Household portfolio choices and nonlinear income risk*

Julio Gálvez
CEMFI

Job Market Paper
This version: September 2017.

Abstract

The recent evidence on income dynamics show that the income risk that households face is highly nonlinear. First, household earnings display varying degrees of persistence that depend on the size of past and current earnings shocks. Second, household earnings distributions exhibit significant asymmetries. In this paper, I study the role of uninsurable income risk on household portfolio decisions over the life cycle when income risk is nonlinear. I motivate the empirical analysis by analyzing the implications of nonlinear earnings dynamics in a portfolio choice model with participation costs. I then develop a flexible, semi-structural framework to empirically quantify the transmission of latent, time-varying income shocks to household stock market participation and portfolio choice decisions. I provide conditions that guarantee nonparametric identification, and propose a tractable, simulation-based estimation algorithm. Using the recent waves of the PSID, I find that differences in income uncertainty across households drive heterogeneous extensive and intensive margin responses on portfolio allocation.

Keywords: Household portfolios, income risk, sample selection, censoring, panel data, quantile regression, latent variables.

*Contact information: CEMFI, Casado del Alisal 5, Madrid, 28014, Spain, galvez@cemfi.edu.es. I am extremely grateful to Manuel Arellano, Javier Mencía, and Enrique Sentana for their invaluable guidance and support. Special thanks goes to Stéphane Bonhomme for numerous discussions. I also thank Tincho Almuzara, Dante Amengual, Diego Astorga, Julio A. Crego, Ivan Fernandez-Val, Charles Gottlieb, Nezih Guner, Christian Julliard, Hamish Low, Michael Manove, Mónica Martinez-Bravo, Pedro Mira, Matthew J. Notowidigdo, Borja Petit, Josep Pijoan-Mas, Rafael Repullo, Juan Carlos Ruiz, Andreas Stegmann, Stijn van Nieuwerburgh, as well as seminar participants at CEMFI and Carlos III, for helpful comments and suggestions. I acknowledge funding from the Spanish Ministry of Economics and Competitiveness, grant no. BES-2014-070515. Finally, I am thankful to Francisco Gomes for sharing his Fortran code. All remaining errors are mine.
1 Introduction

Households invest in financial assets, such as stocks, to transfer wealth across periods, and to pool different risks, with the goal of smoothing consumption. When they make their investment decisions, however, households encounter and manage various idiosyncratic and aggregate risks. The primary and most important source of idiosyncratic risk that households face, which they can neither avoid nor insure themselves against, is on their labor income. As households experience unique earnings histories, their portfolio choices may differ depending on the size and durability of the shocks they receive. In this paper, I empirically assess the asymmetric impact of earnings shocks on household stock market participation and portfolio choices by developing a novel semi-structural framework.

An extensive literature in macroeconomics and finance has studied how uninsurable labor income shocks affect household consumption, saving, and portfolio allocation decisions over the life cycle, as well as its impact on asset prices.\(^1\) To ensure a reasonable standard of living, theory predicts that households respond by accumulating precautionary savings. Moreover, they may reduce their exposure to avoidable risks. For example, they may lower the amount of their wealth invested in equities. The margin of these adjustments, however, depends on the precise nature of earnings dynamics. Previous literature has relied on linear earnings processes as a workhorse model to analyze these decisions. A consensus that has emerged from theoretical and empirical studies is that the effect of labor income risk, while consistent with theory, is quantitatively small. As a result, earnings risk has seemingly lost its appeal as a candidate for explaining household stock market participation and portfolio choice decisions.

Yet recent contributions to the earnings dynamics literature document that household labor income substantially departs from the features that characterize linear earnings processes along two important dimensions.\(^2\) First, household earnings display varying degrees of persistence that depend on the size of past and current earnings shocks. Second, household earnings distributions exhibit significant asymmetries. In contrast, linear processes imply that regardless of households’ earnings histories, all shocks display the same persistence. Moreover, the distributions implied by these processes are typically assumed to be Gaussian, which rules out dynamic skewness in earnings. Due to these features, linear models rule


\(^2\)Arellano et al. (2017) model the persistent component of income as a conditional quantile function of the past persistent component and the current earnings shock. They document nonlinear persistence across households, and asymmetries of the conditional earnings distribution in the US PSID. Guvenen et al. (2015) use US Social Security administrative data and find that earnings exhibit considerable negative skewness and extremely high kurtosis. They propose a parametric earnings process in which earnings is modelled as the sum of two AR(1) processes with Normal mixture innovations. De Nardi et al. (2016) present a method to nonparametrically estimate an age-specific Markov chain directly from US PSID data, and uncover asymmetric income persistence as well.
out nonlinear transmissions of income shocks that are likely to have a first-order effect on household portfolios.

This paper re-examines the role of uninsurable income risk on household investment decisions over the life cycle through the lens of nonlinear earnings dynamics. I develop a panel data-based estimation framework that allows me to study two economic choices that households make with respect to their financial portfolios. First, I assess their willingness to bear financial risk, the so-called extensive margin of stock market participation. Second, conditional on participation, I analyze households’ portfolio allocations, the intensive margin. This modelling decision is justified by two robust empirical findings in the household finance literature: first, stock market participation is limited at all ages, and second, households adjust their financial portfolios as they age.\(^3\)

To motivate my empirical analysis, I study the possible implications of a more flexible specification of labor income dynamics in a standard model of household stock market participation and portfolio choice, both in a two-period set-up (e.g., Campbell and Viceira (2002)) and in a life-cycle one (e.g., Cocco et al. (2005), Alan (2012), and Fagereng et al. (2017a)). In both contexts, I find that an earnings process with varying persistences and income asymmetries results in quantitatively different implications for households’ stock market participation and portfolio choice decisions compared to the predictions implied by a linear earning process. This is because the risks that a household faces in terms of the persistence of its earnings history, combined with the possibility of negative income realizations for a high-income household, result in current and future household labor income becoming more uncertain than otherwise. As a consequence, it becomes more disposed to avoid other risks by not buying stocks. The household will only enter the stock market if its wealth is sufficiently high enough to insure its consumption against potentially large, negative income shocks. In comparison, the implied amount of wealth the household needs to buy stocks under a linear earnings process is smaller.

I then specify a semi-structural representation of the model of household stock market participation and portfolio choice. The model I propose builds on recent contributions in the panel data literature that identify and estimate general dynamic nonlinear systems. Among other features, these models allow for the presence of latent, time-varying variables, such as the persistent and transitory components of income. The econometric framework can also accommodate other components of the economic model, such as households’ wealth dynamics and time-invariant heterogeneity capturing unobserved preferences or discount rates. In my set-up, the empirical portfolio and participation rules are specified as age-specific, nonlinear

\(^3\)See Guiso et al. (2002) and Guiso and Sodini (2013) for excellent surveys of the literature.
functions of the latent earnings components and assets. The recovery of these rules permits the calculation of average derivative effects to an increase in income, or wealth. More importantly, the approach allows to compute objects of interest that reveal empirical nonlinearities. In particular, I can assess the extent to which income shocks influence extensive and intensive margin responses of portfolio allocation.

As households self-select into stock market participation, the resulting econometric model is one in which the dependent variable, the risky asset share, is also latent. I establish nonparametric identification by extending arguments made in the literature on nonlinear models with latent variables. Nonparametric identification in this set-up requires two additional assumptions. First, the mapping between the latent and observed distributions of risky asset shares must be known. Second, I require a variable that shifts participation costs, but not the subsequent portfolio choice. Provided that these hold, I identify the empirical participation and portfolio allocation rules, the average derivative effects, and the impulse responses that correspond to these rules from variation in earnings, assets, and participation data.

To estimate the model, I rely on a simulation-based algorithm that combines recent developments in quantile regression with sieve estimation approaches. Specifically, I take advantage of the stochastic EM algorithm adapted to a time-varying latent variable set-up by Arellano et al. (2017), and combine it with the quantile selection model proposed by Arellano and Bonhomme (2017). The estimation procedure alternates between two steps: first, simulation draws from the posterior distribution of the latent persistent income components, and second, a sequence of likelihood maximization for the participation rule, and quantile regressions for the portfolio rule. An added advantage of the estimation approach is its computational tractability, as the moment conditions for the portfolio rule lead to a convex linear programming problem, which is one of the appealing features of quantile regression methods (Koenker (2005)).

I estimate the semi-structural model using the 1999 to 2009 waves of the US Panel Study of Income Dynamics (PSID), with a particular focus on working-age households. The descriptive statistics indicate that around 53 percent of households transition between stock market entry and stock market exit. These households have higher labor income and wealth than those who never participate in the stock market, but have lower labor income than those who always participate in the stock market. Moreover, there is wide variation across households in their earnings and asset holdings. The estimation results show that nonlinearities in income shocks indeed result in sizeable participation and portfolio allocation responses.

In the appendix to this paper, I consider an alternative procedure based on the censored quantile regression estimator of Buchinsky and Hahn (1998). This procedure is convenient in my setting as it is a special case of the quantile selection model in the absence of an exclusion restriction.
Specifically, the decrease in average stock market participation rates for high income households hit by very negative income shocks is approximately 7 percent; in comparison, low income households who are hit by the same shock have a decrease in participation rates by 3 percent. Meanwhile, the increase in average stock market participation rates for low income households hit by very positive income shocks is approximately 13 percent; in comparison, high income households hit by the same shock increase their participation rates by 4 percent. With respect to the intensive margin, high income households hit by an very negative income shock decrease the proportion of their wealth in stocks to as much as 2 percentage points; in contrast, the decrease in the portfolios of low income households is 0.5 percentage points. Meanwhile, low income households hit by a very positive income shock increase their risky asset shares by as much as 3 percentage points, compared to 2 percentage points from high income households hit by the same shock.

This paper is related to a wide body of empirical literature that studies the impact of labor income risk on household portfolio choices. These include, among others, Guiso et al. (1996), Heaton and Lucas (2000b), Vissing-Jørgensen (2002), Angerer and Lam (2009), Palia et al. (2014), and Fagereng et al. (2017b). Research in this literature has traditionally relied on linear earnings processes and standard econometric methods to investigate the relationship I study here. Relative to these papers, my main contribution is to develop a new empirical methodology that allows for the possibility of studying nonlinear relationships between income risk, stock market participation, and household portfolio choices in a panel data setting. Furthermore, as opposed to previous literature that focuses on measures of income risk, I focus on the impact of earnings shocks, which have a closer connection to life-cycle structural models, such as those by Haliassos and Bertaut (1995), Heaton and Lucas (2000a), Viceira (2001), Cocco et al. (2005), Gomes and Michaelides (2005) and Polkovnichenko (2007), among others.

This paper is also related to more recent empirical work that has looked at the impact of adverse labor market events on stockholding and portfolio choice. The main conclusion of papers in this literature (e.g., Alan (2012), Betermier et al. (2012), Basten et al. (2016), and Knüpfert et al. (2016)) is that households adjust their portfolios in response to events such as unemployment, job switches, and the probability of a zero income realization. To the extent

---

5Briggs et al. (2015), using Swedish data, find that a wealth shock increases participation by 3 percent, and is mainly driven by non-participants. In the case of my paper, both participants and non-participants react similarly.

6In comparison, the predicted decline by Fagereng et al. (2017b) from a one standard deviation increase in income risk is -0.14 percentage points.

7Alan (2012) finds that a positive probability of a zero income realization is needed in order to explain household portfolio decisions of younger, poorer households in a structural model. Betermier et al. (2012), using a panel of Swedish households, find that the more volatile the wage is, the lower the exposure of households to risky assets will be, and the less likely they participate in the stock market. Basten et al. (2016), using Norwegian
that a nonlinear earnings process can be thought of as a parsimonious representation of such adverse events, I complement this literature by considering how households’ participation behavior and portfolio choice decisions change in response to asymmetric earnings shocks over the life-cycle.

My paper is also related to a small but burgeoning literature that studies the implications of the features that are evident in more flexible representations of earnings dynamics on household consumption and savings behavior, and on asset prices. These papers argue that nonlinear features of income result in asymmetries in how households insure their consumption against income shocks (e.g., Guvenen et al. (2015) and Arellano et al. (2017)) or in their wealth accumulation patterns (De Nardi et al. (2016)). Higher-order moments of income have also been shown to be a key driver of asset prices (e.g., Schmidt (2015) and Constantinides and Ghosh (2017)). To the best of my knowledge, this paper is the first to empirically investigate the impact of asymmetric earnings shocks on household portfolio choice decisions.

Finally, my paper is related to recent developments in the nonlinear panel data literature (e.g., Arellano et al. (2017) and Bonhomme et al. (2017)) that propose methods to estimate dynamic systems in the presence of nonlinearities and unobserved heterogeneity, and investigate the nonparametric identification of such models. With respect to this literature, I propose an estimation framework that takes into account situations in which sample selection is paramount. The estimation procedure considered in this paper can also be used to analyze other economic models that exhibit similar features, for example, labor supply (e.g., Heckman (1974) and papers surveyed in French and Taber (2011)), human capital accumulation (e.g., Imai and Keane (2004)) and occupational choice (e.g., Adda et al. (2017)).

The remainder of the paper is organized as follows. Section 2 investigates the potential implications of nonlinear earnings processes on household stock market participation and portfolio choices. Section 3 presents the semi-structural framework that I consider. Section 4 presents the data and some descriptive statistics. Section 5 presents the estimation procedure that I operationalize for my empirical analysis. Section 6 presents the empirical evidence from registry data, find some households who can anticipate job loss prepare for unemployment by increasing their saving and shifting toward riskless assets leading up to unemployment, and by depleting their savings after the job loss. Around two years after unemployment, however, they begin to rebalance their portfolio toward risky assets. Finally, Knüper et al. (2016) find, within the context of the Finnish Great Depression, that adverse labor market conditions affect both stock market participation and household portfolio choice.

*The notion of disasters that this paper considers, which can be thought of as “microeconomic disasters”, is different from those considered by Alan (2012), whose original motivation was to understand whether the aggregate disasters channel proposed by Barro (2006) can explain limited stock market participation. Arguably, individual disasters happen more frequently to individuals, and their consequences have a clear-cut empirical content. Meanwhile, the main difference between the notion of disasters in this paper and that of Fagereng et al. (2017a), who look at individual stock market disasters as a rationalization of Norwegian household portfolios, is that the risk I consider relates to a household’s human capital, as opposed to the risk faced with respect to financial wealth.*
the PSID data. Finally, section 7 concludes.

2 Theoretical motivation

To motivate my empirical investigation, I explore the possible implications of nonlinear income risk on household portfolios. I begin by studying a two-period model with participation costs that closely follows Campbell and Viceira (2002). Then, I calibrate a life-cycle model to illustrate the differences between a linear and a nonlinear earnings process that is mainly based on Cocco et al. (2005).\(^9\)

2.1 Stylized two-period model

Consider a household with a utility function \( U \) and wealth \( W_t \) that makes a financial portfolio decision at time \( t \). It consumes the liquidation value of the portfolio plus labor income \( Y_{t+1} \) one period later. Labor income is stochastic, and follows a distribution \( H(Y) \). The household cannot borrow against future labor income, thereby making it non-tradeable. It has access to two assets for investment: a riskless asset that has a certain return \( r_f \) and a risky asset with a constant expected log excess return \( E_t(r_{t+1} - r_f) \equiv \mu \). To invest in the stock market, the household must pay a fixed cost of participation \( q \).\(^{10}\)

To make its optimal decision, the household considers two subproblems. The first corresponds to the situation in which it does not participate in the stock market. In this case, it calculates optimal consumption, \( C_{np,t+1} \), via the following maximization problem:

\[
V_{np} = \max_{C_{np,t}} \mathbb{E}[U(C)] \\
\text{s.t. } C_{np,t} = W_t(1 + R_f) + Y_{t+1}
\]

In the second subproblem, the household invests part or all of its wealth in stocks. It then solves for the optimal portfolio share, \( \alpha_t \), via the following maximization problem:

\[
V_p = \max_{\alpha_t} \mathbb{E}[U(C)] \\
\text{s.t. } C_{p,t} = W_{c,t}[\alpha_t(R_{t+1} - R_f) + R_f + 1] + Y_{t+1}
\]

in which I have defined \( W_{c,t} = W_t - q \). As the household cannot borrow, nor can it short-sell, the optimal portfolio share is constrained to be in between zero and one. Moreover, it neither

\(^9\)Alan (2012), Briggs et al. (2015), and Fagereng et al. (2017a) also simulate a life-cycle model with fixed participation costs in a partial equilibrium set-up. Favilukis (2013), meanwhile, builds a general equilibrium model that is based on an overlapping generations framework.

\(^{10}\)One can rationalize this cost as a way of capturing several explanations proposed for limited participation in financial markets. These include the presence of trading costs (e.g., Vissing-Jørgensen (2002)), financial sophistication and financial literacy, or the lack of it (e.g., Calvet et al. (2007), Van Rooij et al. (2011)), and trust in financial markets (e.g., Guiso et al. (2008)).
knows the realization of future stock market returns nor future labor income when it makes the portfolio choice decision.

Although a closed form solution generally does not exist, if utility is CRRA with parameter $\gamma$, labor income is lognormal, and the stock return is lognormal with variance $\sigma_u^2$, an approximate formula for the optimal portfolio shares in the participation problem can be obtained as a function of expected future labor income $\tilde{H} \equiv E_t(Y_{t+1})$ and wealth net of participation costs, $W_{c,t}$:

$$\alpha_t = \left(1 + \frac{\tilde{H}}{W_{c,t}}\right) \left[\frac{E_t(r_{t+1} - r_f) + \frac{1}{2}\sigma_u^2}{\gamma \sigma_u^2}\right]$$

where I assume that the covariance between the stock return and labor income is zero; that is, under idiosyncratic risk. To solve the optimization problem, the household compares the indirect utilities calculated from the two subproblems. The household will clearly participate if the expected utility from equity investment is at least as high as that of non-investment. This condition is equivalent to asserting that to find it worthwhile to take on risk, the household’s future consumption if it invested part of its wealth in stocks the previous period should be at least the same as when it did not invest. Thus, if it decides to participate in the stock market, the optimal portfolio rule is characterized by equation (1). Otherwise, the optimal portfolio rule is characterized by $\alpha_t = 0$.

**Comparative statics.** Equation (1) allows me to study the effect of increases in wealth and labor income, respectively. First, keeping labor income constant, an exogenous increase in wealth reduces the portfolio share, as total household wealth becomes a more important source to draw consumption from than labor income. Hence, the household will not invest in stocks, and might prefer to save in riskless bonds, or to spend part of the wealth gain on goods. This result is true regardless of whether labor income is lognormal or not.

Second, an increase in labor income has an ambiguous effect on household portfolio shares when I relax lognormality. This is because now, labor income will be affected by higher-order moments. For example, an increase in labor income might yield the household to invest less in stocks if the distribution of its earnings is negatively skewed. In this case, even if the household experiences an increase in labor income, the possibility that the realization is extremely negative might lead the household to lose money from taking on more risk. In

---

11In the appendix that corresponds to this subsection, I provide a full exposition of the portfolio choice problem. A more formal treatment of portfolio choice necessitates working with higher-order cumulants as in Martin (2012), but this is left for further research.

12The comparative statics results I discuss here apply to a household who is currently a stock investor. In the appendix, I derive results for the marginal investor who is indifferent between entering and exiting the stock market. I provide conditions under which the marginal investor will continue to participate in the stock market with respect to increases in wealth and labor income.
contrast, under lognormality, an increase in labor income will lead the household to invest aggressively in stocks if the expected labor income is greater than the risk the household faces, which is captured by its variance.

2.2 Life-cycle model

The previous subsection highlights the effects of human capital and wealth on stock market participation and portfolio choice decisions in a two-period set-up. Its static nature, however, prevents the analysis of the dynamic effects of income shocks on portfolio allocation.

To this end, I consider the problem of a household that maximizes the expected utility of its consumption over the life-cycle. The household works up until retirement, and dies with certainty at the end of its life.\textsuperscript{13} I describe in detail the labor income process and the results of the model simulation in this subsection. Additional details are outlined in Appendix A.3.

2.2.1 Nonlinear earnings dynamics

Before retirement, the household’s labor income is:

\[ y_{it} \equiv \log(Y_{it}) = f(t, X_{it}) + v_{it} + \epsilon_{it} \]

where \( f(t, X_{it}) \) is a deterministic function of age and other characteristics, \( v_{it} \) is the persistent component, and \( \epsilon_{it} \) is the transitory component of income. At age \( t \), it knows the present realizations of \( v_{it} \) and \( \epsilon_{it} \), and their past values, but not \( v_{it+1} \) and \( \epsilon_{it+1} \).

I consider the Arellano et al. (2017) earnings process in modelling persistent income. Denoting by \( Q_t(v_{it}, \tau) \) the \( \tau \)th conditional quantile of \( v_{it} \) given \( v_{it-1} \) for each \( \tau \in (0, 1) \), a representation of the persistent component is:

\[ v_{it} = Q_t(v_{it-1}, u_{it}) \]

where \((u_{it}|v_{it-1}, v_{it-1}, \ldots) \sim U[0, 1]\) for all \( t \). The distribution of the initial condition \( v_{i0} \) is left unrestricted. Meanwhile, the transitory component \( \epsilon_{it} \) is assumed to have zero mean, is independent over time, and independent of all realizations of the persistent component.

Figure 1 presents the features inherent in the nonlinear earnings process. Panel (a) illustrates persistence calculated as a function of the past component of persistent income, and the shock that the households receive, computed at mean age (44.7 years). As the figure shows, income persistence depends on the direction and magnitude of current and future earnings shocks. It is high for low-earnings households who are hit by an extremely bad shock, and

\textsuperscript{13}Viceira (2001) considers an infinite-horizon life-cycle model, and derives approximate analytical expressions for portfolio shares both in retirement and working age periods.
high-earnings households who are hit by an extremely good shock. In contrast, it is low for high-earnings households who are hit by an extremely bad shock, and low-earnings households who are hit by an extremely good shock.

Figure 1: Nonlinear earnings process

(a) Nonlinear persistence
(b) Conditional skewness

Note: This figure illustrates the features of the nonlinear earnings process. Panel (a) depicts the persistence of the nonlinear earnings process of Arellano et al. (2017); that is, the average derivative of the conditional quantile function of $v_{it}$ on $v_{it-1}$ with respect to $v_{it-1}$. I calculate persistence at different percentiles of the past persistent component $\tau_{init}$ and the current earnings shock $\tau_{shock}$. Panel (1b) presents the conditional skewness of the distribution of $v_{it}$ given $v_{it-1}$. Both results are estimated from my PSID subsample.

Panel (b), meanwhile, indicates the presence of conditional asymmetries in the conditional distribution of $v_{it}$. Specifically, the distribution of $v_{it}$ is positively skewed for low values of $v_{it-1}$ and negatively skewed for high values of $v_{it-1}$. This implies that when low-income households are hit by a good earnings shock, there is a large probability of getting outcomes far to the right of their earnings distribution. Likewise, high-income households who are hit by a bad earnings shock have a large probability of getting extremely negative outcomes. In Appendix A.2, I compare the fit of the linear and nonlinear earnings processes to the data, and I find that the nonlinear earnings process fits the data well.

2.2.2 Simulation results (in progress)

Figure 2 illustrates the simulated portfolio rules of households, which were computed by taking expectations over all income realizations. I compare three models. The first is the benchmark Cocco et al. (2005) model without participation costs. The second is the benchmark Cocco et al. (2005) model, with the inclusion of a fixed cost of participation, but with a linear earnings process. Finally, the third is a model that includes not only the fixed cost of
participation but also the nonlinear earnings process.

Figure 2: Portfolio rules, ages 45 and 70

(a) Portfolio rules during working age

(b) Portfolio rules at retirement

Note: This figure illustrates the difference between the linear earnings process and the nonlinear earnings process in terms of their impact on household portfolios. Panel (a) depicts the portfolio rules for a household who is of working age. Meanwhile, panel (a) depicts the portfolio rules for a household at retirement. The dotted blue line corresponds to the model of Cocco et al. (2005) without fixed participation costs, the green line corresponds to the model with fixed participation costs, and the red line corresponds to the model with fixed participation costs and the nonlinear earnings process of Arellano et al. (2017).

I find the following results. First, the portfolio rules for all three models are decreasing functions of cash-on-hand. The intuition behind this can be connected to the portfolio rule (1) in the two period model; that is, the key driver is the importance of human capital (i.e., the discounted stream of future labor income), which mimics the pay-off of a risk-free bond, relative to total accumulated household wealth. At lower levels of wealth, households have a relatively larger amount of future labor income, and thus, become inclined to invest more aggressively in stocks. As households accumulate wealth, however, the relative importance of human capital becomes smaller. This leads households to invest less heavily on stocks, as now they have a sizeable amount of wealth to draw consumption from.

Second, the introduction of a participation cost introduces a wealth participation threshold for households in working age. However, the size of that threshold is relatively small. Moreover, once households hit that threshold, the model predicts that they will invest up to 100 percent of their wealth in stocks. This is at odds with the usual empirical findings that state that households do not fully allocate their wealth in stocks.

Third, the nonlinear earnings process introduces two new results. First, I find that the wealth participation threshold on average increases relative to the previous model. Second, for some ages (such as the one I show here), households do not fully allocate their wealth into
stocks. This is due to the fact that, with a nonlinear earnings process, human capital becomes riskier than before. This implies that households at low wealth levels become less inclined to participate in the stock market. Even when they do, however, households do not allocate their wealth fully to stocks.\footnote{In simulations that I do not present here, I find that as households age, the allocation to stocks in the threshold rises up to a point when they achieve full allocation. The participation threshold also decreases.}

### 3 A flexible, semi-structural approach

The results in the previous section suggest that introducing a non-standard earnings process into an otherwise standard structural model reveals qualitatively different implications for household stock market participation and portfolio choice behavior. Empirically characterizing the nonlinear relationships between the variables in the model is challenging, as economic theory typically does not suggest a particular functional form.

