Uncovering Asset Market Participation from Household Consumption and Income

Veronika Czellar, René Garcia, and François Le Grand*

November 20, 2023

Abstract

We propose an asset pricing model featuring time-varying limited participation in both bond and stock markets and household heterogeneity. Households participate in financial markets with a certain probability that depends on their individual income and on asset market conditions. We use indirect inference to uncover individual asset market participation from individual consumption data and asset prices. Our model very accurately reproduces the proportions of stockholders in the Survey of Consumer Finances

*Czellar: FAIRR Center, SKEMA Business School - Université Côte d'Azur, Campus Grand Paris, 5 quai Marcel Dassault, 92156 Suresnes, France, veronika.czellar@skema.edu. Garcia: Department of Economics, Université de Montréal, Pavillon Lionel-Groulx, 3150 rue Jean-Brillant, Montréal QC H3T 1N8, Canada, rene.garcia@umontreal.ca, Toulouse School of Economics. He is a research fellow of CIRANO and CIREQ. Le Grand: Rennes School of Business, 2 rue Robert d'Arbrissel, 35065 Rennes, France, francois.le-grand@rennes-sb.com, ETH Zurich (Chair of Integrative Risk Management and Economics).
We received helpful comments from Torben Andersen, Abraham Lioui, Christian Gouriéroux, Alain Monfort, Kevin Sheppard, George Tauchen, Raman Uppal, and participants at the EMLYON Research Workshop in May 2015, the Workshop on Methodological Advances in Statistics Related to Big Data in June 2015, the EDHEC finance seminar in November 2015, the Computational and Financial Econometrics Conference in December 2015, the Toulouse Financial Econometrics Conference in May 2016, the 10$^{\text{th}}$ Annual SOFIE meeting in June 2017, the CREST Financial Econometrics seminar in April 2019, and the Finance seminar at the University of Luxembourg in December 2020. The second author thanks the NSERC (RGPIN-2018-04891), the SSHRC (Insight 435-2019-1216), and the FQRSC (2020-SE2-269954) research grant agencies for their financial support as well as the ANR for support under the Investissements d'Avenir program, grant ANR-17- EURE-0010 and the project ANR-17-CE03-0010 - LONGTERMISM: Evaluation des investissements ultra-longs.

over three-year intervals, provides a reasonable estimate of stock market participation costs, and is able to price characteristic-based stock portfolios with the top decile of households identified as stockholders.

Keywords: limited participation, household-level consumption distribution, heterogeneity, indirect inference, marginal propensity to consume, stock market participation cost.

JEL Classification: C53, C58, E21, E47, G17, G51.

1 Introduction

The seminal paper of Mehra and Prescott (1985) makes it clear that aggregate consumption growth does not fluctuate enough to explain the equity premium in a model with time-additive utility and frictionless complete markets, unless households are extremely risk averse. While the main stream in the literature has focused on elaborating economically-reasonable preferences for the representative household that would make the consumption-based stochastic discount factor more variable,[1] another fruitful research effort has relaxed the assumption of full insurance against individual income risk and introduced consumer heterogeneity. By imposing certain conditions on individual income processes relative to aggregate income, Constantinides and Duffie (1996) show that asset prices can be supported by an exchange economy equilibrium.[2] Limited participation in financial markets is another important economic reason why household heterogeneity influences asset prices. Not all US households have significant amounts of savings in financial markets, and among those who do, not all trade stocks.[3]

In this paper, we construct an asset pricing model featuring time-varying limited participation in both bond and stock markets, and household heterogeneity. Our model considers an economy populated by a large number of households who face idiosyncratic risks that

[1]The two benchmark models feature habit formation (Campbell and Cochrane, 1999) and recursive preferences à la Epstein and Zin (1989) for the representative household. Recent contributions have combined recursive utility with long-run risk to capture several stylized facts regarding the equity premium, its volatility, and the predictability of asset returns (Bansal and Yaron 2004; Bonomo et al. 2011).

[2]An essential feature is that labor income must be a unit-root process with innovations that become more volatile during aggregate downturns. Mankiw (1986) introduces the idea that aggregate shocks and the volatility of idiosyncratic shocks are negatively related. Storesletten et al. (2007) add life-cycle effects and capital accumulation to the Constantinides and Duffie (1996) model.

[3]See supporting references in Guvenen (2009). Favilukis (2013) provides percentages of stockholders over time.

affect their income. Riskless bonds and risky stocks are available but not all households participate in financial markets.

We model financial market participation as a two-stage risk. A first shock makes households enter or quit the financial markets. When outside the financial markets, households cannot save and are hand-to-mouth consumers. When participating in financial markets, households trade bonds – and bonds only. These participating households face a second shock that drives their stock market participation. If the shock is positive, households can hold stocks in addition to bonds. The probabilities of participating in the two financial markets depend on households' individual income and on asset market conditions.

We endow the nonparticipants, bondholders, and stockholders with different preference and shock-exposure parameters and rely on estimation to uncover the probabilities, at each date, of each household belonging to the three groups. The presence of stochastic volatility, an important and well-established feature of stock-return dynamics, means that the likelihood function cannot be available in closed form. In addition, the substantial heterogeneity in individual consumption data may hinder the numerical computation of the likelihood function.

We propose an indirect inference (Gouriéroux, Monfort, and Renault 1993; Smith 1993) method to estimate the parameters of our model. Indirect inference is a simulation-based estimation method that is increasingly used in the financial economics literature,[4] and in the macroeconomics literature with household heterogeneity.[5] This estimation method is particularly well-suited to problems where the structural model – our heterogenous-household model with limited participation in financial markets – is hard to estimate with standard methods such as maximum likelihood (ML), but easy to simulate.

