Household portfolio choices and nonlinear income risk

Julio Gálvez†
CEMFI
<galvez@cemfi.edu.es>

Work in progress. Please do not cite without permission.
This version: May 31, 2017.

Abstract

I revisit the role of uninsurable income risk on household stock market participation and portfolio choices. In particular, I consider the impact of earnings asymmetries and heterogeneity in income persistence, which are captured by more flexible representations of earnings dynamics. I motivate the empirical analysis by studying the implications of nonlinear income risk in a standard life-cycle model of portfolio choice. I then develop a flexible, semi-structural methodology to empirically quantify the transmission of income shocks to household stock market participation and portfolio choices. Participation and portfolio rules are specified as age-dependent nonlinear functions of wealth, and the persistent and transitory earnings components. I provide a tractable simulation-based estimation algorithm based on recent developments in the nonlinear panel data and the sample selection literature that allow the estimation of models with time-varying latent variables. Using the enhanced consumption, earnings and asset information in the 1999 to 2009 waves of the PSID, I find that differences in income uncertainty across households drive heterogeneous participation and portfolio allocation responses. Specifically, young low-income, low-wealth households, are the most (least) likely to participate when they receive a very positive (negative) income shock, with the probability increasing (decreasing) by as much as 10 percent. Similarly, market participants from these households increase (decrease) their portfolio allocation to stocks by as much as 7 percentage points when they receive a very positive (negative) income shock.

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

*I am extremely grateful to Manuel Arellano, Javier Mencia, and Enrique Sentana for their invaluable guidance and unwavering support. I also thank Dante Amengual, Stéphane Bonhomme, Julio A. Crego, Ivan Fernandez-Vál, Charles Gottlieb, Christian Julliard, Hamish Low, Michael Manove, Costas Meghir, Borja Petit, Josep Pijoan-Mas, 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.

†Address: CEMFI, Casado del Alisal 5, Madrid, 28014, Spain. Phone: +34661784067.
1 Introduction

Households utilize financial assets, such as stocks, to pay for current and future 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 is that on their labor income, which they can neither avoid nor insure themselves against. As households experience unique earnings histories, their investment decisions may differ depending on the size and persistence of the income shock they receive. In turn, households’ responses to these shocks are key components in the design of optimal social insurance and taxation policies. Understanding households’ financial portfolio choices in response to income risk is all the more important as the instruments that they hold constitute a sizeable portion of the financial industry relative to nonfinancial corporations and the government in many advanced economies.1 In this paper, I empirically assess the nonlinear impact of earnings shocks on household stock market participation and portfolio choices over the life-cycle by developing a new semi-structural framework.

An extensive literature in macroeconomics and finance has studied how uninsurable labor income shocks affects household consumption, saving, and portfolio allocation patterns over the life cycle, as well as its impact on asset prices.2 To ensure a reasonable standard of living, standard theory predicts that households respond by accumulating precautionary savings. Moreover, they reduce their exposure to avoidable risks. In the case of portfolio choice, they reduce 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, both theoretically and empirically. A common conclusion of both strands of literature is that the effect of labor income risk, while consistent with theory, is quantitatively small. Because of this, uninsurable income risk has seemingly lost its appeal as a candidate for explaining household stock market participation and portfolio choice decisions.

Yet recent important contributions to the earnings dynamics literature document that household labor income exhibits two substantial departures from features that characterize linear processes.3 First, household earnings display varying degrees of persistence that de-

1 As an example, according to the US Federal Reserve, total household wealth in the US as of 3rd quarter 2016 is around US$90.2 trillion, of which $494 million is held in stocks, either directly or indirectly.


3 Arellano et al. (2017) model the persistent component of income as a conditional quantile function of the past persistent component. They allow the persistence of income to depend on the size and the sign of the shock that households receive. The earnings process they utilize is able to uncover nonlinear persistence across households,
pend on the size of past and current earnings shocks. Second, household earnings distributions exhibit significant asymmetries. In contrast, linear processes imply that regardless of the households’ earnings histories, all shocks display the same persistence. Moreover, the distributions implied by these processes are Gaussian, which rules out dynamic skewness in earnings. Most importantly, linear models rule out nonlinear transmissions of income shocks that are likely to have a first-order effect on household portfolios.4

This paper re-examines the role of uninsurable income risk on household investment decisions by considering nonlinear earnings dynamics over the life cycle. I develop a novel, 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, or the extensive margin of stock market participation. Second, conditional on participation, I analyze households’ portfolio allocations, that is, the intensive margin. This modelling decision is justified by two robust empirical findings in the household finance literature: first, that stock market participation is limited at all ages, and second, that households adjust their financial portfolios as they age5.

To motivate my empirical study, I first analyze the possible implications of a more flexible specification of labor income dynamics6 in a standard model of household stock market participation and portfolio choice, both in a two-period (as in Campbell and Viceira (2002)) and in a life-cycle set-up (as in Cocco et al. (2005), Alan (2012), and Fagereng et al. (2017a), among other papers). In both contexts, I find that an earnings process with varying persistences and income asymmetries results in qualitatively different implications for households’ stock market participation and portfolio choice decisions, in comparison to a linear earnings model. The reason behind the results I uncover is precisely due to the nature of income uncertainty that a nonlinear earnings process is able to capture that a linear process cannot. Specifically, the varying degrees of risk that a household faces in terms of nonlinear persistence and the

---

4 As an example, Alan (2012) shows in her study of whether macroeconomic disasters can explain household stock market participation and portfolio choices that it is only through the inclusion of a probability of zero labor income that she is able to match the observed participation and portfolio allocations of young households in the US.

5 An excellent survey of the results in the household finance literature can be found in Guiso and Sodini (2013).

6 In particular, I present results using estimations of the Arellano et al. (2017) earnings process. The choice of this stochastic representation over those proposed by Guvenen et al. (2015) and De Nardi et al. (2016) is mainly motivated by the fact that this earnings process is relatively easy to incorporate into the simulations I perform and into the estimation framework that I propose in this paper.
possibility of very negative income realizations lead to (both current and future) labor income becoming more uncertain than what it would have been otherwise. Hence, the household becomes more disposed to avoid other risks by non-participation in the stock market. It would only be inclined to buy stocks if the expected returns would compensate for the potentially substantial risks it faces with its labor income.

I then specify a semi-reduced form representation of the dynamic structural 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 household portfolio choice model presents two complications that constitute the departure of the model that I advance from those considered previously. Specifically, the economic problem translates into a dynamic econometric model that presents self-selection into stock market participation, thereby yielding censoring in the dependent variable of interest, that is, the share of wealth invested in risky assets.

The empirical portfolio and participation rules are age-specific, nonlinear functions of the latent earnings components and assets. Moreover, the rules I estimate are nonparametric, which permits me to capture the extent by which nonlinear earnings dynamics influence the extensive margin of participation, and the intensive margin of portfolio allocation observed in the data. The econometric framework can also accommodate other components of the economic model, such as households’ wealth dynamics, structural state dependence in participation and time-invariant heterogeneity that captures unobserved preferences or discount rates. Most importantly, the methodology allows to reveal empirical nonlinearities in the data, and to calculate empirical objects of interest. In particular, I am able to calculate extensive and intensive margin responses to the impact of an earnings or a wealth shock. The flexibility of the approach I pursue makes the empirical objects I calculate consistent with several structural models of portfolio choice.

In order to estimate the model, I rely on a simulation-based algorithm that combines recent developments in quantile regression methods 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).\(^7\) The sequential estimation procedure alternates between two steps: first, simulation draws from the posterior distribution of the latent persistent in-

\(^7\) 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.
come components, and second, a sequence of likelihood maximization for the participation rule, and quantile regressions for the portfolio rule. An added advantage of my 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 more 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. A distinct advantage of the PSID, relative to recent empirical studies in the household portfolio choice literature have used detailed administrative data, is the possibility of exploiting information on earnings, consumption and asset holdings across the life-cycle for a representative sample of households. The descriptive statistics I calculate indicates that around 40 percent of households transition in between entries into and exits from the stock market. 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, or those who do not transition between entry and exit. Moreover, there is wide variation across households in their earnings, consumption and asset holdings. Preliminary results from my estimations show that nonlinearities in income shocks indeed result in heterogeneous participation and portfolio choice responses. In fact, for a very positive (negative) shock, a young, low-income, low-wealth household will be 10 (7.2) percent more (less) likely to invest in stocks. In comparison, an old, high-income, high-wealth household will not even adjust his holdings in response to 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-Jorgensen (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 novel empirical methodology that allows for the possibility of studying nonlinear relationships between income risk and household portfolio choices in a panel data setting.

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

---

8Alan (2012) finds that a positive probability of a zero income realization is needed in order to explain house-
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. Moreover, despite the fact that my framework is highly flexible, the results that I find have a natural connection to structural models of household portfolio choice considered by the theoretical literature, such as those considered by Viceira (2001), Cocco et al. (2005), Heaton and Lucas (2000a) and Polkovnichenko (2007), among others, as the dynamic system that I estimate emanates from, and is consistent with, a wide class of structural models.

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., Arellano et al. (2017) and Guvenen et al. (2015)) 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 nonlinear relationship between labor income risk, household stock market participation, and their subsequent portfolio choices.

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 wherein sample selectivity and censoring are important. While the empirical context I consider is one that is within the realm of household finance, the estimation procedure considered in this paper can be used to analyze other economic problems that exhibit similar features, such as models of 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)). As 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 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 depletion of savings after job loss. Around two years after unemployment, however, they begin to rebalance their portfolio toward risky assets. Finally, Knüfer 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.

Benzoni et al. (2007) studies the implications of uninsurable labor market income risk when stock market returns and labor income are cointegrated.
the estimation framework is fully flexible, the model I propose also allows for the possibility of incorporating structural state dependence in these intertemporal models, such as the one considered by Hyslop (1999) in the context of female labor market participation.

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 the PSID data. Finally, section 7 concludes.

2 What happens in an otherwise standard portfolio choice model?

To motivate my empirical investigation, I explore the possible implications of nonlinear income risk on household portfolios. I begin by providing an approximate analytical expression for optimal portfolio shares in a stylized two-period model that follows the set-up and the solution method in Campbell and Viceira (2002). Then, I calibrate a life-cycle model of portfolio choice to illustrate the differences between a standard linear earnings process, and a nonlinear earnings process, which is mainly based on Cocco et al. (2005).

I build on both set-ups, however, by introducing two ingredients. First, I introduce a fixed participation cost, which in the life-cycle model, is paid per period. Second, I relax the distributional assumption on labor income by considering a more general labor income process. I outline all derivations and specific details on the calibration exercise in Appendix A.

2.1 Stylized two-period model

I consider a theoretical framework where a household acts as a single agent. It is endowed with wealth $W_t$ and makes a financial portfolio decision at time $t$. The household consumes the liquidation value of the portfolio $W_{t+1}$, plus labor income $L_{t+1}$ one period later. Its objective is to maximize expected utility:

$$E_t \left( \delta^{1-\gamma} \frac{C_{t+1}}{1 - \gamma} \right)$$

10 Alan (2012) and Fagereng et al. (2017a) also calibrate a life cycle model that builds on the benchmark model of Cocco et al. (2005). A main difference between this paper and Alan (2012) is that she considers a one-time entry cost, while I consider a fixed, per-period participation cost. The aim of her paper is to empirically investigate whether macroeconomic disasters (represented by a disastrous stock return) can explain household portfolios in the US. Fagereng et al. (2017a), meanwhile, rationalizes life-cycle participation and portfolio choice patterns of Norwegian households through the “macroeconomic disasters” channel, but with a fixed per-period participation cost.
where $\delta$ is a discount factor, and $\gamma$ is the risk aversion parameter. I assume that utility is of the constant relative rate of risk aversion (CRRA) type.

There are two assets that are available for investment: a riskless asset and a risky asset. The riskless asset has a fixed simple return $R_f$, with $r_f \equiv \log(1 + R_f)$. The risky asset has a random return $R_t$, with a constant expected log excess return $E_t(R_{t+1} - r_f) \equiv \mu$. The unexpected log return on the risky asset, denoted by $u_{t+1}$, is conditionally Normal, with mean zero and variance $\sigma_u^2$. The return on the risky asset can be correlated with labor income, with $\text{cov}_t(l_{t+1}, r_{t+1}) \equiv \sigma_{lu}$.

To invest in the stock market, the household must pay a fixed cost of participation $F$. 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-Jorgensen (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)).

Finally, labor income $L_t$ is stochastic, and follows a distribution $H(L)$. The household cannot borrow against future labor income, thereby making it non-tradeable. In contrast to previous literature which assumes that labor income is lognormal, with mean $\mu_l$ and variance $\sigma_l^2$, I do not explicitly make a distributional assumption to highlight how different moments of labor income have an impact on household portfolio choices.

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, $C_{np,t+1}$, subject to the following budget constraint:

$$C_{np,t+1} = W_t(1 + R_f) + L_{t+1}. \quad (2)$$

The second corresponds to a situation in which the household participates in the stock market. In this problem, it solves for the optimal portfolio share, $\alpha_t$, subject to the following budget constraint:

$$C_{t+1} = (W_t - F)(1 + R_{p,t+1}) + L_{t+1} \quad (3)$$

where

$$R_{p,t+1} = \alpha_t(R_{t+1} - R_f) + R_f \quad (4)$$

Because 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 knows the realization of future stock market returns nor future labor income when it makes the portfolio choice decision. Hence, the household takes expectations over all future realizations of labor income and stock returns.
2.1.1 Optimal portfolio share and a participation condition

I employ log-linearization techniques to obtain approximate analytical expressions for optimal consumption and portfolio shares, and the participation rule.

Non-participation subproblem. As the household consumes all its resources in period $t+1$, I can log-linearize the budget constraint by dividing both sides of equation (2) by $L_{t+1}$. Taking logs and a first-order Taylor expansion around the mean of the (log) risk-free rate, $r_f$, and the (log) wealth-labor income ratio, $E\left[\log\left(\frac{W_t}{L_{t+1}}\right)\right] = w - l$, results in the following expression for optimal consumption:

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

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.

Participation subproblem. I write a log-linear approximation to the log portfolio return, equation (4), to facilitate the solution of the participation subproblem:

$$r_{p,t+1} = \alpha_t(r_t+1 - r_f) + r_f + \frac{1}{2}\alpha_t(1 - \alpha_t)\sigma_t^2.$$ 

To obtain a formula for optimal portfolio shares, I define the variable $W_{c,t} = W_t - F$, which is total household wealth net of participation costs. Taking logs of equation (3) and calculating a first-order Taylor expansion around the mean portfolio return $E[\log(1 + R_{p,t+1})] = R_{p,t+1}$ and the wealth-labor income ratio $E\left[\log\left(\frac{W_{c,t}}{L_{t+1}}\right)\right] = w_{c,t} - l$, the portfolio budget constraint, I obtain the following expression for optimal consumption:

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

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

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

11This simplification implies that for households who cannot pay the fixed cost, the optimal solution is $\alpha_t = 0$. Thus, the formulas that I derive later in this section apply to households who are able to pay this cost. This restriction is sensible, as households who are at the margin of participating in the stock market, are able to pay this cost. This simplification is not uncommon either. Alan (2012) normalizes wealth net of participation costs in her structural estimations. Brunnermeier and Nagel (2008) redefine consumption net of habits to obtain approximate analytical expressions in their empirical study of the impact of wealth on household portfolios.

8
The magnitude of $\phi_p$ relative to $\phi_{np}$ depends on the size of the wealth that it will accrue at $t+1$ in the participation subproblem compared to the wealth it will accrue in the nonparticipation subproblem.

Finally, a log-linearization of the Euler equation implied by the optimization problem yields the following expression for optimal portfolio shares:

**Proposition 2.1.** The optimal share of stocks out of total household wealth, net of participation costs, in the participation subproblem at time $t$ is to a log-linear approximation,

$$\alpha_t = \frac{1}{\phi_p} \left[ \frac{E_{t+1}(r_{t+1} - r_f) + \frac{1}{2}\sigma_u^2}{\gamma \sigma_u^2} \right] + \left(1 - \frac{1}{\phi_p}\right) \frac{\sigma_u}{\sigma_u^2}. \quad (7)$$

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.

Note that labor income $L_{t+1}$ and household wealth $W_{c,t}$ mainly affect the portfolio share through the wealth elasticity of consumption, $\phi_p$. Alternatively, I can express portfolio rule (7) as a function of expected future labor income $\tilde{H} \equiv E_t(L_{t+1})$ and wealth, $W_{c,t}$. Focusing on the first component, the optimal share takes the following expression:

**Corollary 2.1.1.** Expressed as a function of the ratio of human wealth to total household wealth, the optimal portfolio share (under idiosyncratic labor income risk) in the participation subproblem is:

$$\alpha_t = \left(1 + \frac{\tilde{H}}{W_{c,t}} \right) \left[ \frac{E_{t+1}(r_{t+1} - r_f) + \frac{1}{2}\sigma_u^2}{\gamma \sigma_u^2} \right]. \quad (8)$$

**Participation condition.** Finally, to solve the optimization problem, the household calculates the expected utility gained from non-participation and that gained from participation, and compares the result. 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 consumption at time $t+1$ when it invests in stocks should be at least the same as when it does not invest in stocks, which I formalize below.
Proposition 2.2. The household will participate in the stock market when

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

in which \( c_{p,t+1} \) is the optimal consumption under the participation subproblem, and \( c_{np,t+1} \) is the optimal consumption under the non-participation subproblem.