The goal of this section, hence, is to develop an estimation framework that is consistent with several dynamic structural models of household portfolio choice. The first subsection outlines the nonlinear reduced form model. An important advantage of this approach is the calculation of objects of interest that reveal empirical nonlinearities, which I describe in the second subsection. I then briefly discuss the nonparametric identification of the objects at hand. Finally, I comment on several extensions of the baseline model.

In what follows, I consider a cohort of households $i = 1, \ldots, N$ that act as single agents, and denote household age by $t$.

#### 3.1 Empirical portfolio choice and participation rules

The semi-reduced form of the life-cycle portfolio choice model with a fixed cost of participation is represented by the following system of equations:

\begin{align*}
\alpha_{it}^* &= g_t(v_{it}, \varepsilon_{it}, w_{it}, X_{it}, u_{it}) \\
\alpha_{it} &= \alpha_{it}^* \cdot d_{it} \\
d_{it} &= \begin{cases} 
1, & \text{if } \tilde{g}_t(v_{it}, \varepsilon_{it}, w_{it}, q(Z_{it})) \leq v_{it} \\
0, & \text{otherwise} 
\end{cases} \\
w_{it} &= h_t(v_{it-1}, \varepsilon_{it-1}, w_{it-1}, \alpha_{it-1}, X_{it}, \zeta_{it}) \\
w_{i0} &= \tilde{h}_{i0}(v_{i0}, X_{i0}, \zeta_{i0})
\end{align*}

Equation (3) is the solution of the participation subproblem; that is, the share of wealth the household will invest in the stock market. The portfolio rule is an age-dependent function
$g_t$ of the persistent and transitory components of income, wealth, and a vector $X_{it}$ of characteristics that control for observed life-cycle or preference shifters. In the baseline model, $u_{it}$ is an unobserved preference shifter that increases households’ marginal utility; this implies that $g_t$ is monotone in $u_{it}$. Equation (4), meanwhile, corresponds to the optimal solution of the economic problem; as in the model, the household either invests a proportion $\alpha_{it}^*$ into stocks, or zero.

The household’s decision depends on the participation rule summarized by equation (5). It is a reduced-form characterization of the comparison of the indirect utilities (i.e., value functions) the household gains from participation and non-participation, respectively. The arguments of this rule are the persistent and transitory components of income, wealth, and the participation cost $q(Z_{it})$, which is proxied by function of observed characteristics $Z_{it}$. In this model, $Z_{it} = \{X_{it}, p'_{it}\}$, in which $X_{it}$ are the same characteristics described in the previous paragraph, and $p'_{it}$ is a vector that corresponds to variables that can potentially affect participation, but not the portfolio choice decision. $v_{it}$ is an error term that captures unobserved characteristics that affect households’ participation decisions.

Equation (6) characterizes the household’s wealth dynamics that are summarized by the household’s budget constraint. It is a function of the previous period’s realization of latent earnings components, wealth, the risky asset share, and the current period’s demographic characteristics. The error term $\zeta_{it}$ is a catch-all for aspects of the model that I do not explicitly specify, such as the consumption rule, or the return that the household gains from the investment decision it makes, regardless of its particular choice. Finally, equation (7) is the wealth of the household during the first period of observation. As can be seen, it is a function of the persistent component of income and the characteristics of the households.

The dynamic econometric model represented by equations (3)-(7), added with the non-linear earnings dynamics representation in equation (2), permits the direct estimation of the household’s portfolio and the participation rules under labor income risk. It is compatible with several classes of structural economic models, as it does not impose a specific functional form. Furthermore, the model’s flexibility permits interactions between the different state variables of the economic problem at hand. This stands in contrast to linear reduced form models, which come from first-order approximations of the economic model. A drawback, though, of the nonlinear semi-structural model that I outline here compared to dynamic structural models is that it cannot be used to analyze counterfactual scenarios. However, structural estimation approaches require the researcher to specify all aspects of the model; in my model,

---

15 This will be true in the participation subproblem if $\frac{\partial u(C, u')}{\partial u} > \frac{\partial u(C, u)}{\partial u}$ where $u' > u$. This implies, hence, that the Bellman equation of the participation subproblem is monotonic.
an example of this is specific functional form assumptions for utilities. The approach that I take here can provide guidance through the calculation of quantities that can serve as robust targets for a structural estimation exercise.\footnote{Equations (2)-(7) could be the auxiliary equations to a structural model estimated via indirect inference.}

### 3.2 Objects of interest

The model described by equations (3)-(7) can be used to calculate the following quantities of interest. To fix ideas, I calculate all objects with respect to the persistent component of income $v_{it}$ in the following paragraphs. However, I can also calculate these objects with respect to wealth $w_{it}$.

First, the model can be used to compute average derivative effects. In particular, for a specific realization $(v_{it}, \varepsilon_{it}, w_{it}, Z_{it}) = (v, \varepsilon, w, z)$, I can calculate the following quantities:

\[
\frac{\partial \Pr(\bar{g}(v_{it}, \varepsilon_{it}, w_{it}, q(Z_{it})) \leq v_{it}|v, \varepsilon, w, z)}{\partial v_{it}} = \frac{\partial \Pr(d_{it} = 1|v, \varepsilon, w, z)}{\partial v_{it}} \tag{8}
\]

and

\[
\mathbb{E} \left( \frac{\partial \alpha_{it}^*}{\partial v_{it}} \bigg| v_{it} = v, \varepsilon_{it} = \varepsilon, w_{it} = w, X_{it} = x, d_{it} = 1 \right) = \mathbb{E} \left( \frac{\partial g_{it}(v, \varepsilon, w, x, \tau)}{\partial v} \bigg| d_{it} = 1 \right) \tag{9}
\]

Expressions (8) and (9) correspond to the marginal effects of an increase in the persistent component of income on portfolio shares and participation, respectively. Note that expression (8) considers all households, regardless of participation status, while (9) considers the subset of households who participated in the stock markets. An interpretation of expression (8) is that it corresponds to extensive margin responses, while equation (9) corresponds to intensive margin responses.

Second, the model estimates can be used to compute “impulse response”-like functions with respect to a shock to the persistent component of income, $v'$. With respect to the conditional probability of participation, I can calculate the following empirical object:

\[
\Delta_{E}(v + v', v) = \Pr(d_{it} = 1|v + v', \varepsilon, w, z) - \Pr(d_{it} = 1|v, \varepsilon, w, z) \tag{10}
\]

Similarly, I can calculate the following “impulse response”-like function for the average portfolio share conditional on participation:

\[
\Delta_{I}(v + v', v) = \mathbb{E}(\alpha_{it}|d_{it} = 1, v + v', \varepsilon, w, x) - \mathbb{E}(\alpha_{it}|d_{it} = 1, v, \varepsilon, w, x) \tag{11}
\]

Finally, combining the two expressions, I can calculate the total change in aggregate portfolio shares. However, it is more computationally tractable to calculate “log changes”:

\[
\log \Delta_{T}(v + v', v) = \log \left[ \frac{\mathbb{E}(\alpha_{it}|d_{it} = 1, v + v', \varepsilon, w, x)}{\mathbb{E}(\alpha_{it}|d_{it} = 1, v, \varepsilon, w, x)} \right] + \log \left[ \frac{\Pr(d_{it} = 1|v + v', \varepsilon, w, z)}{\Pr(d_{it} = 1|v, \varepsilon, w, z)} \right]
\]
As can be observed, the aggregate change in portfolio shares as a result of an income shock is composed of two parts: the “impulse response” associated with participation in the stock markets, and the “impulse response” associated with portfolio adjustments in response to an income shock.

3.3 Nonparametric identification

The main challenge inherent in identifying the portfolio and participation rules comes from the fact that \( \alpha_{it} \) and the earnings components are latent.\(^{17}\) The semi-structural representation of the portfolio choice model with participation, however, takes the form of a nonlinear state-space model. Recent papers (which are surveyed in Hu (2017)) have established conditions that guarantee nonparametric identification of the joint dynamic distributions of the observed and latent variables in these nonlinear models. Hence, I leverage techniques developed in this literature and outline a formal argument of nonparametric identification in Appendix B.

As I show, the empirical participation and portfolio rules, the average derivative effects, and the impulse response functions are nonparametrically identified given at least two periods of earnings, assets, the observed participation and portfolio choices, provided that two assumptions are satisfied. First, the mapping between the latent and observed distributions of risky asset shares must be known. Second, there should be a variable that affects participation, but not the subsequent portfolio allocation. Within the context of the model, the exclusion restriction can be thought of as a cost shifter. It is crucial that both assumptions must be satisfied. Knowledge of the mapping between the latent and observed distributions, which is represented by the conditional copula, permits the calculation of the latent quantiles of risky asset shares from the observed data. However, this function will not be informative of the extent of participation in the stock market without the presence of an exclusion restriction. Otherwise, the quantiles, and subsequently, the distribution of risky asset shares, are only set identified (Arellano and Bonhomme (2017)).

The intuition behind the identification argument comes from the connection to nonparametric instrumental variable problems (see, e.g., Newey and Powell (2003) and Blundell et al. (2007)). In this context, the endogenous variable is the persistent component of income, which is unobserved. As I argue in the appendix, given the assumptions of the model, the “excluded instruments” are the lagged portfolio choices, participation indicators, assets, and the leads and lags of earnings. The availability of these instruments allows me to identify aver-
age derivative effects of the persistent component of income, and participation and portfolio choice responses with respect to an income shock.

Using leads and lags of earnings is a common strategy in identifying consumption responses (see, e.g., Blundell et al. (2008)) with respect to an income shock. This has not been used to identify the impact of income shocks on portfolio allocation. The usual approach is to estimate measures of income risk (such as the variance of labor income) or to use information on subjective income expectations, and use these as a dependent variable in a linear regression. The approach taken here provides the possibility of directly estimating these responses from the available data on earnings, assets and participation.

3.4 Model extensions

One virtue of the estimation framework is its flexibility in incorporating some extensions that are empirically and economically relevant. In particular, I discuss extensions that consider state dependence in participation, household unobserved heterogeneity, the estimation of the consumption function, and the presence of advance information in earnings.

State dependence in participation. The baseline model does not allow for potential state dependence in participation. Including this via a lagged participation indicator in equation (5) permits me to study the dynamics of stock market participation. In particular, I can calculate stock market entry and exit rates, and dynamic impulse responses with respect to the extensive margin.

Household unobserved heterogeneity. Accounting for unobserved heterogeneity might be important, as it could represent latent characteristics that do not change over time, and cannot be controlled for with observable proxies. These could include preferences and discount rates. To develop this extension of the baseline model, I introduce a household-specific fixed effect $\xi_i$ into all of the equations in the system (3)-(7). Moreover, I model the distribution of $\xi_i$ conditional on the initial values of the state variables in the model, and the corresponding household portfolio choices and participation decisions.

Consumption function. The third extension I consider is the introduction of the consumption function into the system of equations. This is primarily because of the link between

---

18Measures of income risk are calculated by either recovering the persistent component from a linear earnings process (as in Angerer and Lam (2009) and Fagereng et al. (2017b)), or from a distributional assumption on income (as in Vissing-Jørgensen (2002)). Fagereng et al. (2017b) goes further and argues that identifying the impact of labor income risk requires a variable that is plausibly exogenous. Guiso et al. (1996) and Hochguertel (2003) are examples of papers that use subjective expectations.

19Some papers that have considered panel data estimators for household portfolios include Brunnermeier and Nagel (2008), Chiappori and Paiella (2011), Calvet and Sodini (2014), and Fagereng et al. (2017b), who take a fixed-effects approach, and Angerer and Lam (2009), which, to the best of my knowledge, is the only paper that takes a random-effects approach.
consumption volatility and stockholding, which has been forcefully argued, by among others, Attanasio et al. (2002). Estimating the consumption function permits the calculation of marginal propensities to consume out of wealth and income. This allows me to quantify, in a more direct manner, the influence of consumption in household portfolio choices (or alternatively, the influence of a more diversified portfolio on consumption insurance).

Advance information in earnings. Finally, households may have advance information about future earnings shocks, which might have an impact on their participation and portfolio choice decisions (see Blundell et al. (2008) for an example in the context of consumption and savings decisions). In this case, I modify the portfolio and participation rules via the inclusion of future values of the persistent component of income.

4 Data and descriptive evidence

4.1 Dataset description and sample selection

The main dataset for my empirical analysis is a balanced panel of households from the 1999 to 2009 waves of the Panel Study of Income Dynamics (PSID). The primary aim of the survey was to study the dynamics of income and poverty of US households. Hence, the original 1968 study was drawn from two independent subsamples: 2,000 poor families that were under the Survey of Economic Opportunity (SEO), and a nationally representative sample of approximately 3,000 families. The survey waves were annual from 1968 until 1997, when the data was collected biennially. A distinct advantage of the PSID is that since the 1999 wave, it collects detailed data on consumption expenditures and asset holdings, in addition to information on household earnings. As I need continuous information on labor earnings and portfolio choices over the life cycle, I focus on the 1999 to 2009 waves, which correspond to calendar years 1998 to 2008. I deflate all of the variables with 2000 as the base year.

Sample selection criteria. In the main text, I focus on non-SEO households with participating and married household heads, who are between 25 to 60 years old. I exclude households who have missing information on key demographic variables (age, race, education, and state of residence) and on the main variables in the study, in logs or in levels. To reduce the influence of measurement error, I remove households who have more than $20 million in total household assets, following Blundell et al. (2016). I also drop households who have “extreme jumps” in their earnings and implied hourly wages, and those who have transfer incomes.

A jump is defined as an extremely positive (negative) change from year $t - 2$ to $t$, followed by an extremely negative (positive) change from year $t$ to $t + 2$. Formally, for each variable, I construct the biennial log difference $\Delta^2 \log(x_{it})$ and drop the relevant variables for observation in the bottom 0.25 percentile of the product $\Delta^2 \log(x_{it})\Delta^2 \log(x_{i,t-2})$, following Blundell et al. (2016).
that are more than twice household labor income. The sample selection criteria results in a balanced panel of 661 households. A detailed description of the data cleaning process is in Appendix C.1. I also calculate summary statistics to compare my baseline sample with a sample of all married household heads (independently of work status) and with a sample of all household heads headed by a male recorded at least once in the 1998-2008 period (again, independently of work status). The results indicate that there does not seem to be substantial differences across samples.

4.2 Main variables

Earnings $Y_{it}$ is total pre-tax household labor earnings. In the estimations, I construct $y_{it}$ as the residuals from regressing log household earnings on a set of demographics, which include cohort dummies interacted with education categories for both husband and wife, race, state, and large city dummies, a family size indicator, number of kids, a dummy for income recipient other than husband and wife, and a dummy for kids out of the household.

I consider risky assets as the sum of two components: (i.) the value of stockholdings held in publicly traded corporations, mutual funds or investment in trusts, which I call stocks, and (ii.) the part of Individual Retirement Accounts (IRAs) that are held in stocks. To identify the part of the IRA allocated in stocks, I follow the treatment of Vissing-Jørgensen (2002), Malmendier and Nagel (2011) and Palia et al. (2014). Specifically, the PSID asks a household about the allocation of its pension account, if it has any. I assume that all investments in IRAs if the household reports that most of the money is invested in stocks. If the household reports that the money in the IRA is split between stocks and interest-earning assets, I assume that half the value is in stocks and half the value is in bonds.

Total household wealth $W_{it}$ is constructed as the sum of financial assets, real estate value, pension funds, and car value, net of mortgage and other debt. Financial assets is the sum of the following sources: stocks; cash, which is defined as checking or savings accounts, money market accounts, or Treasury bills, including those held in individual retirement accounts (IRAs); bonds, which includes bonds, the cash value in life insurance policies, valuable col-

\footnote{Albeit some papers in the empirical literature (such as Brunnermeier and Nagel (2008) and Chiappori and Paiella (2011)) include home equity in the definition of risky wealth, as it can be interpreted as such (see Flavin and Yamashita (2002)), doing so requires a more involved estimation framework in which I would also need to model homeownership, and the evolution of house prices, an aggregate state variable (see Hahn et al. (2015) for a discussion on the challenges of estimating models with aggregate shocks). Moreover, empirically disentangling the effects of house price risk from labor income risk is an arduous task, as underscored by Chetty et al. (2017). This definition of wealth recognizes, however, that households indeed have most of their wealth in housing. In this sense, I can also interpret the error term in the evolution of wealth equation as one that also captures the return process of housing value. Moreover, my sample selection criteria is such that almost all households in my sample are homeowners at any given point in time.}
lections, rights in trusts or estates; and pension funds. All of the estimations that I present use the log of total household wealth, \( w_{it} \), as the relevant independent variable.

Finally, the risky share \( \alpha_{it} \) is computed as the proportion of risky assets to total household wealth.

### 4.3 Descriptive evidence

Table 1 presents pooled cross-section/time-series summary statistics for all relevant variables, grouped by income quartiles.\(^{22}\) The table is divided into two panels. The first panel corresponds to all of the households who satisfy the sample selection criteria. The second panel, meanwhile, corresponds to the subset of risky asset market participants.

<table>
<thead>
<tr>
<th>Variable</th>
<th>TOTAL</th>
<th>First</th>
<th>Second</th>
<th>Third</th>
<th>Fourth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Household income</td>
<td>111,612.60</td>
<td>42,206.75</td>
<td>71,354.16</td>
<td>99,518.06</td>
<td>222,091.50</td>
</tr>
<tr>
<td>Total assets</td>
<td>371,189.60</td>
<td>184,597.80</td>
<td>212,986.20</td>
<td>332,963.40</td>
<td>719,382.40</td>
</tr>
<tr>
<td>Liquid wealth</td>
<td>95,529.77</td>
<td>43,204.98</td>
<td>48,433.76</td>
<td>73,681.34</td>
<td>206,119.10</td>
</tr>
<tr>
<td>Risky wealth</td>
<td>81,353.17</td>
<td>35,366.25</td>
<td>36,714.01</td>
<td>60,835.09</td>
<td>182,757.20</td>
</tr>
<tr>
<td>Stocks</td>
<td>51,547.79</td>
<td>23,514.87</td>
<td>19,774.00</td>
<td>34,475.20</td>
<td>121,831.80</td>
</tr>
<tr>
<td>Share of stocks in total wealth</td>
<td>0.077</td>
<td>0.048</td>
<td>0.057</td>
<td>0.073</td>
<td>0.124</td>
</tr>
<tr>
<td>Share of risky assets in total wealth</td>
<td>0.148</td>
<td>0.087</td>
<td>0.120</td>
<td>0.154</td>
<td>0.224</td>
</tr>
<tr>
<td>Ownership of stocks</td>
<td>0.410</td>
<td>0.231</td>
<td>0.343</td>
<td>0.444</td>
<td>0.602</td>
</tr>
<tr>
<td>Ownership of risky assets</td>
<td>0.591</td>
<td>0.372</td>
<td>0.526</td>
<td>0.644</td>
<td>0.793</td>
</tr>
<tr>
<td>Risky market participants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household income</td>
<td>130,668.70</td>
<td>44,692.65</td>
<td>71,214.30</td>
<td>99,955.57</td>
<td>226,578.30</td>
</tr>
<tr>
<td>Total assets</td>
<td>512,872.20</td>
<td>317,459.30</td>
<td>289,769.80</td>
<td>399,048.60</td>
<td>818,180.50</td>
</tr>
<tr>
<td>Liquid wealth</td>
<td>145,598.80</td>
<td>101,981.00</td>
<td>73,786.93</td>
<td>97,593.12</td>
<td>244,521.50</td>
</tr>
<tr>
<td>Risky wealth</td>
<td>137,676.70</td>
<td>95,196.34</td>
<td>69,818.05</td>
<td>93,735.98</td>
<td>230,589.40</td>
</tr>
<tr>
<td>Stocks</td>
<td>87,236.05</td>
<td>63,295.65</td>
<td>37,603.70</td>
<td>53,120.10</td>
<td>153,718.30</td>
</tr>
<tr>
<td>Ownership of stocks</td>
<td>0.695</td>
<td>0.621</td>
<td>0.653</td>
<td>0.684</td>
<td>0.759</td>
</tr>
<tr>
<td>Share of stocks in total wealth</td>
<td>0.130</td>
<td>0.129</td>
<td>0.108</td>
<td>0.113</td>
<td>0.156</td>
</tr>
<tr>
<td>Share of risky assets in total wealth</td>
<td>0.250</td>
<td>0.233</td>
<td>0.229</td>
<td>0.237</td>
<td>0.283</td>
</tr>
</tbody>
</table>

*Note: Data from 1999 to 2009 PSID waves. This table presents sample means of the main economic variables related to this empirical study, calculated across income quartiles. The first column calculates the mean across all households in the sample, while the second to fifth columns calculate the mean for different income quartiles. The first panel presents results across all households. The second panel presents results for risky market participants, defined as households who have direct and indirect stockholdings in stocks, mutual funds, and who have part of their pension funds invested in stocks.

The table indicates a wide dispersion across households in their earnings and assets. The average participation rate in risky wealth is around 59.1 percent for the households in my sample, slightly higher from the participation rates observed in other observational studies, such as the US Survey of Consumer Finances (SCF).\(^{23}\) Furthermore, once I condition on the

---

\(^{22}\) In Appendix C.2 I present the same table, but for age and wealth quartiles.

\(^{23}\) The proportions in the SCF are 48.9 (1998), 52.2 (2001), 50.2 (2003), and 51.1 (2007). (Bucks et al. (2009))
subset of risky asset market participants, and stock market participants, I find that household assets are not monotonic in income. In fact, households at the lowest and highest income quartiles have higher liquid wealth, risky wealth and stockholdings than households at the middle income quintiles. While there may be a host of other reasons why this could be the case, one can surmise that differences in the income risks that these households face could possibly drive this phenomenon.

Figure 3: Participation and the conditional risky share, by income and wealth quartiles

(a) Risky market participation rates

(b) Conditional risky shares

Note: Data from 1999 to 2009 PSID waves. The following figures show the average stock market participation rates and the average conditional risky share for households of different income and wealth quartiles. The x-axis corresponds to the income quartiles, while the y-axis corresponds to the average participation rate or risky share. The blue line corresponds to households in the poorest wealth quartile; the red line corresponds to households in the second wealth quartile; the green line corresponds to households in the third wealth quartile; and the orange line corresponds to households in the richest wealth quartile.

Figure 3 presents the average stock market participation and conditional risky share for households of different income and wealth quartiles. As the graphs indicate, with the exception of the highest wealth quartile, participation rates increase as income increases. For households in the highest wealth quartile, however, the graphs suggest that households at the extreme income quartiles have slightly higher participation rates than those in the middle income quartiles.

The conditional risky shares, meanwhile, reveal non-monotonic patterns across income and wealth quartiles. The graphs also indicate that among households in the lowest income quartile, households in the highest wealth quartile tend to be the most aggressive in their risky asset investment. At the same time, among households in the highest income quartile, households in the lowest wealth quartile tend to invest the most in risky assets, on average. Overall, these plots suggest an interaction between wealth and income in terms of risky asset
market investment.

Table 2 presents the frequency distribution of household stock market participation sequences disaggregated by age quartiles. I distinguish between the following groups of households: (i.) those who have never participated in the stock market; (ii.) those who have always participated in the stock market; (iii.) “pure entry” and “pure exit” households, that is, those who have entered or exited at most once into the stock market; and (iv.) households who transition from one participation state to another more than once.

Table 2: Frequency distribution of stock market participation sequences

<table>
<thead>
<tr>
<th>Age quartile</th>
<th>TOTAL</th>
<th>First</th>
<th>Second</th>
<th>Third</th>
<th>Fourth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never participated</td>
<td>110</td>
<td>15</td>
<td>32</td>
<td>52</td>
<td>11</td>
</tr>
<tr>
<td>Always participated</td>
<td>199</td>
<td>11</td>
<td>40</td>
<td>108</td>
<td>40</td>
</tr>
<tr>
<td>Pure entry households</td>
<td>51</td>
<td>9</td>
<td>9</td>
<td>29</td>
<td>4</td>
</tr>
<tr>
<td>Pure exit households</td>
<td>43</td>
<td>1</td>
<td>15</td>
<td>25</td>
<td>2</td>
</tr>
<tr>
<td>Households with transitions</td>
<td>258</td>
<td>17</td>
<td>70</td>
<td>141</td>
<td>30</td>
</tr>
<tr>
<td>TOTAL</td>
<td>50</td>
<td>166</td>
<td>365</td>
<td>87</td>
<td></td>
</tr>
</tbody>
</table>

Note: Data from 1999 to 2009 PSID waves. This table presents the frequency distribution of household stock market participation sequences disaggregated by age quartiles. In this table, households who have never bought stocks are those who have participation sequences (0,0,0,0,0,0), while households who have always participated are those who have participation sequences (1,1,1,1,1,1). Pure entry and pure exit households are those who have the sequences that are described in the text. Households who have transitions are those who enter or exit more than once.

Looking at the panel of risky market participation sequences, households who move in between entry and exit comprise the majority of households in the sample, at 39.03 percent. This is followed by households who always participated, who are approximately 30.10 percent of the sample. Households who have never participated in the stock markets comprise 16.64 percent, and the rest are the “pure entry” and “pure exit” households. Households with multiple transitions also comprise the majority of households who participate in the stock markets, at 35.41 percent. This is followed by those who have never entered the stock market, at 32.89 percent. Households who have always participated in the stock markets comprise 15.27 percent; “pure entry” and “pure exit” households comprise the rest.

Table 3 presents summary statistics according to stock market participation status. Households who have never participated tend to be less educated, have lower household incomes, and have lower household wealth. The rest of the households in my sample tend to have studied at least a year of university education and are roughly of the same age category.