We estimate the model using quarterly individual consumption and income data provided in the Consumer Expenditure Survey (CEX) conducted in the United States, together with the returns on the three-month Treasury bill (T-bill) rates and the S&P 500 equity index. The CEX is a rotating panel of households that collects consumption data over a maximum of four consecutive quarters for each household. We consider all available data on consumption and income for each quarter. A noteworthy feature of the CEX is that it includes limited and imprecise information on the asset holdings of the household – that we choose not use

[4]See, for instance, Alperovych et al. (2021), Calvet and Czellar (2015), Calzolari et al. (2004), Czellar et al. (2007), Garcia et al. (2011), and Sentana et al. (2008), among many others.

[5]See, in particular, Guvenen and Smith (2014) and Berger and Vavra (2015).

to determine participation in financial markets.[6] Prior to 2013, the questions regarding asset holdings were only asked once in the fifth interview.[7] In particular, this would imply a massive shrink in our data sample and would not allow us to analyze time-variation in individual financial market participation – which is at the core our model. Furthermore, the data have limited precision as they do not allow for properly isolating stockholders. As a consequence, participation in financial markets is determined endogenously through the model estimation.[8]

The estimated model provides strong support for our specification. Time-varying limited participation and stochastic volatility are two key ingredients of our consumption-based asset pricing model. While dropping one of them comes at the expense of a poorer empirical fit, including both in an otherwise standard model (standard additive preferences, expected utility model, standard market arrangement) turns out to be sufficient to capture consumption and asset price dynamics.

Furthermore, our estimation yields a limited participation that is consistent with Euler conditions. Our estimation involves jointly estimating individual consumption growths and asset prices at a quarterly frequency without imposing Euler conditions for either bonds or stocks. Propositions 1 and 2 explain that Euler conditions can be tested in a straightforward way by comparing the expected discounted intertemporal marginal rate of substitution, for bonds and stocks, to 1. We show that these conditions hold for the stock index we use in the estimation. In addition, we conduct the same tests on other test portfolios sorted by size, value, investment, long-run reversal, profitability, and industry by maintaining the preference and participation parameters at their optimal estimated values. Overall, the tests for Euler conditions are consistent with each other and confirm that, in the time series, our estimated model properly isolates three categories of households: stockholders, bondholders, and nonparticipants.

Having successfully tested the theoretical soundness of the model, we were able to assess

[6]However, we verify that our model predictions are consistent with these participation data.

[7]The two questions are: "Did you (or any members of your CU) own any securities, such as stocks, mutual funds, private bonds, government bonds, or Treasury notes on the last day of last month?" and if the answer to the latter is positive"Estimated market value of all stocks, bonds, mutual funds, and other such securities held by CU on the last day of the previous month".

[8]In order to keep as many data points as possible, we only discard households that fail to report two consecutive periods of consumption or that report strictly negative revenues. In the first case, since we are estimating consumption growth processes, we need consumption-level data for at least two consecutive periods. In the second case, because probabilities of participating in financial markets depend directly on household revenues, these revenues must be positive. Beyond these two criteria, we discard no further data.

its empirical implications with actual data. The model very accurately replicates the proportions of stockholders reported in the Survey of Consumer Finances over three-year intervals, especially after 1995, where our CEX sample of households increases significantly with respect to the first part of the sample. We also check that this accurate participation rate at the aggregate level results from correct household-level participation prediction. To do so, we compute the share of households that our model classifies as stockholders with respect to the identified stockholders in the CEX – despite the disputable quality of this information (see our discussion in Section 4.4.1). The proportion of correctly predicted stock market participants amounts to 93.4%. The model also allows us to compute the stock market participation cost. We measure it as the transfer, in consumption units, that would make a household participating in bond markets indifferent between participating in the stock market or not. Over the whole period, the cost amounts to approximately 5.65% of the average quarterly consumption, or approximately 351.20 USD (in 2000 values) per year. This is in line with the estimates of Vissing-Jorgensen (2002b) and the values chosen by Gomes and Michaelides (2008) and Favilukis (2013) in their respective calibrations. We also find that the per-period cost exhibits a decreasing temporal pattern from 6.4% in 1984 to 5.0% in 2017. This is consistent with the rise in financial innovation and the fall in transaction costs over the period.

We conclude our empirical analysis by illustrating the implications of limited participation for macroeconomics and asset pricing. First, for macroeconomics, we quantify propensities to consume for the nonparticipants, bondholders, and stockholders. We find that the nonparticipants' propensity to consume (0.14) is about twice as large as that of bondholders and stockholders. This is consistent with the fact that non-participating households are credit constrained and hence unlikely to be able to smooth out their consumption, contrary to participating households who can save through bonds or stocks. This result is also in line with the empirical literature (e.g., Aguiar et al. 2020). Second, for asset pricing, we use individual asset market participation status to test whether the top decile of households identified as stockholders by our estimated model can price the characteristic-based portfolios formerly used in the time series tests of Euler conditions. We employ the model proposed by Lettau et al. (2019). These authors use a single macroeconomic factor, based on growth in the capital share of aggregate income, to price the same cross-section of equity-characteristic portfolios. The prices of risk obtained for the various sets of portfolios are similar to those obtained in Lettau et al. (2019) and are close to each other. Overall, this gives us confidence that the financial market participation uncovered by our model is accurate and that it provides relevant information, both for macroeconomics and for asset pricing.