Thus, if it decides to participate in the stock market, the optimal portfolio rule is characterized by equation (7) and (8), and optimal consumption will then be equation (6). Otherwise, the optimal portfolio rule is characterized by \( \alpha_t = 0 \), and optimal consumption will then be equation (5).

2.1.2 Comparative statics

Equations (7), (8) and the participation condition (9) allow me to study the effect of increases in wealth and labor income, respectively.\(^\text{12}\)

First, keeping labor income constant, an exogenous increase in wealth reduces the portfolio share. The intuition is related to the wealth elasticity of consumption. An exogenous increase leads to total household wealth becoming a more important source to draw consumption from than labor income. Hence, it will not invest in stocks, and might prefer to save in riskless bonds, or to spend part of the wealth gain on goods.

Second, an increase in labor income has an ambiguous effect on household portfolio shares when I relax lognormality. This is because, to an approximation, labor income is a function of its variance, which measures risk, and skewness, which measures downside risk. This implies that even if the household receives an increase in labor income, its riskiness might induce the investor to be less bullish in buying stocks. In contrast, under lognormality, an increase in labor income will lead the household to allocate part of its wealth in stocks, holding all other things constant.

2.2 Life-cycle model

The previous subsection develops a standard two-period model to highlight the effects of human and financial wealth on stock market participation and portfolio choice decisions. As it is static in nature, however, the model does not provide insights on the dynamic effects of income shocks on stock market participation and portfolio choices.

\(^{12}\)The 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.
To this end, I extend the two-period model into a multi-period setting. In what follows, I consider a household that maximizes the expected utility of its consumption over a finite horizon.\textsuperscript{13} The household works up until retirement, and dies with certainty at the end of its life.\textsuperscript{14} I describe in detail the labor income processes, the formal statement of the intertemporal optimization problem, and the results of the model calibration in this subsection. Additional details related to the parameterization of the model, the solution method, and the discretization of the nonlinear earnings process, are outlined in Appendix A.2.

2.2.1 Household labor income

To describe its human capital, I specify the household’s labor income, and the associated earnings processes. Before retirement, its labor income can be expressed as:

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

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

At retirement, the household’s labor income can be expressed as:

\[ y_{it} = f(v_{it_r}) + \ln(\phi_{t_r}) + \varepsilon_{t_r}, \quad t > t_r. \]

In words, household labor income is a fraction of the persistent income that it earned in the last year of its working life.

2.2.2 Earnings dynamics: linear vs. nonlinear earnings processes

I consider two income processes: a linear earnings process, which is the standard assumption used in the literature, and a nonlinear earnings process, the implications of which I am interested in.

Linear earnings process. 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

\[ v_{it} = \rho v_{it-1} + \eta_{it} \]

\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. The model does not consider the participation decision of households, however.

\textsuperscript{14}Notice that I assume that there is no bequest motive in this paper. Though it might be likely that income asymmetries across households might affect this decision, this is likely to be of less importance. Moreover, in the benchmark model of Cocco et al. (2005), the main results do not assume a bequest motive.
where $\rho$ is a persistence parameter. The idiosyncratic part of the persistent component of income, $\eta_{it}$, is also i.i.d. Gaussian, with $\eta_{it} \sim N(0, \sigma_\eta^2)$. I assume that households have different initial conditions; that is, $v_{i0} \sim i.i.d.N(0, \sigma_v^2)$. Moreover, I assume that the initial condition, the idiosyncratic component of the persistent shock, and the transitory shock are independent of each other. Given these assumptions, I prove in Appendix A.2.1 that the parameters associated with this process are identified, and provide its estimates.

Nonlinear earnings process. Meanwhile, the nonlinear earnings process I consider is that of Arellano et al. (2017), who rely on a conditional quantile specification. In this earnings process, the transitory component $\varepsilon_{it}$ is assumed to have zero mean, is independent over time, and independent of all realizations of the persistent component. The persistent component $\nu_{it}$, on the other hand, is modelled as a general first-order Markov process. Denoting by $Q_t(\nu_{it}, \tau)$ the $\tau^{th}$ conditional quantile of $v_{it}$ given $v_{it-1}$ for each $\tau \in (0, 1)$, a representation of the persistent earnings process is:

$$v_{it} = Q_t(\nu_{it-1}, u_{it})$$

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. In Appendix A.2.2, I describe the full specification of the nonlinear earnings process that I estimate with my data, the estimation procedure, and a detailed discussion of the results I obtained.

Discussion. To illustrate the difference between the two earnings processes, I compare two key ingredients: the persistence of labor income, and the resulting densities of the persistent and transitory components.

Figure 1 illustrates the differences in persistence of the two earnings processes, which are calculated as a function of the past persistent component of income, and the shock that the household receives at the average age of the household in the sample (44.7 years). In the case of the linear earnings process, persistence is the same for all households, regardless of their realizations of $v_{it-1}$ and the shock that they receive, as shown in panel (a). In contrast, the persistence of $v_{it-1}$ in the nonlinear earnings process depends on the direction and magnitude of current and future earnings shocks. As panel (b) depicts, persistence is high for low earnings households who are hit by an extremely bad shock and high earnings households who are hit by an extremely good shock. In contrast, persistence is low for high earnings households who are hit by an extremely bad shock and high earnings households who are hit by an extremely good shock. In contrast, persistence is low for high earnings households who are hit by an extremely bad shock and high earnings households who are hit by an extremely good shock.
households who are hit by an extremely bad shock and low earnings households who are hit by an extremely good shock.

Figures 2 and 3, meanwhile, illustrate the estimated kernel densities of the persistent and transitory components of income as implied by the linear and the nonlinear earnings processes at mean age. As can be observed from the figures, the densities of the persistent and transitory components of the nonlinear earnings process show clear departures from Gaussianity. In particular, the densities present high kurtosis and fat tails. These features have been shown to present substantially different implications for consumptions and savings decisions. In what follows, I examine what their implications are for household portfolio choice decisions over the life cycle.

Figure 1: Difference in persistence, linear vs. nonlinear earnings process

Note: This figure illustrates the difference between the linear earnings process and the nonlinear earnings process in terms of their persistence. Panel (a) shows the persistence parameter of the linear earnings process. Panel (b) meanwhile, 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 $v_{i,t-1}$ with respect to $v_{i,t-1}$. In both cases, I calculate persistence at different percentiles of the past persistent component $\tau_{init}$ and the current earnings shock $\tau_{shock}$. Both results are estimated from my PSID subsample.
Figure 2: Densities of the persistent component of income

(a) Linear earnings process  
(b) Nonlinear earnings process

Note: This figure illustrates estimated kernel densities of the marginal distribution of the persistent component of income. Panel (a) provides the plot of the linear earnings process, while panel (b) depicts the conditional density of the nonlinear earnings process. Both densities were calculated using a Gaussian kernel, with the optimal bandwidth, and were estimated from my PSID subsample.

Figure 3: Densities of the transitory component of income

(a) Linear earnings process  
(b) Nonlinear earnings process

Note: This figure illustrates estimated kernel densities of the marginal distribution of the transitory component of income. Panel (a) provides the plot of the linear earnings process, while panel (b) depicts the conditional density of the nonlinear earnings process. Both densities were calculated using a Gaussian kernel, with the optimal bandwidth, and were estimated from my PSID subsample.

2.2.3 Intertemporal optimization problem

I make the same assumptions as in the stylized two-period model on the assets available for investment and their corresponding returns. That is, 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:

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

Note that the return on the stock is still potentially correlated with labor income, with

$$\text{cov}_t(r_{t+1}, y_{t+1}) = \sigma_{yu}.$$ 

The timing of the model is as follows: a household starts with a certain amount of wealth $w_t$. 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:

$$V_I^t(x_{it}, v_{it}) = \max_{c_{it}, \alpha_{it}} u(c_{it}) + \beta E_t(V_{t+1}(x_{it+1}, v_{it+1}))$$

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:

$$V_O^t(x_{it}, v_{it}) = \max_{c_{it}} u(c_{it}) + \beta E_t(V_{t+1}(x_{it+1}, v_{it+1}))$$

such that

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

The Bellman equation $V_I^t(x, v)$ corresponds to the indirect utility of a household of age $t$ who participates in the stock market, has a persistent income realization of $v$ and financial wealth $x$. Meanwhile, the Bellman equation $V_O^t(x, v)$ 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 $v$ and financial wealth $x$.

\textsuperscript{17}That is, the non-participation problem boils down to a consumption-savings problem, while the participation problem boils down to a portfolio choice problem with a fixed participation cost.
The Bellman equation $V_t(x, v)$ for the household pins down the participation decision:

$$V_t(x, v) = \max(V^I_t(x, v), V^O_t(x, v))$$

Finally a formalization of the participation decision is:

$$1'(x, v) = \begin{cases} 1, & \text{if } x \in X_P(v), \\ 0, & \text{otherwise} \end{cases}, \quad X_P(v) = \{ x \in \mathbb{R}^n : V^I_t(x, v) > V^O_t(x, v) \}$$

The details related to the parameterization of the structural model, and the discretization of the nonlinear earnings process that I use here, are in Appendix A.2.3.

### 2.2.4 Simulation results (in progress)

Figure 4 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 4: 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
(8) 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.\(^{18}\)

3 A flexible semi-structural approach

In the previous section, I showed 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 to characterize these rules.

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 of this paper outlines the nonlinear reduced form model. An important advantage of this approach is the calculation of objects of interest that reveal empirical nonlinearities. I describe the moments and other features that can be recovered from the model at hand in the second subsection. In the third part, I briefly discuss how these objects can be nonparametrically identified. Finally, I comment on several extensions of the baseline model that I discuss here.

\(^{18}\)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.
A more detailed discussion of the extensions can be found in Appendix B.

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 given by the following system of equations:

\begin{align*}
\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})) \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}) \\
m_{i0} &= \tilde{h}_{i0}(v_{i0}, \zeta_{i0})
\end{align*}

Equation (10) 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 a function 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.\(^{19}\) It is indexed by \( t \) since the household solves this problem at every age. \( u_{it} \) is an error term that captures unobserved characteristics, such as preference shifters, or discount rates. Equation (11), 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 (12). The arguments of this rule are the persistent and transitory components of income, wealth, and the fixed cost of participation \( q(Z_{it}) \), which is a function of characteristics \( Z_{it} \). In this model, \( Z_{it} = (X_{it}, p'_{it}) \), in which \( X_{it} \) are the same characteristics that I have 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 (13) characterizes the household’s wealth dynamics. It is a function of the previous period’s realization of latent earnings components, wealth, the risky asset share, and the current period’s socio-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.

\(^{19}\)In practice, \( X_{it} \) contains variables such as household size, education, and time effects.
Finally, equation (14) is the wealth of the household during the first period that I observe it. As can be seen, it is a function of the persistent component of income.

The dynamic econometric model represented by equations (10)-(14) is compatible with several classes of structural economic models, as it does not impose a specific functional form. Because of the model’s flexibility, it 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. One 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; more often than not, one would need to make specific functional form assumptions. 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.\textsuperscript{20}

\subsection*{3.2 Objects of interest}

The model described by equations (10)-(14) 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 $\nu_{it}$ in the following paragraphs. However, I also calculate these objects with respect to the transitory component of income $\varepsilon_{it}$ and wealth $w_{it}$.

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

\[ E \left( \frac{\partial \alpha_{it}}{\partial \nu_{it}} \bigg| \nu_{it} = \nu, \varepsilon_{it} = \varepsilon, w_{it} = w, X_{it} = x, d_{it} = 1 \right) = E \left( \frac{\partial g_{it}(\nu, \varepsilon, w, x, \tau)}{\partial \nu} \bigg| d_{it} = 1 \right) \tag{15} \]

and

\[ \frac{\partial \Pr(\tilde{g}_{it}(\nu_{it}, \varepsilon_{it}, w_{it}, q(Z_{it})) \leq v_{it} | \nu, \varepsilon, w, z)}{\partial \nu_{it}} = \frac{\partial \Pr(d_{it} = 1 | \nu, \varepsilon, w, z)}{\partial \nu_{it}} \tag{16} \]

Expressions (15) and (16) correspond to the marginal effects of an increase in the persistent component of income on portfolio shares and participation, respectively. Note that expression (15) considers the subset of households who participated in the stock markets, while expression (16) considers all households, regardless of participation status.

Second, the model estimates can be used to calculate intensive and extensive margins, and finally, an aggregate effect, which is similar to those calculated in the literature that estimates

\textsuperscript{20} As an example, one can think of equations (10)-(14) as the auxiliary equations to a structural model that I could estimate via indirect inference.
labor supply elasticities\(^{21}\). To see this more clearly, take the average of \(\alpha_{it}\):

\[
\mathbb{E}(\alpha_{it}|v, \varepsilon, w, x) = \mathbb{E}(\alpha_{it}|d_{it} = 1, v, \varepsilon, w, x) \cdot \Pr(d_{it} = 1|v, \varepsilon, w, z)
\]

(17)

Taking logs of both sides, the expression above is equivalent to:

\[
\log \mathbb{E}(\alpha_{it}) = \log \mathbb{E}(\alpha_{it}|d_{it} = 1, v, \varepsilon, w, x) + \log \Pr(d_{it} = 1|v, \varepsilon, w, z)
\]

Finally, taking partial derivatives of both sides with respect to \(\nu\), I have:

\[
\frac{\partial \log \mathbb{E}(\alpha_{it})}{\partial \nu} = \frac{\partial \log \mathbb{E}(\alpha_{it}|d_{it} = 1, v, \varepsilon, w, x)}{\partial \nu} + \frac{\partial \log \Pr(d_{it} = 1|v, \varepsilon, w, z)}{\partial \nu}
\]

(18)

The first component on the right-hand side can be thought of as the intensive margin of portfolio choice, which I denote as \(\theta_I\); like expression (15), the intensive margin is calculated for the subset of participants. The second component on the right-hand side, \(\theta_E\), is the equivalent of the extensive margin of participation. Finally, I can compute the aggregate elasticity, which is the sum of the two margins.

Third, and finally, I can compute “impulse-response” like functions for participation, the intensive margin of portfolio choice, and the aggregate effect\(^{22}\). To see this, consider equation (17), the average portfolio share function for a household with a given realisation of the income components, wealth, and a given demographic characteristic. Notice that each component on the right hand side can be thought of as a function. In this sense, I can calculate the effect of a shock to the persistent component of income, \(\nu'\). With respect to the conditional probability of participation, I can calculate the following function:

\[
\Delta_E(\nu + \tilde{\nu}, \nu) = \Pr(d_{it} = 1|\nu + \nu', \varepsilon, w, z) - \Pr(d_{it} = 1|\nu, \varepsilon, w, z)
\]

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

\[
\Delta_I(\nu + \tilde{\nu}, \nu) = \mathbb{E}(\alpha_{it}|d_{it} = 1, \nu + \nu', \varepsilon, w, x) - \mathbb{E}(\alpha_{it}|d_{it} = 1, \nu, \varepsilon, w, x)
\]

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

\[
\log \Delta_T(\nu + \nu', \nu) = \log \left[ \frac{\mathbb{E}(\alpha_{it}|d_{it} = 1, \nu + \nu', \varepsilon, w, x)}{\mathbb{E}(\alpha_{it}|d_{it} = 1, \nu, \varepsilon, w, x)} \right] + \log \left[ \frac{\Pr(d_{it} = 1|\nu + \nu', \varepsilon, w, z)}{\Pr(d_{it} = 1|\nu, \varepsilon, w, z)} \right]
\]

(19)

\(^{21}\)Logerson and Wallenius (2009) provide a tractable analytical foundation of the micro and macro elasticities of labor supply. Chetty (2012), meanwhile, derives bounds on labor supply elasticities.

\(^{22}\)To the best of my knowledge, the closest paper that has considered impulse response exercises is Calvet and Sodini (2014), who look at aggregate implications of a wealth shock on aggregate portfolio shares. However, they only consider the subset of market participants.
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

In general, the nonlinear reduced form participation and portfolio rules I present here are not nonparametrically point identified (Matzkin (2013)). While the statistical representation I present here attributes separate errors for the participation and portfolio rules, the economic model suggests that the same unobserved error drives both decisions simultaneously. One could think, however, that the errors $u_{it}$ and $v_{it}$ in models (10) and (12) come from a vector-valued error term $u_{it}$.

The participation and portfolio rules I present here can be thought of as nonlinear state space models. Hence, it is possible to use techniques developed in the literature (see Hu (2015) for an excellent survey) to establish conditions that guarantee nonparametric identification of the joint dynamic distributions of the observed and latent variables. This permits the identification of the average derivative effects (15) and (16) in the absence of restrictions on the econometric model. Under some conditions, it is also possible to nonparametrically identify the participation and portfolio rules.

3.4 Model extensions

One virtue of the estimation framework I propose is its flexibility in incorporating some extensions that are empirically and economically relevant. In particular, I explicitly 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. I briefly comment on each extension in turn.

State dependence in participation. The baseline model that I outlined does not allow for potential state dependence in participation. Introducing this via a lagged participation indicator in equation (12) 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. Operationalizing this, however, requires me to model the initial participation and portfolio choice decisions of these households.