---

24 A pure entry household is one with a participation sequence of (0,1,1,1,1,1), (0,0,1,1,1,1), etc.
Table 3: Summary statistics, by stock market participation sequence

<table>
<thead>
<tr>
<th>Participation sequence status</th>
<th>Risky market participation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age of household head</td>
<td>43.32</td>
</tr>
<tr>
<td>Education of household head</td>
<td>12.24</td>
</tr>
<tr>
<td>Household income</td>
<td>61,808.22</td>
</tr>
<tr>
<td>Net household wealth</td>
<td>110,145.50</td>
</tr>
<tr>
<td>Liquid wealth</td>
<td>8,921.91</td>
</tr>
<tr>
<td>Risky wealth</td>
<td>196,810.00</td>
</tr>
<tr>
<td>Ownership of risky assets</td>
<td>1.00</td>
</tr>
<tr>
<td>Share of wealth in risky assets</td>
<td>0.29</td>
</tr>
</tbody>
</table>

Note: Data from 1999 to 2009 PSID waves. This table presents summary statistics of the main variables in the PSID subsample that I consider, disaggregated by stock market participation status. The first column corresponds to households who have never participated in the stock market. The second column corresponds to households who have always participated in the stock market. The third and fourth columns correspond to households who have either “purely entered” or “purely exited” the stock market. Finally, the last column corresponds to households who transition in and out of the stock market.

Interestingly, those who have at transitioned at least once between entry and exit have labor incomes that are less than those of who have always participated in the stock markets. Moreover, households who transition more than once have lower household incomes, lower household net wealth, lower risky household wealth, and lower risky asset shares. Arguably, these results suggest that households who transition at least once are those who are at the margin between participation and non-participation in the stock markets. Furthermore, the labor incomes these households have suggest that they face substantial earnings risk.

5 Estimation strategy

In section 3, I showed that the life-cycle model of household stock market participation and portfolio choice can be translated into a dynamic, nonlinear econometric model. I also illustrated the empirical objects of interest that can be recovered, and discussed how they can be identified with the data. Armed with these results, I specify each component of the system of equations that describe the econometric model, which I describe in the first part of this section. To do so, I rely on sieve estimation approaches, such as those outlined in Chen (2007); that is, each equation is specified as an age-dependent, nonlinear function of the latent earnings components and wealth.

There are two main challenges in estimating this system of equations. First, the reduced form portfolio choice and participation rules are functions of latent variables; in particular, the
arguments of the estimating equations include, apart from wealth, the persistent and transitory components of income. Second, the economic model suggests that the dependent variable of interest, which is the risky asset share, is latent; this is because in my model, households select themselves into buying stocks. To this end, I propose an estimation procedure based on recent developments in the panel data and the sample selection literature that deal with these econometric issues.

5.1 Model specification

In what follows, let \( \varphi_k \), for \( k = 1, \ldots, K \), denote a dictionary of functions, with \( \varphi_0 = 1 \).

**Portfolio rule.** I first discuss the specification for the portfolio rule (3). Letting \( \text{age}_{it} \) denote the age of the household head \( i \) at period \( t \), I specify the portfolio rule as:

\[
\alpha_{it}^* = g_t(v_{it}, \varepsilon_{it}, w_{it}, X_{it}, \tau) \\
= g(v_{it}, \varepsilon_{it}, w_{it}, X_{it}, \text{age}_{it}, \tau) \\
= \sum_{k=0}^{K} b_k(\tau) \varphi_k(v_{it}, \varepsilon_{it}, w_{it}, \text{age}_{it}) + \gamma^\alpha(\tau)'X_{it} 
\]

(12)

In practice, \( \varphi_k(\cdot) \) is a product of Hermite polynomials. The function depends on the quantiles of the distribution of risky shares, which implies that I consider a series quantile model.

The empirical specification is composed of two parts: a nonlinear part that corresponds to the state variables of the economic model, and a linear part that corresponds to variables that proxy for preference shifters/life-cycle controls.\(^{25}\) It is conceptually straightforward to allow all variables to interact with each other, but this would result to a less parsimonious specification. Moreover, as I am interested in the average derivative effects of the state variables, I reduce the dimensions of the nonlinear function I aim to approximate by introducing the life-cycle controls linearly. I allow for flexibility by specifying \( \gamma^\alpha(\tau) \) as a function of \( \tau \).

Because recovering the predicted \( \alpha \)'s might result into portfolio shares that are smaller than zero or larger than one, I introduce a logit transformation. As is well known, quantiles are invariant to monotonic transformations (Koenker and Bassett (1978)).\(^{26}\) Hence, the empirical portfolio model that I take to the data is:

\[
\Lambda^{-1}(\alpha_{it}^*) \equiv \log \left( \frac{\alpha_{it}^*}{1 - \alpha_{it}^*} \right) = \sum_{k=0}^{K} b_k(\tau) \varphi_k(v_{it}, \varepsilon_{it}, w_{it}, \text{age}_{it}) + \gamma^\alpha(\tau)'X_{it} 
\]

(13)

in which \( \Lambda^{-1}(\alpha_{it}) \) is the logit function.

\(^{25}\)In the implementation, I control for education, household size, geographical dummies and cohort effects.

\(^{26}\)Chamberlain (1994) and Buchinsky (1995) apply Box-Cox transformations in a censored quantile model where they study female wage distributions, while Bottai et al. (2010) use the logistic transformation in the context of studying adolescent depression.
Participation rule and the exclusion restriction. I specify the participation rule given current earnings components, assets, age and life-cycle controls as follows:

$$\Pr(d_{it} = 1|v_{it}, \epsilon_{it}, w_{it}, age_{it}, \mathbf{z}_{it}) = \Lambda \left( \sum_{k=0}^{K} b_k^{(v)} \varphi_k(v_{it}, \epsilon_{it}, w_{it}, age_{it}) + \gamma^p \mathbf{z}_{it} \right)$$

(14)

in which $\Lambda(\cdot)$ is the logistic function. Equation (14) corresponds to a sieve logit specification.

Notice that the model relies on an exclusion restriction. In this regard, I consider the lagged value of lifetime wealth\textsuperscript{27}, following Vissing-Jørgensen (2002), Bonaparte et al. (2014), and Fagereng et al. (2017a). The motivation behind this can be easily seen in the portfolio rule of the two-period model. As shown in equation (1), the optimal portfolio rule is a function of the ratio between human and total household wealth, and not on the level of lifetime wealth.\textsuperscript{28}

Evolution of wealth. I specify the distribution of wealth $w_{i1}$ conditional on the persistent component $v_{i1}$, age at the start of the period $age_{i1}$ and life-cycle controls during the initial period when I observe them as follows:

$$Q_w(v_{i1}, age_{i1}) = \sum_{k=0}^{K} b_k^{(w)}(\tau) \tilde{\varphi}_k(v_{i1}, age_{i1}) + \gamma^w(\tau) \mathbf{x}_{i1}$$

(15)

for different choices of $K$ and $\tilde{\varphi}_k$.

Meanwhile, I specify household wealth dynamics via the following equation:

$$w_{it} = h_t(v_{it}, \epsilon_{it}, w_{it-1}, age_{it}, \mathbf{x}_{it}, \tau)$$

$$= h(v_{it-1}, \epsilon_{it-1}, w_{it-1}, age_{it-1}, \mathbf{x}_{it}, age_{it}, \tau)$$

$$= \sum_{k=1}^{K} b_k^{(w)} \tilde{\varphi}_k(v_{it-1}, \epsilon_{it-1}, w_{it-1}, age_{it-1}, age_{it}) + \gamma^w \mathbf{x}_{it} + b_0^{(w)}(\tau)$$

(16)

for some $K$ and $\tilde{\varphi}_k$. In contrast with (15), I specify equation (16) as a nonlinear regression model. Notice as well that the model is additive in $\tau$. In principle, it can also be specified as a series quantile model; in light of sample size, I resort to this model specification.

Implementation. The functions $b_k^d$, $\gamma^d$, $b_k^w$, $\gamma^w$ and $b_0^w$ are indexed by a finite dimensional parameter vector $\theta$, which also contains the coefficients $b_k^{p's}$, $b_k^{m's}$, $\gamma^{p's}$, and $\gamma^{m's}$. I model the functions $b_k^d$ as piecewise-polynomial interpolating splines on a grid $[\tau_1, \tau_2], [\tau_2, \tau_3], \ldots, [\tau_{L-1}, \tau_L]$, contained in the unit interval. I extend the specification of the intercept coefficient $b_0^d$ on $(0, \tau_1]$ and $[\tau_L, 1)$ using a parametric model that is indexed by $\lambda^d$. All other $b_k^d$ for $k \geq 1$ are constant on these two intervals. Thus, denoting $b_k^d = b_k^d(\tau_l)$, the functions $b_k^d$ depend on $\{b_{11}^d, \ldots, b_{KL}^d, \lambda^d\}$. I implement the same modelling for the other functions.

\textsuperscript{27}I explain how to calculate this variable in the data appendix.

\textsuperscript{28}This characterization is not only present in the two-period model that I outlined in section 2, but is also present in the formulas of Campbell and Viceira (2002) and Merton (1971).
To estimate the portfolio rule, I define a grid from $\tau_1 = 0.20$ to $\tau_L = 0.80$, with a step size equal to 0.10. The functions $b^i_k$ are taken as piecewise-linear, which allows the likelihood to be specified in closed form. In addition, the function $b^O_0$ is specified as the quantile of an exponential distribution on $[0, \tau_1]$ (with parameter $\lambda^+$) and $[\tau_L, 1]$ (with parameter $\lambda_+^-$). Meanwhile, I define $\tau_1 = L/10$ and $L = 10$ for the functions that correspond to the initial wealth distribution. I set the wealth accumulation functions $b^m_i$ equal to $\mu + \sigma \Phi^{-1}(\tau)$, where $(\mu, \sigma)$ are parameters to be estimated. Finally, I use tensor products of Hermite polynomials for $q_k$ and $\hat{\phi}_k$, although in practice, other specifications could be used, such as B-splines or wavelets. Each component of the product takes a standardized variable as an argument.\(^2\)

\section{Overview of the estimation algorithm}

The algorithm is an adaptation of techniques in Arellano et al. (2017) to a setting with time-varying latent variables, sample selection, and/or censoring in the dependent variable. It is a stochastic EM-like algorithm, which is a simulated version of the classical EM algorithm of Dempster et al. (1977). I describe the estimation of the portfolio rule (13) and the participation rule (14). Specific details on the algorithm, which include the likelihood function and the model’s restrictions, are described in Appendix D.

To estimate the parameters $\theta = (\theta_p, \theta_A)$, in which $\theta_p = (b^p_0, \ldots, b^p_K, \gamma^p)'$, and $\theta_A = (b^a_0, \ldots, b^a_K, \lambda^a)'$, I estimate a sequence of logit and quantile regressions. Starting with a parameter vector $\hat{\theta}^{(0)}$, I iterate the following two steps on $s = 0, 1, \ldots$ until convergence of the $\hat{\theta}^{(s)}$ process:

1. **Stochastic E-step**: Draw $v_i^{(m)}$ for $m = 1, \ldots, M$ from $f_i(\cdot; \hat{\theta}^{(s)})$.

2. **M-step**: Estimate

\[
\hat{\theta}_O^{(s+1)} = \arg \max_{\theta_O} \sum_{i=1}^N \sum_{m=1}^M O(v_i^{(m)}, \mu, Z_i; \theta_O),
\]

and

\[
\hat{\theta}_p^{(s+1)} = \arg \min_{\theta_p} \sum_{i=1}^N \sum_{m=1}^M P(v_i^{(m)}, \mu, Z_i; \theta_p).
\]

\(^2\)More specifically,

\[
b^O_0 = \frac{1}{\lambda^-} \log \left( \frac{\tau}{\tau_1} \right) \mathbf{1}\{0 < \tau < \tau_1\} + \sum_{i=1}^{L-1} \left[ b^Q_i + \frac{b^Q_{i+1} - b^Q_i}{\tau_{i+1} - \tau_i} (\tau - \tau_i) \right] \mathbf{1}\{\tau_i \leq \tau < \tau_{i+1}\}
\]

\[+ \frac{1}{\lambda^+} \log \left( \frac{\tau_{L} - \tau}{\tau_{L} - \tau_1} \right) \mathbf{1}\{\tau_L \geq \tau < 1\}.
\]

\(^3\)As an example, the portfolio rule arguments are $(w_{it} - \text{mean}(w))/\text{std}(w)$, $(v_{it} - \text{mean}(v))/\text{std}(v)$, $(\epsilon_{it} - \text{mean}(\epsilon))/\text{std}(\epsilon)$, and $(\text{age}_{it} - \text{mean}(\text{age}))/\text{std}(\text{age})$. 

24
where \( O(\cdot) \) and \( P(\cdot) \) correspond to the loss functions associated with maximum likelihood and quantile regressions, respectively.

As the likelihood function has a closed form, the \( E \) step is straightforward.\(^{31}\) In practice, I use a random-walk Metropolis Hastings sampler that targets an acceptance rate of approximately 30 percent. To update the parameters, the \( M \) step is the three-step estimator for the quantile selection model proposed by Arellano and Bonhomme (2017).

**Participation rule.** First, I estimate the participation rule (14) via sieve maximum likelihood:

\[
\max_{(b_0^c, \ldots, b_K^c, \gamma^c)} \sum_{i=1}^N \sum_{t=1}^T \sum_{m=1}^M d_{it} \log \left[ \Lambda \left( \sum_{k=0}^K b_k^c \varphi_k(v_{it}, \varepsilon_{it}, w_{it}, age_{it}) + \gamma^c Z_{it} \right) \right] + (1 - d_{it}) \log \left[ 1 - \Lambda \left( \sum_{k=0}^K b_k^c \varphi_k(v_{it}, \varepsilon_{it}, w_{it}, age_{it}) + \gamma^c Z_{it} \right) \right].
\]

The resulting estimates allow me to recover the propensity score \( p(z_{it}) \), i.e., the probability that a household participates in the stock market.

**Estimation of the conditional copula parameter.** In the second step, I estimate the correlation parameter between the error terms of participation and portfolio choice, \( \rho_c \), by considering the following objective function:

\[
\rho_c = \arg \min_{c \in C} \left\| \sum_{i=1}^N \sum_{t=1}^T \sum_{m=1}^M d_{it} Y(\tau_t, x_{it}) \left[ 1 \left\{ \Lambda^{-1}(a_{it}^*) \leq \sum_{k=0}^K b_{kl}^c(c) \varphi_k(\cdot) \right\} - G(\tau_t, p(z_{it}); c) \right] \right\|
\]

where \( \tau_1, \ldots, \tau_L \) is a finite grid on \((0,1)\), \( \| \cdot \| \) is the Euclidean norm,

\[
G(\tau_t, p(z_{it}); c) = \frac{C(\tau_t, p(z_{it}); \rho_c)}{p(z_{it})}
\]

is a conditional copula of the error terms of the portfolio and participation rules, \( Y(\cdot) \) are instrument functions, and

\[
b_{kl}^c(c) = \arg \min_{b \in B} \sum_{i=1}^N \sum_{t=1}^T \sum_{m=1}^M d_{it} \left[ G(\tau_t, p(z_{it}); c) \left( \Lambda^{-1}(a_{it}^*) - \sum_{k=0}^K b_k^c(\tau) \varphi_k(v_{it}, \varepsilon_{it}, w_{it}, age_{it}) \right) \right] + (1 - G(\tau_t, p(z_{it}); c)) \left( \Lambda^{-1}(a_{it}^*) - \sum_{k=0}^K b_k^c(\tau) \varphi_k(v_{it}, \varepsilon_{it}, w_{it}, age_{it}) \right) \] \quad (19)

The objective function (18) is the result of the conditional moment restrictions that permit the mapping of the distribution of latent outcomes to the distribution of observed outcomes conditional on participation. In the context of my economic model, the latent outcomes are

\(^{31}\)In my empirical implementation, I conduct log-likelihood evaluations as opposed to likelihood evaluations to prevent numerical underflow.
the portfolio shares of the participation subproblem; meanwhile, while the observed outcomes are the observed portfolio shares, the solution of the entire problem. More formally, the mapping is:

\[
Pr(\alpha^*_it \leq g_l(v_{it}, \epsilon_{it}, w_{it}, age_{it})) = G(\tau_l, p(z_{it}); \rho_c).
\]

(20)

As can be observed from (20), \(G(\tau_l, p(z_{it}); \rho_c)\) is the function that maps the two distributions. Though this step is computationally demanding as it is non-convex, evaluating (18) is usually fast and straightforward since (19) is a linear programming problem, for which there exist numerically reliable computational algorithms.

**Estimation of the portfolio rule.** Finally, given consistent estimators of the parameters of the propensity score \(\hat{p}(x_{it})\) and the correlation parameter \(\rho_c\), the third step requires solving for a given quantile \(\tau_l\):

\[
\min_{(b_0^*...b_K^*)} \sum_{i=1}^N \sum_{t=2}^T \sum_{m=1}^M d_{it} \left[ G(\tau_l, \hat{p}(z_{it}); \hat{\rho}_c) \left( \Lambda^{-1}(\alpha^*_it) - \sum_{k=0}^K b_{it}^{\rho} \phi_k(v_{it}, \epsilon_{it}, w_{it}, age_{it}) + \gamma^a(\tau_l) X_{it} \right) \right. \\
+ \left. (1 - G(\tau_l, \hat{p}(z_{it}); \hat{\rho}_c)) \left( \Lambda^{-1}(\alpha^*_it) - \sum_{k=0}^K b_{it}^{\rho} \phi_k(v_{it}, \epsilon_{it}, w_{it}, age_{it}) + \gamma^a(\tau_l) X_{it} \right) \right]
\]

(21)

The optimization problem (21) is equivalent to minimizing a rotated check function, with an individual-specific perturbed \(\tau\) (Arellano and Bonhomme (2017)). The rotation preserves the linear programming formulation, and thus, the computational simplicity, of quantile regression. To correct for selection, I replace \(\tau\) with the individual specific-perturbed \(\tau\), \(G(\tau_l, \hat{p}(z_{it}); \hat{\rho}_c)\).

**Implementation.** To implement the algorithm, I take \(M = 1\), stop the chain after a large number of iterations, and report an average across the last \(\bar{S}\) values \(\hat{\theta} = \frac{1}{S} \sum_{s=S-S+1}^S \hat{\theta}^{(s)}\), where I take \(\bar{S} = S/2\). Each estimation is based on \(S = 200\) iterations, with 200 random walk Metropolis-Hastings draws per iteration. To ensure the computational tractability of the estimation algorithm, I conducted a simulation experiment using a data generating process that is close to the model that is specified. The detailed description, and the results of the simulation experiments are outlined in Appendix C.2 of the paper.

**Alternative identification scheme.** A potential concern is the validity of the exclusion restriction. In other words, it could be that the variables that determine participation are the same ones that determine the portfolio rule. The model in this case is now a “Tobit-like” model, which has a natural connection to censoring corrections. To address this issue, I consider the censored quantile regression estimator of Buchinsky and Hahn (1998).

The choice of this estimator over other procedures (e.g., Powell (1986), Chernozhukov and Hong (2002), Honore et al. (2002), Khan and Tamer (2009)) is mainly motivated by three
reasons. First, the nonlinear semi-reduced form described by equations (3)-(7) when \( X_{it} = Z_{it} \) whittles down to a model with a random censoring point. This rules out Powell (1986) and Chernozhukov and Hong (2002), who both consider models with a fixed censoring point. Second, Buchinsky and Hahn (1998) propose an estimation method that is computationally tractable, as it also results in a convex optimization problem. Though both Honore et al. (2002) and Khan and Tamer (2009) consider models with random censoring, their proposed estimation methods are computationally more demanding.\(^{32}\) Third, and most importantly, the estimator can be interpreted as the limiting case of the more general quantile selection model of Arellano and Bonhomme (2017). A crucial distinction between the two estimators is that the quantile selection allows the recovery of the latent distribution of risky shares, while the censored regression model allows the recovery of the observed distribution. I outline the model specification, estimation algorithm, and the simulation experiment that corresponds to this estimator in Appendix D.3.

Statistical properties. Nielsen (2000) studies the statistical properties of the stochastic EM algorithm in a likelihood case. Arellano and Bonhomme (2016) provide the asymptotic distribution of \( \hat{\theta} \) when the optimization step is based on a sequence of quantile-based estimating equations. They show that the estimator is root \( N \) consistent and asymptotically normal under correct specification of the parametric model, for fixed \( K, L, \) and \( T \). Arellano and Bonhomme (2017), meanwhile, calculates the asymptotic distribution of the estimates of the quantile selection model in a cross-sectional context. I use the parametric bootstrap for inference. It would be relatively straightforward to prove the asymptotic properties in this context.

6 Stock market participation and portfolio choices in the PSID (in progress)

In this section, I present preliminary empirical results. I first begin by showing average derivative effects of the empirical participation and portfolio rules. I then report simulation exercises based on the estimated model. In the estimation of the participation and portfolio rules, I use tensor products of Hermite polynomials with degrees \((2,1,2,1)\).\(^{33}\)

6.1 Average derivative effects

Figure 4 shows the average derivative effect, with respect to \( v_{it} \), of the propensity score \( p(z_{it}) \) and the conditional of mean of the risky asset share \( \alpha_{it} \) for risky market participants, respec-

\(^{32}\)Honore et al. (2002) propose a procedure that does not yield a convex optimization problem, while Khan and Tamer (2009) propose a moment inequality-based approach.

\(^{33}\)The choice of the following specification is motivated by previous literature (e.g., Guiso et al. (1996) and Vissing-Jørgensen (2002)), among others, who model income and wealth with quadratic terms.
tively, evaluated at percentiles of wealth $\tau_{\text{wealth}}$ and age $\tau_{\text{age}}$, and averaged over $v_{it}$, based on the estimated nonlinear model. The average derivative effects lie between -0.25 to 0.125 for the propensity score, and -0.2 to 0.2 for the risky asset share. The results indicate the presence of interactions between wealth and age with respect to the risky asset share. Figure E1 shows the results with the censored quantile regression estimator of Buchinsky and Hahn (1998). As the graphs indicate, the results are quite similar.

Figure 4: Average derivative effect of the persistent component of income $v_{it}$

(a) Propensity score

(b) Risky asset share

Note: The graphs show average derivatives of the propensity score and the risky asset share of stock market participants, respectively, with respect to the persistent income $v_{it}$, given $w_{it}$, persistent component $v_{it}$, income $y_{it}$, and age $a_{it}$, and evaluated at different values of $w_{it}$ and $a_{it}$ that correspond to their $\tau_{\text{wealth}}$ and $\tau_{\text{age}}$ percentiles. All results are based on estimates from the semi-structural model with the Arellano and Bonhomme (2017) quantile selection model estimator.

Figure 5 shows the average derivative effect, with respect to wealth $w_{it}$, of the propensity score $p(z_{it})$ and the conditional of mean of the risky asset share $\alpha_{it}$ for risky market participants, respectively, evaluated at percentiles of wealth $\tau_{\text{wealth}}$ and age $\tau_{\text{age}}$, and averaged over $v_{it}$, based on the estimated nonlinear model. The average derivative effects lie between 0.09 to 0.3 for the propensity score, and -0.05 to 0.05 for the risky asset share. The results indicate the presence of interactions between wealth and age with respect to the risky asset share. Moreover, the figure indicates that the derivative effects seem to increase with age and assets. Figure E2 shows the results with the censored quantile regression estimator of Buchinsky and Hahn (1998). The results are quite similar, with the exception of the average derivative with respect to the portfolio rule.\footnote{A potential explanation for this difference is the nature of the estimators.}
Figure 5: Average derivative effect of wealth

(a) Wealth

(b) Persistent income $v_{it}$

Note: The graphs show average derivatives of the propensity score and the risky asset share of stock market participants, respectively, with respect to wealth $w_{it}$ given $w_{it}$, persistent component $v_{it}$, income $y_{it}$, and age $age_{it}$, evaluated at different values of $w_{it}$ and age$age_{it}$ that correspond to their $τ_{wealth}$ and $τ_{age}$ percentiles. All results are based on estimates from the semi-structural model with the Arellano and Bonhomme (2017) quantile selection model estimator.

Figure 6: Observed and implied densities of the risky asset share

Note: The graph shows the observed and predicted unconditional densities of the share of household wealth in risky asset share based on the nonlinear model. The blue line corresponds to the density implied by the nonlinear model, while the red line corresponds to the density implied by the data. All results are based on estimates from the semi-structural model with the Arellano and Bonhomme (2017) quantile selection model estimator.

Finally, I evaluate the fit implied by the nonlinear model with that of the data. As the results indicate, 55.3 percent of households participate according to the nonlinear model, which is close to the observed 57.30 percent in the data. Moreover, as the graph of the density of risky asset shares indicates, the nonlinear model provides a close fit with the data. Looking
at the figure implied by the estimation with the Buchinsky and Hahn (1998) estimator (Figure E3), I find that the fit is quite similar.

6.2 What is the impact of a persistent earnings shock?

In this subsection, I simulate the life-cycle portfolio choice model with participation according to the nonlinear model, and show the extensive and intensive margin response with respect to a persistent income shock. In the simulation exercise, I report the difference between two types of households: households who are hit at age 37 by a large negative shock to the persistent component ($\tau_{\text{shock}} = 0.1$), or by a large positive shock ($\tau_{\text{shock}} = 0.9$), and households who are hit by a median shock of 0.5 to the persistent component. I report age-specific means across 250,000 simulations. At the start of the simulation (i.e., at age 35), all households have the same persistent component.