Our paper is related to four main strands of literature. First, it provides a new and important contribution to the equity premium literature. We show that a model where households with power utility and with reasonable discount rates and elasticities of intertemporal substitution (EIS) can be consistent with the observed equity premium level and volatility. Apart from using individual income and consumption data, the main source of heterogeneity comes from the fact that some households do not participate in financial markets, others only participate in the bond market, while others hold bonds and stocks. This is different from the model in Guvenen (2009) where there are only two categories of households with different EIS but with recursive utility preferences. Our EIS estimates for the three groups are consistent with the results of Vissing-Jorgensen (2002), who provides EIS estimates based on the log-linearized Euler conditions of asset holders and non-asset holders. A main difference is that the separation of the groups is based on the financial information contained in the CEX. Brav et al. (2002) also show that asset prices are consistent with the assumptions of limited-participation and market participant heterogeneity.

Second, our article is related to papers linking individual consumption to income. Heathcote et al. (2010) document a continuous and sizable increase in wage inequality over the 1980-2004 period, but find that access to financial markets has curbed the level and growth of consumption inequality. Primiceri and van Rens (2009) use CEX data on consumption and income to decompose idiosyncratic changes in income into predictable life-cycle changes and transitory and permanent shocks, and estimate the contribution of each to total income and consumption inequality. Guvenen and Smith (2014) use the joint dynamics of individuals' labor earnings and consumption-choice decisions to quantify individual income risk and access to informal insurance against this risk. Gourinchas and Parker (2002) use income and consumption data in sequence to explain the hump-shaped consumption profile in a life-cycle consumption-savings model.

Third, our estimation strategy allows us to estimate the stock market participation cost per period. Our results confirm several findings in the literature. Vissing-Jorgensen (2002b), for example, finds evidence of a positive relationship between nonfinancial income and the probability of stock market participation. This is the central ingredient in determining participation in our model. Furthermore, she estimates that a per period stock market participation cost of just 350 USD is sufficient to explain the choices of 75% of stock market nonparticipants. Our estimation is of the same order of magnitude. In an incomplete-market overlapping generations model, Favilukis (2013) shows that lower participation costs can explain the substantial increase in stock market participation and the smaller increase in consumption inequality observed in the last 30 years. The model also tracks the decline in interest rates and the expected equity premium.

Finally, our indirect inference estimation method has been used recently in two macroeconomic models with heterogeneity. Guvenen and Smith (2014) estimate a structural consumption-savings model where they study the joint dynamics of individuals' labor earnings and consumption-choice decisions. Their auxiliary model is based on approximations of the true structural equations. Berger and Vavra (2015) use indirect inference together with calibration to estimate a heterogeneous agent incomplete markets model with fixed costs of durable adjustment to study durable expenditure dynamics during recessions.

The remainder of the paper is organized as follows. In Section 2, we present our theoretical framework. In Section 3.1, we describe the data that we use. In Section 3.2, we provide an estimation method for the model introduced in Section 2. We discuss estimation results in Section 4 and provide several time series tests of the model as well as a historical estimate of the stock market participation cost. In Section 5, we explore the benefits for macroeconomics and asset pricing of uncovering limited participation. Section 6 concludes. Proofs and calculation details can be found in the Online Appendix.

2 Theoretical setup

We present our theoretical setup in two steps. First, we start with a full-fledged incomplete-market model of asset pricing. Second, we discuss the results of the full-fledged model and construct the reduced-form model that we will estimate. We consider a discrete-time economy, where time is denoted by $t = 1, \dots, T$. In our simulations, a period will be a quarter throughout the paper. The economy is populated by N finite-life households indexed by i, who enter the economy at various dates and live at most until date T', with $T' \leq T$. We propose an asset pricing model featuring heterogeneity and market incompleteness which is in the spirit of the seminal papers of Bewley (1983), Huggett (1993), and Aiyagari (1994)

Households face idiosyncratic risks that affect their income. These idiosyncratic risks may cover several standard aspects of individual risk, such as unemployment risk, productivity risk, or health risk, among many others. These shocks are assumed to be uninsurable and uncorrelated to other risks in the economy (including other households' idiosyncratic risks). The idiosyncratic risks can be neither avoided nor insured against. The asset market is assumed to be incomplete with respect to these risks: there is no tradable asset whose payoffs are contingent on a household's individual status. Several reasons justify market incompleteness for individual risks, such as monitoring costs and moral hazard. Moreover, several empirical studies support market incompleteness and the fact that households bear a larger idiosyncratic risk than that implied by the complete market assumption (Zeldes, 1989; Attanasio and Davis, 1996; Jappelli and Pistaferri, 2006). As will be made explicit in the following, idiosyncratic risk will also affect household participation in asset markets, as well as household preferences.

For each household i participating in the economy on both dates t and $t+1$, we denote the household's per capita non-financial income at date t by ζ_t^i and assume that ζ_{t+1}^i follows:[9]

$$\zeta_{t+1}^i = \xi + \rho\zeta_t^i + \sigma_\zeta z_{t+1}^i \,, \tag{2.1}$$

where $|\rho| < 1$, $\sigma_\zeta > 0$ and $\{z_t^i\}_{t=2,\dots,T}^{i=1,\dots,N}$ are independent, identically distributed (henceforth, IID) standard normal variables. For a household i participating in the survey for the first time at date t, ζ_t^i is supposed to be fixed at the empirical income at date t.

2.1 Asset markets

Asset returns. Two securities can be traded: a riskless bond and a risky stock. Financial security returns are real and expressed in consumption units. First, investing one consumption unit in bonds at date t pays off R_{t+1}^f consumption units at date $t+1$. The return R_{t+1}^f, assumed to be deterministic at date t, refers to the riskless savings and will be called bond savings throughout the paper. Similarly, the investment of one unit of consumption in stocks at date t pays off R_{t+1}^s consumption units at date $t+1$. The stock return is assumed to be stochastic and dependent on the aggregate shock. We assume that the equity premium

[9]Per capita non-financial income is obtained by dividing household non-financial income by household size and is expressed in units of 10,000 US dollars with base year 2000.