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 (10)-(14).\(^{23}\) 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. This effectively implies that I take a random effects approach in the empirical model specification.

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

The main dataset that I use for my empirical analysis is a balanced panel of households from the 1999 to 2009 waves of the PSID, which correspond to calendar years 1998 to 2008. Since the 1999 wave, the PSID began collection of detailed data on consumption expenditures and asset holdings, in addition to information on household earnings and demographics. In the following paragraphs, I describe the definition of the variables that I use. I deflate all main variables that I use in 1998 US dollars (USD).

Variable definitions. Household income $Y_{it}$ is the sum of earnings of all household members who are in the labor force, including household transfers. In the estimations, I construct $y_{it}$ as the residuals from regressing log household earnings on a set of demographics, which include education dummies, household size, number of children, race, and cohort dummies.

I use the following definition of total household wealth $W_{it}$. Financial wealth is the sum

\(^{23}\)Some 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. A crucial advantage of the approach I take here is that my framework can simultaneously deal with the latent earnings components, sample selection and unobserved heterogeneity.
of three main sources: cash, which is defined as checking or savings accounts, money market accounts, or Treasury bills, including those held in individual retirement accounts (IRAs); stocks, which comprises the shares of stock in publicly traded corporations, mutual funds, or investment in trusts, including those held in IRAs; and bonds, which includes bonds, the cash value in life insurance policies, valuable collections, rights in trusts or estates. Net wealth is the sum of financial wealth, business value, and home equity, net of non-mortgage debt such as credit card and consumer loans. All of the estimations that I present use the log of total household wealth, $w_{it}$, as the relevant independent variable. I then calculate the risky share $\alpha_{it}$ as the sum of stocks and mutual funds held, divided by net wealth.

Sample selection criteria. I select married households who are between 25 to 60 years old, and who are not part of the SEO subsample. Following Calvet and Sodini (2014), I select households who have at least USD 100 of income, and USD 300 of financial wealth. Because some households in my sample have transfer incomes that are greater than twice their labor income, I remove them from my sample as well, as in Blundell et al. (2016). I exclude households who have inconsistencies in the information they provide: specifically, I remove those who have missing, and/or incomplete information on socio-demographic characteristics (age, race and education), or on key variables (amount of stockholding, financial wealth, consumption, home and mortgage value, business value and household income). Overall, this sample selection criteria leaves me with 608 households who I observe for six years. A table that details the sample selection criteria is in the data appendix.

4.2 Descriptive evidence

Table 1 presents pooled cross-section/time-series summary statistics for all relevant variables, grouped by income quartiles. The table is divided into two panels. The first panel corresponds to all of the households who satisfy the criteria to be included in the sample. The sec-

---

24In order to identify the part of the IRA that is allocated between stocks and bonds, I follow the treatment of Vissing-Jorgensen (2002) and Malmendier and Nagel (2011). Details are in the appendix.

25The question in the PSID asks the household to respond to the following question if they own any of the said assets. The next question asked is how much is left if they have sold and paid off debt related to it.

26Albeit 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. 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.

27I relegate all results related to financial wealth in the appendix.

28To determine whether my results are affected by the sample I selected, I repeat all of the descriptive analysis with the sample selection criteria of Fagereng et al. (2017a). The results are quite similar to what I present here. Results are available upon request.
ond panel, meanwhile, corresponds to the subset of stock market participants, that is, those households who have stock holdings that are greater than zero.

I observe a wide dispersion across households in their earnings, consumption and assets, for both the entire sample, and the subsample of stock market participants. The average participation rate in stock markets is around 64.5 percent in my sample. The table also shows that households’ wealth holdings are not monotonically increasing across income quartiles. In fact, households at the lowest income quartile have higher risky asset holdings and financial wealth than those in the middle income quartiles.

Table 1: Summary statistics, by income quartiles

<table>
<thead>
<tr>
<th>Variable</th>
<th>TOTAL</th>
<th>First</th>
<th>Second</th>
<th>Third</th>
<th>Fourth</th>
<th>Stock market participants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household income</td>
<td>103,844.60</td>
<td>44,892.36</td>
<td>71,282.79</td>
<td>95,775.79</td>
<td>203,603.60</td>
<td>116,047.10</td>
</tr>
<tr>
<td>Total household wealth</td>
<td>413,411.70</td>
<td>276,194.80</td>
<td>255,805.90</td>
<td>291,037.80</td>
<td>831,155.90</td>
<td>548,585.00</td>
</tr>
<tr>
<td>Financial wealth</td>
<td>190,287.10</td>
<td>130,675.70</td>
<td>127,730.70</td>
<td>111,077.80</td>
<td>391,965.80</td>
<td>275,417.30</td>
</tr>
<tr>
<td>Consumption</td>
<td>50,105.66</td>
<td>34,529.70</td>
<td>41,250.87</td>
<td>50,865.62</td>
<td>73,816.34</td>
<td>124,924.10</td>
</tr>
<tr>
<td>Home equity</td>
<td>124,924.10</td>
<td>73,969.33</td>
<td>86,978.80</td>
<td>126,823.10</td>
<td>212,047.30</td>
<td>148,923.50</td>
</tr>
<tr>
<td>Risky wealth ownership</td>
<td>0.645</td>
<td>0.264</td>
<td>0.306</td>
<td>0.358</td>
<td>0.460</td>
<td>0.645</td>
</tr>
<tr>
<td>Risky wealth</td>
<td>123,883.80</td>
<td>88,316.97</td>
<td>81,592.38</td>
<td>57,456.24</td>
<td>268,386.00</td>
<td>123,883.80</td>
</tr>
<tr>
<td>Risky asset share, net wealth</td>
<td>0.158</td>
<td>0.120</td>
<td>0.130</td>
<td>0.159</td>
<td>0.225</td>
<td>0.158</td>
</tr>
<tr>
<td>Risky asset share, financial</td>
<td>0.347</td>
<td>0.264</td>
<td>0.306</td>
<td>0.358</td>
<td>0.460</td>
<td>0.347</td>
</tr>
</tbody>
</table>

Note: This paper presents summary statistics 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 across income quartiles. The sixth column presents the F statistic calculated from a multivariate test for equality across groups. The first panel presents results across all households, while the second panel presents results for stock market participants.

This finding is more apparent once I condition on the subset of stock market participants. Looking at the risky asset shares, I find that it is not monotonically increasing with income quartiles. One might be worried that these sample averages could be the same across in-

---

29 Comparing this number to the relevant participation rate in the US SCF, I have a higher proportion of households in my sample who directly or indirectly hold stocks. As an example, the 2007 SCF reports that 53.2 percent of all US households own stocks, directly or indirectly.

30 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 explain this phenomenon. In the appendix corresponding to this section, I calculate the same summary statistics, but across wealth and age quartiles, and find that I do not
come quartiles. To verify whether the means across different income groups are different, I conduct a series of multivariate tests. The last column displays the $F$ test statistic that results from the tests I performed. I find that, indeed, for all economic variables I consider, the means across groups are different.

**Table 2: Frequency distribution of household stock market participation sequences**

<table>
<thead>
<tr>
<th>Age Quartile</th>
<th>First</th>
<th>Second</th>
<th>Third</th>
<th>Fourth</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never bought stocks</td>
<td>4</td>
<td>12</td>
<td>14</td>
<td>23</td>
<td>53</td>
</tr>
<tr>
<td>Always participated</td>
<td>18</td>
<td>37</td>
<td>49</td>
<td>112</td>
<td>216</td>
</tr>
<tr>
<td>Pure entry households</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0,1,1,1,1,1)</td>
<td>6</td>
<td>0</td>
<td>6</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>(0,0,1,1,1,1)</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>(0,0,0,1,1,1)</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>(0,0,0,0,1,1)</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>(0,0,0,0,0,1)</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Pure exit households</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1,0,0,0,0,0)</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>(1,1,0,0,0,0)</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>(1,1,1,0,0,0)</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>(1,1,1,1,0,0)</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>(1,1,1,1,1,0)</td>
<td>4</td>
<td>3</td>
<td>7</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Households who transition</td>
<td>30</td>
<td>49</td>
<td>53</td>
<td>107</td>
<td>239</td>
</tr>
<tr>
<td>TOTAL</td>
<td>92</td>
<td>128</td>
<td>164</td>
<td>324</td>
<td>608</td>
</tr>
</tbody>
</table>

Note: 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 table. Households who have transitions are those who enter or exit more than once.

Table 2 presents the frequency distribution of household stock market participation sequences disaggregated by age quartiles. I distinguish between the following groups of households: (1.) those who have never participated in the stock market, that is, those who have participation sequences of $(0,0,0,0,0,0)$; (2.) those who have always participated in the stock market, that is, those who have participation sequences of $(1,1,1,1,1,1)$; (3.) households who have “pure entries” or “pure exits”, that is, those households who, once they have decided reproduce the non-monotonic pattern I find here.
to enter into the stock market, stay (or those who, once they have decided to exit, leave forever); and (4.) households who transition from one participation state to another.\textsuperscript{31} I find that, in the sample that I consider, households who transition in and out of the stock market comprise the largest group in my sample, at 39.31 percent. This is followed by households who have always participated in the stock markets, who are approximately 35.52 percent of the subsample. Households who have never entered the stock market, and those who have “pure entries” and “pure exits” comprise the rest of the households in the sample.

Table 3: Summary statistics, by stock market participation status

<table>
<thead>
<tr>
<th>Variable</th>
<th>Never</th>
<th>Participation sequence status</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Age of household head</td>
<td>44.31</td>
<td>45.73</td>
<td>44.27</td>
<td>42.80</td>
<td>44.72</td>
</tr>
<tr>
<td>Education</td>
<td>12.35</td>
<td>15.01</td>
<td>15.02</td>
<td>13.79</td>
<td>14.05</td>
</tr>
<tr>
<td>Household income</td>
<td>65,741.76</td>
<td>131,070.60</td>
<td>87,704.88</td>
<td>148,837.20</td>
<td>83,954.13</td>
</tr>
<tr>
<td>Financial wealth</td>
<td>15,538.20</td>
<td>406,528.40</td>
<td>63,832.13</td>
<td>128,084.50</td>
<td>75,495.18</td>
</tr>
<tr>
<td>Net wealth</td>
<td>111,060.20</td>
<td>726,054.80</td>
<td>240,346.50</td>
<td>471,120.60</td>
<td>230,727.70</td>
</tr>
<tr>
<td>Risky wealth</td>
<td>-</td>
<td>306,104.20</td>
<td>28,282.94</td>
<td>49,454.01</td>
<td>23,039.84</td>
</tr>
<tr>
<td>Consumption</td>
<td>36,618.48</td>
<td>60,117.17</td>
<td>46,394.55</td>
<td>52,643.92</td>
<td>44,595.69</td>
</tr>
<tr>
<td>Ownership of risky assets</td>
<td>-</td>
<td>1.00</td>
<td>0.59</td>
<td>0.63</td>
<td>0.48</td>
</tr>
<tr>
<td>Risky asset share, net wealth</td>
<td>-</td>
<td>0.30</td>
<td>0.12</td>
<td>0.12</td>
<td>0.08</td>
</tr>
<tr>
<td>Risky asset share, financial wealth</td>
<td>-</td>
<td>0.30</td>
<td>0.29</td>
<td>0.30</td>
<td>0.21</td>
</tr>
<tr>
<td>Home equity</td>
<td>69,689.17</td>
<td>174,234.80</td>
<td>129,495.60</td>
<td>124,414.30</td>
<td>91,566.55</td>
</tr>
</tbody>
</table>

Note: 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.

Table 3 presents summary statistics according to stock market participation status. I find that compared to households who have at least participated once in the stock market, those who have never participated tend to be less educated, have lower household incomes, and have lower household wealth, regardless of the definition of wealth I consider. Once I compare those households who have had at least participated once in the stock market, I find that, on average, households who transition in and out of the stock market are those who have lower household incomes, lower household net wealth, lower risky household wealth, and lower risky asset shares.\textsuperscript{32} Arguably, this evidence suggests that these households who are at the margin are those who face substantial risks to their labor income.

\textsuperscript{31}As a specific example, these are households who have participation sequences of \((0,1,1,0,1,0)\), \((1,1,0,1,0,1)\), etc. Hyslop (1999) calculates the same participation sequences, but within the context of structural state dependence in female labor force participation.

\textsuperscript{32}Again, I test for equality of means across these groups, and I find that these means are different.
5 Empirical 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. Armed with these results, I am able to specify each component of the system of equations, 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).

There are three 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 households select themselves into stock market participation. Third, some households might decide on solutions that are at the corner of the feasible set, which suggests the presence of censoring. To this end, I propose an estimation procedure based on recent developments in the panel data and the sample selection literature.

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 (10). Letting \( \text{age}_{it} \) denote the age of the household head \( i \) at period \( t \), I specify the portfolio rule as:

\[
\alpha^*_{it} = g_{it}(v_{it}, \epsilon_{it}, w_{it}, X_{it}, \tau) = g(v_{it}, \epsilon_{it}, w_{it}, X_{it}, \text{age}_{it}, \tau) = \sum_{k=0}^{K} b^k(\tau) \varphi_k(v_{it}, \epsilon_{it}, w_{it}, \text{age}_{it}) + \gamma^a(\tau)'X_{it}
\]

(20)

In practice, \( \varphi_k(\cdot) \) is a product of Hermite polynomials. The function depends on different quantiles of the distribution of risky shares, which implies that I consider a series quantile model. This is mainly due to robustness to distributional assumptions; as I show in the results, the conditional distribution of risky asset shares is not Gaussian. The flexibility I introduce here does not imply that estimating this becomes computationally cumbersome.\(^{33}\)

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 the preference

\(^{33}\)In fact, I take advantage of algorithms for fast computation. For example, the interior point algorithm for nonlinear quantile regression described in Koenker (2005) can be downloaded from http://www.econ.uiuc.edu/roger/research/rq/rq.m.
shifters/life-cycle controls, which in my model, are education, household size, and the number of children. I also include time dummies to control for aggregate effects. 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 do allow for flexibility, however, by allowing the coefficients $\gamma^a(\tau)$’s to differ across quantiles.

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. Hence, the empirical portfolio model that I take to the data is:

$$\Lambda^{-1}(\alpha^*_it) \equiv \log \left( \frac{\alpha^*_it}{1 - \alpha^*_it} \right) = \sum_{k=0}^{K} b_{k}^p(\tau) \phi_k(v_{it}, \epsilon_{it}, \omega_{it}, a_{ge_{it}}) + \gamma^p(\tau)'X_{it}$$

(21)

Participation rule and the exclusion restriction. I specify the participation rule given current earnings components, assets, age and life-cycle controls as follows:

$$\Pr(d_{it} = 1|v_{it}, \epsilon_{it}, \omega_{it}, a_{ge_{it}}, Z_{it}) = \Lambda \left( \sum_{k=0}^{K} b_{k}^p \phi_k(v_{it}, \epsilon_{it}, \omega_{it}, a_{ge_{it}}) + \gamma^p Z_{it} \right)$$

(22)

in which $\Lambda(\cdot)$ is the logistic function. Equation (22) corresponds to a sieve logit specification. Other specifications could be entertained for the estimation of the participation rule, such as a sieve probit. An advantage of the logit specification however, is that once I move to the model with state dependence, it is easier to calculate entry and exit rates.

Notice that the model relies on an exclusion restriction. In this regard, I consider the lagged value of lifetime wealth, following Vissing-Jorgensen (2002), Bonaparte et al. (2014), and Fagereng et al. (2017a). The motivation behind this can be seen in the alternative characterization of the portfolio rule. As shown in equation (8), the optimal portfolio rule is a function of the ratio between human and total household wealth, and not on the level of lifetime wealth. I explain how to calculate this variable in the data appendix. As my exclusion restriction is lagged by one period, however, I specify equations that separately account for participation and portfolio choices during the initial period that I observe these households in order not to lose information.

---

34 As is well known, quantiles are invariant to monotonic transformations (Koenker and Bassett (1978)). In fact, 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.

35 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).

36 In the extension of the paper that considers state dependence, I aim to utilize the lagged participation indicator.
Evolution of wealth. I specify the distribution of wealth $w_{i1}$ conditional on the persistent component $\nu_i$, age at the start of the period $ag_{i1}$ and life-cycle controls during the initial period when I observe them as follows:

$$Q_w(\nu_{i1}, \text{ag}_{i1}) = \sum_{k=0}^{K} b^w_k(\tau) \tilde{\varphi}_k(\nu_{i1}, \text{ag}_{i1}) + \gamma^w(\tau)/X_{i1} \hspace{1cm} (23)$$

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

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

$$w_{it} = h_t(\nu_{it-1}, \epsilon_{it-1}, w_{it-1}, \alpha_{it-1}, X_{it}, \tau)$$

$$= h(\nu_{it-1}, \epsilon_{it-1}, w_{it-1}, \alpha_{it-1}, X_{it}, \text{ag}_{it}, \tau)$$

$$= \sum_{k=1}^{K} b^m_k \tilde{\varphi}_k(\nu_{it-1}, \epsilon_{it-1}, w_{it-1}, \alpha_{it-1}, \text{ag}_{it}) + \gamma^m X_{it} + b^m_0(\tau) \hspace{1cm} (24)$$

for some $K$ and $\tilde{\varphi}_k$. In contrast with (23), I specify equation (24) 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; however, to be parsimonious with the estimation procedure I consider, I resort to this model specification.