**Extensive margin responses.** In Figure 7, I report the results with respect to the conditional probability of participation, i.e., the extensive margin. The results show asymmetric extensive margin responses to large income shocks, whether negative or positive. The results also highlight the interaction between the rank of the household in the distribution of the initial earnings ($\tau_{\text{init}}$) and the size of the shock received ($\tau_{\text{shock}}$). In particular, a large negative shock results in a decrease in participation of as much as 7 percent for high income households, and a 3.1 percent decrease for low income households. Meanwhile, a large positive shock yields an increase in participation of as much as 13.2 percent for low-income households, compared to 4 percent for low-income households. I observe similar patterns for the Buchinsky and Hahn (1998) censored quantile regression estimator, as can be observed in Figure E4 in Appendix E of the paper.

I then calculate whether the impulse responses to the conditional probability of stock market participation differ depending on participation status at age 35. Figure 8 graphs the age-specific conditional probability for a high-income household hit by a large, negative income shock, and a low-income household hit by a large, positive income shock. As the results indicate, while there are differences in magnitudes between the set of participants and non-participants, the observed extensive margin responses are quite similar. This is in contrast to the findings of Briggs et al. (2015), who look at the impact of a wealth shock to the stock market participation behavior of Swedish households; in their paper, they find that the increase in participation (which was at 3 percent for the total population) was due to entry by non-participants. The results that I provide here suggest that the income shocks that households receive lead to entries to or exits from the stock market, depending on the type and size of shock received.
Figure 7: Impulse response, participation rule

(a) $\tau_{\text{shock}} = 0.1$

(b) $\tau_{\text{shock}} = 0.9$

Note: The graphs show the difference in average participation rates between a household hit by a shock $\tau_{\text{shock}}$ at age 37, and a household hit by 0.5 shock at the same age. The blue line corresponds to low-income households (i.e., rank of $\tau_{\text{init}} = 0.1$ in the income distribution). The red line corresponds to middle-income households (i.e., rank of $\tau_{\text{init}} = 0.5$ in the income distribution). The green line corresponds to middle-income households (i.e., rank of $\tau_{\text{init}} = 0.9$ in the income distribution). All results are based on estimates from the semi-structural model with the Arellano et al. (2017) quantile selection estimator.

Figure 8: Impulse response, participation rule, by participation status

(a) $\tau_{\text{init}} = 0.9$, $\tau_{\text{shock}} = 0.1$

(b) $\tau_{\text{init}} = 0.1$, $\tau_{\text{shock}} = 0.9$

Note: The graphs show the difference in average participation rates between a household with rank $\tau_{\text{init}}$ hit by a shock $\tau_{\text{shock}}$ at age 37, and a household hit by 0.5 shock at the same age, conditional on participation status at age 35. The blue line corresponds to stock market participants. The red line corresponds to stock market non-participants. All results are based on estimates from the semi-structural model with the Arellano et al. (2017) quantile selection estimator.
**Intensive margin responses.** In Figure 9, I look at the differences with respect to the average risky asset share for stock market participants at age 35. As the results indicate, a large negative income shock yields a 2.1 percentage point decrease in average risky asset shares for high income households. In comparison, the same shock yields a 0.45 percentage point decrease for low income households. A large positive income shock, meanwhile, results in a 3 percentage point increase in the share of wealth invested in risky assets for low income households. In comparison, the same shock results in a 1 percentage point increase for high income households. Looking at the corresponding figure in the appendix (Figure E6) for the censored quantile regression estimator, I observe similar results. These results suggest that the persistence of earnings histories is crucial in understanding changes in risky asset shares with respect to asymmetries in income risk.

![Figure 9: Impulse response, portfolio rule](image)

**Note:** The graphs show the difference in average portfolio shares conditional on participation between a household hit by a shock $\tau_{\text{shock}}$ at age 37, and a household hit by 0.5 shock at the same age. The blue line corresponds to low-income households (i.e., rank of $\tau_{\text{init}} = 0, 1$ in the income distribution). The red line corresponds to middle-income households (i.e., rank of $\tau_{\text{init}} = 0.5$ in the income distribution). The green line corresponds to middle-income households (i.e., rank of $\tau_{\text{init}} = 0.9$ in the income distribution). All results are based on estimates from the semi-structural model with the Arellano et al. (2017) quantile selection estimator.

**Interactions with wealth and age.** Finally, Figures 10 and 11 present similar simulation exercises, but varying the timing of the shocks and the amount of wealth that the households possess. Comparing the magnitudes to households who are hit by the same shock when they are old, as shown in Figure 11, I find that the extensive margin responses are stronger for younger households than for older households. Meanwhile, intensive margin responses seem to be similar for both households. The results suggest the importance of human capital as a factor in portfolio choice decisions. Figures E7 and E8 in the appendix reveal similar
patterns when the relevant estimator is the censored regression estimator of Buchinsky and Hahn (1998).

Figure 10: Extensive and intensive margin responses to an income shock, by income and wealth at age 35

(a) Extensive margin, $\tau_{\text{shock}} = 0.1$
(b) Extensive margin, $\tau_{\text{shock}} = 0.9$
(c) Intensive margin, $\tau_{\text{shock}} = 0.1$
(d) Intensive margin, $\tau_{\text{shock}} = 0.9$

Note: The graphs show the difference between a household hit by a shock $\tau_{\text{shock}}$ at age 37, and a household hit by a 0.5 shock at the same age, by income and wealth categories. The blue line corresponds to low income, low wealth households. The red line corresponds to low income, high wealth households. The green line corresponds to high income, low wealth households. The orange line corresponds to high income, high wealth households. All results are based on estimates from the semi-structural model with the Arellano et al. (2017) quantile selection estimator.
Figure 11: Extensive and intensive margin responses to an income shock, by income and wealth at age 51

(a) Extensive margin, $\tau_{\text{shock}} = 0.1$

(b) Extensive margin, $\tau_{\text{shock}} = 0.9$

(c) Intensive margin, $\tau_{\text{shock}} = 0.1$

(d) Intensive margin, $\tau_{\text{shock}} = 0.9$

Note: The graphs show the difference between a household hit by a shock $\tau_{\text{shock}}$ at age 53, and a household hit by a 0.5 shock at the same age, by income and wealth categories. The blue line corresponds to low income, low wealth households. The red line corresponds to low income, high wealth households. The green line corresponds to high income, low wealth households. The orange line corresponds to high income, high wealth households. All results are based on estimates from the semi-structural model with the Arellano et al. (2017) quantile selection estimator.

7 Conclusions

In this paper, I develop a semi-structural framework to understand the nonlinear transmission of income shocks on household investment behavior. I model stock market participation and portfolio rules as age-dependent functions of the persistent and transitory components of income, and of assets. The model reveals asymmetric participation and portfolio adjustment responses with respect to “unusual” income shocks, which to the best of my knowledge, has not been uncovered in previous literature.
The results that I present in this paper suggest the presence of stock market participation costs for households. This is evident in the impulse response-like analyses that I conduct. For instance, I find that a large negative income shock is associated with a modest decrease in participation for low-income households, while it leads to a sizeable drop in participation rates for high-income households. Likewise, I find that a large positive income shock is associated with a considerable increase in participation rates for low-income households. A natural extension of this paper, hence, would be to combine the framework that I introduced here with structural approaches to quantify participation costs, such as those in Alan (2006) and Khorunzhina (2013).

I have also abstracted from aggregate shocks, such as those related to the volatility of stock market returns and its correlation with labor income. On the one hand, as Benzoni et al. (2007), Betermier et al. (2012), and Bonaparte et al. (2014) show, the correlation between the stock market return and human capital might induce households to hedge against their income risks through the use of stocks. On the other hand, Schmidt (2015) emphasizes that investing in stocks is a poor hedge against potentially disastrous income shocks. A rigorous analysis of the income hedging motives requires extending the empirical framework I present here to incorporate macroeconomic effects.

Finally, it might be paramount to consider the demand for other assets; in particular, housing. As underscored by a large literature, for most households, their abode constitutes the biggest share of their wealth (see Davis et al. (2015)). To this end, extending the framework I present here to consider features inherent in studying models of housing demand, such as movements from renting to homeownership (e.g., Han (2010)), might be important to understanding household investment decisions. All these constitute future avenues in my research agenda.

References


Lawrence DW Schmidt. Climbing and falling off the ladder: Asset pricing implications of labor market event risk. *Available at SSRN 2471342*, 2015.


Appendices to the paper
"Household portfolio choices and nonlinear income risk"
Julio Gálvez

The appendices follow the organization of the paper. Appendix A outlines the derivations associated with the two-period model, and details related to the calibration exercise in the life-cycle model. Appendix B outlines the proof of nonparametric identification, and details on the semi-structural model’s extensions. Appendix C provides more details about the data, and further descriptive evidence. Appendix D provides further details on the empirical strategy, and simulation evidence related to the nonlinear model. It also provides a brief description of the Buchinsky and Hahn (1998) estimation procedure, and related model specification and simulation evidence. Appendix E provides additional empirical evidence.

A Derivations and calibration details
A.1 Two-period model derivations
A.1.1 Model derivations and proofs

Consumption in the non-participation subproblem. As the household consumes all of its resources in period $t + 1$, I can write a log-linear approximation to the budget constraint. Dividing both sides of the budget constraint by $L_{t+1}$, and taking logs yields the following equation:

$$
\frac{C_{t+1}}{L_{t+1}} = \frac{W_t}{L_{t+1}}(1 + R_f) + 1 \rightarrow \log(C_{t+1}) - \log(L_{t+1}) = \log \left( \exp \left( \log \left( \frac{W_t}{L_{t+1}} (1 + R_f) \right) \right) + 1 \right)
$$

A first-order Taylor expansion around the mean of the (log) risk-free rate $E(\log(1 + R_f)) = r_f$ and the wealth-labor income ratio $E(\log(\frac{W_t}{L_{t+1}})) = w - l$, yields the following expression:

$$
c_{t+1} - l_{t+1} = \log(\exp(\log(1 + R_f + \log(w_t + r_f - l_{t+1})))) 
\approx k_{np} + \phi_{np}(w_t + r_f - l_{t+1}),
$$

where, after rearranging both sides of the equation, yields the following expression for optimal consumption:

$$
c_{np,t+1} \approx k_{np} + \phi_{np}(w_t + r_f) + (1 - \phi_{np})l_{t+1} \quad (A1)
$$

in which $k_{np}$ is a log-linearization constant, and $\phi_{np}$ is a function that takes the following form:

$$
\phi_{np} = \frac{\exp(w_t + r_f - l_{t+1})}{1 + \exp(w_t + r_f - l_{t+1})}.
$$

The function $\phi_{np}$ can be interpreted as the elasticity of consumption with respect to wealth; similarly, the expression $(1 - \phi_{np})$ can be interpreted as the elasticity of consumption with respect to income. Because $0 < \phi_{np} < 1$, the consumption function is a strictly concave function of wealth and income.
Consumption in the participation subproblem. Similarly, as in the non-participation sub-
problem, I obtain optimal consumption by log-linearization of the budget constraint.\footnote{The log-linear approximation to the portfolio return is derived in the appendix of Campbell and Viceira (2002) that is available here: https://scholar.harvard.edu/campbell.} Dividing both sides by $L_{t+1}$:

$$\frac{C_{t+1}}{L_{t+1}} = \frac{W_t - F}{L_{t+1}} (1 + R_{p,t+1}) + 1$$

Denoting $W_{c,t} = W_t - F$, taking logs of both sides, and a first-order Taylor expansion around the wealth-income ratio and the mean portfolio return yields the following expression:

$$c_{t+1} - l_{t+1} = \log(\exp(w_{c,t} + r_{p,t+1} - l_{t+1}) + 1) \approx k_p + \phi_p(w_{c,t} + r_{p,t+1} - l_{t+1})$$

Again, after re-arranging both sides of the equation, I end up with the following expression

$$c_{p,t+1} \approx k_p + \phi_p(w_{c,t} + r_{p,t+1}) + (1 - \phi_p)l_{t+1}$$

where $\phi_p$, the wealth elasticity of consumption, now takes the following form:

$$\phi_p = \frac{\exp(w_{c,t} + r_{p,t+1} - l_{t+1})}{1 + \exp(w_{c,t} + r_{p,t+1} - l_{t+1})}.$$ 

Proof of the optimal portfolio share in the participation subproblem. To solve the optimal portfolio shares, I write the first-order condition of the subproblem:

$$E_t[\delta C_t^{-\gamma} (1 + R_{t+1})] = E_t[\delta C_t^{-\gamma} (1 + R_f)]$$

To obtain the optimal portfolio share, I can take a second-order Taylor expansion of the Euler equation. To see this, note that I can rewrite the left-hand side as:

$$E_t[\delta C_t^{-\gamma} (1 + R_{t+1})] = E_t[\exp\{\log(\delta C_t^{-\gamma} (1 + R_{t+1}))\}]$$

$$= E_t[\exp\{\log(\delta) - \gamma c_{t+1} + r_{t+1}\}]$$

$$= E_t[\exp\{x_{t+1}\}]$$

where the notational correspondence between the second and third line is obvious. Taking a second-order Taylor expansion around the mean $\overline{x}_{t+1} = E_t(x_{t+1})$, I obtain the following:

$$E_t(\exp(x_{t+1})) \approx E_t \left[ \exp(\overline{x}_{t+1}) \left( 1 + (x_{t+1} - \overline{x}_{t+1}) + \frac{1}{2} (x_{t+1} - \overline{x}_{t+1})^2 \right) \right]$$

$$\approx \exp(\overline{x}_{t+1}) \left( 1 + \frac{1}{2} \text{Var}(x_{t+1}) \right)$$

Taking another first-order Taylor expansion around zero, I get:

$$E_t(\exp(x_{t+1})) \approx 1 + \overline{x}_{t+1} + \frac{1}{2} \text{Var}(x_{t+1}).$$
From here, it can be shown that the first-order condition becomes:

\[ E_t(r_{t+1} - r_f) + \frac{1}{2} \text{Var}_t(r_{t+1}) = \gamma \text{Cov}_t(r_{t+1}, c_t) \]

\[ = \gamma \text{Cov}_t(r_{t+1}, k_p) + \phi_p(w_{c,t} + r_{p,t+1}) + (1 - \phi_p)I_t + 1 \]

\[ = \gamma [\phi_p \alpha_t \sigma_u^2 + (1 - \phi_p)\text{Cov}_t(I_t, r_{t+1})] \]

where in the second line I substituted the consumption function (A2) and in the third line I substituted the log portfolio return. Rearranging this equation yields the optimal portfolio rule:

\[ \alpha_t = \frac{1}{\phi_p} \left[ \frac{E_t(r_{t+1} + \frac{1}{2}\sigma_u^2)}{\gamma \sigma_u^2} + (1 - \frac{1}{\phi_p}) \frac{\sigma_{lt}^2}{\sigma_u^2} \right] \]

(A3)

where \( \phi_p \) is the wealth elasticity of consumption. The optimal share has two components. The first component describes the optimal allocation when labor income risk is idiosyncratic; that is, when it is uncorrelated with the risky asset. The second component is the income hedging component. This means that the demand depends not only on the expected excess return of the stock relative to its variance, but also on its ability to hedge against a bad realization of labor income. If the covariance is negative, then the risky asset offers a good hedge against negative income shocks. Meanwhile, if the covariance is positive, then one cannot hedge against bad income shocks. Finally, if the covariance is zero, I am left with the first component. In the rest of the discussion, I will assume that \( \sigma_{lt}^2 = 0 \).

Note that labor income \( L_{t+1} \) and household wealth net of participation costs \( W_{c,t} \) mainly affect the portfolio share through the wealth elasticity of consumption, \( \phi_p \). Alternatively, I can express portfolio rule (A3) as a function of expected future labor income \( \tilde{H} \equiv E_t(L_{t+1}) \) and wealth, \( W_{c,t} \). To see this, write the wealth elasticity of consumption as:

\[ \frac{1}{\phi_p} = 1 + \frac{1}{\exp(w_{c,t} + r_{p,t+1} - I)} \]

\[ = 1 + \frac{\exp(l)}{\exp(w_{c,t} + r_{p,t+1})} \]

\[ = 1 + \frac{H}{W_{c,t}} \]

which, when substituted to the portfolio rule with idiosyncratic labor income risks, yields equation (1), the portfolio rule in the main text.

**Participation condition.** The participation condition can be derived from the following inequality:

\[ E_t \left( \frac{c_{p,t+1}^{1-\gamma}}{1-\gamma} \right) \geq E_t \left( \frac{c_{np,t+1}^{1-\gamma}}{1-\gamma} \right) \]

where \( C_{i,t}, i = p, np \) denotes the consumption if the household bought stocks or not, respectively. Taking logs:

\[ E_t \left( \frac{c_{p,t+1}^{1-\gamma}}{1-\gamma} \right) = E_t \left[ \exp \left( \log \left( \frac{c_{p,t+1}^{1-\gamma}}{1-\gamma} \right) \right) \right] \approx E_t(\exp((1-\gamma)c_{i,t+1})) \]
Because the risk aversion parameter is a constant, I focus on 

\( E_t(\exp\{c_{i,t+1}\}) \). Taking a first order Taylor expansion around zero, I obtain:

\[
E_t(\exp\{c_{i,t+1}\}) \approx 1 + E_t(c_{i,t+1}),
\]

which, when substituted to the inequality yields the condition:

\[
E_t(c_{p,t+1}) \geq E_t(c_{np,t+1}). \tag{A4}
\]

### A.1.2 Comparative statics in the two-period model

I calculate two comparative statics results. In the first set of comparative statics results, I compute the average derivatives from the alternative characterization of the optimal portfolio share. This will correspond to households who are not in the margin. The second set of comparative statics are calculated for the marginal investor.

**Average derivatives from the portfolio share.** In this case, it will be useful to restate the alternative characterization:

\[
\alpha_t = \frac{1}{\phi_p} \left( \frac{E_{t+1}(r_{t+1} - r_f) + \frac{1}{2} \sigma_u^2}{\gamma \sigma_u^2} \right)
\]

Taking the total derivative with respect to wealth, I have the following expression:

\[
\frac{d\alpha_t}{dW_t} = \frac{\frac{d\alpha_t}{d\phi_p} \frac{d\phi_p}{dW_t} \frac{dw_c}{dW_t}}{1 - \frac{d\alpha_t}{d\phi_p} \frac{d\phi_p}{dW_t} \frac{dr_p}{d\alpha_t}}
\]

I can now turn to the calculation of the components in the numerator, as the denominator is always positive.\(^{36}\) I compute each of the following components, starting with \(dw_c/dW_t\):

\[
\frac{dw_c}{dW_t} = \frac{d\log(W_t - F)}{dW_t} = \frac{1}{W_t - F} \geq 0
\]

where the last expression is non-negative because for households who have participated, \(W_t \geq F\). Now, for the rest of the exercise, I assume that this is positive.

\(^{36}\)To show that this is positive, compute the following components of the denominator:

\[
\frac{d\alpha_t}{d\phi_p} = -\phi_p^{-2} \left[ \frac{\mu + \frac{1}{2} \sigma_u^2}{\gamma \sigma_u^2} \right] < 0
\]

The second component, meanwhile, is:

\[
\frac{d\phi_p}{dr_p} = \frac{\exp(w_{c,t} + r_{p,t+1} - l_{t+1})}{(1 + \exp(w_{c,t} + r_{p,t+1} - l_{t+1}))^2} > 0
\]

Finally,

\[
\frac{dr_p}{d\alpha_t} = \left( E_t(r_{t+1}) - r_f + \frac{1}{2} \sigma_u^2 \right) \left( 1 - \frac{1}{\phi_p^2 \gamma} \right)
\]

A sufficient condition to prove that the denominator is positive is to show that \(\gamma > \frac{1}{\phi_p^2} > 1\). Now, when \(\gamma = \frac{1}{\phi_p^2}\), this will correspond to an investor who will hold the growth-optimal portfolio with the highest return (Campbell and Viceira (2002)). Hence, for this to be positive, the intuition is that the investor should be one who is sufficiently conservative than the growth-optimal investor.
Next, I obtain \( \frac{d\phi_p}{d\omega_c} \):

\[
\frac{d\phi_p}{d\omega_c} = \frac{\exp(w_{c,t} + r_{p,t+1} - l_{t+1})}{(1 + \exp(w_{c,t} + r_{p,t+1} - l_{t+1}))^2} > 0
\]

Finally, I calculate \( \frac{d\alpha_t}{d\phi_p} \):

\[
\frac{d\alpha_t}{d\phi_p} = -\phi_p^{-2} \left[ \frac{\mu + \frac{1}{2} \sigma^2_u}{\gamma \sigma^2_u} \right] < 0
\]

Putting all of these together:

\[
\text{sign} \left( \frac{d\alpha_t}{d\omega_t} \right) = \left( \frac{(-)}{(+)} \frac{d\alpha_t}{d\phi_p} \frac{d\phi_p}{d\omega_c} \right) \frac{d\omega_c}{dW_t} < 0.
\]

That is, an increase in wealth decreases the share invested in risky assets.

To calculate the effect of income, I need to compute the following components:

\[
\frac{d\alpha_t}{dl_t} = \frac{d\alpha_t}{d\phi_p} \frac{d\phi_p}{dl_t}
\]

Because I have showed earlier that the denominator is positive, I focus on calculating the terms in the numerator. I first calculate \( \frac{d\alpha_t}{d\phi_p} \):

\[
\frac{d\alpha_t}{d\phi_p} = -\phi_p^{-2} \left[ \frac{\mu + \frac{1}{2} \sigma^2_u}{\gamma \sigma^2_u} \right] < 0
\]

Next, I obtain \( \frac{d\phi_p}{dl_t} \):

\[
\frac{d\phi_p}{dl_t} = \frac{\exp(w_{c,t} + r_{p,t+1} - l_{t+1})}{(1 + \exp(w_{c,t} + r_{p,t+1} - l_{t+1}))^2} < 0
\]

Finally, to compute \( dl_t / dL_t \), I take a third-order Taylor expansion of \( l_t \) about the mean value, \( \mu_L \):

\[
l_t \equiv \log(L_t) \approx \log(\mu_L) + (L - \mu_L) \frac{d\log(L_t)}{dL} + \frac{1}{2} (L - \mu_L)^2 \frac{d^2 \log(L_t)}{dL^2} + \frac{1}{6} (L - \mu_L)^3 \frac{d^3 \log(L_t)}{dL^3}.
\]

After taking expectations and computing the derivatives, (because the household does not know the realisation of \( L \)), I obtain the following:

\[
l_t \approx \log(\mu_L) - \frac{E_t[(L - \mu_L)^2]}{L^2} + \frac{1}{3} \frac{E_t[(L - \mu_L)^3]}{L^3}
\]

Now, computing \( dl_t / dL_t \):

\[
\frac{dl_t}{dL} = \frac{E_t[(L - \mu_L)^2]}{L^3} - \frac{E_t[(L - \mu_L)^3]}{L^4}
\]
where, clearly, the magnitudes of the variance, \( E_t[(L - \mu_L)^2] \), and skewness, \( E_t[(L - \mu_L)^3] \), will determine the sign of \( d\alpha / dL \):

\[
\text{sign} \left( \frac{d\alpha_t}{dL_t} \right) = \frac{(-) (-) (?)}{d\phi_p / dL_t} \geq 0.
\]

To compare with the lognormal distribution:

\[
\frac{dl}{dL} = \frac{1}{L^3},
\]

which is true given that I can express \( l \) as\(^{37}\):

\[
l = \log(\mu_L) - \frac{\sigma_L^2}{2L^2}
\]

This means that unambiguously, an increase in labor income leads to an increase in the allocation to risky assets.

**Comparative statics for the marginal investor.** To calculate comparative statics, it is helpful to substitute the consumption functions (A1) and (A2) into the inequality I have introduced in equation (A4). Rearranging terms, I obtain the following expression:

\[
\phi_p w_{t,c} - \phi_{np} w_t \geq (\phi_p - \phi_{np}) l_{t+1} + \phi_{np} r_f - \phi_p r_{p,t+1}
\]

Moreover, it is useful to think about a marginal investor for whom the participation condition is binding with equality:

\[
\phi_p w_{t,c} - \phi_{np} w_t = (\phi_p - \phi_{np}) l_{t+1} + \phi_{np} r_f - \phi_p r_{p,t+1}
\]

I can then derive how changes in labor income and wealth affect the marginal investor. I appeal to the Implicit Function Theorem to calculate how a change in wealth will affect the marginal investor. Taking derivatives with respect to \( w_{c,t} \):

\[
\frac{\partial \phi_p}{\partial w_{c,t}} w_{c,t} + \phi_p = \frac{\partial \phi_p}{\partial w_{c,t}} l_{t+1} - \frac{\partial \phi_p}{\partial w_{c,t}} r_{p,t+1} - \phi_p \frac{\partial r_{p,t+1}}{\partial w_{c,t}}
\]

Rearranging the previous equation yields the following expression:

\[
\frac{\partial r_{p,t+1}}{\partial w_{c,t}} = -\frac{\partial \phi_p}{\partial w_{c,t}} (w_{c,t} + r_{p,t+1} - l_{t+1}) + \frac{\phi_p}{\partial w_{c,t}}
\]

I can expand \( \partial r_{p,t+1} / \partial w_{c,t} \) into:

\[
\frac{\partial r_{p,t+1}}{\partial w_{c,t}} = \frac{\partial \alpha_i}{\partial w_{c,t}} (r_{p,t+1} - r_f) + \frac{1}{2} \frac{\partial \alpha_i}{\partial w_{c,t}} (1 - \alpha_i) \sigma_u^2 - \frac{1}{2} \frac{\partial \alpha_i}{\partial w_{c,t}} \sigma_u^2
\]

\(^{37}\)Campbell and Viceira (2002) express \( l \) as the following function:

\[
l_t = \log(L) - \frac{1}{2} \sigma_t^2
\]

where the idea is that one can express \( l \) as a linear function of variance, but the mean is preserved.
Plugging this in into the equation earlier, I obtain:

$$\frac{\partial \alpha_t}{\partial w_{c,t}} = -\frac{\partial \phi_p}{\partial w_{c,t}} \left( w_{c,t} + r_{p,t+1} - l_{t+1} \right) \phi_p \left( r_{p,t+1} - r_f + \frac{1}{2} \sigma^2_u - \alpha_t \sigma^2_u \right)$$

Now, the only thing left to be checked is the sign of the numerator.\(^{38}\) Now, \(\partial \phi_p / \partial w_{c,t}\) is just:

$$\frac{\partial \phi_p}{\partial w_{c,t}} = \frac{\exp(w_{c,t} + r_{p,t+1} - l_{t+1})}{1 + \exp(w_{c,t} + r_{p,t+1} - l_{t+1})^2} = \frac{\phi_p (r_{p,t+1} + 1 - r_f)}{1 + \exp(w_{c,t} + r_{p,t+1} - l_{t+1})} \geq 0$$

The calculations yield

$$\frac{\phi_p (w_{c,t} + r_{p,t+1} - l_{t+1})}{1 + \exp(w_{c,t} + r_{p,t+1} - l_{t+1})} + \phi_p \geq 0$$

which will mean that:

$$w_{c,t} + r_{p,t+1} - l_{t+1} \geq -1 - \exp(w_{c,t} + r_{p,t+1} - l_{t+1})$$

This inequality is generally true whenever \(w_{c,t} + r_{p,t+1} \geq l_{t+1}\). This characterization simply means that an increase in wealth will make the marginal investor inclined to invest more of his wealth in risky assets if the returns he receives is more than sufficient enough to compensate for the potentially risky labor income in the future. Otherwise, the marginal investor will opt out of the financial markets.

