)	(0.38)
kids	0.0109	0.0114	0.00985	0.00989
	(0.83)	(0.89)	(0.77)	(0.77)
ΔlogTW		0.0720^{***}	0.0896^{***}	0.0898^{***}
		(8.34)	(8.92)	(8.89)
ΔPVYW			0.000337^{***}	
			(2.87)	
ξ^1				0.000389^{**}
				(2.49)
ξ^2				0.000435
				(0.31)
N	5094	5041	5027	5027
adj. R^2	0.007	0.033	0.039	0.039
F	1.452	2.645	2.928	2.883

cluster robust t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 12: Regression results for change in risky share (with liquid financial assets)

Table 12 reports results of OLS regressions of change in conditional risky share on liquid financial assets, human capital variables and controls. Sample includes only those, who are stock market participants in current and previous waves. Column 2 reports results for Specification 5 which include changes in liquid financial assets. Specification 6 in Column 3 includes liquid financial assets and home equity. Columns 4 and 5 show Specifications 7 and 8 with liquid financial assets and indicators for the ratio of the present-value of human capital to liquid financial assets. All Specifications include change in ownership and change in kids as well as change in age (starting with 24 to 26 age change and to 83 to 85 age change) and year dummies. The coefficients for wealth indicators are the 2-year growth rates.

	(5)	(6)	(7)	(8)
owner	0.0253	0.0268	0.0719***	0.0727***
	(1.39)	(0.39)	(2.95)	(2.99)
kids	0.0103	0.0115	0.0128	0.0128
	(0.82)	(0.85)	(0.95)	(0.95)
ΔlogLFA	0.0663***	0.0735***	0.0819***	0.0820***
	(10.40)	(10.44)	(11.07)	(11.08)
ΔlogHE		0.00369		
		(0.62)		
ΔPVYW			0.000113	
			(1.55)	
ξ^1				0.0000916
				(1.19)
ξ^2				-0.000524
				(-1.05)
N	5094	4366	4356	4356
adj. R^2	0.058	0.070	0.075	0.075
F	3.257	3.330	3.616	3.620

cluster robust t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 13: Regression results for active change in risky share (with total wealth)

Table 13 reports results of OLS regressions of active change in conditional risky share on total wealth, human capital variables and controls. Sample includes only those, who are stock market participants in current and previous waves. Active changes in the risky share are computed from equations (6) and (7) with value-weighted return on equities from CRSP, while the return on the riskless asset (Rf) is the real return on the 90-Day T-Bill. Specification 1 in Column 2 includes only two control variables, change in home ownership status and change in the number of children in the household. Column 3 reports results for specification 2 which include changes total wealth. Columns 4 and 5 show Specifications 3 and 4 with indicators for the ratio of the present-value of human capital to total wealth. All Specifications include change in age (starting with 24 to 26 age change and to 83 to 85 age change) and year dummies. The coefficients for wealth variables are the 2-year growth rates.

	(1)	(2)	(3)	(4)
owner	0.00939	-0.00177	0.00705	0.00617
	(0.51)	(-0.09)	(0.37)	(0.32)
kids	0.00961	0.0102	0.00864	0.00868
	(0.74)	(0.79)	(0.68)	(0.68)
ΔlogTW		0.0711***	0.0889***	0.0891***
		(8.23)	(8.87)	(8.84)
ΔPVYW			0.000343***	
			(2.89)	
ξ^1				0.000384**
				(2.45)
ξ^2				0.000273
				(0.20)
N	5094	5041	5027	5027
adj. R^2	0.002	0.029	0.034	0.034
F	1.328	2.391	2.653	2.622

cluster robust t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 14: Regression results for active change in risky share (with liquid financial assets)

Table 14 reports results of OLS regressions of active change in conditional risky share on total wealth, human capital variables and controls. Sample includes only those, who are stock market participants in current and previous waves. Active changes in the risky share are computed from equations (6) and (7) with value-weighted return on equities from CRSP, while the return on the riskless asset (Rf) is the real return on the 90-Day T-Bill. Column 2 reports results for Specification 5 which include changes in liquid financial assets. Specification 6 in Column 3 includes liquid financial assets and home equity. Columns 4 and 5 show Specifications 7 and 8 with liquid financial assets and indicators for the ratio of the present-value of human capital to liquid financial assets. All Specifications include change in ownership and change in kids as well as change in age (starting with 24 to 26 age change and to 83 to 85 age change) and year dummies. The coefficients for wealth indicators are the 2-year growth rates.

	(5)	(6)	(7)	(8)
owner	0.0242	0.0263	0.0717***	0.0725***
	(1.34)	(0.37)	(2.94)	(2.98)
kids	0.00908	0.00999	0.0111	0.0112
	(0.72)	(0.74)	(0.83)	(0.83)
ΔlogLFA	0.0652***	0.0726***	0.0809***	0.0810***
	(10.28)	(10.39)	(11.04)	(11.05)
ΔlogHE		0.00375		
		(0.62)		
ΔPVYW			0.000112	
			(1.54)	
ξ^1				0.0000890
				(1.16)
ξ^2				-0.000563
				(-1.12)
N	5094	4366	4356	4356
adj. R^2	0.052	0.063	0.068	0.068
F	3.079	3.163	3.406	3.397

cluster robust t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 15: Regression results for active change in risky share (by total wealth quartiles)

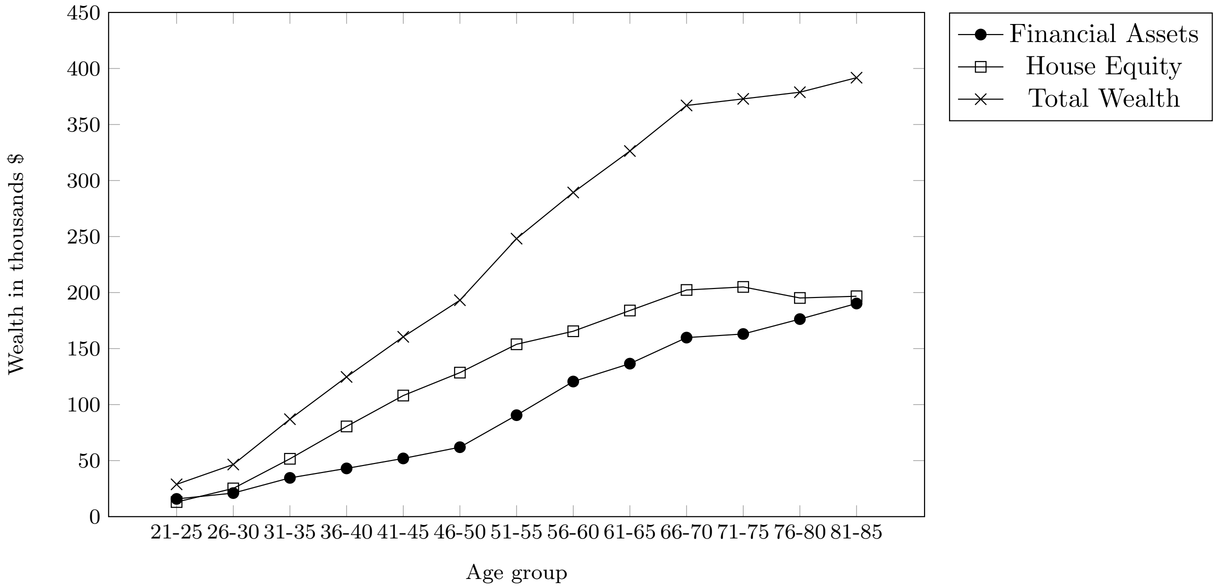
Table 15 reports results of OLS regressions of active change in conditional risky share on total wealth, human capital variables and controls. Sample includes only those, who are stock market participants in current and previous waves. Active changes in the risky share are computed from equations (6) and (7) with value-weighted return on equities from CRSP, while the return on the riskless asset (Rf) is the real return on the 90-Day T-Bill. Columns 2 to 5 report results for quartile regressions for specification 2 (with total wealth). Columns 6 to 9 show estimation results for quartile regressions in specification 3 (with total wealth and the ratio of the present-value of human capital to total wealth). All Specifications include change in home ownership status and change in the number of children in the household as controls. Regression equations also include change in age (starting with 24 to 26 age change and to 83 to 85 age change) and year dummies. The coefficients for wealth variables are the 2-year growth rates.

	(2)				(3)			
	q1	q2	q3	q4	q1	q2	q3	q4
owner	0.00263	-0.0126	-0.0218	0.000282	0.00507	-0.00840	-0.0199	-0.000894
	(0.09)	(-0.35)	(-0.45)	(0.01)	(0.18)	(-0.23)	(-0.41)	(-0.02)
kids	-0.00380	0.0117	0.0232	0.00452	-0.00627	0.0110	0.0235	0.00526
	(-0.15)	(0.48)	(0.82)	(0.16)	(-0.26)	(0.45)	(0.83)	(0.19)
ΔlogTW	0.0123	0.111^{***}	0.0978^{***}	0.118^{***}	0.0212	0.119^{***}	0.100^{***}	0.118^{***}
	(1.16)	(8.16)	(6.43)	(9.07)	(1.43)	(7.08)	(6.00)	(8.32)
ΔPVYW					-0.0000268	0.000301	0.000201	-0.000385
					(-0.19)	(0.81)	(0.33)	(-0.41)
N	1222	1274	1272	1273	1210	1273	1272	1272
adj. R^2	0.025	0.068	0.044	0.077	0.025	0.068	0.044	0.077
F	1.427	2.287	1.818	2.566	1.429	2.263	1.794	2.549