Implementation. The functions $b^a_k, \gamma^a(\tau), b^w_0, b^w_k$ and $\gamma^w(\tau)$ are indexed by a finite dimensional parameter vector $\mu$, which also contains the coefficients $b^p_k's, b^w_k's, \gamma^p's$, and $\gamma^w's$. I model the functions $b^a_k$ 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 $b^a_0$ on $(0, \tau_1]$ and $[\tau_L, 1)$ using a parametric model that is indexed by $\lambda^a$. All other $b^a_k$ for $k \geq 1$ are constant on these two intervals. Thus, denoting $b^a_{kl} = b^a_k(\tau_l)$, the functions $b^a_k$ depend on $\{b^a_{11}, \ldots, b^a_{KL}, \lambda^a\}$.

In the estimation of the portfolio rule, I define a grid from $\tau_1 = 0.40$ to $\tau_L = 0.90$, with step size equal to 0.05. This is due to the degree of censoring found in the data. The functions $b^a_k$ are taken as piecewise-linear, which allows the likelihood to be specified in closed form. In addition, the function $a^Q_0$ is specified as the quantile of an exponential distribution on $[0, \tau_1)$ as another exclusion restriction, similar to Vissing-Jorgensen (2002) and Bonaparte et al. (2014). The rationale behind this is that the lagged indicator can represent non-monetary costs of participation in the stock market, such as the time it takes to learn about investments. In contrast to the baseline model, where the exclusion restriction enters linearly, I allow for interactions between past participation and the other state variables. This is crucial as I would like to study stock market entry and exit dynamics.

Indeed, a Jarque-Bera test on (log) wealth reveals that it is in fact, non-normal.
(with parameter $\lambda_-$) and $(\tau_L, 1)$ (with parameter $\lambda_+$).\footnote{More specifically, \[ a_0^Q = \frac{1}{\lambda^Q} \log \left( \frac{\tau}{\tau_{1}} \right) \mathbb{1}\{0 < \tau < \tau_{1}\} + \sum_{l=1}^{L} \left( a_{l+1}^Q - a_{l}^Q \right) \left( \frac{\tau_{l+1} - \tau_{l}}{\tau_{l+1} - \tau_{l}} \right) \mathbb{1}\{\tau_{l} \leq \tau < \tau_{l+1}\} \right) - \frac{1}{\lambda^Q} \log \left( \frac{1 - \tau}{1 - \tau_{L}} \right) \mathbb{1}\{\tau_{L} \geq \tau < 1\} \]}

Meanwhile, I define $\tau_{1} = 1/L$ and $L = 10$ for the functions that correspond to the initial wealth distribution. I set the wealth accumulation functions $b_i^\theta$ equal to $\alpha + \sigma \Phi^{-1}(\tau)$, where $(\alpha, \sigma)$ are parameters to be estimated. Finally, I use tensor products of Hermite polynomials for $\varphi_k$ and $\tilde{\varphi}_k$, 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.\footnote{As an example, the portfolio rule arguments are $(w_u - \text{mean}(w))/\text{std}(w)$, $(v_u - \text{mean}(v))/\text{std}(v)$, $(\epsilon_u - \text{mean}(\epsilon))/\text{std}(\epsilon)$ and $(\text{age}_u - \text{mean}(\text{age}))/\text{std}(\text{age})$.}

5.2 Overview of the estimation algorithm

The algorithm is an adaptation of the techniques in Arellano et al. (2017) to a setting with time-varying latent variables, sample selection, and/or censoring in the dependent variable. I describe the estimation of the portfolio rule (21) and the participation rule (22). Specific details on the algorithm, which include the likelihood function and the model’s restrictions, are described in Appendix C.

A compact notation for the restrictions implied by the portfolio and participation rules is the following:

$$
\hat{\mu} = \arg\min_{\mu} \mathbb{E} \left[ - \int O(v^T_i, \xi_i^T, w^T_i, \mathbf{Z}_i; \mu) f_i(v_i^T | \hat{\mu}) dv_i^T \right] - \int P(v^T_i, \xi_i^T, w^T_i, \mathbf{Z}_i; \mu) f_i(v_i^T | \hat{\mu}) dv_i^T, \right)
$$

(25)

where $O$ and $P$ are known functions and $\hat{\mu}$ is the true value of $\mu$.

The estimation algorithm is a stochastic EM-like algorithm, which is a simulated version of the classical EM algorithm of Dempster et al. (1977). To estimate the parameters $\mu = (\mu_O, \mu_P)$, I consider a sequence of likelihood maximization and quantile regressions. Starting with a parameter vector $\hat{\mu}^{(0)}$, I iterate the following steps on $s = 0, 1, \ldots$ until convergence of the $\hat{\mu}^{(s)}$ process:

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

2. **M-step:** Compute

$$
\hat{\mu}^{(s+1)} = \arg\max_{\mu} \sum_{i=1}^{N} \sum_{m=1}^{M} O(v_{i}^{(m)}; \xi_i^T, w_i^T, \mathbf{Z}_i; \mu),
$$
and

\[
\mu_p^{(s+1)} = \arg \min_{\mu_p} \sum_{i=1}^{N} \sum_{m=1}^{M} P(v_i^{(m)} e_i^T, w_i^T, Z_i; \mu),
\]

Because the likelihood function can be expressed in closed form, the E step is straightforward. In practice, I use a random-walk Metropolis Hastings sampler that targets an acceptance rate of approximately 30 percent. The M step I consider is the three-step estimator for the quantile selection model proposed by Arellano and Bonhomme (2017), which is characterized by the following steps.

First, I estimate the participation rule (22):

\[
\max_{(b_1^p,...,b_K^p)} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{m=1}^{M} d_{it} \log \left[ \Lambda \left( \sum_{k=0}^{K} b_k^p \varphi_k(v_{it}, \epsilon_{it}, w_{it}, age_{it}) + \gamma^p Z_{it} \right) \right] + (1 - d_{it}) \log \left[ 1 - \Lambda \left( \sum_{k=0}^{K} b_k^p \varphi_k(v_{it}, \epsilon_{it}, w_{it}, age_{it}) + \gamma^p Z_{it} \right) \right].
\]

(26)

I can then recover 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 correlation parameter between the error terms of participation and portfolio choice, \( \rho_c \), by considering the following objective function:

\[
\rho_c = \arg \min_{c \in \mathcal{C}} \left\| \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{m=1}^{M} d_{it} Y(\tau_l, x_{it}) \left[ 1 \left\{ \Lambda^{-1}(a_{it}^+) \leq \sum_{k=0}^{K} b_k^p(\tau) \varphi_k(v_{it}, \epsilon_{it}, w_{it}, age_{it}) \right\} - G(\tau_l, p(x_{it}); c) \right] \right\|
\]

(27)

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

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

is the conditional copula that links the participation and the asset allocation rules, \( Y(\cdot) \) are instrument functions, and

\[
b_k^p(c) = \arg \min_{b \in \mathcal{B}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{m=1}^{M} d_{it} \left[ G(\tau_l, p(x_{it}); c) \left( \Lambda^{-1}(a_{it}^+) - \sum_{k=0}^{K} b_k^p(\tau) \varphi_k(v_{it}, \epsilon_{it}, w_{it}, age_{it}) \right) \right] + (1 - G(\tau_l, p(x_{it}); c)) \left( \Lambda^{-1}(a_{it}^+) - \sum_{k=0}^{K} b_k^p(\tau) \varphi_k(v_{it}, \epsilon_{it}, w_{it}, age_{it}) \right) \]

(28)

Finally, given the propensity score, and the correlation parameter, I can calculate the following
rotated quantile regression:

$$\min_{(b_0, \ldots, b_K)} \sum_{i=1}^{N} \sum_{t=2}^{T} \sum_{m=1}^{M} d_{it} \left[ G(\tau_l, \hat{p}(x_{it}); \hat{\rho}_e) \left( \Lambda^{-1}(\alpha_{it}^* K) - \sum_{k=0}^{K} b_k^d(\tau) \varphi_k(v_{it}, \epsilon_{it}, w_{it}, a_{ge_{it}}) + \gamma^d(\tau)^{X_{it}} \right) \right]^{+}$$

\[+ (1 - G(\tau_l, \hat{p}(x_{it}); \hat{\rho}_e))) \left( \Lambda^{-1}(\alpha_{it}^*) - \sum_{k=0}^{K} b_k^d(\tau) \varphi_k(v_{it}, \epsilon_{it}, w_{it}, a_{ge_{it}}) + \gamma^d(\tau)^{X_{it}} \right)^{-} \tag{29} \]

This yields a convex optimization problem, allowing for computationally fast procedures. I proceed in a similar fashion to update the parameters of the other equations in the model, which I outline in the appendix.

Similar to the exclusion restriction, the conditional copula \( G(\tau_l, p(x_{it}); \rho_c) \) plays an important role in the nonparametric identification and estimation of the functions and the objects of interest. Specifically, the copula maps the ranks from 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 the portfolio shares of the participation subproblem, while the observed outcomes are the observed portfolio shares. I specify a Gaussian copula.

To implement the algorithm, I take \( M = 1 \), stop the chain after a large number of iterations, and report an average across the last \( S \) values \( \hat{\mu} = \frac{1}{S} \sum_{s=S}^{S+\tilde{S}+1} \hat{\mu}^{(s)} \), where I take \( \tilde{S} = S/2 \). 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, both for the baseline model and the model with structural state dependence in participation, are outlined in Appendix D.2 of the paper.

**Censoring corrections.** A potential concern is that my exclusion restriction might be invalid. In other words, it could be the case that the variables that determine participation are the same ones that determine the portfolio rule; more formally, \( X_{it} = Z_{it} \). This implies that the empirical counterpart of the model is one that only involves a censoring correction. To address this issue, I consider an alternative estimation procedure. In particular, I utilize the censored quantile regression estimator of Buchinsky and Hahn (1998).

The choice of this estimator over other procedures for censoring corrections (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 (10)-(14) when \( X_{it} = Z_{it} \) whittles down to a model with random censoring point. This

\[40\] In fact, in the absence of an exclusion restriction, and without a specific parametric form for the conditional copula, the distribution is only partially identified. (Arellano and Bonhomme (2017))

\[41\] This also effectively implies that the the error terms \( u_{it} \) and \( v_{it} \) in models (10) and (12) are the same.
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. 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). 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 \( \hat{\mu} \) 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 presenting how participation and portfolio choices respond to income shocks. 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).

6.1 Stock market participation

Figure 5 presents estimates of the average derivative of the participation rule to variations in the persistent component of earnings, \( v_{it} \) (in panels a and b) and wealth \( w_{it} \), and log household wealth \( w_{it} \) (in panels c and d), evaluated at different percentiles of age, \( \tau_{age} \) and \( \tau_{wealth} \), respectively. The left columns correspond to results obtained with the linear earnings process (that is, panels a and c); the right column corresponds to the nonlinear earnings process (that is, panels b and d).

---

\(^{42}\)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
As can be observed, the linear and nonlinear earnings models produce starkly different average derivative effects. Specifically, an increase in the persistent component $v_{it}$ leads to a decrease in the probability of participation for young households with little wealth and old households with high wealth, as shown in panel (a). In contrast, the nonlinear model shows that an increase in the persistent component $v_{it}$ leads to an unambiguous increase in the probability of participation for all households. The average derivative function also shows that the likelihood of participation also increases with age. Turning to the average derivative effects with respect to wealth, I find that the likelihood of participation is positive at all age levels and at all wealth levels, but exhibits a quadratic shape. A difference between those calculated from the linear model and from the nonlinear model is that there is an obvious increasing pattern across ages.

In the appendix to this section, I present the average derivative effects of the participation rule that are evaluated at different percentiles of the persistent component $\tau_v$ and $\tau_{age}$. The results indicate that indeed, there are nonlinear responses to an increase in the persistent component across ages and income. The same does not hold true for increases in wealth, however.

### 6.2 Portfolio allocation

Figure 6 presents estimates of the average derivative of the portfolio rule to variations in the persistent component of earnings, $v_{it}$ (in panels a and b) and $w_{it}$ (in panels c and d), evaluated at different percentiles of age, $\tau_{age}$ and $\tau_{wealth}$, respectively. The results I present here are for stock market participants.

In terms of the portfolio rule, while the shape of the average derivative effects are similar, the magnitudes are quite different. However, a striking result is that for young households with high wealth, an increase in labor income results in a reduction of the share of wealth invested in risky assets. In both cases, though, an increase in wealth leads to an unambiguous increase in the share invested in risky assets, although the magnitudes are less pronounced for the nonlinear model.

In the appendix to this section, I present the average derivative effects of the portfolio rule that are evaluated at different percentiles of the persistent component $\tau_v$ and $\tau_{age}$. The results indicate that indeed, there are nonlinear responses to an increase in the persistent component across ages and income. The same does not hold true for increases in wealth, however.
Figure 5: Average derivatives of the participation rule, by wealth and age

(a) Participation response to $v_{it}$, linear earnings model

(b) Participation response to $v_{it}$, nonlinear earnings model

(c) Participation response to $w_{it}$, linear earnings model

(d) Participation response to $w_{it}$, nonlinear earnings model

Note: Graphs 5a and 5c show estimates of the average derivative effect of the participation rule with respect to the persistent component $v_{it}$ and wealth $w_{it}$, respectively, evaluated at values of $w_{it}$ and age$_{it}$ that correspond to their $\tau_{wealth}$ and $\tau_{age}$ percentiles based on the estimated model with the linear earnings process in section 2 of the paper. Graphs 5b and 5d show estimates of the average derivative effect of the participation rule with respect to the persistent component $v_{it}$ and wealth $w_{it}$, respectively, evaluated at values of $w_{it}$ and age$_{it}$ that correspond to their $\tau_{wealth}$ and $\tau_{age}$ percentiles based on the estimated model with the nonlinear earnings process of Arellano et al. (2017).
Figure 6: Average derivatives of the portfolio rule, by wealth and age, for stock market participants

(a) Portfolio allocation response to $v_{it}$, linear earnings model

(b) Portfolio allocation response to $v_{it}$, nonlinear earnings model

(c) Portfolio allocation response to $w_{it}$, linear earnings model

(d) Portfolio allocation response to $w_{it}$, nonlinear earnings model

Note: Graphs 6a and 6c show estimates of the average derivative effect of the portfolio allocation rule with respect to the persistent component $v_{it}$ and wealth $w_{it}$, respectively, evaluated at values of $w_{it}$ and age$_{it}$ that correspond to their $\tau_{wealth}$ and $\tau_{age}$ percentiles based on the estimated model with the linear earnings process in section 2 of the paper. Graphs 6b and 6d show estimates of the average derivative effect of the participation rule with respect to the persistent component $v_{it}$ and wealth $w_{it}$, respectively, evaluated at values of $w_{it}$ and age$_{it}$ that correspond to their $\tau_{wealth}$ and $\tau_{age}$ percentiles based on the estimated model with the nonlinear earnings process of Arellano et al. (2017).

6.3 Model simulations

I simulate portfolio choices according to the nonlinear model with the linear and nonlinear earnings processes to compare the fit of the models with the data, and to calculate “impulse response”-like functions.

Figure 7 presents the observed densities and the predicted densities from simulations of the nonlinear model. While not obvious, the results indicate that the nonlinear model with
the nonlinear earnings process provides a better fit to the observed data than the nonlinear model with the linear earnings process. Moreover, comparing the predicted proportion of households who participate in the stock markets, it appears to be the case that the nonlinear model with the nonlinear earnings process provides a better fit as the proportion of households who participate is approximately 66.54 percent, close to the observed 64.53 percent in the data. The results with the linear model, meanwhile, show that the predicted proportion is approximately 69.98 percent.

Figure 7: Predicted and observed densities of the share of household wealth in risky assets

(a) Linear earnings model

(b) Nonlinear earnings model

Note: Graph 7a presents the observed and predicted unconditional densities of the share of household wealth in risky assets based on the linear earnings model, while graph 7b presents the observed and predicted unconditional densities of the share of household wealth in risky assets based on the nonlinear earnings model. In both cases, the red line is the simulated unconditional density, while the blue line is the observed density. All results were obtained using a Gaussian kernel with the optimal bandwidth.

Figures 8 and 9 present participation and portfolio allocation responses. In particular, I report the difference between two types of households: households who, at age 36, are hit by either 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. I report age-specific medians across 250,000 simulations of the model.

I first focus on Figure 8, which looks at participation responses. The results suggest the presence of interaction effects between the rank of the household in the distribution of the earnings component ($\tau_{\text{init}}$) and the sign and the size of the shock that the household receives ($\tau_{\text{shock}}$). In particular, the asymmetries in participation responses are much stronger when I look at the effect of large positive shocks ($\tau_{\text{shock}} = 0.9$). A large positive shock yields a 9.6 percent increase in the likelihood of participation for low-earnings households, but a modest 1.7 percent increase in the likelihood of participation for high-earnings households.
A large negative shock, meanwhile, produces modest participation responses. Specifically, a low-earnings household experiences a drop of 2.41 percent in the likelihood of participation, while a high-earnings household hit by the same shock experiences a 1.89 percentage point drop. A potential reason why this is the case could be that the baseline model does not fit well the dynamics of stock market participation, which, as shown by Table 2, seems to be important.