To calculate how a change in income will affect the marginal investor, I again employ the

\(^{38}\)I first check the sign of the denominator. It will only be negative when

$$r_{p,t+1} - r_f + \frac{1}{2} \sigma^2_u - \alpha_t \sigma^2_u \leq 0.$$  

When this is rearranged, the inequality suggests that:

$$r_{p,t+1} - r_f + \frac{1}{2} \sigma^2_u \leq \alpha_t \sigma^2_u$$

which, after rearranging, becomes:

$$\alpha_t \geq \frac{r_{p,t+1} - r_f}{\sigma^2_u} + \frac{1}{2} \geq 0$$

For the right hand side of this inequality to be positive however, it should be the case that

$$\frac{r_{p,t+1} - r_f}{\sigma^2_u} > -\frac{1}{2},$$

which means that:

$$\alpha_t \geq \frac{r_{p,t+1} - r_f}{\sigma^2_u} > -\frac{1}{2},$$

Given that the marginal investor has positive \(\alpha_t\), then the denominator indeed is negative.
Implicit Function Theorem:

\[ \frac{\partial \phi_p}{\partial l_{t+1}} w_{t,c} + \frac{\partial \phi_p}{\partial r_{p,t+1}} \frac{\partial r_{p,t+1}}{\partial \alpha_t} w_{t,c} - \frac{\partial \phi_{np}}{\partial l_{t+1}} w_t = \]

\[ \frac{\partial \phi_p}{\partial l_{t+1}} l_{t+1} + \frac{\partial \phi_p}{\partial r_{p,t+1}} \frac{\partial r_{p,t+1}}{\partial \alpha_t} - \frac{\partial \phi_{np}}{\partial l_{t+1}} l_{t+1} + (\phi_p - \phi_{np}) + \frac{\partial \phi_{np}}{\partial l_{t+1}} r_f \]

\[ - \frac{\partial \phi_p}{\partial l_{t+1}} r_{p,t+1} - \frac{\partial \phi_p}{\partial r_{p,t+1}} \frac{\partial r_{p,t+1}}{\partial \alpha_t} p_{p,t+1} - \phi_p \frac{\partial r_{p,t+1}}{\partial \alpha_t} \frac{\partial \alpha_t}{\partial l_{t+1}} \]

Rearranging this equation yields the following:

\[ \frac{\partial \phi_p}{\partial l_{t+1}} (w_{t,c} + r_{p,t+1} - l_{t+1}) - \frac{\partial \phi_{np}}{\partial l_{t+1}} (w_{t,c} + r_f - l_{t+1}) - (\phi_p - \phi_{np}) = \]

\[ \frac{\partial \alpha_t}{\partial l_{t+1}} \left[ - \frac{\partial \phi_p}{\partial r_{p,t+1}} \frac{\partial r_{p,t+1}}{\partial \alpha_t} (w_{t,c} + r_{p,t+1} - l_{t+1}) - \phi_p \frac{\partial r_{p,t+1}}{\partial \alpha_t} \frac{\partial \alpha_t}{\partial l_{t+1}} \right] \]

Notice though, that:

\[ \frac{\partial \phi_p}{\partial l_{t+1}} = -\frac{\phi_p}{1 + \exp(w_{t,c} + r_{p,t+1} - l_{t+1})} \]

and similarly,

\[ \frac{\partial \phi_{np}}{\partial l_{t+1}} = -\frac{\phi_{np}}{1 + \exp(w_t + r_{f,t+1} - l_{t+1})} \]

Hence, substituting and rearranging, I can express the earlier equality into:

\[ \frac{\partial \alpha_t}{\partial l_{t+1}} = \frac{\phi_p (w_{t,c} + r_{p,t+1} - l_{t+1}) + \phi_{np} (w_t + r_{f,t+1} - l_{t+1}) - (\phi_p - \phi_{np})}{\frac{\phi_p}{1 + \exp(w_{t,c} + r_{p,t+1} - l_{t+1})} + \frac{\phi_{np}}{1 + \exp(w_t + r_{f,t+1} - l_{t+1})} - \phi_p \frac{\partial r_{p,t+1}}{\partial \alpha_t} \frac{\partial \alpha_t}{\partial l_{t+1}}} \]

Removing the negative signs will yield:

\[ \frac{\partial \alpha_t}{\partial l_{t+1}} = \frac{\phi_p (w_{t,c} + r_{p,t+1} - l_{t+1}) + \phi_{np} (w_t + r_{f,t+1} - l_{t+1}) + (\phi_p - \phi_{np})}{\frac{\phi_p}{1 + \exp(w_{t,c} + r_{p,t+1} - l_{t+1})} + \frac{\phi_{np}}{1 + \exp(w_t + r_{f,t+1} - l_{t+1})} + \phi_p \frac{\partial r_{p,t+1}}{\partial \alpha_t} \frac{\partial \alpha_t}{\partial l_{t+1}}} \]

Now, the sign will be determined by the numerator. Notice, however, that I can express

\[ c_{p,t+1} - l_{t+1} \approx \phi_p (w_{t,c} + r_{p,t+1} - l_{t+1}) \]

This expression comes from the consumption functions. Hence,

\[ \frac{c_{p,t+1} - l_{t+1}}{1 + \exp(w_{t,c} + r_{p,t+1} - l_{t+1})} - \frac{c_{np,t+1} - l_{t+1}}{1 + \exp(w_t + r_{f,t+1} - l_{t+1})} + (\phi_p - \phi_{np}) \geq 0 \]

which in the end, will result in this expression:

\[ \frac{c_{p,t+1} - l_{t+1}}{1 + \exp(w_{t,c} + r_{p,t+1} - l_{t+1})} \geq \frac{c_{np,t+1} - l_{t+1}}{1 + \exp(w_t + r_{f,t+1} - l_{t+1})} - (\phi_p - \phi_{np}) \]
Finally, assuming that the elasticities of consumption for this consumer are relatively small, I can obtain the following inequality:

$$\frac{c_{p,t+1} - I_{t+1}}{1 + \exp(w_{t,c} + r_{p,t+1} - I_{t+1})} \geq \frac{c_{np,t+1} - I_{t+1}}{1 + \exp(w_l + r_{f,t+1} - I_{t+1})}$$

If I express the denominators as constants, the inequality above is approximately:

$$c_{p,t+1} \geq c_{np,t+1}$$

Clearly, this means that an increase in labor income will only induce the marginal investor to continue participating in the stock market if his consumption in the participation state is at least as high as that in the non-participation state.

### A.2 Linear and nonlinear earnings processes comparison

#### A.2.1 Linear earnings process estimation

The linear earnings process I specify is similar to those considered by Storesletten et al. (2004) and Kaplan and Violante (2010), among others. In this earnings process, the transitory component $\varepsilon_{it}$ is assumed to be independently and identically distributed (i.i.d.) Gaussian, with $N(0, \sigma^2_\varepsilon)$. The persistent component, meanwhile, is modelled as

$$v_{it} = \rho v_{i(t-1)} + \eta_{it}$$

where $\rho$ is a persistence parameter. The idiosyncratic part of the persistent component of income, $\eta_{it}$, is also i.i.d. Gaussian, with $\eta_{it} \sim N(0, \sigma^2_\eta)$. I assume that households have different initial conditions; that is, $v_{i0} \sim i.i.d. N(0, \sigma^2_v)$. Moreover, I assume that the initial condition, the idiosyncratic component of the persistent shock, and the transitory shock are independent of each other.

The standard estimation strategy is minimum distance estimation. An alternative, which I implement here, is to estimate the parameters via pseudo maximum likelihood estimation. That is, if $u_i \sim N(0, \Omega(\theta))$, then the pseudo maximum likelihood estimator of $\theta$ solves:

$$\hat{\theta}_{PML} = \arg\min_{\theta} \left\{ \log \det(\Omega(c)) + \frac{1}{N} \sum_{i=1}^{N} \hat{u}_i \Omega(c)^{-1} \hat{u}_i \right\}.$$  

This is equivalent to:

$$\hat{\theta}_{PML} = \arg\min_{\theta} \left\{ \log \det(\Omega(c)) + \text{tr}(\Omega(c)^{-1} \hat{\Omega}) \right\},$$

where $\text{tr}$ is the trace of the resulting matrix, and $\hat{\Omega} = \sum \hat{u}_i' \hat{u}_i$. The parameter estimates are reported in the table below.
### Table A1: Parameter estimates, linear earnings process

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimates</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autoregressive parameter</td>
<td>0.8784</td>
<td>(0.0987)</td>
</tr>
<tr>
<td>SD of transitory component</td>
<td>0.2333</td>
<td>(0.0550)</td>
</tr>
<tr>
<td>SD of initial persistent component</td>
<td>0.3148</td>
<td>(0.0666)</td>
</tr>
<tr>
<td>SD of idiosyncratic persistent component</td>
<td>0.1968</td>
<td>(0.0365)</td>
</tr>
</tbody>
</table>

Note: These are estimates of the parameters of the linear earnings process. Standard errors are in parentheses, and are calculated using the asymptotic covariance matrix.

### A.2.2 Specification and estimation of Arellano et al. (2017) earnings process

**Persistent component.** Denote the persistent component of the household head $i$ at period $t$ by $v_{it}$ and $age_{it}$ the age of the household head. Then, the conditional quantile of the persistent component as a function of the past persistent component and age is:

$$Q_t(v_{it-1}, \tau) = \sum_{k=0}^{K} a_k^P(\tau) \varphi_k(v_{it-1}, age_{it})$$

In practice, I estimate this function using tensor products of Hermite polynomials, which are of the order (3,3).

**Initial condition of the persistent component.** The conditional quantile function of the initial persistent component as a function of age is:

$$Q_{v_{i1}}(age_{i1}, \tau) = \sum_{k=0}^{K} a_k^I(\tau) \tilde{\varphi}_k(age_{i1})$$

**Transitory component.** The conditional quantile function of the transitory component as a function of age is:

$$Q_{\varepsilon}(age_{it}, \tau) = \sum_{k=0}^{K} a_k^T(\tau) \bar{\varphi}_k(age_{it})$$

The estimation of this earnings process follows the stochastic EM algorithm described in Arellano et al. (2017). I refer the reader to Arellano et al. (2017) for a full description of the estimation procedure, and the likelihood function of the earnings process.

### A.2.3 Estimation results

Figure A1 presents the results of the estimation of the earnings process, and the estimation of a quantile autoregression of earnings. Specifically, the first three graphs are plots of the average derivative of the conditional quantile function $y_{it}$ on $y_{it-1}$, with respect to the simulated
earnings data. Meanwhile, the last graph is the average derivative of the conditional quantile function $\nu_{it}$ on $\nu_{it-1}$ using the nonlinear earnings process.

To check whether the results that I have obtained fit the model well, I compare the estimated persistence from the true data, which is depicted in panel (a), with the estimated persistence that comes from simulated data according to the nonlinear and the linear earnings models, which are in panels (b) and (c), respectively. What I find is that the nonlinear earnings process is able to fit the data well in terms of persistence. Panel (d) describes the results of the persistence in the persistent component $\nu_{it}$ in the nonlinear earnings model. As the results indicate, there seems to be heterogeneity in persistence of income shocks.

Figure A1: Nonlinear persistence

Note: Panel (a), (b), and (c) show estimates of the conditional quantile function of $y_{it}$ given $y_{it-1}$ with respect to $y_{it-1}$, evaluated at $\tau_{\text{shock}}$ and at a value of $y_{it-1}$ that corresponds to the $\tau_{\text{init}}$ percentile of the distribution of $y_{it-1}$. Panel (a) is based on the PSID data in my subsample, panel (b) comes from the simulated data based on the nonlinear earnings process, and panel (c) comes from the simulated data based on the linear earnings process. Finally, panel (d) is the average derivative of the the conditional quantile function of $\nu_{it}$ on $\nu_{it-1}$ with respect to $\nu_{it-1}$, based on estimates from the nonlinear earnings model.
Figure A2 shows estimates of conditional skewness, which were calculated using quantile-based skewness measures. The results in panel (b) indicate evidence of conditional asymmetry, which goes in the same direction as in Arellano et al. (2017). There is less evidence of it though in the simulated data and in the earnings data of the PSID.

Figure A2: Conditional skewness

Note: These are nonparametric kernel estimates of densities of the persistent and transitory components of income in the nonlinear earnings model of Arellano et al. (2017). The results are based on simulated data according to a Gaussian kernel, with the optimal bandwidth.

A.3 Life-cycle model details

In this subsection, I summarize the details related to the life-cycle model that I simulate to show some possible implications of nonlinear earnings processes on stock market participation and household portfolio choices.

Model set-up. In what follows, a household head enters the labor market at age $t_0$, works until age $t_r$, and dies at age $T$. The household maximizes

$$E_0 \left( \sum_{t=1}^{T} \beta^{t-1} \frac{C_{t+1}^{1-\gamma}}{1-\gamma} \right)$$

The household has available at its disposal a riskless asset that pays a constant return $r_f$ and a risky asset $r_t$ that earns a constant risk premium $\mu$ and evolves according to the following process:

$$r_{t+1} = r_f + \mu + u_{t+1}, \quad u_{t+1} \sim N(0, \sigma_u^2)$$

Note that the return on the stock is still potentially correlated with labor income, with $\text{cov}_t(r_{t+1}, y_{t+1}) = \sigma_{lu}$.

The timing of the model is as follows: a household starts with a certain amount of wealth $w_{it}$. Following Cocco et al. (2005), I denote cash-on-hand by $x_{it} = w_{it} + y_{it}$. Each period, the household solves the participation and the non-participation subproblems, with the same
objectives as before. Finally, to solve its optimal choice, the household has to compare the indirect utility gained from participation and non-participation, and then, decide whether it enters (or stays) or exits the stock market.

Given these assumptions, next period’s wealth before the realization of labor income, then, is the following:

\[ x_{it+1} = (x_{it} - c_{it} - 1(\alpha_{it} > 0)F) (\alpha_{it} R_t + (1 - \alpha_{it}) R_f) \]

in which \( 1(\cdot) \) is an indicator function that is equal to one when the household participates, and zero otherwise.

The participation subproblem that the household faces then, is the following:

\[ V^I_t(x_{it}, \nu_{it}) = \max_{c_{it}} u(c_{it}) + \beta \mathbb{E}_t(V_{t+1}(x_{it+1}, \nu_{it+1})) \]

such that

\[ x_{it+1} = [x_{it} - c_{it} - F] [\alpha_{it} R_t + (1 - \alpha_{it}) R_f] + y_{it+1} \]

Meanwhile, the non-participation subproblem that the household faces is:

\[ V^O_t(x_{it}, \nu_{it}) = \max_{c_{it}} u(c_{it}) + \beta \mathbb{E}_t(V_{t+1}(x_{it+1}, \nu_{it+1})) \]

such that

\[ x_{it+1} = [x_{it} - c_{it}] (R_f) + y_{it+1} \]

The Bellman equation \( V^I_t(x, \nu) \) corresponds to the indirect utility of a household of age \( t \) who participates in the stock market, has a persistent income realization of \( \nu \) and financial wealth \( x \). Meanwhile, the Bellman equation \( V^O_t(x, \nu) \) corresponds to the indirect utility of a household of age \( t \) who does not participate in the stock market, has a persistent income realization of \( \nu \) and financial wealth \( x \).

The Bellman equation \( V_t(x, \nu) \) for the household pins down the participation decision:

\[ V_t(x, \nu) = \max(V^I_t(x, \nu), V^O_t(x, \nu)) \]

Finally a formalization of the participation decision is:

\[ 1'(x, \nu) = \begin{cases} 1, & \text{if } x \in X^P(\nu), \text{ } X^P(\nu) = \{ \nu \in \mathbb{R} : V^I_t(x, \nu) > V^O_t(x, \nu) \} \\ 0, & \text{otherwise} \end{cases} \]

Parameterization. I choose the following parameters, which are either estimated or taken from the literature.

1. Demographics. In this model, people enter the market at age 25, retire at age 60, and die with certainty at age 95. I use survival probabilities from the National Center for Health Statistics.
2. Preferences, discounting and participation costs. I assume, following Cocco et al. (2005), that $\gamma$, the risk aversion parameter, takes a value of 10, and that $\beta$, the discount rate, is 0.96. The participation cost $q$ is set as 0.300, following Fagereng et al. (2017a).

3. Interest rates. With respect to interest rates, I set the risk-free interest rate to 2 percent, following Cocco et al. (2005), Alan (2012) and Fagereng et al. (2017a). With respect to the excess return, I obtain the mean and the standard deviation from the Fama-French website. The mean equity premium is 4 percent, while the standard deviation is 0.174, which are similar to parameterizations of Cocco et al. (2005) and Alan (2012).

Solution method. The model is solved using backward induction. Given the terminal condition, the policy functions and the value function are trivial: households will consume all wealth, and the value function will be equal to the indirect utility function. I substitute this value function, and compute the subsequent policy functions backward. I do this for 60 periods, from age 95 to age 25. I discretize the state space for the cash-on-hand state variable and iterate on the value function.

Discretization of the earnings process. I discretize the nonlinear earnings process through the following simulation approach:

1. I simulate one million households throughout their life-cycle. I then rank households by their earnings and put them into 50 bins for the persistent component, and 25 bins for the transitory component, at a given age.

2. To estimate transition probabilities for the persistent component, I count the transitions between any two bins from neighboring ages, and use this information to estimate transition probabilities.

B A flexible, semi-structural approach

B.1 Nonparametric identification

It is useful to review the model to facilitate discussion of nonparametric identification.

$$\alpha_{it}^* = g_t(v_{it}, \epsilon_{it}, w_{it}, X_{it}, u_{it})$$  \hspace{1cm} (B5)

$$\alpha_{it} = \alpha_{it}^* \cdot d_{it}$$  \hspace{1cm} (B6)

$$d_{it} = \begin{cases} 1, & \text{if } \tilde{g}_t(v_{it}, \epsilon_{it}, w_{it}, q(Z_{it})) \leq \nu_{it} \\ 0, & \text{otherwise} \end{cases}$$  \hspace{1cm} (B7)

$$w_{it} = h_t(v_{it-1}, \epsilon_{it-1}, w_{it-1}, \alpha_{it-1}, X_{it}, \zeta_{it})$$  \hspace{1cm} (B8)

$$w_{i0} = \tilde{h}_0(v_{i0}, \zeta_{i0})$$  \hspace{1cm} (B9)

The following are assumptions on the semi-structural model.

Assumption 1. For all $t \geq 1$:

a. The unobserved errors $(u_{i1}, \ldots, u_{iT}, v_{i1}, \ldots, v_{iT}, \epsilon_{i1}, \ldots, \epsilon_{iT}, \zeta_{i2}, \ldots, \zeta_{iT})$ are mutually independently distributed.
b. The bivariate distribution of \( u_{it} = (u_{it}, v_{it})' \) given \( w_{it}', a_{it}', d_{it}', y_{it}' \) and \( \eta_{it}' \) are absolutely continuous with respect to the Lebesgue measure, with standard uniform marginals and rectangular support. I denote the c.d.f. as \( C_x(u, v) \).

c. The conditional c.d.f. \( F_{a_{it}|Z}(a|Z) \) is strictly increasing. Moreover, \( C_x(u, v) \) is strictly increasing in \( u \).

d. \( \Pr(d_{it} = 1|z_{it}) > 0 \) for all \( t \) with probability one.

The first assumption implies the following: first, that current and future earnings shocks are independent of current and past wealth; second, a Markovian assumption on wealth dynamics; and third, that the unobserved errors of the portfolio and participation rules are independent over time, independent of earnings components, and independent of current and past assets. The second assumption states that the source of the dependence between \( (u_{it}, v_{it}) \) is the source of sample selection bias. In the context of the economic model, the dependence is the link between the participation subproblem, the result of which is the latent risky asset shares \( a_{it}^* \), and the economic problem at hand, which results in the participation rule denoted by \( d_{it} \) and the observed outcome \( a_{it} \). The third assumption restricts the analysis to continuous outcomes, and the fourth is a standard assumption in sample selection models.

The proof proceeds sequentially, starting with the first period. Letting \( y_i = (y_{i1}, \ldots, y_{iT}) \), and using \( f \) as a generic notation for a density function:

\[
f(w_1|y) = \int f(w_1|v_1)f(v_1|y)dv_1 \tag{B10}
\]

where, by Assumption 1a, \( f(w_1|v_1, y) \) and \( f(w_1|v_1) \) coincide. Provided that the distribution of \( (v_1|y_i) \) is complete (given that this is identified from the earnings process, see Arellano et al. (2017)), the density \( f(w_1|y) \) is identified.

Next, I turn into the participation rule. Using equation (B7) and assumption 1a, I have:

\[
f(d_1|w_1, q_1, y) = \int f(d_1|w_1, v_1, q_1, y_1)f(v_1|w_1, y)dv_1 \tag{B11}
\]

where under completeness of \( (v_1|y_i) \) in \( y_i \), the density \( f(d_1|w_1, q_1, y_1) \) is identified.

Identification of the joint density of risky shares and the participation indicator is not straightforward, however. To illustrate, write the density \( f(a_{it}^*, d_{it}|w_{it}, y) \) as:

\[
f(a_{it}^*, d_{it}|w_{it}, q_1, y) = \int [\Pr(a_{it}^*, d_{it} = 1|q_1, w_{it}, v_{it}, y)]^{d_{it}} [\Pr(d_{it} = 0|q_1, w_{it}, v_{it}, y)]^{(1-d_{it})} f(v_1|w_1, y)dv_1.
\]

I decompose this further into:

\[
f(a_{it}^*, d_{it}|w_{it}, q_1, y) = \int [f(a_{it}^*|d_{it} = 1, w_{it}, v_{it}, y)]^{d_{it}} \times \left[\Pr(d_{it} = 1|w_{it}, q_1, v_{it}, y)\right]^{d_{it}} [\Pr(d_{it} = 0|w_{it}, q_1, v_{it}, y)]^{(1-d_{it})} f(v_1|w_1, y)dv_1,
\]

which, after, using the definition of the participation rule, yields:

\[
f(a_{it}^*, d_{it}|w_{it}, y) = \int [f(a_{it}^*|d_{it} = 1, w_{it}, v_{it}, y)]^{d_{it}} f(d_{it}|w_{it}, v_{it}, y) f(v_1|w_1, y)dv_1 \tag{B12}
\]
If \( a_i^* \) were not latent, I can proceed with the same identification arguments as those for wealth and the participation rule. To prove that this density is identified under the presence of sample selection, note that, conditional on participation, for all \( \tau \in (0,1) \) and \( t \geq 1 \):

\[
\Pr(a_{it}^* \leq q(\tau, x)) |_{d_{it} = 1, Z_{it} = z_{it}} = G_x(\tau, p(z_{it}))
\]

(B13)

where I have used Assumption 1 and \( G_x(\tau, p(z_{it})) = C(\tau, p) / p \). The mapping \( G_x(\tau, p(z_{it})) \) provides the link between the latent and the observed distribution of risky asset shares in my model. Notably, what equation (B13) implies is that if \( G_x(\cdot, \cdot) \) is known, then I can recover \( q(\tau, x) \), as a quantile of the observed portfolio shares, by suitably shifting percentile ranks. Subsequently, I would be able to recover the portfolio rule. More formally, under Assumption 1, and given that the mapping \( G_x \) is known, Proposition 1 of Arellano and Bonhomme (2017) will hold. I restate the theorem here for completeness:

**Proposition 1** (Proposition 1 of Arellano and Bonhomme (2017)). Let Assumption 1 hold. Let \( x \in X \). Suppose that one of the following conditions holds:

i. (identification at infinity) There exists some \( z_x \in Z_x \) such that \( p(z) = 1 \).

ii. (analytic extrapolation) \( P_x \) contains an open interval, and for all \( \tau \in (0,1) \), the function \( p \rightarrow G_x(\tau, p) \) is real analytic on the unit interval.

Then, the functions \( \tau \mapsto p \) and \( \tau \mapsto q(\tau, x) \) are nonparametrically identified.

The first condition requires that there are some households who will always participate in the stock market. The second condition is closer to the argument I have made for nonparametric identification. Given that this is satisfied, \( F(a_i^* | d_1 = 1, w_1, v_1, y) \) and subsequently, \( f(a_i^* | d_1 = 1, w_1, v_1, y) \) are identified nonparametrically.