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

(a) Mean



(b) Median

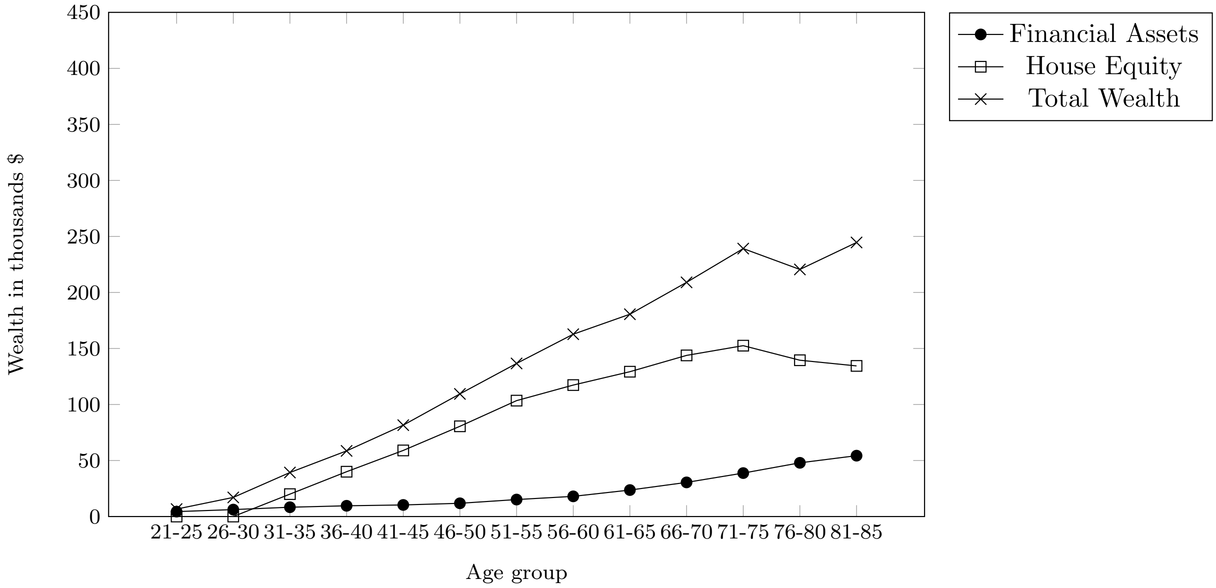


Figure 1: Wealth as function of age

Figure 1 shows mean and median wealth as a function of age, for three wealth categories: total wealth, financial assets and home equity, as defined in section 3.2. The data is taken from the PSID and corresponds to our baseline sample, which is obtained after applying the filters described in section 3.1.

(a) Mean

(b) Median

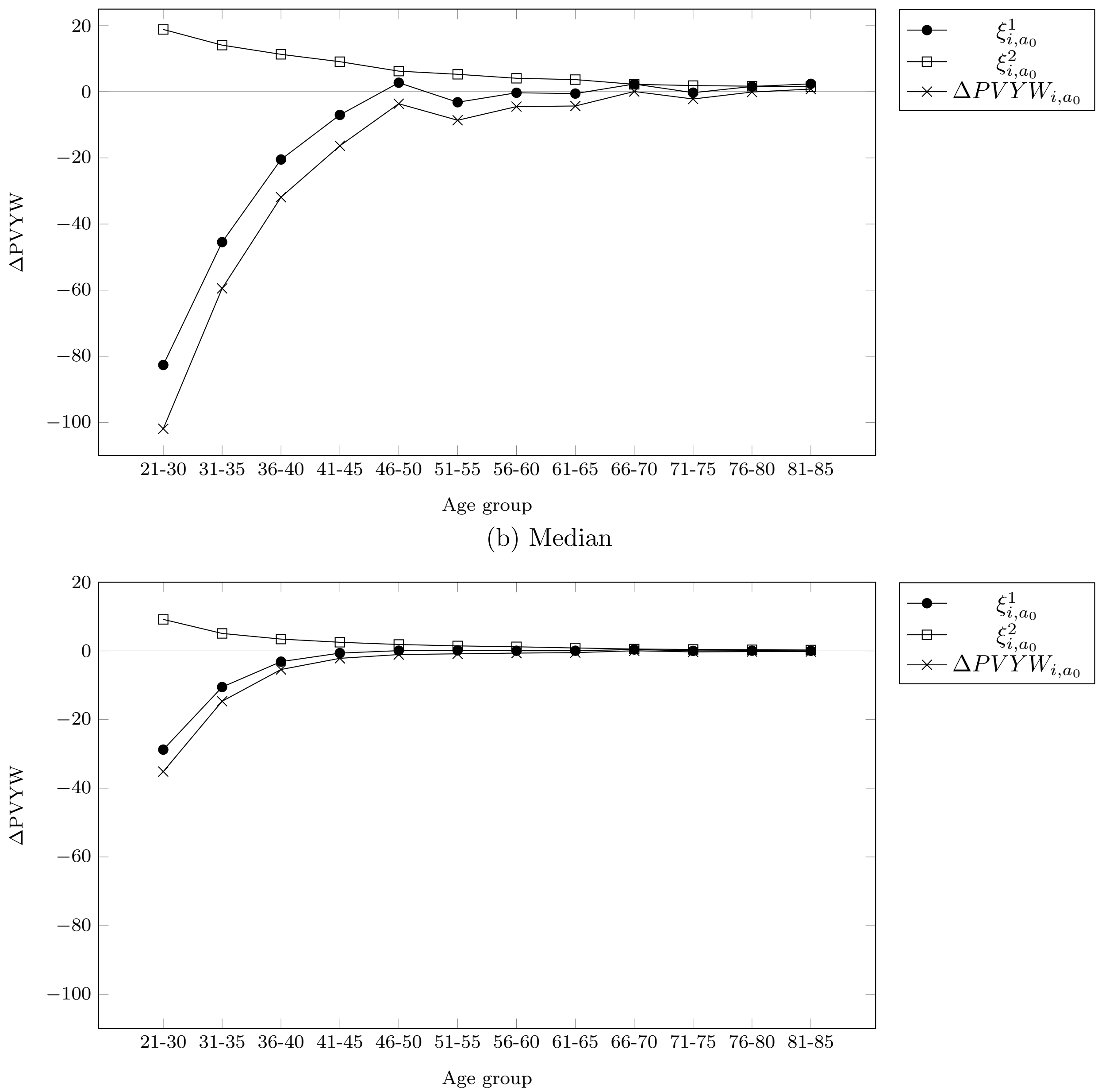
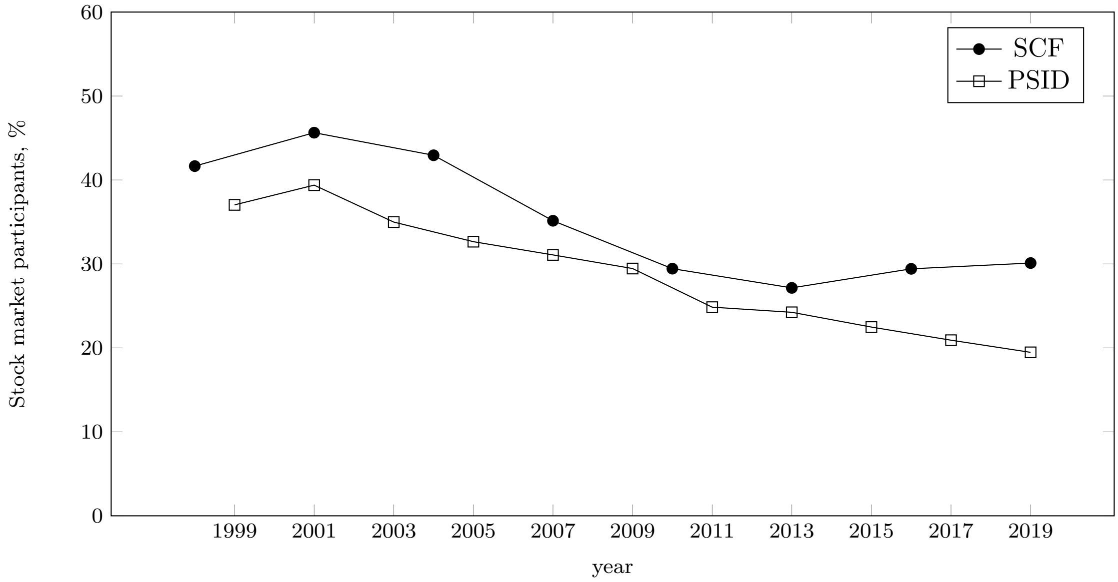


Figure 2: Summary of change in present value of income over wealth over age

Figure 2 shows mean and median values for changes in the ratio of human capital to total wealth ($\Delta PVYW_{i,a_0}$) as a function of age, as well as its two components, ξ^1_{i,a_0} and ξ^2_{i,a_0}. Details on calculations of these variables are in Section 2.3.2. The data is taken from the PSID and corresponds to our baseline sample, which is obtained after applying the filters described in section 3.1.