$\log R_t^s - \log R_t^f$ has the following dynamics:

$$\log R_t^s = \log R_t^f + \mu\,\omega_t + \sqrt{\omega_t} u_t, \quad t = 1, \ldots, T\,, \tag{2.2}$$

$$\tilde{\omega}_t = \omega + \phi(\tilde{\omega}_{t-1} - \omega) + \sigma v_t, \quad t = 1, \ldots, T\,, \tag{2.3}$$

$$\omega_t = \begin{cases} \tilde{\omega}_t & \text{if } \tilde{\omega}_t \geq \omega \\ \omega^2/(2\omega - \tilde{\omega}_t) & \text{if } \tilde{\omega}_t < \omega \end{cases}, \tag{2.4}$$

where $\omega > 0$, $|\phi| < 1$, $\sigma > 0$, $\{u_t\}_{t=1,\ldots,T}$, and $\{v_t\}_{t=1,\ldots,T}$ are mutually and serially IID standard normal variables that are also independent of income shocks $\{z_t^i\}_{t=2,\ldots,T}^{i=1,\ldots,N}$. The initial hidden component $\tilde{\omega}_0$ is assumed to be generated from the stationary normal distribution with mean ω and variance $\sigma^2/(1-\phi^2)$.

Consistent with the data, the equity premium features stochastic volatility, denoted by ω_t. The transformation of $\tilde{\omega}$ into ω is standard and guarantees that the process ω_t is always non-negative.[10] In the Online Appendix, we tried two alternatives of the transformation in (2.4). Our results remain largely unchanged as the parameter estimates and the proportion of correctly identified stockholders are very close to those with our benchmark specification.

Asset market participation. A combination of idiosyncratic risk and credit constraints endogenously affects asset market participation, and saving decisions more generally. The intuition is rather straightforward: a household that experiences a sufficiently long sequence of bad idiosyncratic shocks will end up facing a binding credit constraint and will therefore be prevented from saving and trading assets. We do not model credit constraints explicitly, but directly assume that idiosyncratic risks affect household asset market participation. Borrowing constraints are well-documented in the literature (see Jappelli 1990 or Grant 2007 for a more recent contribution) and have successfully reconciled individual consumption-saving behavior with data. Another gain of this modeling strategy is that it can account for a possible stock market participation cost (see Vissing-Jorgensen 2002 and Fagereng et al. 2017 for empirical support of such a cost).

[10]Note that the function transforming $\tilde{\omega}$ into ω is continuous and continuously differentiable in ω on $\mathbb{R}_+$ and in $\tilde{\omega}_t$ on $\mathbb{R}$. We can check that computing the derivatives yields: $\frac{\partial \omega_t}{\partial \omega} = \begin{cases} 0 \text{ if } \tilde{\omega}_t \geq \omega \\ \frac{4\omega^2 - 2\omega^2 - 2\omega\tilde{\omega}_t}{(2\omega - \tilde{\omega}_t)^2} \text{ if } \tilde{\omega}_t < \omega \end{cases}$ and $\frac{4\omega^2 - 2\omega^2 - 2\omega\tilde{\omega}_t}{(2\omega - \tilde{\omega}_t)^2} = 0$ if $\tilde{\omega}_t = \omega$, as well as $\frac{\partial \omega_t}{\partial \tilde{\omega}_t} = \begin{cases} 1 \text{ if } \tilde{\omega}_t \geq \omega \\ \frac{\omega^2}{(2\omega - \tilde{\omega}_t)^2} \text{ if } \tilde{\omega}_t < \omega \end{cases}$ and $\frac{\omega^2}{(2\omega - \tilde{\omega}_t)^2} = 1$ if $\tilde{\omega}_t = \omega$.

We model financial market participation as a two-stage risk. A first shock makes households enter or quit the financial markets. When outside the financial markets, households cannot save and are hand-to-mouth consumers. They can neither trade bonds nor stocks and are credit-constrained. When participating in financial markets, households trade bonds – and bonds only. These participating households then face a second shock that drives their stock market participation. If the shock is positive, households can hold stocks in addition to the bond, otherwise they are prevented from trading stocks and can only trade bonds.

Formally, we denote the participation process of household i in the bond market by $\{\tilde{h}_t^i\}_{t\geq 1}$. If $\tilde{h}_t^i = 0$, household i is excluded from the bond market at date t. The bond market participations $\{\tilde{h}_t^i\}_{t\geq 1}$ are independent processes defined on $\{0, 1\}$, such that the probability of $\tilde{h}_t^i = 1$ is:

$$\tilde{p}_t^i = (1 - e^{-\tilde{a}\zeta_t^i})^2 \tag{2.5}$$

if the household's per capita non-financial income $\zeta_t^i > 0$ and zero otherwise. The parameter $\tilde{a} > 0$ is driving the strength of the relationship between household income and bond market participation. Similarly, $\{h_t^i\}_{t\geq 1}$ characterizes household i's willingness to participate in the stock market. The $\{h_t^i\}_{t\geq 1}$ processes are defined in the same way as bond market participation, except that the probability of $h_t^i = 1$ is:

$$p_t^i = (1 - e^{-aR_t^s/R_t^f})^2\,, \tag{2.6}$$

where $a > 0$. However, willingness to participate in the stock market is not a sufficient condition for actual participation, since the household needs to be a bond market participant first. Hence, stock market participation is given by $\{\tilde{h}_t^i h_t^i\}_{t\geq 1}$. We assume that $\tilde{h}_t^i$ and h_t^i are independent, which implies that the probability of household i participating in the stock market is:[11]

$$\tilde{p}_t^i p_t^i = (1 - e^{-\tilde{a}\zeta_t^i})^2 (1 - e^{-aR_t^s/R_t^f})^2\,. \tag{2.7}$$

We can also interpret the $\{h_t^i\}_{t\geq 1}$ process as a conditional stock market participation. Since it is conditional on participating in bond markets ($\tilde{h}_t^i = 1$), household i's participation in the stock market will be driven by h_t^i. More precisely, if $\tilde{h}_t^i h_t^i = 1$, household i will trade stocks.