*Second period’s wealth.* Notice that:

\[
f(w_2 | a_1, d_1, w_1, y, q) = \int f(w_2 | a_1, d_1, w_1, \eta_1, y) f(v_1 | a_1, d_1, w_1, y) dv_1
\]

(B14)

Provided that the distribution \( (v_1 | a_1, d_1, w_1, y) \) is complete in \( y_i = (y_{1i}, y_{2i}, \ldots, y_{Ti}) \), I can identify the density \( f(w_2 | a_1, d_1, w_1, v_1, y) \).

Via Bayes’ rule and Assumption 1a,

\[
f(v_2 | w_2, w_1, a_1, d_1, y, b) = \frac{f(v_1, v_2 | w_1, w_2, a_1, d_1, y_1) f(y | y_1, v_1, v_2)}{f(y | w_2, w_1, a_1, y_1)} dv_2
\]

(B15)

\(^{40}\)Alternatively, if the economic problem is such that there is no selection into stock market participation, the same arguments will still apply. This is because, in the absence of the participation rule, the density of risky asset shares is simply:

\[
f(a_1 | w_1, y) = \int [f(a_1 | w_1, v_1, y)]^{1 \{a_1 > 0\}} [F(0 | w_1, v_1, y)]^{1 \{a_1 = 0\}} f(v_1 | w_1, y) dv_1
\]

It can then be shown that the density is nonparametrically identified under completeness of \( (v_{1i} | w_{1i}, y_i) \) on \( (y_{1i}, y_{2i}, \ldots, y_{Ti}) \), following similar arguments.
As \( f(v_1, v_2|w_1, w_2, a_1, d_1, y_1) = f(v_1|w_1, w_2, a_1, d_1, y_1)f(v_2|v_1) \) is identified (given that \( f(v_2|v_1) \) is identified from the earnings process, and \( f(v_1|w_1, w_2, a_1, d_1, y_1) \) is identified from above), it follows that \( f(v_2|w_2, w_1, a_1, d_1, y) \) is identified.

**Subsequent periods.** To prove the participation rule, notice that:

\[
f(d_2|w_2, w_1, a_1, y) = \int f(d_2|w_2, v_2, y, q)f(v_2|w_2, w_1, a_1, d_1, y)dv_2. \tag{B16}
\]

Given that \( (v_2|w_2, w_1, a_1, d_1, y) \) (which is identified from the previous paragraph) is complete in \( (a_1, d_1, w_1, y_1, y_3, \ldots, y_T) \), the density \( f(d_2|w_2, v_2, y, q) \) is nonparametrically identified.

In the case of the portfolio rule,

\[
f(a_2^*, d_2|w_2, w_1, a_1) = \int [f(a_2^*|d_2 = 1, w_2, v_2, y)]d_2 f(v_2|d_2, b_2, w_2, w_1, a_1, d_1, y)dv_2. \tag{B17}
\]

As the distribution of \( (v_2|w_2, w_1, a_1, d_1, y) \) is complete in \( (d_2, a_1, d_1, w_1, y_1, y_3, \ldots, y_T) \) (which is identified from the participation rule) and under Proposition 1, the distribution \( F(a_2^*|d_2 = 1, w_2, v_2, y) \), and subsequently, the density \( f(a_2^*|d_2 = 1, w_2, v_2, y) \) is nonparametrically identified.

Finally, by induction, and using assumptions 1 and Proposition 1 from the third period onward, the joint density of \( v_1 \)'s, assets, earnings, portfolio choices, and participation decisions are nonparametrically identified. This is provided that, for all \( t \geq 1 \), the distributions of \( (v_i|a_i^t, w_i^t, d_i^t, y_i) \) are complete in \( (a_i^{t-1}, w_i^{t-1}, d_i^{t-1}, y_i^{t-1}, y_{i+1}, y_T) \).

Intuitively, the identification argument comes from the link to nonparametric instrumental variables problems. To see this, note that I can rewrite equation (B10) as

\[
f(w_1|y) = \mathbb{E}[f(w_1|v_1)|y = y] \tag{B18}
\]

where I am taking the expectation of the distribution of \( v_1 \) conditional on \( y_i \) for a given fixed value of \( w_1 \). This is analogous to a nonparametric IV problem where the endogenous regressor is \( v_1 \) and \( y_i \) are the excluded instruments. Likewise, I can rewrite equations (B11) and (B12) as the following functional equations:

\[
f(d_1|w_1, y) = \mathbb{E}[f(d_1|v_1, w_1, y)|w_1 = w, y_i = y] \tag{B19}
\]

\[
f(a_1^*, d_1|w_1, y) = \mathbb{E}[\{f(a_1^*|v_1, w_1, d_1 = 1, y_1)\}^1|d_1 = d, q_1 = q, w_1 = w, y_i = y] \tag{B20}
\]

In these particular cases, conditional on \( (w_1, y_1), (y_2, \ldots, y_T) \) are the “excluded instruments” for \( v_1 \) with respect to the functional equation that corresponds to the participation rule. With respect to the portfolio rule, however, I require not only these “excluded instruments”, but also a participation cost shifter that does not affect the subsequent portfolio choice.

The identification arguments also rely on completeness conditions, which relate to the relevance of the excluded instruments. The notion of completeness that I refer to here relates to the concept of operator injectivity. More formally, a linear operator \( L \) is injective if the only
solution \( h \in \mathcal{H}_1 \) to the equation \( \mathcal{L} h = 0 \) is \( h = 0 \). In this paper, \( \mathcal{L} \) is the conditional expectation operator. For example, in equation (B18), completeness implies that the only solution to the equation \( [\mathcal{L} h](y_i) = \mathbb{E} [f(w_1|\nu_{11})|y_i = y] \) is \( h = 0 \).

### B.2 Nonlinear semi-reduced form model extensions

In this subsection, I discuss the full specification of the model extensions. First, I discuss state dependence in stock market participation. In the second part, I proceed with discussing household unobserved heterogeneity. Third, I discuss the inclusion of the consumption function. Fourth, and finally, I explain advance information in earnings.

**State dependence in participation.** To develop the model extension, I augment the baseline model specification and consider the following system of equations:

\[\alpha^*_{it} = g_t(v_{it}, \epsilon_{it}, w_{it}, X_{it}, u_{it})\]  
\[\alpha_{it} = \alpha^*_{it} \cdot d_{it}\]  
\[d_{it} = \begin{cases} 1, & \text{if } \tilde{g}_t(v_{it}, \epsilon_{it}, w_{it}, q(Z_{it}), d_{it-1}) \leq v_{it} \\ 0, & \text{otherwise} \end{cases}\]  
\[w_{it} = h_t(v_{it-1}, \epsilon_{it-1}, w_{it-1}, \alpha_{it-1}, X_{it}, w_{it})\]  
\[w_{i0} = \tilde{h}_{i0}(v_{i0}, \zeta_{i0})\]  
\[\zeta_{i} = \phi(v_{i0}, w_{i0}, v_{i1})\]

Notice that in equation (B23), I permit past participation decisions to have an influence on current participation decisions. This permits the study of dynamics in stock market participation. To be more specific, I re-define the conditional probability that a household participates in the stock markets as:

\[\delta_{dd'}(v) = \Pr(d_{it} = d'|Z_{it} = z, v_{it} = v, \epsilon_{it} = \epsilon, w_{it} = w, d_{it-1} = d)\]

in which \( d = \{0, 1\}, d' = \{0, 1\} \). It follows that I can then compute the following functions:

\[\Delta_{dd'}(v + v', v) = \delta_{dd'}(v + v') - \delta_{dd'}(v)\]

which provides me a way of calculating entry and exit rates. The calculation, meanwhile, of the intensive margin of portfolio choice decisions remains the same.

**Household unobserved heterogeneity.** The nonlinear reduced form model in the case of household unobserved heterogeneity is as follows:

\[\alpha^*_{it} = g_t(v_{it}, \epsilon_{it}, w_{it}, X_{it}, \zeta_{it}, u_{it})\]  
\[\alpha_{it} = \alpha^*_{it} \cdot d_{it}\]  
\[d_{it} = \begin{cases} 1, & \text{if } \tilde{g}_t(v_{it}, \epsilon_{it}, w_{it}, q(Z_{it}), \zeta_{it}) \leq v_{it} \\ 0, & \text{otherwise} \end{cases}\]  
\[w_{it} = h_t(v_{it-1}, \epsilon_{it-1}, w_{it-1}, \alpha_{it-1}, X_{it}, \zeta_{it}, w_{it})\]  
\[w_{i0} = \tilde{h}_{i0}(v_{i0}, \tilde{\zeta}_{i0})\]  
\[\tilde{\zeta}_{i} = \phi(v_{i0}, w_{i0}, v_{i1})\]
In this model, I consider a scalar unobserved heterogeneity, which I denote by $\xi_i$. Notice that $\xi_i$ enters in all of the equations in the baseline specification, and in particular, in the portfolio and participation rules. Though the economic model suggests that there could possibly be an additional fixed effect in the participation rule (B28), modelling unobserved heterogeneity in this manner results in a more parsimonious specification.

Another detail worth noting is that in this model, I take a random-effects approach in that I model the distribution of $\xi_i$ conditional on the initial observations of the households. This is in contrast to a fixed-effects approach (see, e.g., Kyriazidou (1997) and Kyriazidou (2001)) in which I condition on the household-specific fixed effect, the distribution of which is left unspecified. An advantage of this approach is that the estimation algorithm that I use is still applicable to this model. The calculation of intensive and extensive margins, in this case, have to be modified as now I condition on $\xi_i$. Nonparametric identification also becomes more challenging in that I would need to take care of the initial distribution of unobserved heterogeneity.

Consumption function. The introduction of the consumption function changes the estimating equations into the following:

$$\alpha^*_{it} = g_t(v_{it}, \epsilon_{it}, w_{it}, X_{it}, u_{it}) \quad (B32)$$
$$\alpha_{it} = \alpha^*_{it} \cdot d_{it} \quad (B33)$$
$$d_{it} = \begin{cases} 1, \text{ if } \tilde{g}_t(v_{it}, \epsilon_{it}, w_{it}, q(Z_{it})) \leq v_{it} \\ 0, \text{ otherwise} \end{cases} \quad (B34)$$
$$c_{it} = n_t(v_{it}, \epsilon_{it}, w_{it}, X_{it}, u_{it}) \quad (B35)$$
$$w_{it} = h_t(v_{it-1}, \epsilon_{it-1}, w_{it-1}, \alpha_{it-1}, Z_{it}, c_{it-1}, \xi_{it}) \quad (B36)$$
$$w_{t0} = \tilde{h}_{t0}(v_{t0}, \bar{\xi}_{t0}) \quad (B37)$$

in which I introduce the consumption rule (B35) and past consumption in equation (B36), the budget constraint. In this case, the error term in (B36) can now be thought of as one that captures the returns realized on the asset investments. Introducing the consumption function permits the calculation of the marginal propensity to consume out of income and wealth.

Advance information in earnings. To introduce advanced information in earnings, I include $v_{it+1}$ in the portfolio and participation rules:

$$\alpha^*_{it} = g_t(v_{it}, v_{it+1}, \epsilon_{it}, w_{it}, X_{it}, u_{it}) \quad (B38)$$
$$\alpha_{it} = \alpha^*_{it} \cdot d_{it} \quad (B39)$$
$$d_{it} = \begin{cases} 1, \text{ if } \tilde{g}_t(v_{it}, v_{it+1}, \epsilon_{it}, w_{it}, q(Z_{it})) \leq v_{it} \\ 0, \text{ otherwise} \end{cases} \quad (B40)$$
$$w_{it} = h_t(v_{it-1}, \epsilon_{it-1}, w_{it-1}, \alpha_{it-1}, X_{it}, \xi_{it}) \quad (B41)$$
$$m_{t0} = \tilde{h}_{t0}(v_{t0}, \bar{\xi}_{t0}) \quad (B42)$$
C Data and descriptive statistics

C.1 Sample selection criteria in detail

Table C1 shows the detailed sample selection criteria that was implemented for the main sample in the empirical study that largely follows the criteria of Blundell et al. (2016). In the sample selection, I first construct a subsample of all households who I can follow for the entire six waves of the PSID. I then apply each of the criteria in Blundell et al. (2016).

Table C1: Sample selection criteria, PSID

<table>
<thead>
<tr>
<th>Less:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total households eligible from the PSID interviews</td>
<td>4304</td>
</tr>
<tr>
<td>Households who were not continuously married</td>
<td>2064</td>
</tr>
<tr>
<td>Households with missing information on state</td>
<td>20</td>
</tr>
<tr>
<td>Households with missing information on age</td>
<td>0</td>
</tr>
<tr>
<td>Households with missing information on education</td>
<td>292</td>
</tr>
<tr>
<td>Non-SEO households</td>
<td>273</td>
</tr>
<tr>
<td>Households with missing information on race</td>
<td>7</td>
</tr>
<tr>
<td>Households who are not within the age range</td>
<td>561</td>
</tr>
<tr>
<td>Households with higher than $20,000,000 assets</td>
<td>1</td>
</tr>
<tr>
<td>Households with missing information on assets</td>
<td>192</td>
</tr>
<tr>
<td>Households with wages less than half the state minimum wage</td>
<td>187</td>
</tr>
<tr>
<td>Households with missing labor income</td>
<td>26</td>
</tr>
<tr>
<td>Households with jumps on their labor income</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>661</td>
</tr>
</tbody>
</table>

Note: Data from 1999 to 2009 PSID waves. This table presents the sample selection criteria I operationalized for the paper. The sample selection criteria mostly follows that of Blundell et al. (2016).

Table C2 compares the households in my baseline sample with a sample of all married male household heads (regardless of participation status) and with a sample of all male household heads who were recorded as married at least once in the 1998 to 2008 period (again, independently of participation status). The table shows very small differences in the observables across households. Household earnings are only slightly smaller for the more comprehensive households than for the baseline sample, and total household wealth is smaller for the more comprehensive samples than for the baseline sample. The proportion of stock market participants is also roughly similar across these households, which suggests that concerns on whether I am removing households whose heads are facing extremely large income shocks that result in long unemployment spells seem to be disspelled.
Table C2: Baseline sample comparisons

<table>
<thead>
<tr>
<th></th>
<th>Baseline sample</th>
<th>With nonworking males</th>
<th>All ever married</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Median</td>
<td>Mean</td>
</tr>
<tr>
<td>All households</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household income</td>
<td>82,699.47</td>
<td>54,000.00</td>
<td>81,744.46</td>
</tr>
<tr>
<td>Total assets</td>
<td>371,189.60</td>
<td>164,043.40</td>
<td>367,263.70</td>
</tr>
<tr>
<td>Liquid wealth</td>
<td>95,529.77</td>
<td>17,083.33</td>
<td>97,067.34</td>
</tr>
<tr>
<td>Risky wealth</td>
<td>81,353.17</td>
<td>4,741.95</td>
<td>81,998.92</td>
</tr>
<tr>
<td>Stocks</td>
<td>51,547.79</td>
<td>-</td>
<td>52,983.55</td>
</tr>
<tr>
<td>Share of stocks in total wealth</td>
<td>0.077</td>
<td>-</td>
<td>0.077</td>
</tr>
<tr>
<td>Share of risky assets in total wealth</td>
<td>0.148</td>
<td>0.030</td>
<td>0.148</td>
</tr>
<tr>
<td>Ownership of stocks</td>
<td>0.410</td>
<td>0.408</td>
<td>0.408</td>
</tr>
<tr>
<td>Ownership of risky assets</td>
<td>0.591</td>
<td>0.587</td>
<td>0.586</td>
</tr>
<tr>
<td>Risky market participants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household income</td>
<td>98,524.98</td>
<td>65,000.00</td>
<td>97,221.93</td>
</tr>
<tr>
<td>Total assets</td>
<td>512,872.20</td>
<td>269,187.60</td>
<td>510,379.10</td>
</tr>
<tr>
<td>Liquid wealth</td>
<td>145,598.80</td>
<td>41,201.23</td>
<td>148,873.50</td>
</tr>
<tr>
<td>Risky wealth</td>
<td>137,676.70</td>
<td>43,994.58</td>
<td>139,767.20</td>
</tr>
<tr>
<td>Stocks</td>
<td>87,236.06</td>
<td>9,576.4</td>
<td>90,310.51</td>
</tr>
<tr>
<td>Share of stocks in total wealth</td>
<td>0.130</td>
<td>0.046</td>
<td>0.131</td>
</tr>
<tr>
<td>Share of risky assets in total wealth</td>
<td>0.251</td>
<td>0.192</td>
<td>0.252</td>
</tr>
</tbody>
</table>

Note: Data from 1999 to 2009 PSID waves. This table presents summary statistics for different subsamples. The first two columns correspond to the baseline sample that I use in my empirical study. The next two columns correspond to all male married households heads (independently of work status). The last two columns correspond to all male households heads who have been married at least once, again independently of work status. The first panel corresponds to all households in a given sample; the second panel corresponds to the risky market participants in a given sample.

C.2 Additional descriptive evidence

Tables C3 and C4 present summary statistics that are grouped by age and wealth quartiles, respectively. Looking at the tables, it is interesting to note that even for these categories, there appears to be non-monotonicities in the proportion of wealth invested in risky assets, regardless of the definition of risky assets I apply. Note, however, that looking at the other variables, the non-monotonicities that I find with respect to income quartiles does not appear.

Figures C1 and C2 present plots of the average stock market participation and the conditional risky shares for households of different age and wealth quartiles, and age and income quartiles, respectively. As the graphs indicate, participation rates in risky asset markets are increasing in wealth and income. The same observation does not hold true across age groups, however. As an example, Figure C1a indicates that young households in the poorest wealth quartiles are the most aggressive in investing in stocks, while middle-age households tend to be more aggressive in the richer wealth quartiles. However, the same figure indicates that older households participate the least in stock markets across wealth quartiles. Meanwhile, old households tend to invest the most in risky assets, across income quartiles, as indicated in Figure C2a. While the conditional risky shares do not show clear patterns, what I find is that for all age groups, there are non-monotonicities in the average conditional risky share, across

41 To group the households by age, I divide them into four categories: age less than 35, between 35 and 44, between 45 to 54, and age greater than 55.
wealth and income quartiles.

Table C3: Sample means of main variables, by age quartiles

<table>
<thead>
<tr>
<th>Age quartile</th>
<th>TOTAL</th>
<th>First</th>
<th>Second</th>
<th>Third</th>
<th>Fourth</th>
<th>All participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Household income</td>
<td>111,612.60</td>
<td>76,593.25</td>
<td>100,028.50</td>
<td>121,430.40</td>
<td>135,466.00</td>
<td></td>
</tr>
<tr>
<td>Total assets</td>
<td>371,189.60</td>
<td>137,241.40</td>
<td>277,322.60</td>
<td>422,290.60</td>
<td>645,399.90</td>
<td></td>
</tr>
<tr>
<td>Liquid wealth</td>
<td>95,529.77</td>
<td>31,349.16</td>
<td>64,892.01</td>
<td>106,183.50</td>
<td>200,763.70</td>
<td></td>
</tr>
<tr>
<td>Risky wealth</td>
<td>81,353.17</td>
<td>14,929.38</td>
<td>49,670.31</td>
<td>95,809.68</td>
<td>175,255.80</td>
<td></td>
</tr>
<tr>
<td>Stocks</td>
<td>51,547.79</td>
<td>10,620.69</td>
<td>31,041.50</td>
<td>59,315.04</td>
<td>117,464.00</td>
<td></td>
</tr>
</tbody>
</table>

| Share of stocks in total wealth | 0.077 | 0.064 | 0.069 | 0.083 | 0.085 |
| Share of risky assets in total wealth | 0.148 | 0.091 | 0.131 | 0.161 | 0.194 |
| Ownership of stocks | 0.410 | 0.319 | 0.403 | 0.428 | 0.440 |
| Ownership of risky assets | 0.591 | 0.452 | 0.561 | 0.614 | 0.703 |

| Risky market participants | Household income | 130,668.70 | 94,579.77 | 117,560.50 | 137,812.00 | 156,376.00 |
| Total assets              | 512,872.20      | 221,466.30 | 397,915.30 | 557,513.20 | 792,848.60 |
| Liquid wealth             | 145,598.80      | 56,364.76  | 103,700.40 | 154,777.00 | 265,198.90 |
| Risky wealth              | 137,676.70      | 33,058.69  | 88,490.72  | 156,039.50 | 249,176.00 |
| Stocks                    | 87,236.05       | 23,517.80  | 55,302.41  | 96,602.85  | 167,008.50 |
| Ownership of stocks       | 0.695 | 0.706 | 0.717 | 0.696 | 0.625 |
| Share of stocks in total wealth | 0.130 | 0.141 | 0.124 | 0.135 | 0.120 |
| Share of risky assets in total wealth | 0.250 | 0.201 | 0.234 | 0.263 | 0.276 |

Note: Data from 1999 to 2009 PSID waves. This table presents sample means of the main economic variables related to this empirical study, calculated across age quartiles. The first column calculates the mean across all households in the sample, while the second to fifth columns calculate the mean for different age quartiles. The first age quartile corresponds to households whose heads are less than 35 years old. The second age quartile corresponds to households whose heads are between 35 to 44 years old. The third age quartile corresponds to households between 45 to 54 years old. The last age quartile corresponds to households whose heads are greater than 54 years old. The first panel presents results across all households. The second panel presents results for risky market participants, defined as households who have direct and indirect stockholdings in stocks, mutual funds, and who have part of their pension funds invested in stocks.
Table C4: Sample means of main variables, by wealth quartiles

<table>
<thead>
<tr>
<th></th>
<th>TOTAL</th>
<th>First</th>
<th>Second</th>
<th>Third</th>
<th>Fourth</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Wealth quartile</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>All participants</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household income</td>
<td>111,612.60</td>
<td>67,004.64</td>
<td>86,454.48</td>
<td>101,647.70</td>
<td>181,489.00</td>
</tr>
<tr>
<td>Total assets</td>
<td>371,189.60</td>
<td>37,420.39</td>
<td>114,656.10</td>
<td>251,108.60</td>
<td>999,550.10</td>
</tr>
<tr>
<td>Liquid wealth</td>
<td>95,529.77</td>
<td>7,146.83</td>
<td>20,027.36</td>
<td>48,636.80</td>
<td>283,161.70</td>
</tr>
<tr>
<td>Risky wealth</td>
<td>81,353.17</td>
<td>2,949.97</td>
<td>13,611.84</td>
<td>37,937.66</td>
<td>250,221.60</td>
</tr>
<tr>
<td>Stocks</td>
<td>51,547.79</td>
<td>10,620.69</td>
<td>31,041.53</td>
<td>59,315.04</td>
<td>117,464.00</td>
</tr>
<tr>
<td>Share of stocks in total wealth</td>
<td>0.077</td>
<td>0.028</td>
<td>0.054</td>
<td>0.066</td>
<td>0.149</td>
</tr>
<tr>
<td>Share of risky assets in total wealth</td>
<td>0.148</td>
<td>0.066</td>
<td>0.113</td>
<td>0.146</td>
<td>0.252</td>
</tr>
<tr>
<td>Ownership of stocks</td>
<td>0.410</td>
<td>0.126</td>
<td>0.313</td>
<td>0.451</td>
<td>0.699</td>
</tr>
<tr>
<td>Ownership of risky assets</td>
<td>0.591</td>
<td>0.248</td>
<td>0.490</td>
<td>0.701</td>
<td>0.868</td>
</tr>
</tbody>
</table>

|                      | Risky market participants |
|                      | TOTAL   | First  | Second | Third  | Fourth |
|                      | Wealth quartile |
| Household income     | 130,668.70 | 78,603.50 | 94,826.08 | 103,112.20 | 182,725.30 |
| Total assets         | 512,872.20 | 44,418.93 | 118,376.50 | 255,587.80 | 1,025,478.00 |
| Liquid wealth        | 145,598.80 | 12,775.20 | 27,607.04 | 55,214.89 | 307,467.50 |
| Risky wealth         | 137,676.70 | 11,914.06 | 27,805.73 | 54,154.86 | 288,411.80 |
| Stocks               | 87,236.05 | 5,219.48 | 13,131.40 | 24,473.62 | 193,159.00 |
| Ownership of stocks  | 0.695 | 0.510 | 0.640 | 0.644 | 0.806 |
| Share of stocks in total wealth | 0.130 | 0.112 | 0.111 | 0.094 | 0.171 |
| Share of risky assets in total wealth | 0.250 | 0.267 | 0.230 | 0.208 | 0.291 |

Note: Data from 1999 to 2009 PSID waves. This table presents sample means of the main economic variables related to this empirical study, calculated across age quartiles. The first column calculates the mean across all households in the sample, while the second to fifth columns calculate the mean for different wealth quartiles. The first panel presents results across all households. The second panel presents results for risky market participants, defined as households who have direct and indirect stockholdings in stocks, mutual funds, and who have part of their pension funds invested in stocks.
Figure C1: Participation and the conditional risky share, by wealth and age quartiles

(a) Risky market participation rates

(b) Conditional risky shares

Note: Data from 1999 to 2009 PSID waves. The following figures show the average stock market participation rates and the average conditional risky share for households of different wealth and age quartiles. The x-axis corresponds to the wealth quartiles, while the y-axis corresponds to the average participation rate or risky share, depending on the definition. The blue line corresponds to households whose heads are less than 35 years old; the red line corresponds to households whose heads are 35 to 44 years old; the green line corresponds to households whose heads are 45 to 54 years old; and the orange line corresponds to households who are greater than 54 years old.