(a) Average stock market participation, PSID sample and the full SCF



(b) Average change in stock market participation over time, PSID sample

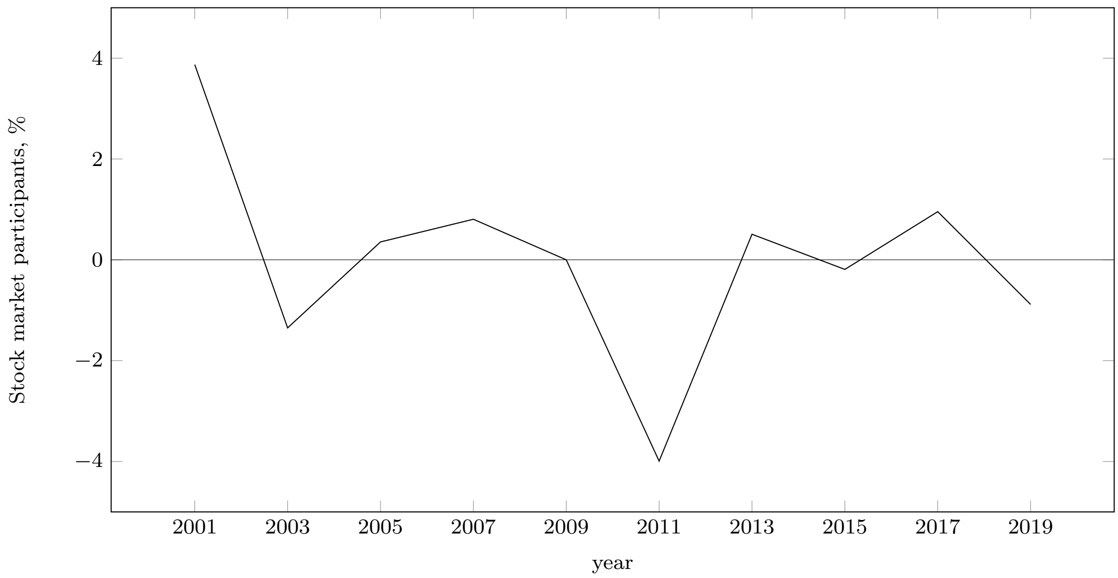


Figure 3: Stock market participation dynamics

Panel (a) of Figure 3 shows average stock market participation from the filtered PSID data and from the SCF. Stock market participation equals to 1 if the individual has positive value of risky financial assets and equals to 0 otherwise. For a clean comparison, we exclude wealth in retirement accounts when considering the SCF data. Panel (b) plots the average change in participation calculated as difference between current participation and previous wave participation at the individual level and then averaged across the years.

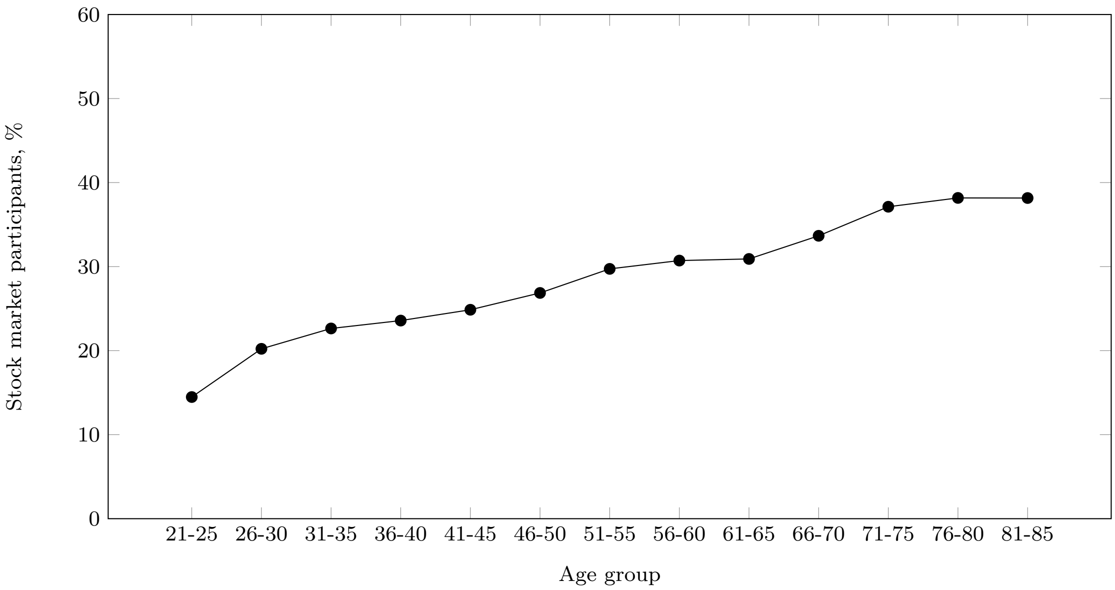


Figure 4: Average stock market participation over the life-cycle

Figure 4 shows average stock market participation in our filtered PSID sample by age cohort. Stock market participation equals to 1 if the individual has positive value of risky financial assets and equals to 0 otherwise. Risky financial assets are shares of stock in publicly held corporations, stock mutual funds, or investment trusts (not including stocks in employer-based pensions or I.R.A.s).

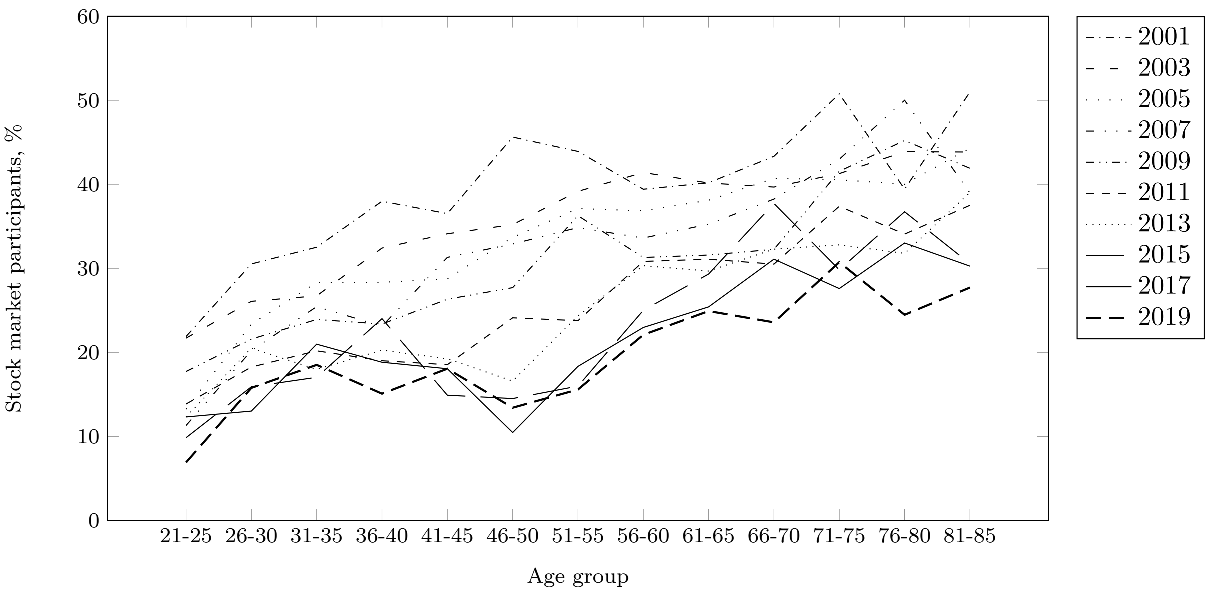


Figure 5: Average stock market participation over the life-cycle

Figure 5 shows average stock market participation in the filtered PSID sample by age cohort in each separate wave of the data. Stock market participation equals to 1 if the individual has positive value of risky financial assets and equals to 0 otherwise. Risky financial assets are shares of stock in publicly held corporations, stock mutual funds, or investment trusts (not including stocks in employer-based pensions or I.R.A.s).

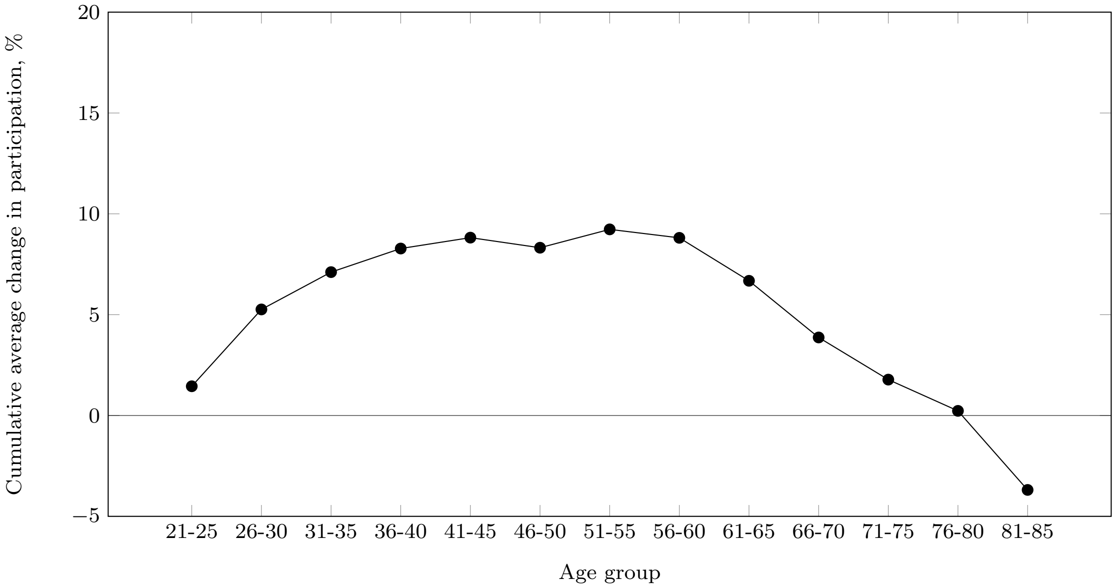


Figure 6: Cumulative average change in participation over the life-cycle

Figure 6 shows the life-cycle profile of stock market participation based on first computing changes in participation at the individual level, and then averaging these by age. The life-cycle profile is generated by computing the cumulative changes over age.