[11] There is obviously no uniqueness for the functional form in (2.5) or (2.6). To keep the indirect inference estimation tractable, we have chosen a simple functional form featuring an increasing and smooth relationship between income, the R_t^s/R_t^f ratio, and participation probabilities.

Our choice of h_t^i and $\tilde{h}_t^i$ enables us to model conditional stock market participation using two independent processes.[12] Similarly to our robustness checks of the transformation in (2.4), we also try several alternatives to the transformation $x \mapsto (1 - e^{-x})^2$ of equations (2.5) and (2.6) in the Online Appendix and our results remain unchanged to a large extent.[13]

To summarize, there are three possible combinations for asset market participation. First, households trading neither bonds nor stocks will be called nonparticipants (subscript n). Second, households trading only bonds will be called bondholders (subscript b). Finally, for simplicity, households trading both bonds and stocks will be called stockholders (subscript s).[14]

A final noteworthy remark is in order regarding the benefits of our specification for asset market participation. Combined with our estimation method, this will allow us to model participation shocks as hidden processes. The participation – which is difficult to assess in the data and in particular in the CEX (stock market participation is observed in 4.35% of cases starting year 1996 and bond market participation is unavailable) – can therefore be recovered through the estimation of asset prices and household income and consumption data.

2.2 Household preferences

Households are expected utility maximizers endowed with additive time-separable preferences. Households enjoy instantaneous utility by consuming a single good. We assume that household preferences are state-dependent and that they depend on household asset market participation status (and thus ultimately on the idiosyncratic state). See Karni (1990) for a presentation of state-dependent preferences. In other words, nonparticipants, bondholders, and stockholders will be endowed with different utility function parameters, though with the same functional form. The motivation for this is twofold. First, from an empirical perspective, household preferences have been shown to depend on asset market participation. In her

[12] The literature on the determinants of stock ownership reports that the probability of holding stocks increases with wealth and education. So income should be a good proxy given its high correlation with wealth and education. Moreover, the financial wealth data in the CEX cannot be used to perfectly separate stockholders and non-stockholders. See the data Section 3.1 for more details.

[13] The polynomial alternative does not markedly change the results, while the inverse tangent one slightly deteriorates them.

[14] Note that our asset market participation processes could allow for an additional combination where households trade stocks but not bonds. This could be a solution in line with the so-called wealthy hand-to-mouth households of Kaplan and Violante (2014). We leave this route to be explored in future research.

seminal paper, Vissing-Jorgensen (2002) has shown that the EIS differs according to asset market participation. Second, from a theoretical point of view, this variation in preference attributes may be due to non-modeled elements. A typical example of such an element is a credit constraint that may affect consumption smoothing from one period to another.

All households are endowed with a utility function featuring a constant EIS that does not depend on consumption levels but on the household asset market participation status. We denote the utility function by u_x and the inverse of EIS by γ_x, where $x \in \{n, b, s\}$ represents the household's participation status. It will henceforth be referred to as the household's type. Formally, for $x \in \{n, b, s\}$ we have:

$$u_x(c) = \begin{cases} \frac{c^{1-\gamma_x}-1}{1-\gamma_x} \text{ if } \gamma_x \neq 1, \\ \ln(c) \text{ otherwise.} \end{cases} \tag{2.8}$$

Note that the case where $\gamma_x = 1$ in (2.8) is a continuous extension of the case where $\gamma_x \neq 1$.

Similar to the EIS, the household discount factor is constant for each type and is denoted by β_x.

A household of type $x = n, b, s$ at date t will choose its consumption stream $(c^i_{x,\tau})_{\tau \geq t}$, its bond holdings $(b^i_{x,\tau})_{\tau \geq t}$, and its stock holdings $(s^i_{x,\tau})_{\tau \geq t}$, so as to maximize its intertemporal utility function, subject to budget and financial market participation constraints. The maximization program will imply Euler conditions that characterize household behaviors with respect to bond and stock holdings. Before discussing Euler conditions in Section 2.4, we need to specify a general consumption process that will apply distinctly to each group of households in order to estimate the overall model.

2.3 Consumption processes

We aim to quantify the extent to which the features of the above model – market incompleteness, combination of individual and aggregate risks, possible stock market participation costs – are jointly compatible with the asset prices and longitudinal data of individual income and consumption. In Section 3.1, we explain that the CEX survey has very limited data on asset holdings. To proceed with our estimation of the model, we therefore need to specify the dynamics of individual consumption processes. We denote by $\Delta \log c^i_{x,t+1} = \log(c^i_{x,t+1}/c^i_{x,t})$ the process for the consumption growth rate of household i of type x between dates t and

$t+1$. For any date $t = 1, \ldots, T-1$, and any household i of type $x = n, b, s$ at date t, we assume that the consumption process satisfies the following dynamics:

$$\Delta \log c^i_{x,t+1} = \frac{1}{\gamma_x} \left(\log(\beta_x) + \log(R^f_{t+1}) + \sigma_x \varepsilon^i_{t+1} + \kappa_x \omega_{t+1} + \psi_x \sqrt{\omega_{t+1}} u_{t+1} \right), \tag{2.9}$$

where $\{\varepsilon^i_t\}$ are IID standard normal variables, mutually and serially independent of previously introduced shocks, whether aggregate such as $\{u_t\}$ and $\{v_t\}$ or idiosyncratic such as $\{z^i_t\}$.