Figure C2: Participation and the conditional risky share, by income and age quartiles

(a) Risky market participation rates

(b) Conditional risky shares

Note: Data from 1999 to 2009 PSID waves. The following figures show the average stock market participation rates and the average conditional risky share for households of different income and age quartiles. The x-axis corresponds to the income quartiles, while the y-axis corresponds to the average participation rate or risky share, depending on the definition. The blue line corresponds to households whose heads are less than 35 years old; the red line corresponds to households whose heads are 35 to 44 years old; the green line corresponds to households whose heads are 45 to 54 years old; and the orange line corresponds to households who are greater than 54 years old.
C.3 Calculating human wealth

I follow the definition of Calvet and Sodini (2014) and Fagereng et al. (2017a) in the calculation of human wealth, which is crucial in the calculation of lifetime wealth, the exclusion restriction I follow in the paper. To be specific, the formula is the following:

\[ HW_{i,t} = L_{i,t} + \sum_{\tau=1}^{T-t} \pi_{\tau+1|\tau} \frac{\mathbb{E}_t(L_{i,t+\tau})}{(1+r)^\tau} \]  \hspace{1cm} (C43)

in which \( HW_{i,t} \) denotes human wealth, \( L_{i,t+\tau} \) is the labor income of the household at age \( t + \tau \) and \( \pi_\tau \) is the survival probability of the household head being alive at age \( t + \tau \) given that he was alive at age \( t \). I set the discount rate to the one I use in the calibrations, approximately 2 percent. Lifetime wealth, in this case, is the sum of human wealth and accumulated assets during that period.

D Estimation strategy

D.1 Model estimation: details

D.1.1 Likelihood function

The likelihood function is:

\[
f(\alpha_{it}^T, \nu_{it}, \epsilon_{it}, w_{it}, z_{it}, d_{it}; \mu_i) = \prod_{t=1}^{T} \left[ f(\alpha_{it}^T|\nu_{it}, \epsilon_{it}, w_{it}, x_{it}) p(d_{it} = 1|\nu_{it}, \epsilon_{it}, w_{it}, z_{it}) \nabla C(u, \nu; c) \right]^{d_{it}} 
\times \prod_{t=1}^{T} [p(d_{it} = 0|\nu_{it}, \epsilon_{it}, w_{it}, z_{it})]^{-1-d_{it}} \prod_{t=2}^{T} f(w_{it}|w_{it-1}, \nu_{it-1}, \epsilon_{it-1}, \alpha_{it-1}, x_{it}) 
\times f(w_{it}|\nu_{it}, x_{it}) \prod_{t=1}^{T} f(y_{it}|\nu_{it}) \prod_{t=2}^{T} f(y_{it}|y_{it-1}) f(v_{it}) \]  \hspace{1cm} (D44)

where \( u = F(\alpha_{it}^T|\nu_{it}, \epsilon_{it}, w_{it}, x_{it}) \), \( v = p(d_{it} = 1|\nu_{it}, \epsilon_{it}, w_{it}, z_{it}) \) and \( \nabla C(\cdot, \cdot; \cdot) \) is the first derivative of the conditional copula with respect to the first argument.

I can simplify the likelihood function further by noting that I can rewrite the conditional copula as follows:

\[
C(u, \nu; c) = \frac{G(F(\alpha_{it}^T|\nu_{it}, \epsilon_{it}, w_{it}, x_{it}), p(d_{it} = 1|\nu_{it}, \epsilon_{it}, w_{it}, z_{it}); \rho_c)}{p(d_{it} = 1|\nu_{it}, \epsilon_{it}, w_{it}, z_{it})} \]  \hspace{1cm} (D45)

where \( G(\cdot, \cdot; \rho_c) \) is the Gaussian copula. It follows that the first derivative of this function with respect to the first argument is:

\[
\nabla G(u, \nu; c) = \Phi \left( \frac{\Phi^{-1}(v) - \rho_c \Phi^{-1}(u)}{\sqrt{1-\rho_c^2}} \right) \]  \hspace{1cm} (D46)
Substituting the resulting expression for $\nabla C(u, v; c)$ to the expression above, the likelihood function simplifies to:

\[
f(a^*_i | v_i, e_i, w_i, x_i) = \prod_{l=1}^{T} \left[ f(a^*_l | v_{i_l}, e_{i_l}, w_{i_l}, x_{i_l}) \Phi \left( \frac{\Phi^{-1}(v) - \rho_c \Phi^{-1}(u)}{\sqrt{1 - \rho_c^2}} \right) \right]^{d_{i_l}} \times \prod_{l=1}^{T} \left[ f(d_{i_l} = 0 | v_{i_l}, e_{i_l}, w_{i_l}, z_{i_l}) \right]^{1 - d_{i_l}} \prod_{l=2}^{T} f(w_{i_l} | w_{i_l-1}, v_{i_l-1}, y_{i_l-1}, a_{i_l-1}, x_{i_l}) f(w_{i_l} | v_{i_l}, x_{i_l}) \times \prod_{l=1}^{T} f(y_{i_l} | v_{i_l}) \prod_{l=2}^{T} f(v_{i_l} | v_{i_l-1}) f(v_{i_l})
\]

As the model is fully specified, I can write the likelihood function in closed form. To focus my discussion, I illustrate the specification for $f(a^*_l | v_{i_l}, e_{i_l}, w_{i_l}, x_{i_l})$ and $F(a^*_l | v_{i_l}, e_{i_l}, w_{i_l}, x_{i_l})$. Notice that I can write the approximating outcome density as:

\[
f(a^*_l | v_{i_l}, e_{i_l}, w_{i_l}, x_{i_l}) = \prod_{l=1}^{L-1} \left[ \sum_{k=0}^{K} b^{g}_{kl} \varphi_k(v_{i_l}, e_{i_l}, w_{i_l}, a_{i_l}, \epsilon_{i_l}) (b^{g}_{i_l+1} - b^{g}_{ik}) \right] \times 1 \left\{ \frac{\sum_{k=0}^{K} b^{g}_{kl} \varphi_k(v_{i_l}, e_{i_l}, w_{i_l}, a_{i_l}, \epsilon_{i_l})}{\sum_{k=0}^{K} b^{g}_{kl} \varphi_k(v_{i_l}, e_{i_l}, w_{i_l}, a_{i_l}, \epsilon_{i_l}) (b^{g}_{i_l+1} - b^{g}_{ik})} \right\}^{\lambda_a} \left[ 1 - \frac{\varphi_k(v_{i_l}, e_{i_l}, w_{i_l}, a_{i_l}, \epsilon_{i_l})}{\sum_{k=0}^{K} b^{g}_{kl} \varphi_k(v_{i_l}, e_{i_l}, w_{i_l}, a_{i_l}, \epsilon_{i_l})} \right]^{\lambda_a} \left[ 1 - \frac{\sum_{k=0}^{K} b^{g}_{kl} \varphi_k(v_{i_l}, e_{i_l}, w_{i_l}, a_{i_l}, \epsilon_{i_l})}{\sum_{k=0}^{K} b^{g}_{kl} \varphi_k(v_{i_l}, e_{i_l}, w_{i_l}, a_{i_l}, \epsilon_{i_l})} \right]
\]

where I proceed with an exponential modelling of the tails.

The approximating conditional distribution functions are:

\[
F(a^*_l | v_{i_l}, e_{i_l}, w_{i_l}, x_{i_l}) = \prod_{l=1}^{L-1} \left[ \sum_{k=0}^{K} b^{g}_{kl} \varphi_k(v_{i_l}, e_{i_l}, w_{i_l}, a_{i_l}, \epsilon_{i_l}) (b^{g}_{i_l+1} - b^{g}_{ik}) \right] \times 1 \left\{ \frac{\sum_{k=0}^{K} b^{g}_{kl} \varphi_k(v_{i_l}, e_{i_l}, w_{i_l}, a_{i_l}, \epsilon_{i_l})}{\sum_{k=0}^{K} b^{g}_{kl} \varphi_k(v_{i_l}, e_{i_l}, w_{i_l}, a_{i_l}, \epsilon_{i_l}) (b^{g}_{i_l+1} - b^{g}_{ik})} \right\}^{\lambda_a} \left[ 1 - \frac{\sum_{k=0}^{K} b^{g}_{kl} \varphi_k(v_{i_l}, e_{i_l}, w_{i_l}, a_{i_l}, \epsilon_{i_l})}{\sum_{k=0}^{K} b^{g}_{kl} \varphi_k(v_{i_l}, e_{i_l}, w_{i_l}, a_{i_l}, \epsilon_{i_l})} \right]^{\lambda_a} \left[ 1 - \frac{\sum_{k=0}^{K} b^{g}_{kl} \varphi_k(v_{i_l}, e_{i_l}, w_{i_l}, a_{i_l}, \epsilon_{i_l})}{\sum_{k=0}^{K} b^{g}_{kl} \varphi_k(v_{i_l}, e_{i_l}, w_{i_l}, a_{i_l}, \epsilon_{i_l})} \right]
\]

\[
D.1.2 \quad \text{Model restrictions}
\]

The semi-structural model implies the following restrictions. Let $a^+ = \max\{a, 0\}$ and $a^- = \max\{-a, 0\}$ denote the check function of quantile regression. (Koenker and Bassett (1978)). Let $\tilde{\theta}$ denote the true value of $\theta$, and let

\[
f_i(v_{i_l}^T | a^*_l, \tilde{\theta})
\]
denote the posterior density of the persistent component \( v_i^T = (v_{i1}, \ldots, v_{iT})' \) given portfolio shares, earnings, wealth and socio-demographic data. Because the model is fully specified, this is a known function of \( \bar{\theta} \).

The participation rule implies the following restrictions (removing \( Z_i \) for conciseness):

\[
\begin{align*}
(\bar{b}_0^p, \ldots, \bar{b}_K^p) &= \arg \max_{(b_0^p, \ldots, b_K^p, \nu^p)} \sum_{t=1}^{T} \mathbb{E} \left[ d_{it} \log \Lambda \left( \sum_{k=0}^{K} b_k^p \varphi_k(v_{it}, \epsilon_{it}, w_{it}, a\epsilon_{it}) \right) \right] \\
&\quad \times (1 - d_{it}) \log \left( 1 - \Lambda \left( \sum_{k=1}^{K} b_k^p \varphi_k(v_{it}, \epsilon_{it}, w_{it}, a\epsilon_{it}) \right) \right) f_i(v_i^T | \alpha_i^T, \epsilon_i^T, w_i^T, Z_i, d_i^T; \bar{\theta}) dv_i^T \\
&= \frac{\partial}{\partial \nu^p} \sum_{t=1}^{T} \mathbb{E} \left[ \int d_{it} \log \Lambda \left( \sum_{k=0}^{K} b_k^p \varphi_k(v_{it}, \epsilon_{it}, w_{it}, a\epsilon_{it}) \right) \right] \\
&\quad \times (1 - d_{it}) \log \left( 1 - \Lambda \left( \sum_{k=1}^{K} b_k^p \varphi_k(v_{it}, \epsilon_{it}, w_{it}, a\epsilon_{it}) \right) \right) f_i(v_i^T | \alpha_i^T, \epsilon_i^T, w_i^T, Z_i, d_i^T; \bar{\theta}) dv_i^T \\
&= (D49)
\end{align*}
\]

The copula linking the portfolio and participation rules implies the following restrictions:

\[
\tilde{\rho}_c = \arg \min_{\rho_c} \sum_{t=1}^{T} \mathbb{E} \left[ \int d_{it} \mathbb{1} \left\{ \Lambda^{-1}(a_{it}^\nu) - \sum_{k=0}^{K} b_k^u \varphi_k(v_{it}, \epsilon_{it}, w_{it}, a\epsilon_{it}) \right\} \right] \\
&\quad - G(\tau, p(x_i); \rho_c) f_i(v_i^T | \alpha_i^T, \epsilon_i^T, w_i^T, Z_i, d_i^T; \bar{\theta}) dv_i^T \\
&= (D50)
\]

The portfolio rules implies the following restrictions, for all \( l \in \{1, \ldots, L\} \),

\[
\begin{align*}
(\bar{b}_{0l}^p, \ldots, \bar{b}_{KL}^p) &= \arg \min_{(b_{0l}^p, \ldots, b_{KL}^p)} \sum_{t=1}^{T} \mathbb{E} \left[ \int d_{it} \mathbb{1} \left\{ \Lambda^{-1}(a_{it}^\nu) - \sum_{k=0}^{K} b_k^u \varphi_k(v_{it}, \epsilon_{it}, w_{it}, a\epsilon_{it}) \right\} \right] \\
&\quad + (1 - G(\tau, p(x_i); \rho_c)) \left\{ \Lambda^{-1}(a_{it}^\nu) - \sum_{k=0}^{K} b_{kl+1}^u \varphi_k(v_{it}, \epsilon_{it}, w_{it}, a\epsilon_{it}) \right\} - f_i(v_i^T | \alpha_i^T, \epsilon_i^T, w_i^T, Z_i, d_i^T; \bar{\theta}) dv_i^T \\
&= (D51)
\end{align*}
\]

The initial wealth condition implies the following restriction:

\[
\begin{align*}
(\bar{b}_{0l}^w, \ldots, \bar{b}_{KL}^w) &= \arg \min_{(b_{0l}^w, \ldots, b_{KL}^w)} \mathbb{E} \left[ \int \tau_i \left( w_{i1} - \sum_{k=0}^{K} b_{kl}^w \varphi_k(v_{i1}, a\epsilon_{i1}) \right) \right] \\
&\quad + (1 - \tau_i) \left( w_{i1} - \sum_{k=0}^{K} b_{kl}^w \varphi_k(v_{i1}, a\epsilon_{i1}) \right) - f_i(v_i^T | \alpha_i^T, \epsilon_i^T, w_i^T, Z_i, d_i^T; \bar{\theta}) dv_i^T \\
&= (D52)
\end{align*}
\]

Turning to the wealth accumulation rule, I have the following restrictions:

\[
\begin{align*}
(\bar{b}_{0l}^m, \ldots, \bar{b}_{KL}^m) &= \arg \min_{(b_{0l}^m, \ldots, b_{KL}^m)} \sum_{t=2}^{T} \mathbb{E} \left[ \int \left( w_{it} - \sum_{k=1}^{K} b_k^m \varphi_k(v_{it-1}, \epsilon_{it-1}, w_{it-1}, \alpha_{it-1}, a\epsilon_{it}) \right)^2 \right] \\
&\quad \times f_i(v_i^T | \alpha_i^T, \epsilon_i^T, w_i^T, Z_i, d_i^T; \bar{\theta}) dv_i^T \\
&= (D53)
\end{align*}
\]

The variance of the evolution of wealth, meanwhile, follows:

\[
\begin{align*}
\tilde{\sigma}^2 &= \frac{1}{T} \sum_{t=2}^{T} \mathbb{E} \left[ \int \left( w_{it} - \sum_{k=1}^{K} b_k^m \varphi_k(v_{it-1}, \epsilon_{it-1}, w_{it-1}, \alpha_{it-1}, a\epsilon_{it}) \right)^2 \right] \\
&\quad \times f_i(v_i^T | \alpha_i^T, \epsilon_i^T, w_i^T, Z_i, d_i^T; \bar{\theta}) dv_i^T \\
&= (D54)
\end{align*}
\]

68
Finally, the tail parameters satisfy the following moment conditions. For example, in the case of the tail parameters of the portfolio rule, I have:

\[
\lambda^a = -\sum_{t=1}^T \mathbb{E} \left[ \mathbf{1} \{ \Lambda^{-1}(a^a_t) \leq \sum_{k=0}^K b^a_k \varphi_k(v_t, \epsilon_t, w_t, a \epsilon_t) \} f_t(v^T_t) \right]
\]

\[
\lambda^a = \sum_{t=1}^T \mathbb{E} \left[ \mathbf{1} \{ \Lambda^{-1}(a^a_t) \geq \sum_{k=0}^K b^a_k \varphi_k(v_t, \epsilon_t, w_t, a \epsilon_t) \} f_t(v^T_t) \right]
\]

(D55)

and

\[
\lambda^a = \sum_{t=1}^T \mathbb{E} \left[ \mathbf{1} \{ \Lambda^{-1}(a^a_t) \leq \sum_{k=0}^K b^a_k \varphi_k(v_t, \epsilon_t, w_t, a \epsilon_t) \} f_t(v^T_t) \right]
\]

\[
\lambda^a = \sum_{t=1}^T \mathbb{E} \left[ \mathbf{1} \{ \Lambda^{-1}(a^a_t) \geq \sum_{k=0}^K b^a_k \varphi_k(v_t, \epsilon_t, w_t, a \epsilon_t) \} f_t(v^T_t) \right]
\]

(D56)

with similar model restrictions for the other tail parameters.

D.1.3 Estimation algorithm: details

Start with \( \theta^{(0)} \). Then, iterate on \( s = 0, 1, 2, \ldots \) the following two steps:

**Stochastic E-step:** Draw values \( v_i^{(m)} = (v_{i1}^{(m)}, \ldots, v_{iT}^{(m)}) \) from

\[
f(a_t^T, v_t^T, \epsilon_t^T, w_t^T, z_t^T, d_t, \tilde{\theta}^{(s)}), = \prod_{t=1}^T f(a_t^T | v_t, \epsilon_t, w_t, x_t; \tilde{\theta}^{(s)}) \Phi \left( \frac{\Phi^{-1}(v) - \rho_c \Phi^{-1}(u)}{\sqrt{1 - \rho_c^2}} \right)
\]

\[
\times \prod_{t=1}^T p(d_{it} = 0 | v_t, \epsilon_t, w_t, x_t; \tilde{\theta}^{(s)}) \prod_{t=2}^T f(w_t | w_{t-1}, v_{t-1}, y_{t-1}, a_{t-1}, x_t; \tilde{\theta}^{(s)}) f(v_{t1} | v_{t1}, x_t; \tilde{\theta}^{(s)})
\]

\[
\times \prod_{t=1}^T f(y_t | v_t, \tilde{\theta}) \prod_{t=2}^T f(v_t | v_{t-1}, \tilde{\theta}) f(v_{t1})
\]

(D57)

**M-step:** Compute, for all \( l = 1, \ldots, L \):

\[
\hat{b}_0^{P(s+1)}, \ldots, \hat{b}_K^{P(s+1)} = \arg \max_{(\hat{b}_0, \ldots, \hat{b}_K)} \sum_{l=1}^L \sum_{i=1}^N \sum_{m=1}^M d_{it} \log \Lambda \left( \sum_{k=0}^K b_k^a \varphi_k(v_{i1}^{(m)}, \epsilon_t, w_t, a \epsilon_t) \right)
\]

\[
+ (1 - d_{it}) \log \left( 1 - \Lambda \left( \sum_{k=0}^K b_k^a \varphi_k(v_{i1}^{(m)}, \epsilon_t, w_t, a \epsilon_t) \right) \right)
\]

(D58)

\[
\hat{\rho}_{c}^{(s+1)} = \arg \min_{\rho_c} \left\| \sum_{t=1}^T \sum_{i=1}^N \sum_{m=1}^M d_{it} Y(t_i, x_t) \left[ 1 \{ \Lambda^{-1}(a^a_t) \leq \sum_{k=0}^K b_k^a \varphi_k(\cdot) \} - G(t_i, p(x_t); \rho_c) \right] \right\|
\]

(D59)

\[
\hat{a}_0^{P(s+1)}, \ldots, \hat{a}_K^{P(s+1)} = \arg \min_{(a_0, \ldots, a_K)} \sum_{l=1}^L \sum_{i=1}^N \sum_{m=1}^M d_{it} \left[ G(t_i, p(x_t); \rho_c) \left( \Lambda^{-1}(a_t^a) - \sum_{k=0}^K b_k^a \varphi_k(v_{i1}^{(m)}, \epsilon_t, w_t, a \epsilon_t) \right) \right]
\]

\[
+ (1 - G(t_i, p(x_t); \rho_c)) \left( \Lambda^{-1}(a_t^a) - \sum_{k=0}^K b_k^a \varphi_k(v_{i1}^{(m)}, \epsilon_t, w_t, a \epsilon_t) \right)
\]

(D60)
such that there is an exclusion restriction, represented by variable, the observed counterpart of which, \( x \), in which \( y \) is measured with error. The model set-up is as follows:

\[
\sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{m=1}^{M} \{ \lambda^{-1}(a^*_\alpha) \leq \sum_{k=0}^{K} \hat{\beta}_{k1} \varphi_k(v_{it}^{(m)}, \epsilon_{it}, w_{it}, a g e_{it}) \} \]

For the tail parameters, I calculate the following:

\[
\lambda_{\alpha}^{(s+1)} = \frac{\sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{m=1}^{M} \{ \lambda^{-1}(a^*_\alpha) \leq \sum_{k=0}^{K} \hat{\beta}_{k1} \varphi_k(v_{it}^{(m)}, \epsilon_{it}, w_{it}, a g e_{it}) \}}{\sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{m=1}^{M} (\lambda^{-1}(a^*_\alpha) - \sum_{k=0}^{K} \hat{\beta}_{k1} \varphi_k(v_{it}^{(m)}, \epsilon_{it}, w_{it}, a g e_{it}))} \}
\]

with similar updating rules for the other tail parameters.

D.2 Simulation experiment

I outline the results of a small-scale exercise that is a representation of the economic model that I take to the data. The data generating process is as follows:

\[
y^*_it = \alpha_0 + \alpha_1 x^*_it + \alpha_2 z^*_it + \alpha_3 b^*_it + \sigma(x^*_it, z^*_it)u^*_it \\
y^*_it = y^*_it \cdot d^*_it \\
d^*_it = 1(\beta_0 + \beta_1 x^*_it + \beta_2 z^*_it + \beta_3 b^*_it + \beta_4 w^*_it + \nu^*_it \geq 0) \\
z^*_it = \gamma_0 + \gamma_1 x^*_i(t-1) + \gamma_2 z^*_i(t-1) + \gamma_3 y^*_i(t-1) + \gamma_4 b^*_i + \sigma(x^*_i(t-1), z^*_i(t-1), y^*_i(t-1)) + \xi^*_it \\
z^*_it = \gamma_0 + \gamma_1 x^*_i(t-1) + \gamma_2 b^*_i + \xi^*_it
\]

in which \( y^*_it \) is a latent variable that is a function of variables \( x^*_it, z^*_it \) and \( b^*_it \), and \( x^*_it \) is a latent variable, the observed counterpart of which, \( x_{it} \), is measured with error. The model set-up is such that there is an exclusion restriction, represented by \( w_{it} \). I model the joint distribution of \( u_{it} \) and \( v_{it} \) conditional on the observed and latent variables as:

\[
\begin{pmatrix} u_{it} \\ v_{it} \end{pmatrix} | x_{it} \sim N\left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_c \\ \rho_c & 1 \end{pmatrix} \right)
\]

The equations below describe the evolution of the latent variable.

\[
x_{it} = x_{it}^* + \epsilon_{it} \\
x^*_{it} = \rho x^*_{it-1} + \nu_{it}
\]
The following are further assumptions on the independence and the distribution of the error terms:

1. (independence of the error terms) \( x_{it}^* \perp \epsilon_{it} \perp \nu_{it} \)

2. (distributional assumptions) \( \epsilon_{it} \sim iidN(0, \sigma_{\epsilon}^2) \), \( \nu_{it} \sim iidN(0, \sigma_{\nu}^2) \) and \( x_{it}^* \sim iidN(0, \sigma_z^2) \)

To generate the data, I draw \( b_{it} \) and \( w_{it} \) from a Normal(0,1). I consider a location-scale model; that is, \( \sigma(x_{it}^*, z_{it}) = 1 + \delta_1 x_{it}^* + \delta_2 z_{it} \). The parameter configurations that I consider are written in the table below.

Table D1: Parameter configurations, Monte Carlo simulations

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True value</th>
<th>Parameter</th>
<th>True value</th>
<th>Parameter</th>
<th>True value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha_0 )</td>
<td>1.25</td>
<td>( \gamma_0 )</td>
<td>1.0</td>
<td>( \delta_1 )</td>
<td>0.1</td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>0.5</td>
<td>( \gamma_1 )</td>
<td>0.5</td>
<td>( \delta_2 )</td>
<td>0.1</td>
</tr>
<tr>
<td>( \alpha_2 )</td>
<td>0.5</td>
<td>( \gamma_2 )</td>
<td>-0.2</td>
<td>( \sigma_\epsilon )</td>
<td>0.2</td>
</tr>
<tr>
<td>( \alpha_3 )</td>
<td>0.25</td>
<td>( \gamma_3 )</td>
<td>0.1</td>
<td>( \sigma_{\nu} )</td>
<td>0.5</td>
</tr>
<tr>
<td>( \beta_0 )</td>
<td>1.0</td>
<td>( \gamma_4 )</td>
<td>0.2</td>
<td>( \sigma_z )</td>
<td>0.3</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>-0.25</td>
<td>( \gamma_0 )</td>
<td>1.0</td>
<td>( \rho_c )</td>
<td>0.8</td>
</tr>
<tr>
<td>( \beta_2 )</td>
<td>-0.25</td>
<td>( \gamma_1 )</td>
<td>1.5</td>
<td>( \rho )</td>
<td>0.95</td>
</tr>
<tr>
<td>( \beta_3 )</td>
<td>0.25</td>
<td>( \gamma_2 )</td>
<td>0.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \beta_4 )</td>
<td>-0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( N )</td>
<td>1000</td>
<td>( T )</td>
<td>6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: This table describes the parameter configurations of the small-scale simulation exercise that I perform to verify the finite-sample performance of the stochastic EM algorithm of the nonlinear reduced form model with the Arellano and Bonhomme (2017) quantile selection estimator.

I simulate 100 datasets, each with \( N = 1,000 \) cross-sectional units and \( T = 6 \) time series observations. I take \( M = 1 \), stop the chain after a large number of iterations, and report an average across the last \( \tilde{S} \) values \( \hat{\theta} = \frac{1}{\tilde{S}} \sum_{s=S-S+1}^{S} \hat{\theta}^{(s)} \), where I take \( \tilde{S} = S/2 \). Each estimation is based on \( S = 100 \) iterations, with 100 random walk Metropolis-Hastings draws per iteration. The parameters of the data generating process are such that the percentage of censoring is approximately 22 percent.