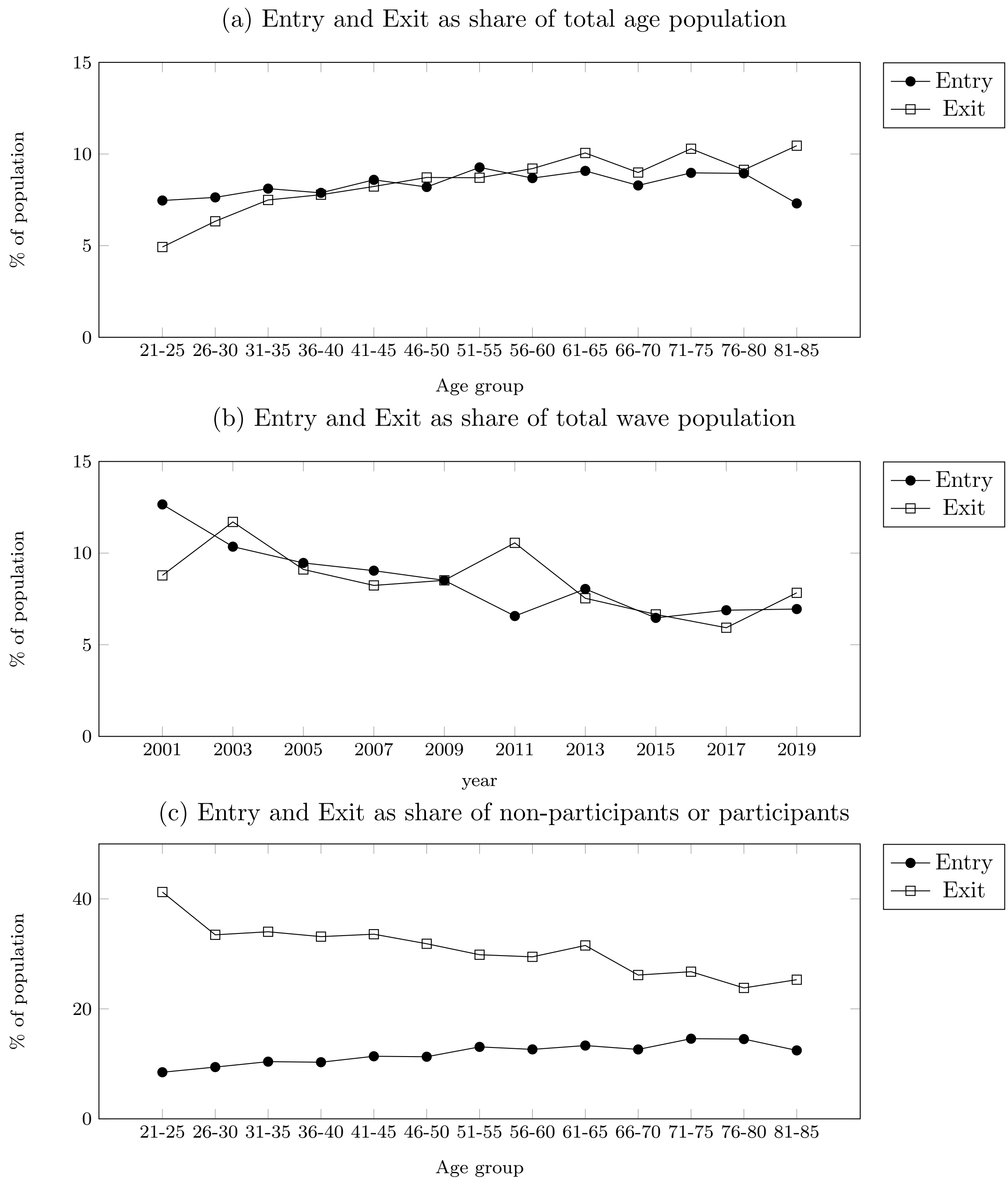


Figure 7: Entry and Exit shares

Figure 7 shows stock market entry and exit shares in the filtered PSID sample. In Panel (a) entry and exit rates are scaled by the total population in each age group. In Panel (b) shows entry and exit rates are scaled by the total population in each wave. Panel (c) plots entry and exits relative to the fraction of non-participants and participants, respectively.

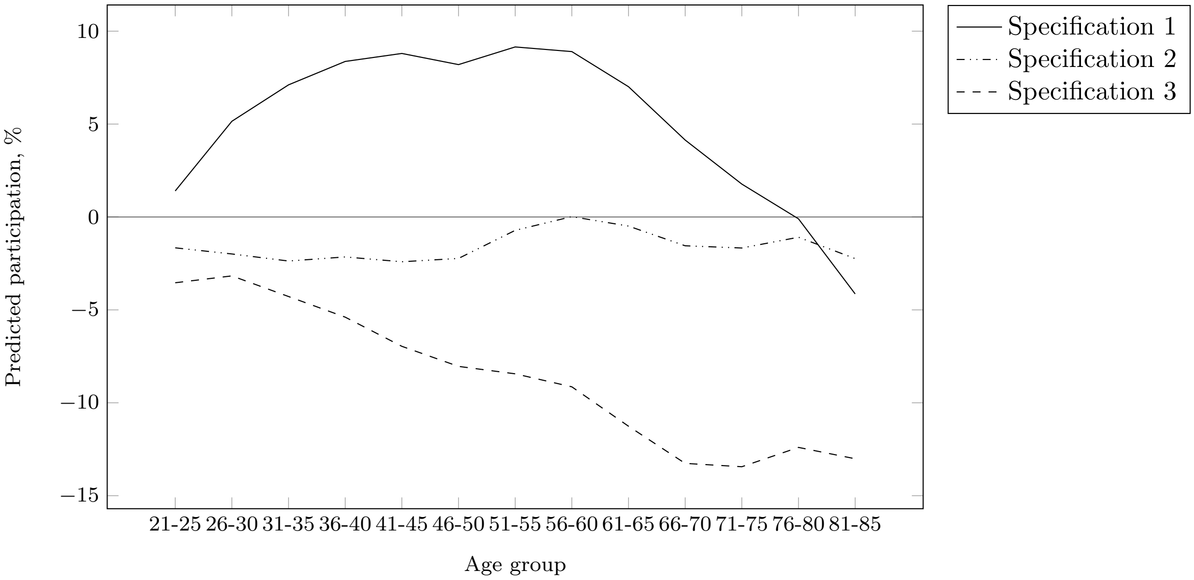


Figure 8: Predicted cumulative change in participation

Figure 8 shows predicted cumulative change in stock market participation from OLS regressions of change in participation on age and time dummies, wealth, human capital variables and controls. For the estimation we considered our filtered PSID sample. Specification 1 includes only two control variables, change in home ownership status and change in the number of children in the household. Specification 2 also includes change in total wealth, while specification 3 further includes changes in the present-value of human capital to total wealth. Results are presented in 5-year age groups to facilitate the visual interpretation, but regression estimation considers each year-age as the unit of analysis.

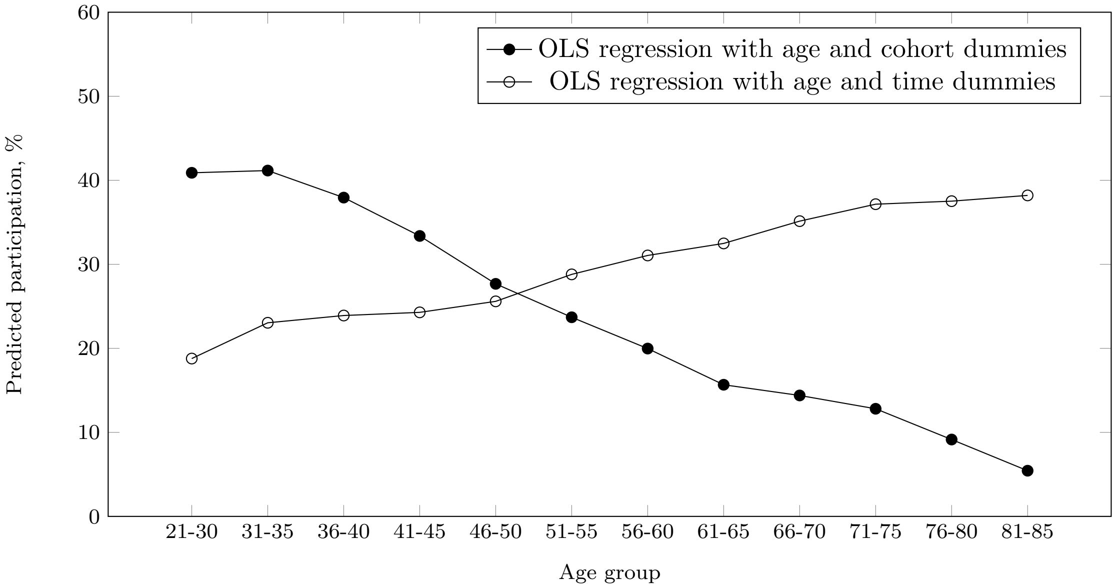
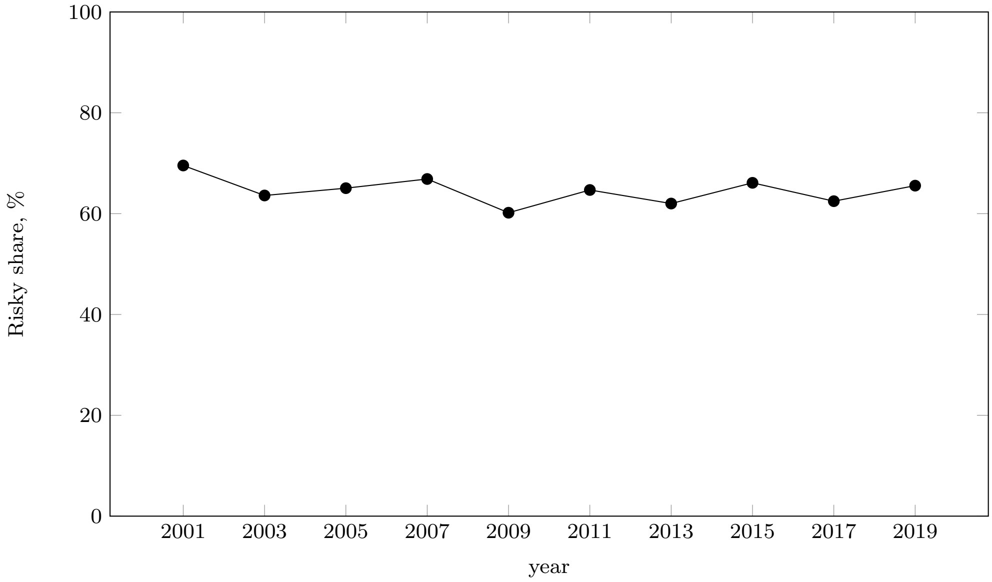


Figure 9: Time and cohort effects for predicted participation over age

Figure 9 shows predicted values for the cumulative change in participation implied by OLS regressions of changes in participation, considering either age and cohort dummies only, or age and time dummies only. For the estimations we use our filtered PSID sample, and predictions are made for the full filtered sample using actual observed values.