In comparison, I compute the same responses with a standard linear model of household earnings, the results of which are in panels (g) and (h) of Figure 8. As can be observed, the differences in persistences across households, and the presence of interaction effects between income shocks and the households’ positions in the income distribution is not completely captured by the results.

Figure 9, meanwhile, show the portfolio allocation responses for the subset of stock market participants. As can be observed from the figure, again there are presence of interaction effects between the size and the sign of the shock received, and the household’s position in the earnings distribution. Similarly, as in the impulse response for stock market participation, the asymmetries are more pronounced for positive earnings shocks, with low-earnings households increasing their portfolio allocation to stocks by as much as 9 percentage points, while high-earnings households increase their portfolio allocation to stocks by as much as 1.7 percentage points.

In contrast, high-earnings households decrease the share of their wealth in stocks by 4 percentage points when hit by a low earnings shock, while low-earnings households decrease their allocation by as much as 2.46 percentage points. Comparing the results that I present to those obtained from the linear earnings process, I find that the persistence of the effects appear to be different across households than that implied by the linear earnings process.

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. Looking at Figure 10, the results suggest that both positive and negative income shocks have a stronger impact on households who have little wealth ($\tau_{wealth} = 0.1$) relative to those who have high wealth ($\tau_{wealth} = 0.9$). The results also suggest an age effect, as households who are hit when they are young with a high earnings shock are more likely to participate, then households who are hit with the same shock when they are old. Similarly, households who are hit with a negative earnings shock when they are old are less likely to buy stocks than those who are hit by the same shock when young. Households who have high wealth, however, seem not to move that much. I observe similar responses to portfolio allocation in Figure 11.
Figure 8: Impulse responses, participation rule

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

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

(c) $\tau_{init} = 0.5, \tau_{shock} = 0.1$

(d) $\tau_{init} = 0.5, \tau_{shock} = 0.9$

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

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

(g) $\tau_{shock} = 0.1$, linear model

(h) $\tau_{shock} = 0.9$, linear model

Note: The graphs show the difference between a household with a given persistent component $\tau_{init}$ at age 35, who is hit by a shock $\tau_{shock}$ at age 36, and a household hit by a 0.5 shock at the same age. Graphs (8a) to (8f) correspond to the nonlinear earnings model. Graphs (8g) to (8h) correspond to the linear earnings model.
Figure 9: Impulse responses, portfolio allocation rule for market participants

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

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

(c) $\tau_{\text{init}} = 0.5, \tau_{\text{shock}} = 0.1$

(d) $\tau_{\text{init}} = 0.5, \tau_{\text{shock}} = 0.9$

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

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

(g) $\tau_{\text{shock}} = 0.1$, linear model

(h) $\tau_{\text{shock}} = 0.9$, linear model

Note: The graphs show the difference between a household with a given persistent component $\tau_{\text{init}}$ at age 35, who is hit by a shock $\tau_{\text{shock}}$ at age 36, and a household hit by a 0.5 shock at the same age. Graphs (9a) to (9f) correspond to the nonlinear earnings model. Graphs (9g) to (9h) correspond to the linear earnings model.
Figure 10: Impulse responses, participation rule, by wealth and age

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

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

(c) $\tau_{\text{init}} = 0.1$, $\tau_{\text{shock}} = 0.9$, $\tau_{\text{wealth}} = 0.1$, old

(d) $\tau_{\text{init}} = 0.1$, $\tau_{\text{shock}} = 0.9$, $\tau_{\text{wealth}} = 0.9$, old

(e) $\tau_{\text{init}} = 0.9$, $\tau_{\text{shock}} = 0.1$, $\tau_{\text{wealth}} = 0.1$, young

(f) $\tau_{\text{init}} = 0.9$, $\tau_{\text{shock}} = 0.1$, $\tau_{\text{wealth}} = 0.9$, young

(g) $\tau_{\text{init}} = 0.9$, $\tau_{\text{shock}} = 0.1$, $\tau_{\text{wealth}} = 0.1$, old

(h) $\tau_{\text{init}} = 0.9$, $\tau_{\text{shock}} = 0.1$, $\tau_{\text{wealth}} = 0.9$, old

Note: See the note to Figure (8). Young households refer to households who are hit by a shock at age 35, while old households refer to households who are hit by a shock at age 50.
Figure 11: Impulse responses, portfolio allocation rule, by wealth and age

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

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

(c) \( \tau_{init} = 0.1, \tau_{shock} = 0.9, \tau_{wealth} = 0.1, \) old

(d) \( \tau_{init} = 0.1, \tau_{shock} = 0.9, \tau_{wealth} = 0.9, \) old

(e) \( \tau_{init} = 0.9, \tau_{shock} = 0.1, \tau_{wealth} = 0.1, \) young

(f) \( \tau_{init} = 0.9, \tau_{shock} = 0.1, \tau_{wealth} = 0.9, \) young

(g) \( \tau_{init} = 0.9, \tau_{shock} = 0.1, \tau_{wealth} = 0.1, \) old

(h) \( \tau_{init} = 0.9, \tau_{shock} = 0.1, \tau_{wealth} = 0.9, \) old

Note: See the note to Figure (9). Young households refer to households who are hit by a shock at age 35, while old households refer to households who are hit by a shock at age 50.
7 Conclusions

This paper develops a semi-structural framework that sheds new light on the transmission of income risk into households’ investment behavior. In my model, stock market participation and portfolio rules are modelled as age-dependent nonlinear functions of the persistent and transitory earnings components, and of wealth. The estimation framework reveals new asymmetric participation and portfolio adjustment responses with respect to “unusual” income shocks.

In a future version of this paper, I aim to provide additional empirical evidence through extensions of the semi-structural model that have natural economic counterparts, such as state dependence in participation and time-invariant unobserved heterogeneity that can represent discount rates or preference shifters. I also aim to assess, through a structural estimation exercise, the extent to which a standard model can rationalize the effects I uncover in the data. Specifically, I aim to quantify the magnitude of participation costs that households need to pay to participate in the stock market.

References


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


Appendices for the paper “Household portfolio choices and non-linear income risks”

The appendices follow the organization of the paper. Appendix A outlines the derivations associated with the two-period simple model, and details related to the calibration exercise in the life-cycle model. Appendix B presents the model 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. To obtain optimal consumption in the non-participation sub-problem, I write a log-linear approximation to the budget constraint. To see this, divide both sides of equation (2) by \( L_{t+1} \). 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 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:

\[
c_{t+1} - l_{t+1} = \log(\exp(w_t + r_f - l_{t+1}) + 1)
\]

\[
\approx k_{np} + \phi_{np}(w_t + r_f - l_{t+1}),
\]

where, after rearranging both sides of the equation, yields the expression in equation (5).

Consumption in the participation subproblem. Similarly, as in the non-participation sub-problem, I obtain optimal consumption by log-linearization of the budget constraint (3).\(^{43}\)

Dividing both sides by \( L_{t+1} \):

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

\(^{43}\)The log-linear approximation to the portfolio return, equation is derived in the appendix of Campbell and Viceira (2002) that is available here: https://scholar.harvard.edu/campbell.
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 expression in equation (6).

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

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

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

\[
E_t[\delta C_{t+1}^{-\gamma}(1 + R_{t+1})] = E_t[\exp\{\log(\delta C_{t+1}^{-\gamma}(1 + R_{t+1}))\}] = E_t[\exp\{\log(\delta) - \gamma c_{t+1} + r_{t+1}\}] = E_t[\exp\{x_{t+1}\}]
\]

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

\[
E_t(\exp(x_{t+1})) \approx E_t\left[\exp(\bar{x}_t)\left(1 + (x_{t+1} - \bar{x}_t) + \frac{1}{2}(x_{t+1} - \bar{x}_t)^2\right)\right]
\approx \exp(\bar{x}_t)\left(1 + \frac{1}{2}\text{Var}_t(x_{t+1})\right)
\]

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

\[
E_t(\exp(x_{t+1})) \approx 1 + \bar{x}_t + \frac{1}{2}\text{Var}_t(x_{t+1}).
\]

From here, it can be shown that the first-order condition becomes:

\[
E_t(r_{t+1} - r_f) + \frac{1}{2}\text{Var}_t(r_{t+1}) = \gamma\text{Cov}_t(r_{t+1}, c_t) = \gamma\text{Cov}_t(r_{t+1}, k_p + \phi_p(w_{c,t} + r_{p,t+1}) + (1 - \phi_p)l_{t+1}) = \gamma[\phi_p \alpha \sigma_u^2 + (1 - \phi_p)\text{Cov}_t(l_{t+1}, r_{t+1})]
\]

where in the second line I substituted the consumption function (6) and in the third line I substituted the log portfolio return. Rearranging this equation yields the optimal portfolio rule (7).
**Alternative expression of the portfolio rule.** The alternative expression of the portfolio rule follows from the wealth elasticity of consumption. To see this, write the wealth elasticity of consumption as:

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

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

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

which, when substituted to the portfolio rule with idiosyncratic labor income risks, yields equation (8).

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

\[
E_t \left( \delta \frac{C_{i,t+1}}{1 - \gamma} \right) \geq E_t \left( \delta \frac{C_{np,t+1}}{1 - \gamma} \right)
\]

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

\[
E_t \left( \delta \frac{C_{i,t+1}}{1 - \gamma} \right) = E_t \left[ \exp \left( \log \left( \delta \frac{C_{i,t+1}}{1 - \gamma} \right) \right) \right] \approx E_t(\exp\{(1 - \gamma)c_{i,t+1}\})
\]

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 in equation (9).

**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:

\[
a_t = \frac{1}{\phi_p} \left[ \frac{E_{t+1}(r_{t+1} - r_f) + \frac{1}{2}\sigma_u^2}{\gamma\sigma_u^2} \right]
\]
Taking the total derivative with respect to wealth, I have the following expression:

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

I can now turn to the calculation of the components in the numerator, as the denominator is always positive.\(^{44}\) 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.

Next, we obtain \(d\phi_p/dw_c\):

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

Finally, I calculate \(d\alpha_t/d\phi_p\):

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

Putting all of these together:

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

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:

\[
\frac{d\alpha_t}{dl_t} = \frac{\frac{d\alpha_t}{d\phi_p} \frac{dl_t}{d\phi_p} dl_t}{1 - \frac{d\alpha_t}{d\phi_p} \frac{d\phi_p}{dr_p} dr_p}.
\]

\(^{44}\)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:

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

Finally,

\[
\frac{dr_{p,t+1}}{d\alpha_t} = \left( E_t(l_{t+1}) - r_f + \frac{1}{2} \sigma_u^2 \right) \left( 1 - \frac{1}{\phi_p^2} \right)
\]

A sufficient condition to prove that the denominator is positive is to show that \(\gamma > \frac{1}{\phi_p} > 1\). Now, when \(\gamma = \frac{1}{\phi_p}\), 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.
Because I have showed earlier that the denominator is positive, I focus on calculating the terms in the numerator. I first calculate $d\alpha_t / d\phi_p$:

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

Next, I obtain $d\phi_p / dl$:

$$d\phi_p / dl = -\exp(w_{it} + \mu + \sigma_u u) / (1 + \exp(w_{it} + \mu + \sigma_u u))^2 < 0$$

Finally, to compute $dl / dL$, 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 $L_t$), I obtain the following:

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

Now, computing $dl / dL$:

$$dl / dL = \frac{E_t[(L - \mu_L)^2]}{L^3} - \frac{E_t[(L - \mu_L)^3]}{L^4}$$

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 $d\alpha / dL$:

$$\text{sign} \left( \frac{d\alpha_t}{dl} \right) = \frac{(-\phi_p) (-\phi_p)}{\mu + \sigma_u^2} \frac{d\phi_p}{dl} \frac{dl}{dL_t} \propto 0.$$  

To compare with the lognormal distribution:

$$dl / dL = \frac{1}{L^3},$$

which is true given that I can express $l$ as\textsuperscript{45}:

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

\textsuperscript{45}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.

53
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 (5) and (6) into the inequality I have introduced in equation (9). 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 \( w_{c,t} \):

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

Rearranging the previous equation yields the following expression:

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

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

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

Plugging this in into the equation earlier, I obtain:

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

Now, the only thing left to be checked is the sign of the numerator.\(^{46}\) Now, \( \frac{\partial \phi_p}{\partial w_{c,t}} \) is just:

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

\(^{46}\) I first check the sign of the denominator. It will only be negative when

\[
r_{p,t+1} - r_f + \frac{1}{2} \sigma_u^2 - \alpha_t \sigma_u^2 \leq 0.
\]

When this is rearranged, the inequality suggests that:

\[
r_{p,t+1} - r_f + \frac{1}{2} \sigma_u^2 \leq \alpha_t \sigma_u^2
\]
The calculations yield
\[
\frac{\phi_p (w_{t,t} + r_{p,t+1} - l_{t+1})}{1 + \exp(w_{t,t} + r_{p,t+1} - l_{t+1})} + \phi_p \geq 0
\]
which will mean that:
\[
w_{t,t} + r_{p,t+1} - l_{t+1} \geq -1 - \exp(w_{t,t} + r_{p,t+1} - l_{t+1})
\]
This inequality is generally true whenever \(w_{t,t} + r_{p,t+1} \geq l_{t+1}\). 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 Implicit Function Theorem:
\[
\frac{\partial \phi_p}{\partial l_{t+1}} w_{t,c} + \frac{\partial \phi_p}{\partial r_{p,t+1}} \frac{\partial r_{p,t+1}}{\partial l_{t+1}} w_{t,c} + \frac{\partial \phi_p}{\partial l_{t+1}} \phi_l = \frac{\partial \phi_p}{\partial r_{p,t+1}} \frac{\partial r_{p,t+1}}{\partial l_{t+1}} w_{t,c} - \frac{\partial \phi_p}{\partial l_{t+1}} w_{t,c} + \frac{\partial \phi_p}{\partial l_{t+1}} \phi_l + \frac{\partial \phi_p}{\partial l_{t+1}} r_f - \frac{\partial \phi_p}{\partial r_{p,t+1}} r_{p,t+1} - \frac{\partial \phi_p}{\partial l_{t+1}} \phi_l
\]

Rearranging this equation yields the following:
\[
\frac{\partial \phi_p}{\partial l_{t+1}} (w_{t,c} + r_{p,t+1} - l_{t+1}) - \frac{\partial \phi_p}{\partial l_{t+1}} (w_{t,c} + r_f - l_{t+1}) - (\phi_p - \phi_n) = \frac{\partial \phi_l}{\partial r_{p,t+1}} (w_{t,c} + r_{p,t+1} - l_{t+1}) - \phi_p \frac{\partial r_{p,t+1}}{\partial l_{t+1}}
\]

Notice though, that:
\[
\frac{\partial \phi_p}{\partial l_{t+1}} = -\frac{\phi_p}{1 + \exp(w_{t,c} + r_{p,t+1} - l_{t+1})}
\]
which, after rearranging, becomes:
\[
a_l \geq \frac{r_{p,t+1} - r_f}{\sigma_n^2} + \frac{1}{2} \geq 0
\]
For the right hand side of this inequality to be positive however, it should be the case that
\[
\frac{r_{p,t+1} - r_f}{\sigma_n^2} > -\frac{1}{2},
\]
which means that:
\[
a_l \geq \frac{r_{p,t+1} - r_f}{\sigma_n^2} > -\frac{1}{2},
\]
Given that the marginal investor has positive \(a_l\), then the denominator indeed is negative.
and similarly,

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

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

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

Removing the negative signs will yield:

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

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:

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

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 Life-cycle model details

A.2.1 Identification and estimation of the linear earnings process

A more formal statement of the assumptions behind the linear earnings process that I use here are the following:

1. $|\rho| < 1$.
2. $\eta_{it} \perp \nu_{it} \perp \epsilon_{it}$.
3. $\eta_{it} \sim iidN(0, \sigma_\eta^2)$, $\nu_{it} \sim iidN(0, \sigma_\nu^2)$, $\epsilon_{it} \sim iidN(0, \sigma_\epsilon^2)$

I can formally identify the parameters of interest in this model from the autocovariance function alone, following standard arguments, and the assumptions I have made here. First, I can identify $\rho$ from the slope:

$$\frac{\text{Cov}(v_{i0}, v_{i3}) - \text{Cov}(v_{i0}, v_{i2})}{\text{Cov}(v_{i0}, v_{i2}) - \text{Cov}(v_{i0}, v_{i1})} = \frac{\rho^3 \sigma_\nu^2 - \rho^2 \sigma_\nu^2}{\rho^2 \sigma_\nu^2 - \rho \sigma_\nu^2} = (\rho^3 - \rho^2)(\sigma_\nu^2) = \rho.$$

The difference between the covariances allows me to obtain $\sigma_\nu$:

$$\text{Cov}(v_{i0}, v_{i2}) - \text{Cov}(v_{i0}, v_{i1}) = \rho^2 \sigma_\nu^2 - \rho \sigma_\nu^2 = (\rho^2 - \rho)(\sigma_\nu^2).$$

The difference between the variances allows me to obtain $\sigma_\eta$:

$$\var(v_{i1}) - \var(v_{i0}) = (\rho \sigma_\nu^2 + \sigma_\eta^2 + \sigma_\epsilon^2) - (\sigma_\nu^2 + \sigma_\eta^2) = (\rho - 1)\sigma_\nu^2 + \sigma_\eta^2.$$

Finally, the variance allows me to identify $\sigma_\epsilon$:

$$\var(v_{i0}) = \sigma_\nu^2 + \sigma_\epsilon^2.$$

The standard estimation strategy is to use minimum distance estimation, where the goal is to choose the parameters that minimise the distance between the empirical and theoretical moments. 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:
\[ \hat{\theta}_{PML} = \arg \min_c \left\{ \log \det(\Omega(c)) + \text{tr}(\Omega(c)^{-1} \hat{\Omega}) \right\}, \]
where \( \text{tr} \) is the trace of the resulting matrix, and \( \hat{\Omega} = \sum \hat{u}_i \hat{u}_i' \).