The model specification is close to that outlined in the main text. As an example, the estimating equation for the latent outcome \( y_{it}^* \) takes the following form:

\[
y_{it}^* = g(x_{it}^*, x_{it}, z_{it}, b_{it}, \tau) = \sum_{k=0}^{K} a_k(\tau)g_k(x_{it}^*, x_{it}, z_{it}) + \gamma y b_{it}
\]

where \( \tau \in (0, 1) \), where \( g_k(\cdot) \) is a dictionary of functions, with \( g_0 = 1 \).42

---

42 In the simulations, I approximate the participation and portfolio rules via tensor products of Hermite polynomials of order (2,1,1).
Table D2 presents the results of the simulation experiments. I calculate the mean of the average derivative effects across all simulations, for each of the quantiles in which I estimate the model. As can be observed, the biases are moderate, and are larger at the extremes of the quantiles estimated. Moreover, the biases are larger for the mismeasured variable and the constant; meanwhile, biases seem to be smaller for the observed variables. Finally, the estimation results are relatively precise outside of the tails, with smaller standard deviations.

Table D2: Simulation results, Arellano and Bonhomme (2017) set-up

<table>
<thead>
<tr>
<th>Quantile</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.726</td>
<td>0.997</td>
<td>1.250</td>
<td>1.503</td>
<td>1.774</td>
<td>2.092</td>
<td>2.532</td>
</tr>
<tr>
<td>$x_{it}$</td>
<td>0.448</td>
<td>0.475</td>
<td>0.500</td>
<td>0.552</td>
<td>0.584</td>
<td>0.628</td>
<td></td>
</tr>
<tr>
<td>$z_{it}$</td>
<td>0.448</td>
<td>0.475</td>
<td>0.500</td>
<td>0.552</td>
<td>0.584</td>
<td>0.628</td>
<td></td>
</tr>
<tr>
<td>$b_{it}$</td>
<td>0.250</td>
<td>0.250</td>
<td>0.250</td>
<td>0.250</td>
<td>0.250</td>
<td>0.250</td>
<td>0.250</td>
</tr>
</tbody>
</table>

Monte Carlo means

<table>
<thead>
<tr>
<th>Quantile</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.688</td>
<td>0.970</td>
<td>1.223</td>
<td>1.481</td>
<td>1.743</td>
<td>2.073</td>
<td>2.564</td>
</tr>
<tr>
<td>$x_{it}$</td>
<td>0.418</td>
<td>0.445</td>
<td>0.468</td>
<td>0.492</td>
<td>0.513</td>
<td>0.547</td>
<td>0.595</td>
</tr>
<tr>
<td>$z_{it}$</td>
<td>0.451</td>
<td>0.482</td>
<td>0.510</td>
<td>0.535</td>
<td>0.562</td>
<td>0.599</td>
<td>0.644</td>
</tr>
<tr>
<td>$b_{it}$</td>
<td>0.264</td>
<td>0.268</td>
<td>0.279</td>
<td>0.285</td>
<td>0.286</td>
<td>0.282</td>
<td>0.285</td>
</tr>
</tbody>
</table>

Monte Carlo standard deviations

<table>
<thead>
<tr>
<th>Quantile</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.087</td>
<td>0.067</td>
<td>0.063</td>
<td>0.060</td>
<td>0.070</td>
<td>0.082</td>
<td>0.129</td>
</tr>
<tr>
<td>$x_{it}$</td>
<td>0.052</td>
<td>0.041</td>
<td>0.039</td>
<td>0.040</td>
<td>0.041</td>
<td>0.052</td>
<td>0.083</td>
</tr>
<tr>
<td>$z_{it}$</td>
<td>0.029</td>
<td>0.025</td>
<td>0.023</td>
<td>0.021</td>
<td>0.021</td>
<td>0.026</td>
<td>0.041</td>
</tr>
<tr>
<td>$b_{it}$</td>
<td>0.017</td>
<td>0.013</td>
<td>0.010</td>
<td>0.010</td>
<td>0.008</td>
<td>0.011</td>
<td>0.012</td>
</tr>
</tbody>
</table>

Note: This table describes the result of the simulation experiment where the relevant estimator of the equations that correspond to the portfolio and participation rules is the Arellano and Bonhomme (2017) quantile selection model estimator. This is an average across 100 datasets with $N = 1,000$ and $T = 6$.

D.3 Buchinsky and Hahn (1998)

D.3.1 Nonlinear reduced form and model specification

The equivalent nonlinear reduced form that corresponds to the Buchinsky and Hahn (1998) model is the following:

\[
\alpha_{it}^* = g_t(v_{it}, \varepsilon_{it}, w_{it}, X_{it}, u_{it})
\]

(D64)

\[
\alpha_{it} = \alpha_{it}^* \cdot d_{it}
\]

(D65)

\[
d_{it} = \begin{cases} 
1, & \text{if } g_t(v_{it}, \varepsilon_{it}, w_{it}, q(X_{it})) \leq u_{it} \\
0, & \text{otherwise}
\end{cases}
\]

(D66)

\[
w_{it} = \tilde{h}_t(v_{it-1}, \varepsilon_{it-1}, w_{it-1}, \alpha_{it-1}, X_{it}, \zeta_{it})
\]

(D67)

\[
w_{i0} = \tilde{h}_{i0}(v_{i0}, \zeta_{i0})
\]

(D68)
The specification outlined here is similar to the one outlined in the main text. There are two main differences: the first is that the variables that determine participation are the same as the ones that determine the outcome, and the second is that the error terms of equations (D64) and (D66) are the same.

### D.3.2 Model specification and estimation algorithm

**Participation rule.** Most of the model specifications outlined in the main text remain to be the same when I move to the model of Buchinsky and Hahn (1998); the main difference is in the participation rule, equation (D66). The specification now becomes:

\[
\Pr(d_{it} = 1 | v_{it}, \varepsilon_{it}, w_{it}, age_{it}, X_{it}) = \Lambda \left( \sum_{k=0}^{K} b_k^p \phi_k(v_{it}, \varepsilon_{it}, w_{it}, age_{it}) + \gamma^p X_{it} \right)
\]

(D69)

where \( \Lambda(\cdot) \) is the logistic function and \( \phi_k \) is a dictionary of functions.\(^{43}\)

**Overview of the estimation algorithm.** The M-step that corresponds with Buchinsky and Hahn (1998) is characterized by the following steps. First, I estimate the participation rule:

\[
\max_{(b_0, \ldots, b_K, \gamma^p)} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{m=1}^{M} d_{it} \log \left[ \Lambda \left( \sum_{k=0}^{K} b_k^p \phi_k(v_{it}, \varepsilon_{it}, w_{it}, age_{it}) + \gamma^p X_{it} \right) \right] + (1 - d_{it}) \log \left[ 1 - \Lambda \left( \sum_{k=0}^{K} b_k^p \phi_k(v_{it}, \varepsilon_{it}, w_{it}, age_{it}) + \gamma^p X_{it} \right) \right].
\]

(D70)

From here, I can compute the propensity score \( p(x_{it}) \); that is, the probability that a household participates in the stock market. In the second step, I estimate the following censored quantile regression, which updates the parameters of the portfolio rule:

\[
\min_{(b_0, \ldots, b_K, \gamma^p)} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{m=1}^{M} d_{it} 1\{h_{\tau}(x_{it}) > 0\} \left[ h_{\tau}(x_{it}) \left( \Lambda^{-1}(a_{it}^\tau) - \sum_{k=0}^{K} b_k^p(\tau) \phi_k(v_{it}, \varepsilon_{it}, w_{it}, age_{it}) + \gamma^p(\tau)'X_{it} \right)^+ + (1 - h_{\tau}(x_{it})) \left( \Lambda^{-1}(a_{it}^\tau) - \sum_{k=0}^{K} b_k^p(\tau) \phi_k(v_{it}, \varepsilon_{it}, w_{it}, age_{it}) + \gamma^p(\tau)'X_{it} \right)^- \right]
\]

(D71)

where \( h_{\tau}(x_{it}) = \frac{\tau + p(x_{it}) - 1}{p(x_{it})} \). The role of this function is to “shift” the mass from the unobserved to the observed part of the distribution of portfolio shares. In fact, \( h_{\tau}(x_{it}) \) provides the link between Buchinsky and Hahn (1998) and Arellano and Bonhomme (2017). This is because the conditional copula of the error terms of the participation and portfolio rules when there is no exclusion restriction and where the error terms are the same is the lower Fréchet bound, i.e., \( G^-(\tau, p) = \max \left\{ \frac{\tau + p(x_{it}) - 1}{p(x_{it})}, 0 \right\} \).

\(^{43}\)Buchinsky and Hahn (1998) propose to estimate the propensity score with a nonparametric kernel density estimator, as the propensity score depends on the latent distribution of outcomes. However, this leads to a less computationally tractable estimation procedure in the context of the nonlinear reduced form model. Hence, I specify the propensity score with this model. An added advantage is the possibility of calculating extensive margins of income components and wealth.
As the model restrictions and implementation are similar as in the main text, I do not outline them here. I show, however, the likelihood function implied by the model.

**Likelihood function.** The corresponding likelihood function has the following form:

\[
f(\alpha_i^T, v_i^T, \epsilon_i^T, m_i^T, X_i, \tau_i; \Theta) = \prod_{t=1}^{T} \left[ f(\alpha_i^* | v_{it}, \epsilon_{it}, m_{it}, x_{it}) \right]^{d_{it}} \prod_{t=1}^{T} \left[ p(d_{it} = 0 | v_{it}, \epsilon_{it}, m_{it}, x_{it}) \right]^{1-d_{it}} \]

\[
\times \prod_{t=1}^{T} \prod_{t=2}^{T} \left[ f(m_{it} | m_{it-1}, v_{it-1}, y_{it-1}, x_{it}, \alpha_{it-1}, x_{it}) \right] \times f(m_{i1} | v_{i1}, x_{i1}) \prod_{t=1}^{T} f(v_{it} | v_{it-1}) f(v_{i1}) \]  \hspace{1cm} (D72)

I can simplify the likelihood function further by noting that I can rewrite the lower Fréchet bound as follows:

\[
C(u, v; c) = \begin{cases} 
\frac{\tau + p(x_0) - 1}{p(x_0)} & \text{if } p(x_0) > 1 - \tau \\
0 & \text{otherwise}
\end{cases}
\]

It follows that the first derivative of this function with respect to the first argument is:

\[
\nabla C(u, v; c) = \begin{cases} 
\frac{1}{p(x_0)} & \text{if } p(x_0) > 1 - \tau \\
0 & \text{otherwise}
\end{cases}
\]

Substituting this, the likelihood function above simplifies to:

\[
f(\alpha_i^T, v_i^T, \epsilon_i^T, m_i^T, X_i, \tau_i; \Theta) = \prod_{t=1}^{T} \left[ f(\alpha_i^* | v_{it}, \epsilon_{it}, m_{it}, x_{it}) \right]^{d_{it}} \prod_{t=1}^{T} \left[ p(d_{it} = 0 | v_{it}, \epsilon_{it}, m_{it}, x_{it}) \right]^{1-d_{it}} \]

\[
\times \prod_{t=1}^{T} \prod_{t=2}^{T} \left[ f(m_{it} | m_{it-1}, v_{it-1}, y_{it-1}, x_{it}, \alpha_{it-1}, x_{it}) \right] \times f(m_{i1} | v_{i1}, x_{i1}) \prod_{t=1}^{T} f(v_{it} | v_{it-1}) f(v_{i1}) \]  \hspace{1cm} (D73)

**D.3.3 Simulation evidence**

Similar to the Arellano and Bonhomme (2017) estimator, I conduct a small-scale Monte Carlo simulation exercise. The data generating process is as follows:

\[
y_{it}^* = \alpha_0 + \alpha_1 x_{it}^* + \alpha_2 z_{it} + \alpha_3 b_{it} + \sigma(x_{it}^*, z_{it}) u_{it}
\]

\[
y_{it} = y_{it}^* \cdot d_{it}
\]

\[
d_{it} = 1(\beta_0 + \beta_1 x_{it}^* + \beta_2 z_{it} + \beta_3 b_{it} \geq 0)
\]

\[
z_{it} = \gamma_0 + \gamma_1 x_{it-1}^* + \gamma_2 z_{it-1} + \gamma_3 y_{it-1} + \gamma_4 b_{it} + \sigma(x_{it-1}^*, z_{it-1}, y_{it-1}) v_{it}
\]

\[
z_{i1} = \tilde{\gamma}_0 + \tilde{\gamma}_1 x_{i1}^* + \tilde{\gamma}_2 b_{i1} + \tilde{\nu}_{i1}
\]

in which $y_{it}^*$ is a latent variable that is a function of variables $x_{it}^*$, $z_{it}$ and $b_{it}$, and $x_{it}^*$ is a latent variable, the observed counterpart of which, $x_{it}$, is measured with error. The equations below
describe the evolution of the latent variable.

\[ x_{it} = x^*_it + \epsilon_{it} \]
\[ x^*_it = \rho x^*_{i,t-1} + \nu_{it} \]

The following are further assumptions on the independence and the distribution of the error terms:

1. (independence of the error terms) \( x^*_i0 \perp \epsilon_{it} \perp \nu_{it} \)

2. (distributional assumptions) \( \epsilon_{it} \sim iidN(0, \sigma^2_\epsilon) \), \( \nu_{it} \sim iidN(0, \sigma^2_\nu) \) and \( x^*_i0 \sim iidN(0, \sigma^2_z) \).

To generate the data, I draw \( b_{it} \) from a Normal(0,1). \( u_{it} \) is also drawn from Normal(0,1). I consider a location-scale model; that is, \( \sigma(x^*_it, z_{it}) = 1 + \delta_1 x^*_it + \delta_2 z_{it} \). The parameter configurations that I consider are written in the table below.

Table D3: Parameter configurations, Monte Carlo simulations

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True value</th>
<th>Parameter</th>
<th>True value</th>
<th>Parameter</th>
<th>True value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha_0 )</td>
<td>1.25</td>
<td>( \gamma_0 )</td>
<td>1.0</td>
<td>( \sigma_\epsilon )</td>
<td>0.2</td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>0.5</td>
<td>( \gamma_1 )</td>
<td>0.5</td>
<td>( \sigma_\nu )</td>
<td>0.5</td>
</tr>
<tr>
<td>( \alpha_2 )</td>
<td>0.5</td>
<td>( \gamma_2 )</td>
<td>-0.2</td>
<td>( \sigma_z )</td>
<td>0.3</td>
</tr>
<tr>
<td>( \alpha_3 )</td>
<td>0.25</td>
<td>( \gamma_3 )</td>
<td>0.1</td>
<td>( \delta_1 )</td>
<td>0.1</td>
</tr>
<tr>
<td>( \beta_0 )</td>
<td>0.5</td>
<td>( \gamma_4 )</td>
<td>0.2</td>
<td>( \delta_2 )</td>
<td>0.1</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>0.25</td>
<td>( \gamma_{0} )</td>
<td>1.0</td>
<td>( \rho )</td>
<td>0.95</td>
</tr>
<tr>
<td>( \beta_2 )</td>
<td>0.25</td>
<td>( \gamma_{1} )</td>
<td>1.5</td>
<td>( \tilde{\gamma}_0 )</td>
<td>0.4</td>
</tr>
</tbody>
</table>

| \( N \) | 1000       | \( T \) | 6          |

Note: This table describes the parameter configurations of the small-scale simulation exercise that I perform to verify the finite-sample performance of the stochastic EM algorithm of the nonlinear reduced form model with the Buchinsky and Hahn (1998) censored quantile regression estimator.

I simulate 100 datasets in this simulation experiment. I take \( M = 1 \), stop the chain after a large number of iterations, and report an average across the last \( \hat{S} \) values \( \hat{\theta} = \frac{1}{\hat{S}} \sum_{s=\hat{S}-\hat{S}+1}^{\hat{S}} \hat{\theta}^{(s)} \), where I take \( \hat{S} = S/2 \). Each estimation is based on \( S = 100 \) iterations, with 200 random walk Metropolis-Hastings draws per iteration. The parameters of the data generating process are such that the percentage of censoring is around 22 percent.

The model specification is close to that outlined in the main text. As an example, the estimating equation for the latent outcome \( y^* \) takes the following form:

\[ y_{it}^* = g(x^*_it, x_{it}, z_{it}, b_{it}, \tau) \]
\[ = \sum_{k=0}^{K} a_k(\tau)g_k(x^*_it, x_{it}, z_{it}) + \gamma_y b_{it} \]

where \( \tau \in (0, 1) \), where \( g_k(\cdot) \) is a dictionary of functions, with \( g_0 = 1 \).
Table D4 presents the results of the simulation experiments. Again, the biases are moderate, and are larger at the extremes of the quantiles estimated. The biases are also larger for the mismeasured variable and the constant. Meanwhile, biases seem to be smaller for the correctly measured variables. Moreover, the estimation results are relatively precise outside of the tails, with smaller standard deviations. Comparing the results to those of the estimation procedure with the Arellano and Bonhomme (2017) quantile selection estimator, I find that the results are closer to the true values, but this is probably due to the fact that there are less number of estimating equations with the Buchinsky and Hahn (1998) estimator.

Table D4: Simulation results, Buchinsky and Hahn (1998) set-up

<table>
<thead>
<tr>
<th></th>
<th>Quantile</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.3</td>
<td>0.4</td>
<td>0.5</td>
<td>0.6</td>
<td>0.7</td>
<td>0.8</td>
<td>0.9</td>
</tr>
<tr>
<td>True value</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.726</td>
<td>0.997</td>
<td>1.503</td>
<td>1.774</td>
<td>2.092</td>
<td>2.532</td>
<td></td>
</tr>
<tr>
<td>$x_{it}$</td>
<td>0.448</td>
<td>0.475</td>
<td>0.500</td>
<td>0.525</td>
<td>0.552</td>
<td>0.584</td>
<td>0.628</td>
</tr>
<tr>
<td>$z_{it}$</td>
<td>0.448</td>
<td>0.475</td>
<td>0.500</td>
<td>0.525</td>
<td>0.552</td>
<td>0.584</td>
<td>0.628</td>
</tr>
<tr>
<td>$b_{it}$</td>
<td>0.250</td>
<td>0.250</td>
<td>0.250</td>
<td>0.250</td>
<td>0.250</td>
<td>0.250</td>
<td>0.250</td>
</tr>
</tbody>
</table>

Monte Carlo means

<table>
<thead>
<tr>
<th></th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.698</td>
<td>0.993</td>
<td>1.251</td>
<td>1.510</td>
<td>1.784</td>
<td>2.119</td>
<td>2.635</td>
</tr>
<tr>
<td>$x_{it}$</td>
<td>0.411</td>
<td>0.438</td>
<td>0.462</td>
<td>0.492</td>
<td>0.518</td>
<td>0.551</td>
<td>0.610</td>
</tr>
<tr>
<td>$z_{it}$</td>
<td>0.458</td>
<td>0.480</td>
<td>0.503</td>
<td>0.529</td>
<td>0.556</td>
<td>0.586</td>
<td>0.637</td>
</tr>
<tr>
<td>$b_{it}$</td>
<td>0.251</td>
<td>0.249</td>
<td>0.248</td>
<td>0.247</td>
<td>0.246</td>
<td>0.248</td>
<td>0.247</td>
</tr>
</tbody>
</table>

Monte Carlo standard deviations

<table>
<thead>
<tr>
<th></th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.087</td>
<td>0.072</td>
<td>0.071</td>
<td>0.073</td>
<td>0.086</td>
<td>0.110</td>
<td>0.147</td>
</tr>
<tr>
<td>$x_{it}$</td>
<td>0.037</td>
<td>0.036</td>
<td>0.036</td>
<td>0.039</td>
<td>0.041</td>
<td>0.047</td>
<td>0.091</td>
</tr>
<tr>
<td>$z_{it}$</td>
<td>0.027</td>
<td>0.026</td>
<td>0.023</td>
<td>0.022</td>
<td>0.023</td>
<td>0.025</td>
<td>0.041</td>
</tr>
<tr>
<td>$b_{it}$</td>
<td>0.017</td>
<td>0.017</td>
<td>0.016</td>
<td>0.016</td>
<td>0.017</td>
<td>0.016</td>
<td>0.020</td>
</tr>
</tbody>
</table>

Note: This table describes the result of the simulation experiment where the relevant estimator of the equations that correspond to the portfolio and participation rules is the Buchinsky and Hahn (1998) censored quantile regression estimator. This is an average across 100 datasets with $N = 1,000$ and $T = 6$. 
E Additional empirical evidence

Figure E1: Average derivative effect of the persistent component of income $v_{it}$

(a) Propensity score

(b) Risky asset share

Note: The graphs show average derivatives of the propensity score and the risky asset share of stock market participants, respectively, with respect to wealth $w_{it}$ given $w_{it}$, persistent component $v_{it}$, income $y_{it}$, and age $age_{it}$, evaluated at different values of $w_{it}$ and $age_{it}$ that correspond to their $\tau_{wealth}$ and $\tau_{age}$ percentiles. All results are based on estimates from the semi-structural model with the Buchinsky and Hahn (1998) censored quantile regression estimator.

Figure E2: Average derivative effect of wealth

(a) Propensity score

(b) Risky asset share

Note: The graphs show average derivatives of the portfolio rule with respect to wealth $w_{it}$ and persistent income $v_{it}$, respectively, given $w_{it}$, persistent component $v_{it}$, income $y_{it}$, and age $age_{it}$, evaluated at different values of $w_{it}$ and $age_{it}$ that correspond to their $\tau_{wealth}$ and $\tau_{age}$ percentiles. All results are based on estimates from the semi-structural model with the Buchinsky and Hahn (1998) censored quantile regression estimator.
Figure E3: Observed and implied densities of the risky asset share

Note: The graph shows the observed and predicted unconditional densities of the share of household wealth in risky asset share based on the nonlinear model. The blue line corresponds to the density implied by the nonlinear model, while the red line corresponds to the density implied by the data. All results are based on estimates from the semi-structural model with the Buchinsky and Hahn (1998) censored quantile regression estimator.

Figure E4: Impulse response, participation rule

Note: The graphs show the difference in average participation rates between a household hit by a shock $\tau_{\text{shock}}$ at age 37, and a household hit by 0.5 shock at the same age. The blue line corresponds to low-income households (i.e., rank of $\tau_{\text{init}} = 0, 1$ in the income distribution). The red line corresponds to middle-income households (i.e., rank of $\tau_{\text{init}} = 0.5$ in the income distribution). The green line corresponds to middle-income households (i.e., rank of $\tau_{\text{init}} = 0.9$ in the income distribution). All results are based on estimates from the semi-structural model with the Buchinsky and Hahn (1998) censored quantile regression estimator.
Figure E5: Impulse response, participation rule, by participation status

(a) $\tau_{init} = 0.9, \tau_{shock} = 0.1$

(b) $\tau_{init} = 0.1, \tau_{shock} = 0.9$

Note: The graphs show the difference in average participation rates between a household with rank $\tau_{init}$ hit by a shock $\tau_{shock}$ at age 37, and a household hit by 0.5 shock at the same age, conditional on participation status at age 35. The blue line corresponds to stock market participants. The red line corresponds to stock market non-participants. All results are based on estimates from the semi-structural model with the Buchinsky and Hahn (1998) censored quantile regression estimator.

Figure E6: Impulse response, portfolio rule

(a) $\tau_{shock} = 0.1$

(b) $\tau_{shock} = 0.9$

Note: The graphs show the difference in average portfolio shares conditional on participation between a household hit by a shock $\tau_{shock}$ at age 37, and a household hit by 0.5 shock at the same age. The blue line corresponds to low-income households (i.e., rank of $\tau_{init} = 0, 1$ in the income distribution). The red line corresponds to middle-income households (i.e., rank of $\tau_{init} = 0.5$ in the income distribution). The green line corresponds to middle-income households (i.e., rank of $\tau_{init} = 0.9$ in the income distribution). All results are based on estimates from the semi-structural model with the Buchinsky and Hahn (1998) censored quantile regression estimator.
Figure E7: Extensive and intensive margin responses to an income shock, by income and wealth at age 35

(a) \( \tau_{\text{shock}} = 0.1 \)

(b) \( \tau_{\text{shock}} = 0.9 \)

(c) \( \tau_{\text{shock}} = 0.1 \)

(d) \( \tau_{\text{shock}} = 0.9 \)

Note: The graphs show the difference between a household hit by a shock \( \tau_{\text{shock}} \) at age 37, and a household hit by a 0.5 shock at the same age, by income and wealth categories. The blue line corresponds to low income, low wealth households. The red line corresponds to low income, high wealth households. The green line corresponds to high income, low wealth households. The orange line corresponds to high income, high wealth households. All results are based on estimates from the semi-structural model with the Buchinsky and Hahn (1998) censored quantile regression estimator.
Figure E8: Extensive and intensive margin responses to an income shock, by income and wealth at age 51

(a) $\tau_{\text{shock}} = 0.1$

(b) $\tau_{\text{shock}} = 0.9$

(c) $\tau_{\text{shock}} = 0.1$

(d) $\tau_{\text{shock}} = 0.9$

Note: The graphs show the difference between a household hit by a shock $\tau_{\text{shock}}$ at age 53, and a household hit by a 0.5 shock at the same age, by income and wealth categories. The blue line corresponds to low income, low wealth households. The red line corresponds to low income, high wealth households. The green line corresponds to high income, low wealth households. The orange line corresponds to high income, high wealth households. All results are based on estimates from the semi-structural model with the Buchinsky and Hahn (1998) censored quantile regression estimator.