(a) Across time



(b) Across age groups

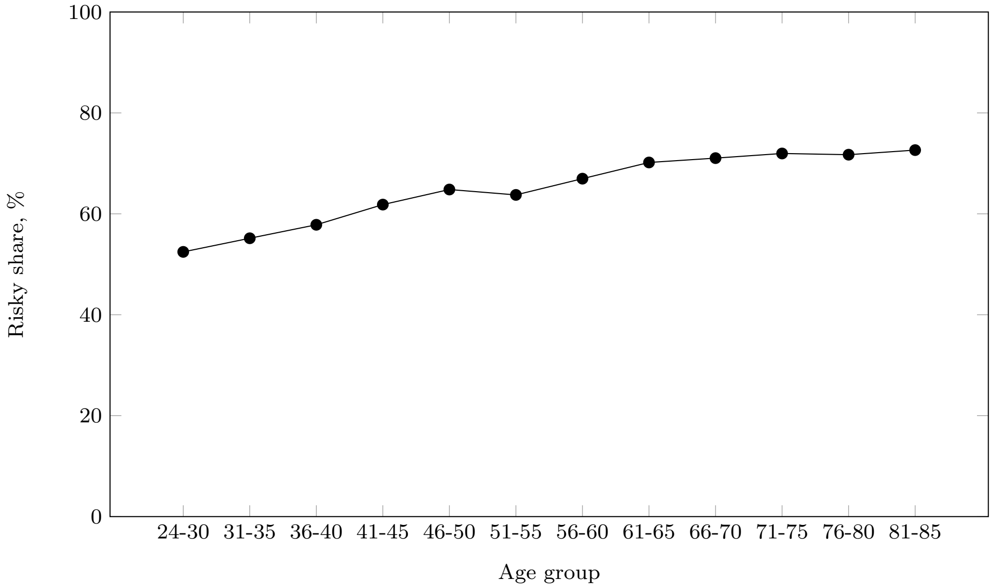


Figure 10: Average cross-sectional conditional risky share

Figure 10 shows the time and age profiles of the conditional risky share in our filtered PSID sample. Conditional risky share is a share of risky assets (shares of stock in publicly held corporations, stock mutual funds, or investment trusts excluding I.R.A.s) in liquid financial assets of stock markets participants.

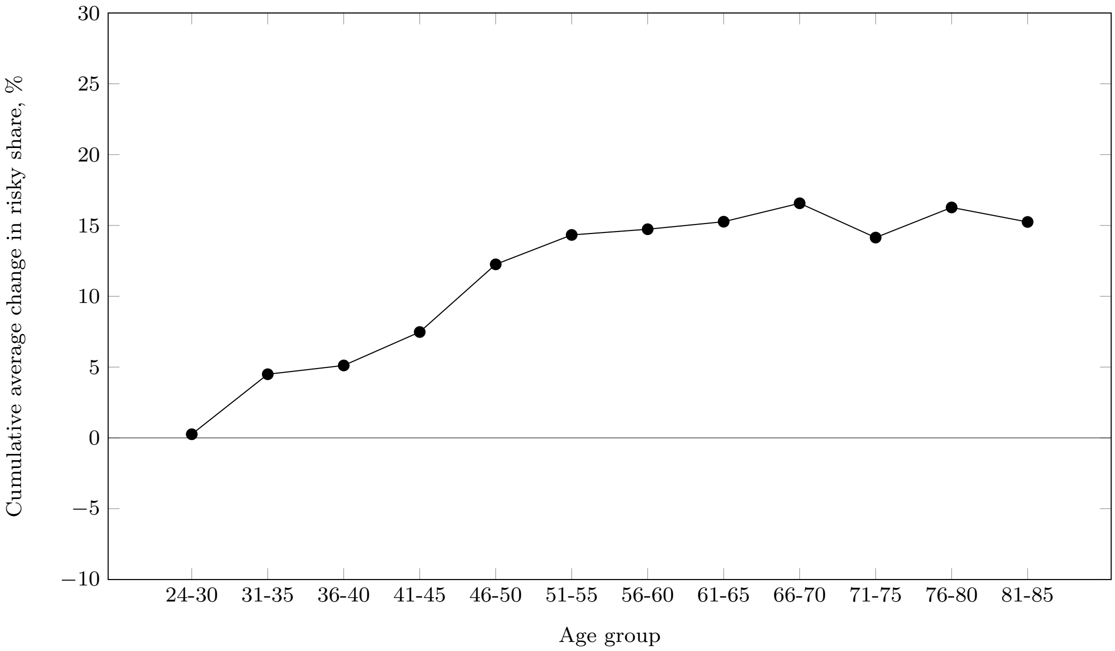


Figure 11: Cumulative average change in risky share over the life-cycle

Figure 11 shows the life-cycle profile of conditional risky share based on changes in risky share at the individual level and then averaging these by age. The life-cycle profile is generated by computing the cumulative changes over age. Results are obtained from our filtered PSID sample.

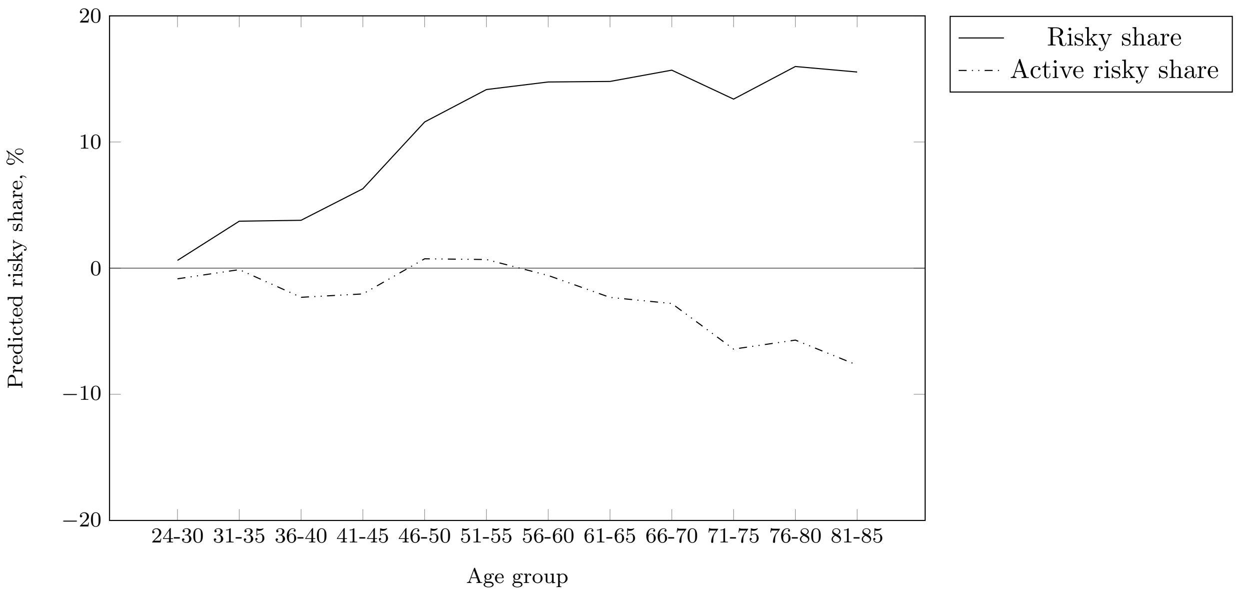


Figure 12: Predicted cumulative change in risky share

Figure 12 shows the predicted cumulative change in risky share from OLS regressions of risky share on age and time dummies, change in home ownership status, and change in the number of children in the household, using our filtered PSID sample. This corresponds to specification 1 in our regression tables. The Figure includes results for both change in risky share, and active change in risky share (based on the value-weighted return on equities from CRSP to compute passive changes). Results are presented in 5-year age groups to facilitate the visual interpretation, but regression estimation considers each year-age as the unit of analysis.