The consumption growth rate of a household i between dates t and $t+1$ in equation (2.9) depends on the asset market participation status $x = n, b, s$ and is the sum of five terms: (a) a constant pure time preference term, related to β_x, (b) a term related to the riskless interest rate between t and $t+1$ characterizing the consumption opportunity cost, (c) an idiosyncratic risk effect related to the shock ε^i_{t+1}, and (d) two terms related to aggregate risk u_{t+1} and its stochastic variance ω_{t+1}.[15] The magnitude of all these effects is affected by a scaling factor that is type-specific. Estimating these different factors will enable us to determine how household types are affected by idiosyncratic and aggregate risks as a function of asset market participation status.

2.4 Euler conditions

In Section 2.1, asset market participation was determined by an exogenous process. However, it is possible to compute household Euler conditions – for all assets, whether they are traded or not – and check whether these Euler conditions are consistent with household asset market participation status. For instance, we can compute the bond and stock Euler conditions of nonparticipants and test whether these equations are consistent with their non-participation in bond and stock trading. We can also verify whether the Euler conditions of stockholders are consistent with the fact that they trade both assets. Furthermore, the specification of the consumption growth rate in equation (2.9) will provide a parametric solution for testing Euler conditions.

[15] We verify in the Online Appendix that our results are robust to measurement errors. More precisely, the standard normal increments ε^i_{t+1} are replaced with $(1-B^i_{t+1})\varepsilon^i_{t+1} + B^i_{t+1} v^i_{t+1}$, where $\{B^i_{t+1}\}$ are IID Bernoulli variables with $\mathbb{P}(B^i_{t+1} = 1) = \epsilon$ and v^i_{t+1} are Student t_3 variables. Our estimation method can withstand measurement errors up to $\epsilon = 20\%$ and considering contaminated Gaussian errors does not improve the accuracy of uncovered stock participation. Overall, the Gaussian distribution is a legitimate assumption for the errors.

We start with the bond Euler conditions. For a household i of type $x = n, b, s$ at date t, the bond Euler equation compares the quantity $\beta_x R^f_{t+1}\mathbb{E}_t\left[u'(c^i_{x,t+1})/u'(c^i_{x,t})\right]$ to 1 (R^f_{t+1} being the bond gross return and $\beta_x u'(c^i_{x,t+1})/u'(c^i_{x,t})$ the stochastic discount factor). If household i trades bonds at date t equation $\beta_x R^f_{t+1}\mathbb{E}_t\left[u'(c^i_{x,t+1})/u'(c^i_{x,t})\right] = 1$ should hold, otherwise inequality $\beta_x R^f_{t+1}\mathbb{E}_t\left[u'(c^i_{x,t+1})/u'(c^i_{x,t})\right] < 1$ should hold. This last inequality reflects the fact that the bond return is not sufficiently high or attractive for household i to purchase them. The following proposition summarizes the constraint imposed by bond holdings.

Proposition 1 (Euler conditions for bonds) *We consider a household of type x.*

- *The bond Euler equation for a bondholder $x = b$ or a stockholder $x = s$ will hold with equality:*

$$\beta_x R^f_{t+1}\mathbb{E}_t\left[u'(c^i_{x,t+1})/u'(c^i_{x,t})\right] = 1 \tag{2.10}$$

  if and only if:

$$E^B_{x,t} \equiv \mathbb{E}_t\left[\exp\left\{\sigma_x^2/2 + (\psi_x^2/2 - \kappa_x)\omega_{t+1}\right\}\right] = 1\,. \tag{2.11}$$

- *The bond Euler inequality for a nonparticipant $x = n$ will hold with strict inequality:*

$$\beta_n R^f_{t+1}\mathbb{E}_t\left[u'(c^i_{n,t+1})/u'(c^i_{n,t})\right] < 1 \tag{2.12}$$

  if and only if:

$$E^B_{n,t} \equiv \mathbb{E}_t\left[\exp\left\{\sigma_n^2/2 + (\psi_n^2/2 - \kappa_n)\omega_{t+1}\right\}\right] < 1\,. \tag{2.13}$$

Proposition 1 enables us to convert individual participation constraints (depending on the household's index i) into conditions on model parameters. We can construct one-sided and two-sided tests to formally verify whether Euler conditions hold with equality or inequality in the dynamics and, therefore, whether our participation model (for bonds in this case) is consistent with the data. These tests are detailed and computed in Section 4.3. Note that due to the presence of stochastic volatility, the conditions are slightly involved. They would be greatly simplified if we assumed constant volatility for the aggregate risk. We could then remove the expectation operator in equations (2.11) and (2.13) and replace the quantity ω_{t+1} by its constant value.

Similarly to the bond market case, we now state our results for stocks. The main difference compared to bonds is the equation for bondholders, which will imply the presence of stock market participation costs. Indeed, in anticipation of our empirical exercise, their Euler equation will be $\beta_x \mathbb{E}_t\left[R^s_{t+1} u'(c^i_{x,t+1})/u'(c^i_{x,t})\right] > 1$. If they trade no stocks, this will reflect the presence of an additional financial market friction, which will be assumed to be the stock market participation cost.[16]

Proposition 2 (Euler conditions for stocks) *We consider a household i of type x.*

- *The stock Euler equation for a stockholder $x = s$ will hold with equality and verify:*

$$\beta_s \mathbb{E}_t\left[R^s_{t+1} u'(c^i_{s,t+1})/u'(c^i_{s,t})\right] = 1 \tag{2.14}$$

  if and only if:

$$E^S_{s,t} \equiv \mathbb{E}_t\left[\exp\left\{\frac{\sigma_s^2}{2} + (\mu + \frac{1}{2} - \psi_s + \frac{\psi_s^2}{2} - \kappa_s)\omega_{t+1}\right\}\right] = 1\,. \tag{2.15}$$