In practice, I first regress (log) income on a second-order polynomial on age, education dummies, household size, number of children, and cohort effects, and obtain the residuals from this estimation. 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>0.9665</td>
<td>(0.0586)</td>
</tr>
<tr>
<td>SD of transitory component</td>
<td>0.2502</td>
<td>(0.0754)</td>
</tr>
<tr>
<td>SD of initial persistent component</td>
<td>0.4196</td>
<td>(0.0309)</td>
</tr>
<tr>
<td>SD of idiosyncratic persistent component</td>
<td>0.1753</td>
<td>(0.0496)</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 \( \nu_{it} \) and \( \text{age}_{it} \) 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:
\[ Q_t(\nu_{it-1}, \tau) = \sum_{k=0}^{K} a^p_k(\tau) \phi_k(\nu_{it-1}, \text{age}_{it}) \]

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:
\[ Q_{\nu_{i1}}(\text{age}_{i1}, \tau) = \sum_{k=0}^{K} a^I_k(\tau) \tilde{\phi}_k(\text{age}_{i1}) \]

Transitory component. The conditional quantile function of the transitory component as a function of age is:
\[ Q_{\epsilon}(\text{age}_{it}, \tau) = \sum_{k=0}^{K} a^T_k(\tau) \bar{\phi}_k(\text{age}_{it}) \]
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.

Figure A1: Nonlinear persistence

(a) Earnings, PSID data
(b) Earnings, nonlinear model
(c) Earnings, linear model
(d) Persistent component $\nu_{it}$, nonlinear model

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 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 seems to describe the data somewhat well, as the fit is close to that of the real data. Panel (d) describes the results of the persistence in the persistent component $v_{it}$ in the nonlinear earnings model. As the results indicate, there seems to be heterogeneity in persistence of income shocks.\footnote{Comparing my estimates to those obtained by Arellano et al. (2017), I find that for my PSID subsample, the estimated persistence for high earnings households hit by an extremely good earnings shock, and low earnings households hit by an extremely bad earnings shock is around 0.94 to 1.31, while Arellano et al. (2017) find an estimated persistence of around 0.9 to 1. Similarly, for high earnings households hit by a extremely bad earnings shock, and for low earnings households hit by an extremely good earnings shock, estimated persistence in my data ranges from -0.22 to 0.34, while in Arellano et al. (2017), the estimated persistence is around 0.3 to 0.4.}

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.

Finally, 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.

A.2.3 Calibration details

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

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.

**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).

**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 discretise the state space for the cash-on-hand state variable and iterate on the value function. The density function for the linear income process and the risky return were both approximated using a three-point Gaussian quadrature.

**Discretizing the nonlinear earnings process.** I discretize the nonlinear earnings process by the following manner:

1. I simulate one million households throughout their life-cycle, using the parameters estimated from the Arellano et al. (2017) earnings process.

2. I discretise the income process by ranking households by their earnings and putting them into 50 bins for the persistent component, and 25 bins for the transitory component, at a given age.

3. To estimate the probabilities for the transitory component, I calculate the proportion of households who fall at a given bin at a certain age.

4. 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 Nonlinear semi-reduced form model extensions

In this section, 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 house-
hold 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:

\[
\alpha^*_i t = g_t(v_{it}, \epsilon_{it}, w_{it}, X_{it}, u_{it}) \quad (B1)
\]

\[
\alpha_i t = \alpha^*_i t \cdot d_{it} \quad (B2)
\]

\[
d_{it} = \begin{cases} 1, & \text{if } \tilde{g}_t(v_{it}, \epsilon_{it}, w_{it}, q(Z_{it}), d_{it-1}) \leq v_{it} \\ 0, & \text{otherwise} \end{cases} \quad (B3)
\]

\[
w_{it} = h_t(v_{it-1}, \epsilon_{it-1}, w_{it-1}, \alpha_{it-1}, X_{it}, w_{it}) \quad (B4)
\]

\[
\alpha^*_{i0} = g(v_{i0}, \epsilon_{i0}, w_{i0}, X_{i0}, u_{i0}) \quad (B5)
\]

\[
\alpha_{i0} = \alpha^*_{i0} \cdot d_{i0} \quad (B6)
\]

\[
d_{i0} = \begin{cases} 1, & \text{if } \tilde{g}(v_{i0}, \epsilon_{i0}, w_{i0}, q(X_{i0})) \leq v_{i0} \\ 0, & \text{otherwise} \end{cases} \quad (B7)
\]

\[
w_{i0} = h_{i0}(v_{i0}, \epsilon_{i0}) \quad (B8)
\]

In this model, equations (B5)-(B8) correspond to the equations that describe the decisions of the household at the first period in which it is observed. Notice that in equation (B3), which corresponds to the participation rule in subsequent periods, 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:

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

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

\[
\delta_{00}(v) = \Pr(d_{it} = 0|Z_{it} = z, v_{it} = v, \epsilon_{it} = \epsilon, w_{it} = w, d_{it-1} = 0)
\]

\[
\delta_{10}(v) = \Pr(d_{it} = 1|Z_{it} = z, v_{it} = v, \epsilon_{it} = \epsilon, w_{it} = w, d_{it-1} = 0)
\]

\[
\delta_{01}(v) = \Pr(d_{it} = 0|Z_{it} = z, v_{it} = v, \epsilon_{it} = \epsilon, w_{it} = w, d_{it-1} = 1)
\]

\[
\delta_{11}(v) = \Pr(d_{it} = 1|Z_{it} = z, v_{it} = v, \epsilon_{it} = \epsilon, w_{it} = w, d_{it-1} = 1)
\]

This effectively implies that I can calculate the following components:

\[
\Delta_{00}(v + v', v) = \delta_{00}(v + v', v) - \delta_{00}(v)
\]

\[
\Delta_{10}(v + v', v) = \delta_{10}(v + v', v) - \delta_{10}(v)
\]

\[
\Delta_{01}(v + v', v) = \delta_{01}(v + v', v) - \delta_{01}(v)
\]

\[
\Delta_{11}(v + v', v) = \delta_{11}(v + v', v) - \delta_{11}(v)
\]
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:

\[ \alpha_{it}^* = g_t(v_{it}, e_{it}, w_{it}, X_{it}, \zeta_i, u_{it}) \] (B9)

\[ \alpha_{it} = \alpha_{it}^* \cdot d_{it} \] (B10)

\[ d_{it} = \begin{cases} 1, & \text{if } \tilde{g}_t(v_{it}, e_{it}, w_{it}, q(Z_{it}), \zeta_i) \leq v_{it} \\ 0, & \text{otherwise} \end{cases} \] (B11)

\[ w_{it} = h_t(v_{it-1}, e_{it-1}, w_{it-1}, e_{it-1}, X_{it}, \zeta_i, w_{it}) \] (B12)

\[ w_{i0} = \tilde{h}_{i0}(v_{i0}, \zeta_{i0}) \] (B13)

\[ \zeta_i = \phi(v_{i0}, w_{i0}, v_i) \] (B14)

In this model, I consider a scalar unobserved heterogeneity, which I denote by \( \zeta_i \). Notice that \( \zeta_i \) 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 (B11), 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 \( \zeta_i \) conditional on the initial observations of the households. This is in contrast to a fixed-effects approach in which I condition on the household-specific fixed effect, the distribution of which is left unspecified.\(^{48}\) 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 \( \zeta_i \). 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 estimat-

ing equations into the following:

\[ a_{it}^* = g_t(v_{it}, \epsilon_{it}, w_{it}, X_{it}, u_{it}) \]  
(B15)

\[ a_{it} = a_{it}^* \cdot d_{it} \]  
(B16)

\[ d_{it} = \begin{cases} 1, & \text{if } \tilde{g}_t(v_{it}, \epsilon_{it}, w_{it}, q(Z_{it})) \leq v_{it} \\ 0, & \text{otherwise} \end{cases} \]  
(B17)

\[ c_{it} = n_t(v_{it}, \epsilon_{it}, w_{it}, X_{it}, u_{it}) \]  
(B18)

\[ w_{it} = h_t(v_{it-1}, \epsilon_{it-1}, w_{it-1}, a_{it-1}, Z_{it}, c_{it-1}, \zeta_{it}) \]  
(B19)

\[ w_{i0} = \tilde{h}_i(v_{i0}, \zeta_{i0}) \]  
(B20)

in which I introduce the consumption rule (B18) and past consumption in equation (B19), the budget constraint. In this case, the error term in (B19) 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:

\[ a_{it}^* = g_t(v_{it}, v_{it+1}, \epsilon_{it}, w_{it}, X_{it}, u_{it}) \]  
(B21)

\[ a_{it} = a_{it}^* \cdot d_{it} \]  
(B22)

\[ d_{it} = \begin{cases} 1, & \text{if } \tilde{g}_t(v_{it}, v_{it+1}, \epsilon_{it}, w_{it}, q(Z_{it})) \leq v_{it} \\ 0, & \text{otherwise} \end{cases} \]  
(B23)

\[ w_{it} = h_t(v_{it-1}, \epsilon_{it-1}, w_{it-1}, a_{it-1}, X_{it}, c_{it}, \zeta_{it}) \]  
(B24)

\[ m_{i0} = \tilde{h}_i(v_{i0}, \zeta_{i0}) \]  
(B25)

C Data and descriptive statistics

C.1 Sample selection criteria

The tables in this subsection detail the removal of households from the 1999 to 2009 waves of the PSID. In the data cleaning process, I first remove inconsistent information from each of the family files, the results of which are in Table A2. I then merge the family files with the individual file. The details of the merge and the removal of other criteria are in Table A3.
**Table A2: Household selection criteria, PSID family files**

<table>
<thead>
<tr>
<th></th>
<th>1999</th>
<th>2001</th>
<th>2003</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FULL SAMPLE</strong></td>
<td>6997</td>
<td>7406</td>
<td>7822</td>
</tr>
<tr>
<td>LESS: work hours reported, no labor income</td>
<td>293</td>
<td>252</td>
<td>68</td>
</tr>
<tr>
<td>No hours reported, with labor income</td>
<td>8</td>
<td>2</td>
<td>15</td>
</tr>
<tr>
<td>No race information</td>
<td>58</td>
<td>135</td>
<td>148</td>
</tr>
<tr>
<td>No home value</td>
<td>47</td>
<td>61</td>
<td>49</td>
</tr>
<tr>
<td>No mortgage value</td>
<td>127</td>
<td>142</td>
<td>139</td>
</tr>
<tr>
<td>No age of head</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Married couples only</td>
<td>2934</td>
<td>3102</td>
<td>3446</td>
</tr>
<tr>
<td>Subtotal before wealth criteria</td>
<td>3470</td>
<td>3697</td>
<td>3866</td>
</tr>
<tr>
<td>Net wealth &lt; 0</td>
<td>239</td>
<td>297</td>
<td>325</td>
</tr>
<tr>
<td>Risky share &gt; 1</td>
<td>66</td>
<td>73</td>
<td>72</td>
</tr>
<tr>
<td>Subtotal</td>
<td>3222</td>
<td>3339</td>
<td>3559</td>
</tr>
<tr>
<td>No consumption</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TOTAL</td>
<td>3222</td>
<td>3339</td>
<td>3559</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>2005</th>
<th>2007</th>
<th>2009</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FULL SAMPLE</strong></td>
<td>8002</td>
<td>8289</td>
<td>8690</td>
</tr>
<tr>
<td>LESS: work hours reported, no labor income</td>
<td>95</td>
<td>81</td>
<td>72</td>
</tr>
<tr>
<td>No hours reported, with labor income</td>
<td>7</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>No race information</td>
<td>52</td>
<td>66</td>
<td>46</td>
</tr>
<tr>
<td>No home value</td>
<td>86</td>
<td>91</td>
<td>208</td>
</tr>
<tr>
<td>No mortgage value</td>
<td>158</td>
<td>176</td>
<td>212</td>
</tr>
<tr>
<td>No age of head</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Married couples only</td>
<td>3499</td>
<td>3678</td>
<td>2445</td>
</tr>
<tr>
<td>Subtotal before wealth criteria</td>
<td>3899</td>
<td>4096</td>
<td>2989</td>
</tr>
<tr>
<td>Net wealth &lt; 0</td>
<td>4103</td>
<td>4193</td>
<td>5701</td>
</tr>
<tr>
<td>Risky share &gt; 1</td>
<td>356</td>
<td>389</td>
<td>1514</td>
</tr>
<tr>
<td>Subtotal</td>
<td>3653</td>
<td>3722</td>
<td>3386</td>
</tr>
<tr>
<td>No consumption</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TOTAL</td>
<td>3653</td>
<td>3722</td>
<td>3386</td>
</tr>
</tbody>
</table>

*Note: This table describes the deletion of households based on sample selection criteria used in the paper. This sample selection removal is performed prior to the merging of the PSID family files with the PSID individual files.*
Table A3: Sample selection criteria, merged PSID individual and family files

<table>
<thead>
<tr>
<th>Eligible households from individual files</th>
<th>2057</th>
</tr>
</thead>
<tbody>
<tr>
<td>LESS:</td>
<td></td>
</tr>
<tr>
<td>HHs missing from 1999 wave</td>
<td>804</td>
</tr>
<tr>
<td>HHs missing from 2001 wave</td>
<td>166</td>
</tr>
<tr>
<td>HHs missing from 2003 wave</td>
<td>89</td>
</tr>
<tr>
<td>HHs missing from 2005 wave</td>
<td>85</td>
</tr>
<tr>
<td>HHs missing from 2007 wave</td>
<td>58</td>
</tr>
<tr>
<td>HHs missing from 2009 wave</td>
<td>79</td>
</tr>
<tr>
<td>Households from merged files</td>
<td>776</td>
</tr>
<tr>
<td>LESS: inconsistent education</td>
<td>44</td>
</tr>
<tr>
<td>LESS: HHs with &lt; $300 financial wealth</td>
<td>124</td>
</tr>
<tr>
<td>TOTAL NUMBER OF HOUSEHOLDS</td>
<td>608</td>
</tr>
</tbody>
</table>

Note: This table describes the deletion of households based on sample selection criteria used in the paper. This sample selection removal is performed at the merging of the PSID family files with the PSID individual files.

C.2 Additional descriptive evidence

Table A4 presents the summary statistics by wealth quartiles. I find, as in the table I present in the main text, that there is wide dispersion across households in their earnings, wealth and consumption. One difference between the results that I obtain with respect to wealth quartiles with those that I find in Table 1 is that all relevant variables are monotonically increasing with respect to the wealth quartile.