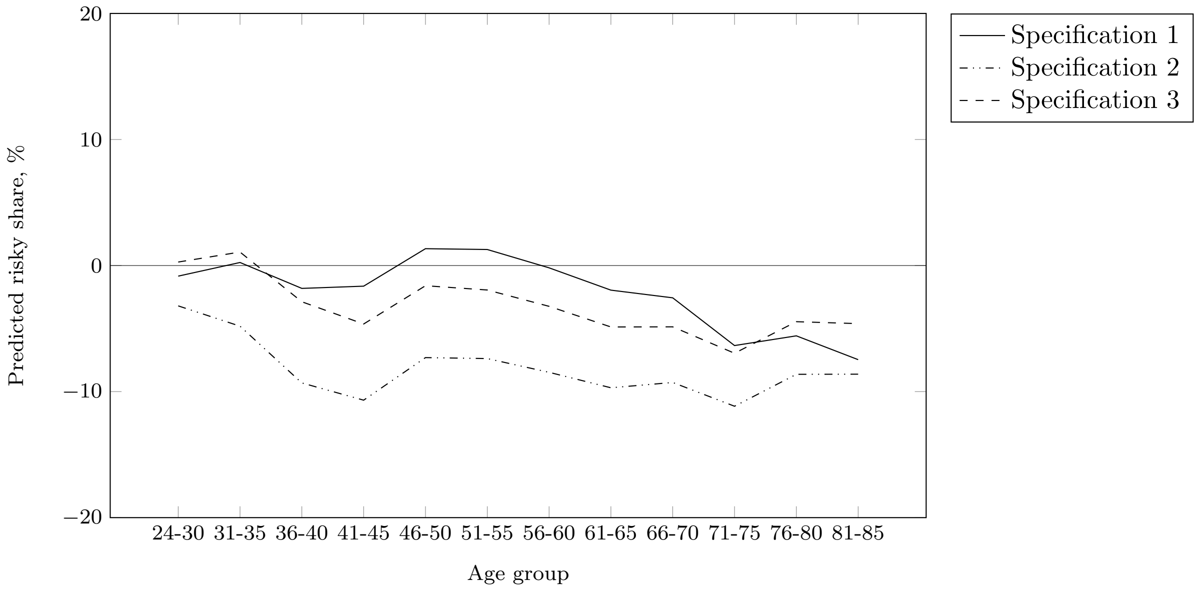


Figure 13: Predicted cumulative active change in risky share

Figure 13 shows the predicted cumulative active change in risky share from OLS regressions of the active risky share on age and time dummies, different wealth and human capital variables, and other controls, using our filtered PSID sample. We use the value-weighted return on equities from CRSP which includes dividends to calculate the active risky share. Specification 1 includes only two control variables, change in home ownership status and change in the number of children in the household. Specification 2 also includes change in total wealth. Specification 3 further includes changes in the present-value of human capital to total wealth. Results are presented in 5-year age groups to facilitate the visual interpretation, but regression estimation considers each year-age as the unit of analysis.

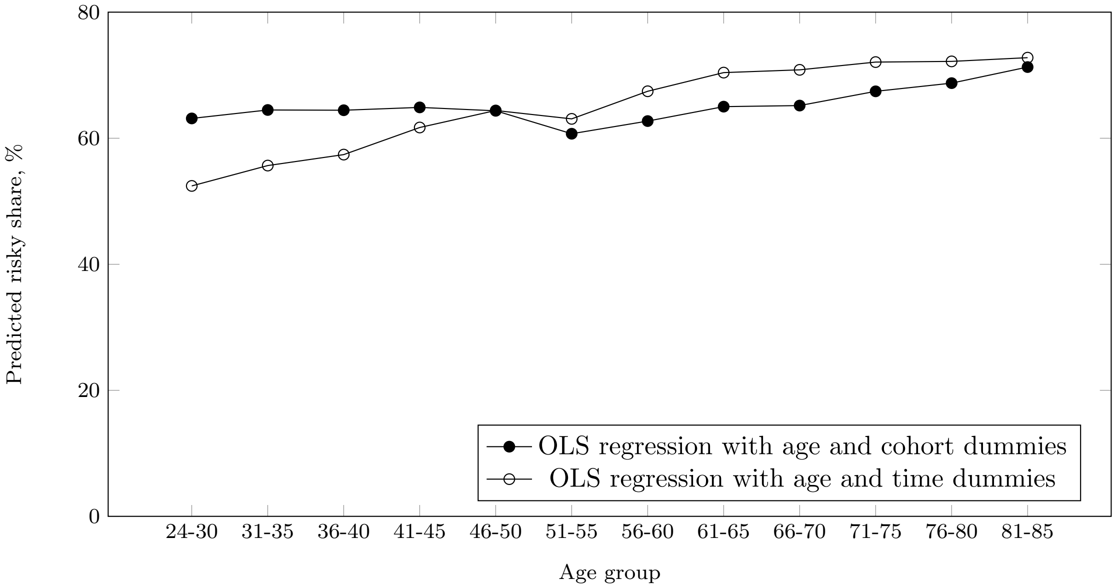


Figure 14: Time and cohort effects in predicted conditional risky share over age

Figure 14 shows the predicted values for conditional risky share from OLS regression of risky share on either age and cohort dummies or age and time dummies, using our filtered PSID sample. Predictions are made for the full filtered sample using actual observed values.

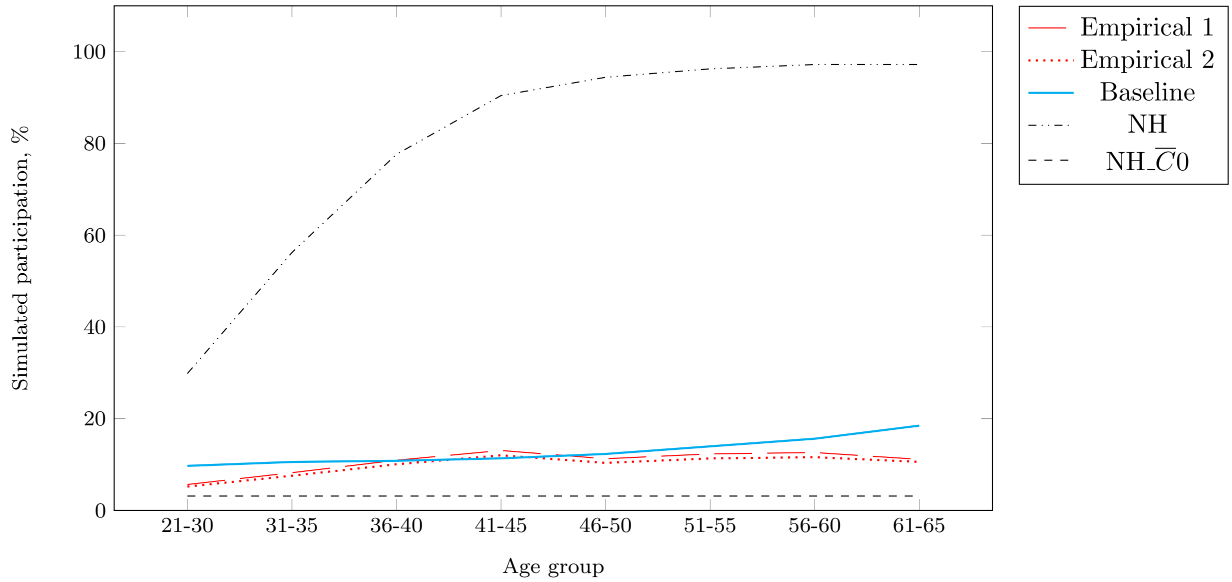


Figure 15: Cumulative change in participation

Figure 15 shows the cumulative change in participation from data and from life-cycle model simulations. The values from the data correspond to our baseline estimation of predictive cumulative changes, adjusted to reflect participation through DC accounts using SCF data. The series "Empirical 1" assumes that participation in DC accounts follows the same life-cycle pattern as direct participation, while the series "Empirical 2" assumes that participation in DC accounts does not fall with age late in life. We plot results for three different versions of the life-cycle model: our baseline specification (model version:"Baseline"), without ex-ante household heterogeneity (model version: "NH"), and both without ex-ante heterogeneity and without a consumption (model version: "NH_$\overline{C}$0").

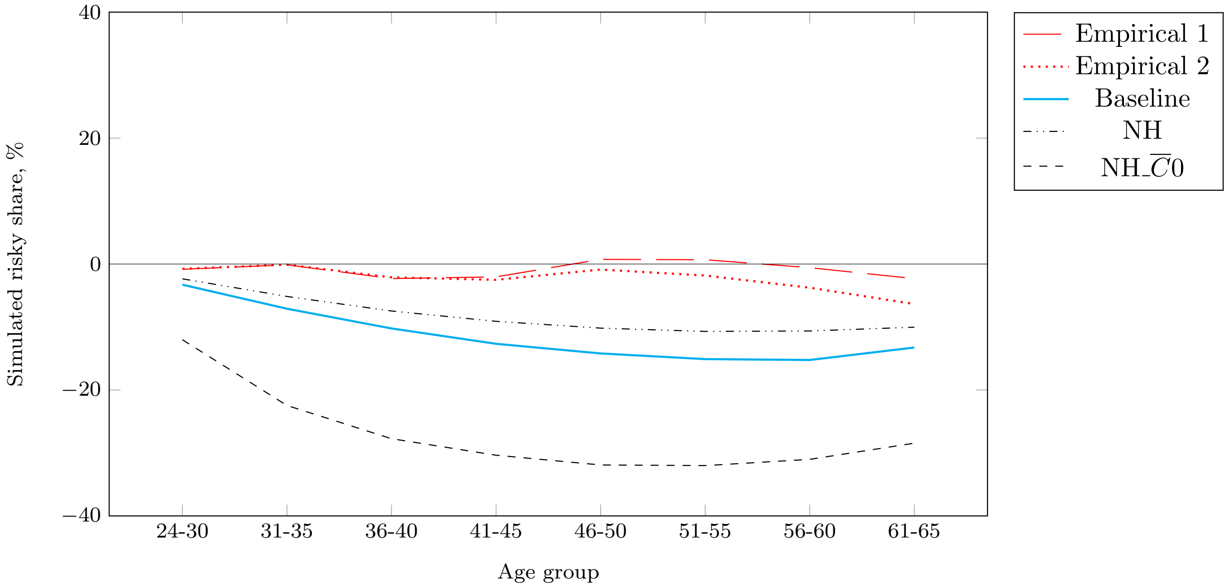


Figure 16: Cumulative change in risky share

Figure 16 shows the cumulative change in risky share from data and from life-cycle model simulations. The series "Empirical 1" corresponds to our baseline estimation of predictive cumulative changes in the risky share. The series "Empirical 2" adjusts for the fact that significant fraction of retirement wealth is being invested in Target-Date Funds (TDF) which have a strong age-profile for their risky share. We plot results for three different versions of the life-cycle model: our baseline specification (model version:"Baseline"), without ex-ante household heterogeneity (model version: "NH"), and both without ex-ante heterogeneity and without a consumption (model version: "NH_$\overline{C}$0").