- *The stock Euler equation for a nonparticipant $x = n$ will hold with strict inequality and verify:*

$$\beta_n \mathbb{E}_t\left[R^s_{t+1} u'(c^i_{n,t+1})/u'(c^i_{n,t})\right] < 1 \tag{2.16}$$

  if and only if:

$$E^S_{n,t} \equiv \mathbb{E}_t\left[\exp\left\{\frac{\sigma_n^2}{2} + (\mu + \frac{1}{2} - \psi_n + \frac{\psi_n^2}{2} - \kappa_n)\omega_{t+1}\right\}\right] < 1\,. \tag{2.17}$$

- *The stock Euler equation for a bondholder $x = b$ will reflect the presence of stock market participation cost and verify:*

$$\beta_b \mathbb{E}_t\left[R^s_{t+1} u'(c^i_{b,t+1})/u'(c^i_{b,t})\right] > 1 \tag{2.18}$$

  if and only if:

$$E^S_{b,t} \equiv \mathbb{E}_t\left[\exp\left\{\frac{\sigma_b^2}{2} + (\mu + \frac{1}{2} - \psi_b + \frac{\psi_b^2}{2} - \kappa_b)\omega_{t+1}\right\}\right] > 1\,. \tag{2.19}$$

Proposition 2 handles three different cases. The first two cases are related to stockholders

[16] Our empirical exercise in Section 4.4.2 will be consistent with the presence of a participation cost. However, other financial market imperfections, such as informational frictions for instance, could also be present.

(holding the stock) and nonparticipants (not participating in financial markets). These two conditions are very similar to the bond Euler conditions in Proposition 1. The third case differs from bond trading due to stock market participation costs. Indeed, equation (2.18) states that bondholders would benefit from participating in stock markets – or put differently, that they perceive stock trading as being profitable. However, they do not trade stocks because of stock market participation costs. These costs stem from a financial market imperfection that must be paid before households can trade stocks. Once the costs have been paid, households can trade stocks such that their Euler conditions hold with equality. Conversely, if stock trading is not attractive per se, as for nonparticipants in equation (2.16), then the presence of a participation cost has no effect on household participation decisions. Regardless of the cost, households decide not to trade the stock. As a result, equations (2.14) and (2.16) are actually silent regarding the presence of costs, since they would also hold in the absence of such costs. Only equation (2.18) is informative regarding the presence of costs. Indeed, it could not hold without cost, in which case we would observe an Euler equation with equality, reflecting the fact that bondholders would then trade the stock. We empirically investigate the conditions of Propositions 1 and 2 in Section 4.3.

3 Estimation

3.1 Data

Our data stem from the Consumer Expenditure Survey (CEX) conducted between the first quarter of 1984 and the first quarter of 2017. The CEX contains quarterly data on US households collected via regular surveys. Each household selected is interviewed for a maximum of four consecutive quarters. In each interview, the household gives a detailed account of its consumption, income, and composition. All of our values are normalized in year 2000 US dollars.

We are mainly interested in consumption data and in particular in total household consumption expressed in real terms. Since our model relies on consumption growth rates, we only eliminate household-quarters for which the consumption growth rate does not exist and household-quarters reporting negative revenues. Concretely, we eliminate 9.89% of the original CEX data and the great majority of these 9.89% eliminated data correspond to households participating only during one single quarter in the survey. We eliminate no further observations as we will use a robust estimation method that can deal with outliers in the data, as detailed in Section 3.2. We have $N = 179,964$ households that are occasionally observed over $T = 133$ periods, providing the consumption matrix:

$$\begin{pmatrix} c_1^1 & c_2^1 & \cdots & c_T^1 \\ c_1^2 & c_2^2 & \cdots & c_T^2 \\ \vdots & \vdots & \ddots & \vdots \\ c_1^N & c_2^N & \cdots & c_T^N \end{pmatrix}. \tag{3.1}$$

Note, however, that the data matrix is sparse and embeds many zeros since every household is observed for at most $T' = 4$ consecutive periods. Overall, our dataset includes 415,324 non-zero household-quarter consumption observations.

Besides total consumption, we also use the weights $\{w_t^i\}$ of households, representing their size in the US population. The weights $\{w_t^i\}$ are obtained by dividing the original weights by the average weight of the whole sample.

Our consumption data choice deserves some further comments. We rely on a measure of total consumption expenditures. These expenditures include spending on non-durable goods (such as food), equipment, entertainment, lodging, and vehicles. The first elements, which can be classified as non-durable or small durable goods and services, are flow expenditures. The two last elements, lodging and vehicle expenditures, are durable goods. However, lodging expenditures mainly include rent paid for the primary residence, which is either the actual rent or an estimate for home owners (both values being provided in the CEX). Lodging expenditures are therefore mainly accounted for as flows.[17] Vehicle expenses include purchases of vehicles, which are therefore the only actual durable good expenditures not accounted for as flows, and other vehicle expenses. All these expenditure data are sourced directly from the CEX. Similarly to our choice to retain as many households as possible in our sample, we have decided not to transform data for consumption expenditures and to rely on raw data from the CEX, even though it includes vehicle purchases,[18] which are not accounted for as flows.

[17]Note that these expenditures also include owned dwellings and other lodging expenditures. They can be considered as small durable goods and are therefore close to a flow expenditure.

[18]Accounting for vehicles as flows, although possible, implies some arbitrariness since the stock value of household vehicles, used for flow accounting, is not observed.