Table A5, meanwhile, calculates the summary statistics by age quartiles. Again, as in the table in the main text, I find that there is some dispersion across households in the relevant economic variables. As in the previous table, I find that most of the wealth variables are monotonic with respect to the age quartiles. Interestingly, however, the risky asset share when defined in terms of net wealth display a non-monotonic pattern similar to that calculated for income quartiles. This is suggestive that perhaps, there are non-monotonicities in this context.
Table A4: Summary statistics, by wealth quartiles

<table>
<thead>
<tr>
<th>Variable</th>
<th>TOTAL</th>
<th>First</th>
<th>Second</th>
<th>Third</th>
<th>Fourth</th>
<th>F statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Household income</td>
<td>103,844.60</td>
<td>67,265.50</td>
<td>82,956.21</td>
<td>99,664.16</td>
<td>165,492.60</td>
<td>103.19</td>
</tr>
<tr>
<td>Total household wealth</td>
<td>413,411.70</td>
<td>61,139.50</td>
<td>82,956.21</td>
<td>99,664.16</td>
<td>165,492.60</td>
<td>150.16</td>
</tr>
<tr>
<td>Financial wealth</td>
<td>190,287.10</td>
<td>38,862.36</td>
<td>71,952.90</td>
<td>127,950.30</td>
<td>272,784.50</td>
<td>534.43</td>
</tr>
<tr>
<td>Consumption</td>
<td>50,105.66</td>
<td>42,275.17</td>
<td>50,981.46</td>
<td>72,855.31</td>
<td>72,855.31</td>
<td>388.20</td>
</tr>
<tr>
<td>Home equity</td>
<td>124,924.10</td>
<td>27,008.92</td>
<td>71,952.90</td>
<td>127,950.30</td>
<td>272,784.50</td>
<td>534.43</td>
</tr>
<tr>
<td>Risky wealth ownership</td>
<td>0.645</td>
<td>0.571</td>
<td>0.750</td>
<td>0.894</td>
<td>249.46</td>
<td></td>
</tr>
<tr>
<td>Risky wealth</td>
<td>123,883.80</td>
<td>3,712.96</td>
<td>16,095.49</td>
<td>48,409.60</td>
<td>427,317.00</td>
<td>30.69</td>
</tr>
<tr>
<td>Risky asset share, net wealth</td>
<td>0.158</td>
<td>0.120</td>
<td>0.171</td>
<td>0.267</td>
<td>163.34</td>
<td></td>
</tr>
<tr>
<td>Risky asset share, financial wealth</td>
<td>0.347</td>
<td>0.289</td>
<td>0.397</td>
<td>0.530</td>
<td>208.20</td>
<td></td>
</tr>
</tbody>
</table>

Note: This paper presents summary statistics of the main economic variables related to this empirical study, calculated across wealth quartiles. The first column calculates the mean across all households in the sample, while the second to fifth columns calculate the mean across income quartiles. The sixth column presents the F statistic calculated from a multivariate test for equality across groups. The first panel presents results across all households, while the second panel presents results for stock market participants.
Table A5: Summary statistics, by age quartiles

<table>
<thead>
<tr>
<th>Variable</th>
<th>TOTAL</th>
<th>First</th>
<th>Second</th>
<th>Third</th>
<th>Fourth</th>
<th>F statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Household income</td>
<td>103,844.60</td>
<td>84,526.57</td>
<td>104,769.00</td>
<td>116,096.10</td>
<td>114,624.70</td>
<td>11.47</td>
</tr>
<tr>
<td>Total household wealth</td>
<td>413,411.70</td>
<td>221,157.30</td>
<td>407,812.60</td>
<td>503,996.40</td>
<td>576,725.60</td>
<td>11.32</td>
</tr>
<tr>
<td>Financial wealth</td>
<td>190,287.10</td>
<td>76,978.27</td>
<td>198,547.20</td>
<td>242,338.20</td>
<td>273,378.40</td>
<td>5.26</td>
</tr>
<tr>
<td>Consumption</td>
<td>50,105.66</td>
<td>42,183.78</td>
<td>49,409.17</td>
<td>54,749.64</td>
<td>56,348.69</td>
<td>46.09</td>
</tr>
<tr>
<td>Home equity</td>
<td>124,924.10</td>
<td>80,435.05</td>
<td>122,665.00</td>
<td>141,023.70</td>
<td>169,709.20</td>
<td>46.87</td>
</tr>
<tr>
<td>Risky wealth ownership</td>
<td>0.645</td>
<td>0.612</td>
<td>0.628</td>
<td>0.660</td>
<td>0.695</td>
<td>5.19</td>
</tr>
<tr>
<td>Risky wealth</td>
<td>123,883.80</td>
<td>36,730.48</td>
<td>135,138.20</td>
<td>173,714.00</td>
<td>169,933.60</td>
<td>3.14</td>
</tr>
<tr>
<td>Risky asset share, net wealth</td>
<td>0.158</td>
<td>0.136</td>
<td>0.161</td>
<td>0.161</td>
<td>0.182</td>
<td>7.69</td>
</tr>
<tr>
<td>Risky asset share, financial wealth</td>
<td>0.347</td>
<td>0.317</td>
<td>0.345</td>
<td>0.354</td>
<td>0.382</td>
<td>5.22</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>TOTAL</th>
<th>First</th>
<th>Second</th>
<th>Third</th>
<th>Fourth</th>
<th>F statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Household income</td>
<td>116,047.10</td>
<td>94,632.87</td>
<td>120,462.40</td>
<td>122,290.10</td>
<td>129,750.90</td>
<td>8.37</td>
</tr>
<tr>
<td>Total household wealth</td>
<td>548,585.00</td>
<td>301,150.50</td>
<td>558,779.10</td>
<td>650,809.40</td>
<td>720,097.70</td>
<td>7.04</td>
</tr>
<tr>
<td>Financial wealth</td>
<td>275,417.30</td>
<td>114,245.50</td>
<td>298,137.40</td>
<td>344,858.70</td>
<td>365,391.20</td>
<td>3.91</td>
</tr>
<tr>
<td>Consumption</td>
<td>54,667.32</td>
<td>46,041.07</td>
<td>54,205.39</td>
<td>58,900.66</td>
<td>60,834.88</td>
<td>25.61</td>
</tr>
<tr>
<td>Home equity</td>
<td>148,923.50</td>
<td>97,749.08</td>
<td>146,945.50</td>
<td>163,169.40</td>
<td>196,867.90</td>
<td>28.14</td>
</tr>
<tr>
<td>Risky wealth</td>
<td>191,983.00</td>
<td>60,056.94</td>
<td>215,104.20</td>
<td>263,069.30</td>
<td>244,380.70</td>
<td>2.72</td>
</tr>
<tr>
<td>Risky asset share, net wealth</td>
<td>0.246</td>
<td>0.222</td>
<td>0.257</td>
<td>0.244</td>
<td>0.262</td>
<td>4.31</td>
</tr>
<tr>
<td>Risky asset share, financial wealth</td>
<td>0.537</td>
<td>0.518</td>
<td>0.549</td>
<td>0.535</td>
<td>0.549</td>
<td>1.55</td>
</tr>
</tbody>
</table>

**Note:** This paper presents summary statistics 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 across age quartiles. The sixth column presents the F statistic calculated from a multivariate test for equality across groups. The first panel presents results across all households, while the second panel presents results for stock market participants.

### 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:

\[
HW_{i,t} = L_{i,t} + \sum_{\tau=1}^{T-t} \pi_{\tau+1|\tau} E_{t}(L_{i,t+\tau}) \left(1 + r^{\tau}\right)^{-1} \tag{C26}
\]

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 Empirical strategy

D.1 Model estimation: details

D.1.1 Likelihood function

The likelihood function is:

\[
f(\mathbf{a}_i^T, \mathbf{v}_i^T, \mathbf{\epsilon}_i^T, w_i^T, \mathbf{Z}_i^T, d_i^T; \mathbf{\mu}) = \prod_{t=1}^T \left[ f(\mathbf{a}_{it}^T | v_{it}, \epsilon_{it}, w_{it}, x_{it}) p(d_{it} = 1 | v_{it}, \epsilon_{it}, w_{it}, z_{it}) \nabla C(u, v; c) \right]^{d_{it}}
\]

\[
\times \prod_{t=1}^T [p(d_{it} = 0 | v_{it}, \epsilon_{it}, w_{it}, z_{it})]^{1-d_{it}} \prod_{t=2}^T f(w_{it} | w_{it-1}, v_{it-1}, y_{it-1}, a_{it-1}, x_{it})
\]

\[
\times f(w_{it} | v_{it}, x_{it}) \prod_{t=2}^T f(v_{it} | v_{it-1}) f(v_{i1})
\]

where \( u = F(a_{it}^T | v_{it}, \epsilon_{it}, w_{it}, x_{it}) \), \( v = p(d_{it} = 1 | v_{it}, \epsilon_{it}, w_{it}, z_{it}) \) and \( \nabla C(\cdot, \cdot; \cdot) \) is the first derivative of the conditional copula with respect to the first argument. As the model is fully specified, I can write the likelihood function in closed form.

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

\[
C(u, v; c) = G(F(a_{it}^T | v_{it}, \epsilon_{it}, w_{it}, x_{it}) \! \!, \! p(d_{it} = 1 | v_{it}, \epsilon_{it}, w_{it}, z_{it}); \rho_c)
\]

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:

\[
\nabla G(u, v; c) = \nabla G(u, v; c) = \Phi \left( \frac{\Phi^{-1}(v) - \rho_c \Phi^{-1}(u)}{\sqrt{1 - \rho_c^2}} \right)
\]

Substituting the resulting expression for \( \nabla C(u, v; c) \) to the expression above, the likelihood function simplifies to:

\[
f(\mathbf{a}_i^T, \mathbf{v}_i^T, \mathbf{\epsilon}_i^T, w_i^T, \mathbf{Z}_i^T, d_i^T; \mathbf{\mu}) = \prod_{t=1}^T \left[ f(\mathbf{a}_{it}^T | v_{it}, \epsilon_{it}, w_{it}, x_{it}) \Phi \left( \frac{\Phi^{-1}(v) - \rho_c \Phi^{-1}(u)}{\sqrt{1 - \rho_c^2}} \right) \right]^{d_{it}}
\]

\[
\times \prod_{t=1}^T [p(d_{it} = 0 | v_{it}, \epsilon_{it}, w_{it}, z_{it})]^{1-d_{it}} \prod_{t=2}^T f(w_{it} | w_{it-1}, v_{it-1}, y_{it-1}, a_{it-1}, x_{it}) f(w_{i1} | x_{i1})
\]

\[
\times \prod_{t=2}^T f(v_{it} | v_{it-1}) \prod_{t=2}^T f(v_{i1} | v_{i1}) f(v_{i1})
\]

D.1.2 Model restrictions

The economic model implies the following restrictions. Let \( a^+ \equiv \max\{a, 0\} \) and \( a^- \equiv \max\{-a, 0\} \) denote the check function of quantile regression. (Koenker and Bassett (1978)).
Let \( \bar{\mu} \) denote the true value of \( \mu \), and let
\[
 f_i(v_i^T|a_i^T, \varepsilon_i^T, w_i^T, Z_i, d_i^T \bar{\mu}) \tag{D31}
\]
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 \( \bar{\mu} \).

The portfolio rule implies the following restrictions, for all \( l \in \{1, \ldots, L\} \) (removing \( Z_{it} \) for conciseness),
\[
 (\bar{b}_0^p, \ldots, \bar{b}_{K_l}^p) = \arg\min_{(b_0^p, \ldots, b_{K_l}^p)} \sum_{l=1}^{T} \mathbb{E} \left[ \int d_{il} \left\{ G(\tau_l, p(x_{il}); \rho_c) \left( \Lambda^{-1}(a_{il}) - \sum_{k=0}^{K_l} b_k^p(\tau) \phi_k(v_{il}, \varepsilon_{il}, w_{il}, a_{ge_{il}}) \right)^+ 
 + (1 - G(\tau_l, p(x_{il}); \rho_c)) \left( \Lambda^{-1}(a_{il}) - \sum_{k=0}^{K_l} b_k^p(\tau) \phi_k(v_{il}, \varepsilon_{il}, w_{il}, a_{ge_{il}}) \right)^- \right\} f_i(v_i^T|a_i^T, \varepsilon_i^T, w_i^T, Z_i, d_i^T \bar{\mu})dv_i^T \right] \tag{D32}
\]

The participation rule, meanwhile, implies the following restrictions:
\[
 (\bar{b}_0^p, \ldots, \bar{b}_{K_l}^p) = \arg\min_{(b_0^p, \ldots, b_{K_l}^p)} \sum_{l=1}^{T} \mathbb{E} \left[ \int \Lambda \left( \sum_{k=0}^{K_l} b_k^p \phi_k(v_{il}, \varepsilon_{il}, w_{il}, a_{ge_{il}}) \right) \right] d_{il} \times \left( 1 - \Lambda \left( \sum_{k=0}^{K_l} b_k^p \phi_k(v_{il}, \varepsilon_{il}, w_{il}, a_{ge_{il}}) \right) \right)^{1-d_{il}} f_i(v_i^T|a_i^T, \varepsilon_i^T, w_i^T, Z_i, d_i^T \bar{\mu})dv_i^T \tag{D33}
\]

The initial wealth condition implies the following restriction:
\[
 (\bar{b}_0^w, \ldots, \bar{b}_{K_l}^w) = \arg\min_{(b_0^w, \ldots, b_{K_l}^w)} \mathbb{E} \left[ \int \tau_l \left( w_{i1} - \sum_{k=0}^{K_l} b_k^w(\tau) \phi_k(v_{i1}, a_{ge_{i1}}) \right)^+ 
 + (1 - \tau_l) \left( w_{i1} - \sum_{k=0}^{K_l} b_k^w(\tau) \phi_k(v_{i1}, a_{ge_{i1}}) \right)^- \right] f_i(v_i^T|a_i^T, \varepsilon_i^T, w_i^T, Z_i, d_i^T \bar{\mu})dv_i^T \tag{D34}
\]

Turning to the wealth accumulation rule, I have the following restrictions:
\[
 (\bar{b}_0^m, \ldots, \bar{b}_{K_l}^m) = \arg\min_{(b_0^m, \ldots, b_{K_l}^m)} \sum_{l=1}^{T} \mathbb{E} \left[ \int \left( w_{it} - \sum_{k=1}^{K_l} b_k^m \phi_k(v_{it-1}, \varepsilon_{it-1}, w_{it-1}, a_{ge_{il}}) \right)^2 \right] \times f_i(v_i^T|a_i^T, \varepsilon_i^T, w_i^T, Z_i, d_i^T \bar{\mu})dv_i^T \tag{D35}
\]
The variance of the evolution of wealth, meanwhile, follows:

$$\sigma^2 = \frac{1}{T} \sum_{t=2}^{T} \mathbb{E} \left[ \int \left( w_{it} - \sum_{k=1}^{K} b_k^T \tilde{\theta}_k(v_{it-1}, \epsilon_{it-1}, w_{it-1}, \alpha_{it-1}, a_{it}) \right)^2 \right] \times f_i(v_i^T | \alpha_i, \epsilon_i, w_i, Z_i, d_i^T \tilde{\mu}) dv_i^T$$

(D36)

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

$$\lambda^x = \frac{\sum_{t=1}^{T} \mathbb{E} \left[ \int 1 \{ a_{it} \leq \sum_{k=0}^{K} b_k^T \tilde{\theta}_k(v_{it}, \epsilon_{it}, w_{it}, a_{it}) \} f_i(v_i^T | \cdot) dv_i^T \right]}{\sum_{t=1}^{T} \mathbb{E} \left[ \int (a_{it} - \sum_{k=0}^{K} b_k^T \tilde{\theta}_k(v_{it}, \epsilon_{it}, w_{it}, a_{it})) 1 \{ a_{it} \leq \sum_{k=0}^{K} b_k^T \tilde{\theta}_k(v_{it}, \epsilon_{it}, w_{it}, a_{it}) \} f_i(v_i^T | \cdot) dv_i^T \right]}$$

(D37)

and

$$\lambda^y = \frac{\sum_{t=1}^{T} \mathbb{E} \left[ \int 1 \{ a_{it} \geq \sum_{k=0}^{K} b_k^T \tilde{\theta}_k(v_{it}, \epsilon_{it}, w_{it}, a_{it}) \} f_i(v_i^T | \cdot) dv_i^T \right]}{\sum_{t=1}^{T} \mathbb{E} \left[ \int (a_{it} - \sum_{k=0}^{K} b_k^T \tilde{\theta}_k(v_{it}, \epsilon_{it}, w_{it}, a_{it})) 1 \{ a_{it} \geq \sum_{k=0}^{K} b_k^T \tilde{\theta}_k(v_{it}, \epsilon_{it}, w_{it}, a_{it}) \} f_i(v_i^T | \cdot) dv_i^T \right]}$$

(D38)

with similar model restrictions for the other tail parameters.