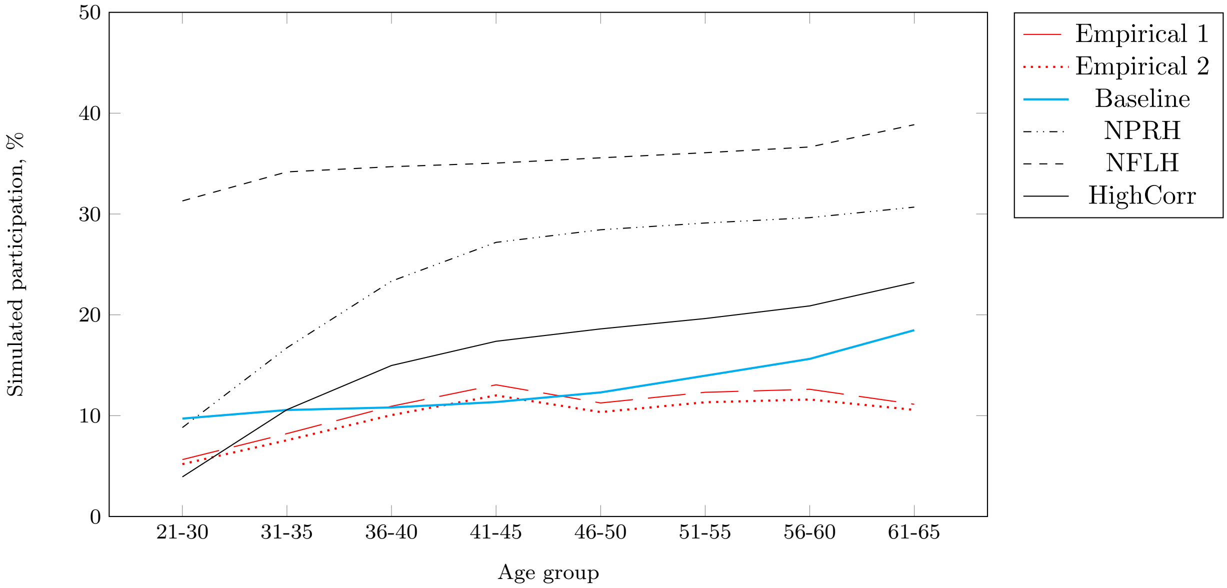


Figure 17: Cumulative change in participation

Figure 17 shows the cumulative change in participation from data and from life-cycle model simulations. The values from the data correspond to our baseline estimation of predictive cumulative changes, adjusted to reflect participation through DC accounts using SCF data. The series "Empirical 1" assumes that participation in DC accounts follows the same life-cycle pattern as direct participation, while the series "Empirical 2" assumes that participation in DC accounts does not fall with age late in life. We plot results for four different versions of the life-cycle model: our baseline specification (model version:"Baseline"), no ex-ante heterogeneity in financial literacy (model version:"NHFL"), no ex-ante heterogeneity in preferences and retirement replacement ratio (model version: "NHPR"), and with a higher value of the correlation between stock returns and income shocks, namely a 0.2 correlation with both transitory and permanent income innovations (model version: "HighCorr").

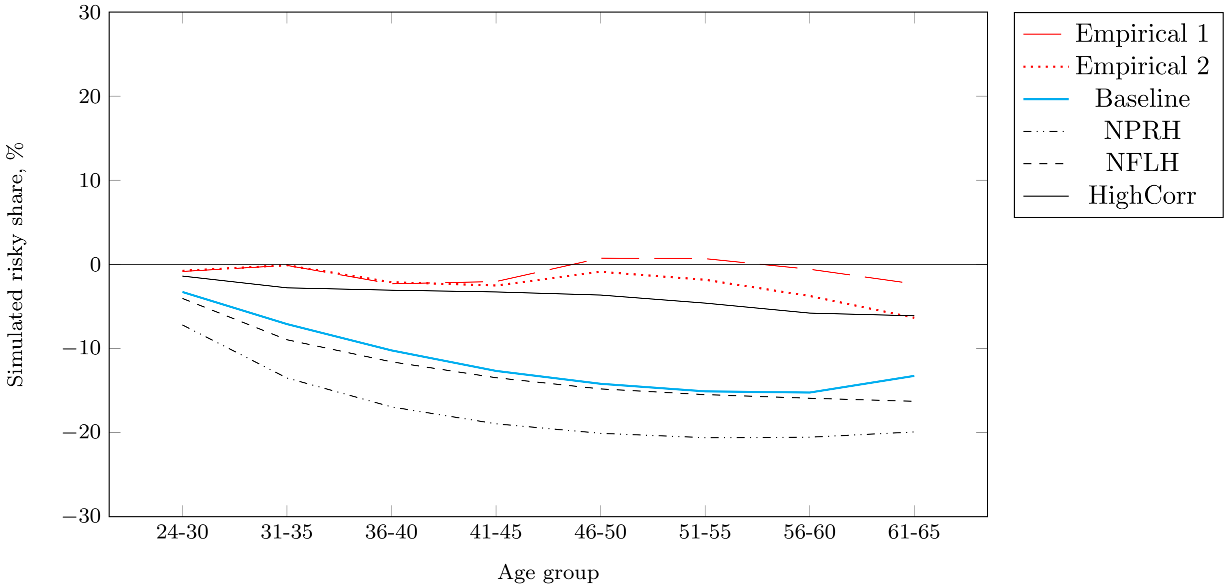


Figure 18: Cumulative change in risky share

Figure 18 shows cumulative change in risky share from data and from life-cycle model simulations. The series "Empirical 1" corresponds to our baseline estimation of predictive cumulative changes in the risky share. The series "Empirical 2" adjusts for the fact that significant fraction of retirement wealth is being invested in Target-Date Funds (TDF) which have a strong age-profile for their risky share. We plot baseline results (model version:"Baseline"), results for no ex-ante heterogeneity in financial literacy (model version:"NHFL"), no ex-ante heterogeneity in preferences and retirement replacement ratio (model version: "NHPR") and results with a higher value of the correlation between stock returns and income shocks, namely a 0.2 correlation with both transitory and permanent income innovations (model version: "HighCorr"). Results are presented in 5-year age groups to facilitate the visual interpretation, but model solution considers each age as the unit of analysis.

Appendix 1: Entry and Exit regressions with alternative definition

Appendix 1 reports the results from regressions of stock market entry and stock market exit using definitions (25) and (26), where entry and exit rates are computed as a share of the total population. We consider the same six regression specifications as in the estimations reported in the main text, where we consider the alternative definitions of entry and exit shares (given by equations ((27) and (28)).

The results between the two sets of regressions are qualitatively identical, and quantitatively very similar. Higher financial or total wealth predicts higher (lower) stock market entry (exit). Likewise an increase in the ratio of human capital to total wealth also predicts higher stock market entry, and lower stock market exit.

Table A1: Regression results for entries (as share of all, with total wealth)

Table A1 reports results of Probit regressions of stock market entries (as share of all) on wealth, human capital variables and controls. The numbers reflect marginal effects on the probability to entry, evaluated at the sample means of the explanatory variables. Specification 1 in Column 2 includes only two control variables, change in home ownership status and change in the number of children in the household. Column 3 reports results for Specification 2 which include changes total wealth. Columns 4 and 5 show Specifications 3 and 4 with indicators for the ratio of the present-value of human capital to total wealth. All Specifications include change in age (starting with 21 to 23 age change and to 83 to 85 age change) and year dummies. The coefficients for wealth variables are the 2-year growth rates.

	(1)	(2)	(3)	(4)
owner	0.00374	-0.00442	-0.0220***	-0.0169**
	(0.62)	(-0.71)	(-2.87)	(-2.21)
kids	-0.00590	-0.00687	-0.00587	-0.00563
	(-1.23)	(-1.40)	(-1.18)	(-1.14)
ΔlogTW		0.00972***	0.0528***	0.0524***
		(12.00)	(20.89)	(20.60)
ΔPVYW			0.0000860***	
			(6.89)	
ξ^1				0.00000560
				(0.29)
ξ^2				-0.00138***
				(-7.51)
N	26861	26078	24134	24134
pseudo R^2	0.011	0.018	0.053	0.058
chi2	174.2	315.5	639.2	633.4

cluster (id) robust t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2: Regression results for entries (as share of all, with financial wealth)

Table A2 reports results of Probit regressions of stock market entries (as share of all) on wealth, human capital variables and controls. The numbers reflect marginal effects on the probability to entry, evaluated at the sample means of the explanatory variables. Column 2 reports results for Specification 5 which include changes financial wealth. Specification 6 in Column 3 includes financial wealth and home equity. Columns 4 and 5 show Specifications 7 and 8 with financial wealth and indicators for the ratio of the present-value of human capital to total wealth. All Specifications include change in ownership and change in kids as well as change in age (starting with 21 to 23 age change and to 83 to 85 age change) and year dummies. The coefficients for wealth variables are the 2-year growth rates.

	(5)	(6)	(7)	(8)
owner	0.00152	-0.0283	0.00401	0.00513
	(0.26)	(-1.61)	(0.70)	(0.92)
kids	-0.00797*	-0.0103*	-0.000523	0.00196
	(-1.70)	(-1.81)	(-0.12)	(0.46)
ΔlogLFA	0.0130***	0.0166***	0.0557***	0.0525***
	(23.01)	(22.15)	(34.86)	(31.22)
ΔlogHE		0.00145		
		(0.92)		
ΔPVYW			0.0000334***	
			(6.41)	
ξ^1				-0.0000208**
				(-2.31)
ξ^2				-0.000743***
				(-10.68)
N	26860	19606	22637	22637
pseudo R^2	0.041	0.052	0.150	0.160
chi2	761.0	700.2	1351.7	1405.9

cluster robust t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3: Regression results for exits (as share of all, with total wealth)

Table A3 reports results of Probit regressions of stock market exits (as share of all) on wealth, human capital variables and controls. The numbers reflect marginal effects on the probability to entry, evaluated at the sample means of the explanatory variables. Specification 1 in Column 2 includes only two control variables, change in home ownership status and change in the number of children in the household. Column 3 reports results for Specification 2 which include changes total wealth. Columns 4 and 5 show Specifications 3 and 4 with indicators for the ratio of the present-value of human capital to total wealth. All Specifications include change in age (starting with 21 to 23 age change and to 83 to 85 age change) and year dummies. The coefficients for wealth variables are the 2-year growth rates.