**D.1.3 Estimation algorithm: details**

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

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

$$f(a_{i1}^T, v_i^T, \epsilon_i, w_i, Z_i, d_i; \tilde{\mu}^{(s)}) = \prod_{t=1}^{T} f(a_{it} | v_{it}, \epsilon_{it}, w_{it}, \bar{x}_{it}; \tilde{\mu}^{(s)}) \Phi \left( \frac{\Phi^{-1}(v) - \rho_c \Phi^{-1}(u)}{\sqrt{1 - \rho_c^2}} \right)$$

$$\times \prod_{t=1}^{T} p(d_{it} = 0 | v_{it}, \epsilon_{it}, w_{it}, \bar{x}_{it}; \tilde{\mu}^{(s)}) \prod_{t=2}^{T} f(w_{it} | w_{it-1}, v_{it-1}, \epsilon_{it-1}, \bar{y}_{it}; \tilde{\mu}^{(s)}) f(w_{it} | v_{it}, \bar{x}_{it}; \tilde{\mu}^{(s)})
$$

$$\times \prod_{t=1}^{T} f(y_{it} | v_{it}; \bar{\theta}) \prod_{t=2}^{T} f(v_{it} | v_{it-1}; \bar{\theta}) f(v_{it}; \bar{\theta})$$

(D39)

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

$$\mathbb{E}^{(s+1)}_{kl} = \arg \min \sum_{l=1}^{N} \sum_{m=1}^{M} \sum_{d_{it}} G(\gamma, p(\bar{x}_i); \rho_c) \left( \Lambda^{-1}(a_{it}) - \sum_{k=0}^{K} b_k^T \tilde{\theta}_k(v_{it}^{(m)}, \epsilon_{it}, w_{it}, a_{it}) \right) +$$

$$+ \left(1 - G(\gamma, p(\bar{x}_i); \rho_c) \right) \left( \Lambda^{-1}(a_{it}) - \sum_{k=0}^{K} b_k^T \tilde{\theta}_k(v_{it}^{(m)}, \epsilon_{it}, w_{it}, a_{it}) \right)$$

(D40)
\[
\begin{align*}
&\left(\bar{\mathbf{b}}_0^{(s+1)}, \ldots, \bar{\mathbf{b}}_K^{(s+1)}\right) = \arg\min_{(\mathbf{b}_0^{(s+1)}, \ldots, \mathbf{b}_K^{(s+1)})} \sum_{t=1}^T \sum_{i=1}^N \sum_{m=1}^M d_{it} \log \Lambda \left( \sum_{k=0}^K b_k^p \varphi_k(v_{it}^{(m)}, \epsilon_{it}, w_{it}, \alpha \gamma_{it}) \right) \\
&\quad + (1 - d_{it}) \log \left( 1 - \Lambda \left( \sum_{k=0}^K b_k^p \varphi_k(v_{it}^{(m)}, \epsilon_{it}, w_{it}, \alpha \gamma_{it}) \right) \right) \tag{D41}
\end{align*}
\]

\[
\begin{align*}
&\left(\bar{\mathbf{m}}_0^{(s+1)}, \ldots, \bar{\mathbf{m}}_K^{(s+1)}\right) = \arg\min_{(\mathbf{m}_0^{(s+1)}, \ldots, \mathbf{m}_K^{(s+1)})} \sum_{t=2}^T \sum_{i=1}^N \sum_{m=1}^M \left( w_{it} - \sum_{k=1}^K b_k^m \tilde{\varphi}_k(v_{it-1}^{(m)}, \epsilon_{it-1}, w_{it-1}, \alpha \gamma_{it-1}, \alpha \gamma_{it}) \right)^2 \tag{D42}
\end{align*}
\]

\[
\begin{align*}
&\left(\bar{\mathbf{w}}_0^{(s+1)}, \ldots, \bar{\mathbf{w}}_K^{(s+1)}\right) = \sum_{i=1}^N \sum_{m=1}^M \arg\min_{(\mathbf{w}_0^{(s+1)}, \ldots, \mathbf{w}_K^{(s+1)})} \tau_i \left( w_{it} - \sum_{k=0}^K b_k^w \tilde{\varphi}_k(v_{it}^{(m)}, \alpha \gamma_{it}) \right) + \\
&\quad + (1 - \tau_i) \left( w_{it} - \sum_{k=0}^K b_k^w \tilde{\varphi}_k(v_{it}^{(m)}, \alpha \gamma_{it}) \right) \tag{D43}
\end{align*}
\]

For the tail parameters, I calculate the following:

\[
\lambda_\alpha^{(s)} = \frac{\sum_{t=1}^T \sum_{i=1}^N \sum_{m=1}^M 1 \{ \alpha_{it} \leq \sum_{k=0}^K \bar{b}_k^p \varphi_k(v_{it}^{(m)}, \epsilon_{it}, w_{it}, \alpha \gamma_{it}) \}}{\sum_{t=1}^T \sum_{i=1}^N \sum_{m=1}^M 1 \{ \alpha_{it} \leq \sum_{k=0}^K \bar{b}_k^p \varphi_k(v_{it}^{(m)}, \epsilon_{it}, w_{it}, \alpha \gamma_{it}) \}}.
\]

with similar updating rules for the other tail parameters.

D.2 Simulation experiments

D.2.1 Baseline model

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:

\[
\begin{align*}
y_{it}^* &= \alpha_0 + \alpha_1 x_{it}^* + \alpha_2 z_{it} + \alpha_3 b_{it} + \sigma(x_{it}^*, z_{it}) u_{it} \\
y_{it} &= y_{it}^* \cdot d_{it} \\
d_{it} &= 1(\beta_0 + \beta_1 x_{it}^* + \beta_2 z_{it} + \beta_3 b_{it} + \beta_4 w_{it} + v_{it} \geq 0) \\
z_{it} &= \gamma_0 + \gamma_1 x_{it-1} + \gamma_2 z_{it-1} + \gamma_3 y_{it-1} + \gamma_4 b_{it} + \sigma(x_{it-1}^*, z_{it-1}, y_{it-1}) \zeta_{it} \\
z_{1i} &= \gamma_0 + \gamma_1 x_{1i}^* + \gamma_2 b_{1i} + \zeta_{1i}
\end{align*}
\]

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.

\[
x_{it} = x_{it}^* + \varepsilon_{it} \\
x_{it}^* = \rho x_{it-1}^* + v_{it}
\]

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

1. (independence of the error terms) \( x_{0i}^* \perp \varepsilon_{it} \perp v_{it} \)

2. (distributional assumptions) \( \varepsilon_{it} \sim i.i.d. N(0,\sigma_{\varepsilon}^2) \), \( v_{it} \sim i.i.d. N(0,\sigma_{v}^2) \) and \( x_{0i}^* \sim i.i.d. N(0,\sigma_{x}^2) \)

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>
<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_{\varepsilon} )</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_{v} )</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 )</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_c )</td>
<td>0.95</td>
</tr>
<tr>
<td>( \beta_3 )</td>
<td>0.25</td>
<td>( \gamma_2 )</td>
<td>0.4</td>
<td>( \rho )</td>
<td>0.95</td>
</tr>
<tr>
<td>( \beta_4 )</td>
<td>-0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: 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 \( \bar{S} \) values \( \hat{\mu} = \frac{1}{\bar{S}} \sum_{s=\bar{S}-S+1}^{\bar{S}} \hat{\mu}^{(s)} \), where I take \( \bar{S} = S/2 \). 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 18 to 22 percent.

The model specification is close to that outlined in the main text. The only difference between the specification outlined and the one considered in this experiment is the number of variables that enter in the approximating function. As an example, the estimating equation for the latent outcome \( y^* \) takes the following form:

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

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

Table A7 presents the results of the simulation experiments. In here, I calculate the average of the 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 A7: 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>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>1.250</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>0.525</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>
<tr>
<td>Monte Carlo means</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<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>
<tr>
<td>Monte Carlo standard deviations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<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.2.2 Model with state dependence

To evaluate the finite sample performance of the estimation procedure in a model with state dependence, I perform a simulation experiment for a dataset with $N = 1,000$ individuals and $T = 6$ periods. The data generating process is similar to the previous simulation experiment, but with additional equations for the initial period that the household is observed, and with the inclusion of the lagged participation indicator in the participation equation. Table A8 below presents results of the simulation experiments. The simulations also seem to be well-behaved in the sense that the average of the derivative effects I compute are close to that of the true values, which are the point estimates I consider.

Table A8: Simulation results, Arellano and Bonhomme (2017) set-up with state dependence

<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><strong>True value</strong></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>1.250</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>0.525</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>
<tr>
<td><strong>Monte Carlo means</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.712</td>
<td>0.993</td>
<td>1.246</td>
<td>1.495</td>
<td>1.753</td>
<td>2.080</td>
<td>2.585</td>
</tr>
<tr>
<td>$x_{it}$</td>
<td>0.424</td>
<td>0.448</td>
<td>0.475</td>
<td>0.502</td>
<td>0.521</td>
<td>0.558</td>
<td>0.610</td>
</tr>
<tr>
<td>$z_{it}$</td>
<td>0.451</td>
<td>0.480</td>
<td>0.508</td>
<td>0.536</td>
<td>0.563</td>
<td>0.599</td>
<td>0.648</td>
</tr>
<tr>
<td>$b_{it}$</td>
<td>0.245</td>
<td>0.259</td>
<td>0.260</td>
<td>0.259</td>
<td>0.262</td>
<td>0.266</td>
<td>0.243</td>
</tr>
<tr>
<td><strong>Monte Carlo standard deviations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.081</td>
<td>0.065</td>
<td>0.060</td>
<td>0.063</td>
<td>0.068</td>
<td>0.088</td>
<td>0.131</td>
</tr>
<tr>
<td>$x_{it}$</td>
<td>0.041</td>
<td>0.041</td>
<td>0.040</td>
<td>0.041</td>
<td>0.038</td>
<td>0.049</td>
<td>0.071</td>
</tr>
<tr>
<td>$z_{it}$</td>
<td>0.028</td>
<td>0.024</td>
<td>0.023</td>
<td>0.024</td>
<td>0.025</td>
<td>0.030</td>
<td>0.044</td>
</tr>
<tr>
<td>$b_{it}$</td>
<td>0.015</td>
<td>0.012</td>
<td>0.012</td>
<td>0.011</td>
<td>0.008</td>
<td>0.010</td>
<td>0.013</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 $T = 6$.  

75
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:

\[
\alpha^*_{it} = g_t(\nu_{it}, \epsilon_{it}, w_{it}, X_{it}, u_{it}) \tag{D45}
\]

\[
\alpha_{it} = \alpha^*_{it} \cdot d_{it} \tag{D46}
\]

\[
d_{it} = \begin{cases} 
1, & \text{if } \tilde{g}_t(\nu_{it}, \epsilon_{it}, w_{it}, q(X_{it})) \leq u_{it} \\
0, & \text{otherwise} \tag{D47}
\end{cases}
\]

\[
w_{it} = h_t(\nu_{it-1}, \epsilon_{it-1}, w_{it-1}, \alpha_{it-1}, X_{it}, \zeta_{it}) \tag{D48}
\]

\[
w_{i0} = \tilde{h}_0(\nu_{i0}, \zeta_{i0}) \tag{D49}
\]

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 (D45) and (D47) 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 (D47). The specification now becomes:

\[
Pr(d_{it} = 1|v_{it}, \epsilon_{it}, w_{it}, \text{age}_{it}, X_{it}) = \Lambda \left( \sum_{k=0}^{K} b_k^p \phi_k(v_{it}, \epsilon_{it}, w_{it}, \text{age}_{it}) + \gamma^p X_{it} \right) \tag{D50}
\]

where \(\Lambda(\cdot)\) is the logistic function. Buchinsky and Hahn (1998) propose to estimate the propensity score with a nonparametric kernel density estimator, as the errors of the participation and outcome equations are the same. However, this leads to a less computationally tractable estimation procedure in the context of the nonlinear reduced form model. Hence, I choose to specify the propensity score with this model. An added advantage is the possibility of calculating extensive margins of income components and wealth with this specification.

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

\[
\max_{(b_0^p,...,b_K^p,\gamma^p)} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{m=1}^{M} d_{it} \log \left[ \Lambda \left( \sum_{k=0}^{K} b_k^p \phi_k(v_{it}, \epsilon_{it}, w_{it}, \text{age}_{it}) + \gamma^p X_{it} \right) \right] \\
+ (1 - d_{it}) \log \left[ 1 - \Lambda \left( \sum_{k=0}^{K} b_k^p \phi_k(v_{it}, \epsilon_{it}, w_{it}, \text{age}_{it}) + \gamma^p X_{it} \right) \right] . \tag{D51}
\]
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:

$$
\min_{(b_1^{it},...b_K^{it})} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{m=1}^{M} d_{it} 1[h_{\tau}(x_{it}) > 0] \left[ h_{\tau}(x_{it}) \left( \Lambda(z_{it}) - \sum_{k=0}^{K} b_k^{it} \phi_k(v_{it}, \epsilon_{it}, w_{it}, a_{ge_{it}}) + \gamma^d(\tau) \right) \right] + (1 - h_{\tau}(x_{it})) \left( \Lambda(z_{it}) - \sum_{k=0}^{K} b_k^{it} \phi_k(v_{it}, \epsilon_{it}, w_{it}, a_{ge_{it}}) + \gamma^d(\tau) \right) - 1 
$$

where $h_{\tau}(x_{it}) = \frac{\tau + p(x_{it}) - 1}{p(x_{it})}$. The role of this function is to “shift” the mass of part of the distribution of portfolio shares that is unobserved, to the observed part of the distribution. 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., $G^-(\tau, p) = \max \left\{ \frac{\tau + p(x_{it}) - 1}{p(x_{it})}, 0 \right\}$.

As the model restrictions and implementation are similar as in the main text, I do not outline them here. I discuss, however, the likelihood function implied by the model.

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

$$
f(\alpha_i^T, v_i^T, \epsilon_i^T, m_{it}, Z_{it}, d_{it}^T; \mu) = \prod_{t=1}^{T} \left[ f(\alpha_{it}^T | v_{it}, \epsilon_{it}, m_{it}, Z_{it}) p(d_{it} = 1 | v_{it}, \epsilon_{it}, m_{it}, Z_{it}) \right]^{d_{it}} \times \prod_{t=1}^{T} [p(d_{it} = 0 | v_{it}, \epsilon_{it}, m_{it}, Z_{it})]^{1-d_{it}} \prod_{t=1}^{T} f(m_{it} | m_{it-1}, v_{it-1}, y_{it-1}, a_{it-1}, Z_{it}) \times f(m_{it} | v_{it}, Z_{it}) \prod_{t=1}^{T} f(y_{it} | v_{it}) \prod_{t=2}^{T} f(v_{it} | v_{it-1}) f(v_{it}) \right) \right) \right) \right) \right) \right) \right) \right)
$$

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

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

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

$$
\nabla C(u, v; c) = \begin{cases} 
\frac{1}{p(x_{it})}, & \text{if } p(x_{it}) > 1 - \tau \\
0, & \text{otherwise}
\end{cases}
$$
Substituting this, the likelihood function above simplifies to:

\[
    f(\alpha_i^T, v_i, \epsilon_i, m_i, Z_i, d_i^T; \mu) = \prod_{t=1}^{T} \left[ f(\alpha_i^T | v_{it}, \epsilon_{it}, m_{it}, Z_{it}) \right]^{d_{it}} \prod_{t=1}^{T} \left[ p(d_{it} = 0 | v_{it}, \epsilon_{it}, m_{it}, Z_{it}) \right]^{1-d_{it}}
    \times \prod_{t=2}^{T} f(m_{it-1}, v_{it-1}, y_{it-1}, \alpha_{it-1}, Z_{it}) f(m_{it} | v_{it}, Z_{it})
    \times \prod_{t=2}^{T} f(\nu_{it} | \nu_{it-1}) f(\nu_{it})
\]  

(D54)

D.3.3 Simulation evidence

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:

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

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

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

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

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

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 equations below describe the evolution of the latent variable.

\[
x_{it} = x_{it}^* + \epsilon_{it}
\]

\[
x_{it}^* = \rho x_{it-1}^* + v_{it}
\]

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

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

2. (distributional assumptions) \( \epsilon_{it} \sim N(0, \sigma_\epsilon^2) \), \( v_{it} \sim N(0, \sigma_v^2) \) and \( x_{it}^* \sim N(0, \sigma_x^2) \).

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.

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 \( \bar{S} \) values \( \hat{\mu} = \frac{1}{\bar{S}} \sum_{s=\bar{S}-\bar{S}+1}^{\bar{S}} \hat{\mu}^{(s)} \),
Table A9: 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>$\sigma_\varepsilon$</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>$\sigma_\zeta$</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></td>
<td></td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>0.1</td>
<td>$\gamma_2$</td>
<td>0.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N$</td>
<td>1000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$T$</td>
<td>6</td>
<td></td>
<td></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 Buchinsky and Hahn (1998) censored quantile regression estimator.

where I take $\tilde{S} = S/2$. 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 18 to 22 percent.

The model specification is close to that outlined in the main text. The only difference between the specification outlined and the one considered in this experiment is the number of variables that enter in the approximating function. As an example, the estimating equation for the latent outcome $y^*$ takes the following form:

$$y^*_{it} = g(x_{it}^*, x_{it}, z_{it}, p_{it}, \tau) = \sum_{k=0}^{K} a_k(\tau) g_k(x_{it}^*, x_{it}, z_{it}) + \gamma y p_{it}$$

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

Table A10 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 A10: Simulation results, Buchinsky and Hahn (1998) 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>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>1.250</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>0.525</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>
<tr>
<td>Monte Carlo means</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<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>
<tr>
<td>Monte Carlo standard deviations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<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 E3: Average derivatives of the participation rule, by persistent component and age

(a) Participation response to $v_{it}$, linear model

(b) Participation response to $v_{it}$, nonlinear model

(c) Participation response to $w_{it}$, linear model

(d) Participation response to $w_{it}$, nonlinear model

Note: Graphs a and c show estimates of the average derivative effect of the participation rule with respect to the persistent component $v_{it}$ and wealth $w_{it}$, respectively, evaluated at values of $v_{it}$ and age$_{it}$ that correspond to their $\tau_{\text{income}}$ and $\tau_{\text{age}}$ percentiles based on the estimated model with the linear earnings process in section 2 of the paper. Graphs b and d show estimates of the average derivative effect of the participation rule with respect to the persistent component $v_{it}$ and wealth $w_{it}$, respectively, evaluated at values of $v_{it}$ and age$_{it}$ that correspond to their $\tau_{\text{income}}$ and $\tau_{\text{age}}$ percentiles based on the estimated model with the nonlinear earnings process of Arellano et al. (2017).
Figure E4: Average derivatives of the portfolio rule, by persistent component and age

(a) Portfolio allocation response to $v_{it}$, linear model

(b) Portfolio allocation response to $v_{it}$, nonlinear model

(c) Portfolio allocation response to $w_{it}$, linear model

(d) Portfolio allocation response to $w_{it}$, nonlinear model

Note: Graphs a and c show estimates of the average derivative effect of the portfolio rule with respect to the persistent component $v_{it}$ and wealth $w_{it}$, respectively, evaluated at values of $v_{it}$ and age$_{it}$ that correspond to their $\tau_{\text{income}}$ and $\tau_{\text{age}}$ percentiles based on the estimated model with the linear earnings process in section 2 of the paper. Graphs b and d show estimates of the average derivative effect of the portfolio rule with respect to the persistent component $v_{it}$ and wealth $w_{it}$, respectively, evaluated at values of $v_{it}$ and age$_{it}$ that correspond to their $\tau_{\text{income}}$ and $\tau_{\text{age}}$ percentiles based on the estimated model with the nonlinear earnings process of Arellano et al. (2017).