	(1)	(2)	(3)	(4)
owner	-0.00524	-0.00304	-0.00623	-0.000297
	(-0.83)	(-0.47)	(-0.82)	(-0.04)
kids	0.00317	0.00445	0.00502	0.00488
	(0.67)	(1.03)	(1.00)	(1.00)
ΔlogTW		-0.0361***	-0.0505***	-0.0508***
		(-20.60)	(-18.82)	(-18.58)
ΔPVYW			-0.0000630***	
			(-5.61)	
ξ^1				-0.000140***
				(-7.53)
ξ^2				-0.00188***
				(-8.74)
N	26861	26078	24188	24188
pseudo R^2	0.014	0.058	0.048	0.057
chi2	233.7	505.6	574.5	608.7

cluster (id) robust t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A4: Regression results for exits (as share of all, with financial wealth)

Table A4 reports results of Probit regressions of stock market exits (as share of total stock market participants in previous wave) on wealth, human capital variables and controls. The numbers reflect marginal effects on the probability to entry, evaluated at the sample means of the explanatory variables. Column 2 reports results for Specification 5 which include changes financial wealth. Specification 6 in Column 3 includes financial wealth and home equity. Columns 4 and 5 show Specifications 7 and 8 with financial wealth and indicators for the ratio of the present-value of human capital to total wealth. All Specifications include change in ownership and change in kids as well as change in age (starting with 21 to 23 age change and to 83 to 85 age change) and year dummies. The coefficients for wealth variables are the 2-year growth rates.

	(5)	(6)	(7)	(8)
owner	-0.00671*	-0.0420***	-0.00988	-0.00916
	(-1.80)	(-3.48)	(-1.52)	(-1.46)
kids	0.00246	0.00143	0.00449	0.00699
	(0.91)	(0.40)	(0.97)	(1.54)
ΔlogLFA	-0.0345***	-0.0428***	-0.0649***	-0.0593***
	(-43.47)	(-41.14)	(-35.30)	(-29.10)
ΔlogHE		0.00227**		
		(2.20)		
ΔPVYW			-0.0000379***	
			(-6.40)	
ξ^1				-0.0000583***
				(-7.66)
ξ^2				-0.000942***
				(-10.21)
N	26860	19606	22687	22687
pseudo R^2	0.180	0.188	0.153	0.165
chi2	1426.0	1298.0	1334.2	1362.4

cluster robust t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix 2: Regression results with alternative total wealth definition

Appendix 2 reports the results from re-estimating our regressions for change in participation, stock market entry, stock market exit and change in risky share with our second definition of total wealth (NTW), where we also consider uncollateralized debt. The results are shown in tables A5, A6, A7 and A8, respectively for the four endogenous variables. To facilitate the comparison each table considers the same set of specifications as the corresponding tables in the main paper. The results are qualitatively identical, and quantitatively very similar, to the ones obtained with the alternative definition of total wealth.

Higher total net wealth predicts an increase in stock market participation, higher stock market entry and lower stock market exit, and an increase in the conditional risky share. Likewise positive changes in the present value of future labor income to total net wealth also predict an increase in stock market participation, higher stock market entry and lower stock market exit, and an increase in the conditional risky share.

Table A5: Regression results for change in participation

Table A5 reports results of OLS regressions of change in participation on different wealth, human capital variables and controls. Compared to Table 4, Total Wealth in this setting also takes non-mortgage debt into account. Specification 1 in Column 2 includes only two control variables, change in home ownership status and change in the number of children in the household. Column 3 reports results for specification 2 which include change in total wealth. Columns 4 and 5 show Specifications 3 and 4 with indicators for the ratio of the present-value of human capital to total wealth. All Specifications include change in age (starting with 21 to 23 age change and to 83 to 85 age change) and year dummies. The coefficients for wealth variables are the 2-year growth rates. Since total wealth does not enter Specification 1, the result in the corresponding column is identical to the one previously reported.

	(1)	(2)	(3)	(4)
owner	0.00948	-0.0163	0.00365	0.00267
	(1.11)	(-1.43)	(0.29)	(0.21)
kids	-0.00875	-0.00932	-0.00912	-0.00906
	(-1.29)	(-1.14)	(-1.07)	(-1.06)
ΔlogTW		0.0400***	0.131***	0.131***
		(17.78)	(22.14)	(22.14)
ΔPVYW			0.000429***	
			(12.10)	
ξ^1				0.000451***
				(10.83)
ξ^2				-0.000208
				(-0.83)
N	26861	21081	20177	20177
adj. R^2	0.002	0.021	0.050	0.050
F	1.801	6.150	8.811	8.717

cluster robust t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A6: Regression results for entries

Table A6 reports results of Probit regressions of stock market entries (as share of total stock market nonparticipants in previous wave) on different wealth, human capital variables and controls. Compared to Table 7, Total Wealth in this setting also takes non-mortgage debt into account. The numbers reflect marginal effects, the effect on the probability to entry, evaluated at the sample means of the explanatory variables. Specification 1 in Column 2 includes only two control variables, change in home ownership status and change in the number of children in the household. Column 3 reports results for Specification 2 which include changes total wealth. Columns 4 and 5 show Specifications 3 and 4 with indicators for the ratio of the present-value of human capital to total wealth. All Specifications include change in age (starting with 21 to 23 age change and to 83 to 85 age change) and year dummies. The coefficients for wealth variables are the 2-year growth rates. Since total wealth does not enter Specification 1, the result in the corresponding column is identical to the one previously reported.

	(1)	(2)	(3)	(4)
owner	-0.00231	-0.0116	-0.00974	0.00539
	(-0.29)	(-1.00)	(-0.70)	(0.38)
kids	-0.0111*	-0.0143*	-0.0139*	-0.0149*
	(-1.72)	(-1.75)	(-1.65)	(-1.81)
ΔlogTW		0.0116***	0.0827***	0.0808***
		(8.18)	(18.27)	(17.38)
ΔPVYW			0.000311***	
			(9.47)	
ξ^1				0.00000168
				(0.03)
ξ^2				-0.00453***
				(-7.99)
N	19461	14328	13449	13449
pseudo R^2	0.022	0.025	0.062	0.074
chi2	282.6	283.3	542.2	529.9

cluster (id) robust t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A7: Regression results for exits

Table A7 reports results of Probit regressions of stock market exits (as share of total stock market participants in previous wave) on different wealth, human capital variables and controls. Compared to Table 9, Total Wealth in this setting also takes non-mortgage debt into account. The numbers reflect marginal effects, the effect on the probability to entry, evaluated at the sample means of the explanatory variables. The numbers reflect marginal effects, the effect on the probability to exit, evaluated at the sample means of the explanatory variables. Specification 1 in Column 2 includes only two control variables, change in home ownership status and change in the number of children in the household. Column 3 reports results for Specification 2 which include changes total wealth. Columns 4 and 5 show Specifications 3 and 4 with indicators for the ratio of the present-value of human capital to total wealth. All Specifications include change in age (starting with 21 to 23 age change and to 83 to 85 age change) and year dummies. The coefficients for wealth variables are the 2-year growth rates. Since total wealth does not enter Specification 1, the result in the corresponding column is identical to the one previously reported.

	(1)	(2)	(3)	(4)
owner	0.0298	0.0343	0.0208	-0.00184
	(1.30)	(1.21)	(0.71)	(-0.06)
kids	0.0298*	0.0231	0.0270	0.0274
	(1.79)	(1.28)	(1.49)	(1.51)
ΔlogTW		-0.110***	-0.159***	-0.156***
		(-11.69)	(-16.18)	(-16.02)
ΔPVYW			-0.000580***	
			(-6.82)	
ξ^1				-0.0000976
				(-0.92)
ξ^2				0.00938***
				(7.12)
N	7400	6753	6705	6705
pseudo R^2	0.019	0.053	0.064	0.071
chi2	166.2	281.8	393.1	440.2

cluster (id) robust t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A8: Regression results for change in risky share

Table A8 reports results of OLS regressions of change in conditional risky share on different wealth, human capital variables and controls. Compared to Table 11, Total Wealth in this setting also takes non-mortgage debt into account. Sample includes only those, who are stock market participants in current and previous waves. Specification 1 in Column 2 includes only two control variables, change in home ownership status and change in the number of children in the household. Column 3 reports results for specification 2 which include change in total wealth. Columns 4 and 5 show Specifications 3 and 4 with indicators for the ratio of the present-value of human capital to total wealth. All Specifications include change in age (starting with 24 to 26 age change and to 83 to 85 age change) and year dummies. The coefficients for wealth variables are the 2-year growth rates. Since total wealth does not enter Specification 1, the result in the corresponding column is identical to the one previously reported.

	(1)	(2)	(3)	(4)
owner	0.0102	-0.00653	0.00581	0.00291
	(0.55)	(-0.31)	(0.27)	(0.14)
kids	0.0109	0.00737	0.00901	0.00852
	(0.83)	(0.54)	(0.66)	(0.62)
ΔlogTW		0.0687***	0.0900***	0.0906***
		(8.46)	(8.69)	(8.68)
ΔPVYW			0.000528***	
			(3.70)	
ξ^1				0.000691***
				(3.51)
ξ^2				0.00196
				(1.06)
N	5094	4756	4740	4740
adj. R^2	0.007	0.035	0.040	0.040
F	1.452	2.641	2.758	2.724

cluster robust t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$