We also use the reported income of each household i at time t. The income measure includes wages, salaries, and government transfers (such as unemployment insurance benefits) for all household members, net of taxes and of social contributions.

Financial returns have been computed on a quarterly basis using Shiller's data.[19] The riskless return is the 3-month US T-Bill rate, while the stock return is the S&P 500 return (dividends included). All returns are computed in real terms using the US CPI index. Financial returns cover the same time period as the CEX data.

We make a final comment regarding the CEX information on asset holdings. In the fifth quarter, the survey asks households whether they hold financial assets in four categories: stocks, bonds, mutual funds, and other such securities; US savings bonds; savings accounts; checking accounts, brokerage accounts, and other similar accounts.[20] Vissing-Jorgensen (2002) uses this information when available, to separate stockholders from bondholders in order to estimate their respective EIS. However, relying on this information raises a number of concerns, including limitations in scope and precision. First, since this information is not available for a large fraction of households, we would have to sizably shrink our dataset and to consider only a much smaller number of data points. Second, this information is not very precise and does enable us to precisely identify stockholders. We would therefore need to introduce further assumptions to identify the three types of agents of our model in the data. Finally, the asset information is at most only available once per household and therefore does not exhibit by construction any time variation. This would not enable us to identify the time-varying asset market participation, which is at the core of our model. We therefore follow a different route by using the more reliable information on consumption and income to infer asset holder status in each quarter, which will be checked for consistency with Euler conditions. We endow the nonparticipants, bondholders, and stockholders with different preference and shock-exposure parameters and rely on indirect inference estimation to uncover, for each household and date, the probabilities of belonging to these three participation groups.

[19] See `http://www.econ.yale.edu/~shiller/data.htm`.

[20] The question has been modified from 2013 on and further details are now asked. However, the new questions only correspond to a dataset with limited depth.

3.2 Indirect inference

Directly estimating the structural model defined in (2.2)–(2.9) of Section 2 is challenging. The presence of stochastic volatility, an important and well-established feature of stock-return dynamics, means that the likelihood function is not available in closed form. In addition, since we eliminate no observations from the original CEX dataset, there is substantial heterogeneity in individual consumption growth. Even if volatility is not stochastic, single outliers in the consumption growth data may prevent the numerical computation of the likelihood (see equation (3.5) and below for a detailed explanation). However, the model is clearly easy to simulate. We therefore introduce an indirect inference method (Gouriéroux, Monfort, and Renault, 1993; Smith, 1993), which is a two-step estimation procedure.

In a first step, a simple, and easy to estimate, approximation of the structural model is chosen (called the auxiliary model). The second step of the estimation involves generating pseudo-data under the structural model and computing auxiliary estimates based on these simulated data. In the third and final step, we need to determine the structural parameter estimates that produce simulated auxiliary estimates similar to the empirical auxiliary estimates. In other words, the estimates of the auxiliary model serve as a metric for comparing the distance between the simulated data and actual data. The estimated structural parameters are those for which the simulated data are the closest to actual data (according to the auxiliary-model metric).

Indirect inference leaves us considerable flexibility in terms of choosing the auxiliary model. An important requirement is that the relationship that binds the structural and the auxiliary parameters must identify the respective effects of each structural parameter on at least one auxiliary parameter.[21] We choose an auxiliary model that is robust to model misspecification but that remains simple to estimate.

3.2.1 A robust auxiliary model

The first step of indirect inference estimation involves choosing an auxiliary model that is different from the structural model, but that is easy to estimate by the least squares, maximum likelihood, or a moment-based method. The parameters of the auxiliary model need to capture the parameters of the structural model but do not need to provide consistent

[21] Formally, the mapping must be injective in the vicinity of the true structural parameters. In our case, we will have as many auxiliary parameters as structural parameters and the relationship will be invertible.

estimators of the structural parameters. Again, the auxiliary model mostly serves as a metric between actual and simulated data. We choose an auxiliary model that remains close to the original one, while addressing its main estimation limitations: stochastic volatility and sensitivity to outliers. For the latter aspect, we opt for a robust auxiliary model to ensure that the indirect inference estimators are not too sensitive to outliers and model misspecifications (see Genton and Ronchetti 2003).

The first characteristic of the auxiliary model is the absence of stochastic volatility. The hidden state ω_{t+1}, corresponding to the variance in the equity premium dynamics of equation (2.3), is replaced by an observable ARCH(1) term. More precisely, the set of equations (2.2)–(2.4) specifying the dynamics of the equity premium in the presence of stochastic volatility is replaced by the following two equations:

$$\log R^s_{t+1} - \log R^f_{t+1} = \mu\tilde{\tilde{\omega}}_{t+1} + \sqrt{\tilde{\tilde{\omega}}_{t+1}}\tilde{u}_{t+1}\,, \tag{3.2}$$

$$\tilde{\tilde{\omega}}_{t+1} = a_\omega + b_\omega\left(\log R^s_t - \log R^f_t\right)^2, \tag{3.3}$$

where $a_\omega, b_\omega > 0$ are two constants, and $\{\tilde{u}_{t+1}\}$ are IID Student-t errors with ν degrees of freedom. The reason for replacing the standard Gaussian process u_t by a Student t is to enable the ARCH(1) process to generate kurtosis in the equity premium, as kurtosis is naturally generated in the structural model with stochastic volatility.

The second characteristic of the auxiliary model is that it needs to be robust to outliers. For the consumption growth process of equation (2.9), this has two implications: (i) the hidden state ω_{t+1} is replaced by the observable variance $\tilde{\tilde{\omega}}_{t+1}$ of equation (3.3), and (ii) the Gaussian innovations are changed into Student-t innovations. Formally, the auxiliary consumption growth process at date $t+1$ for a household of type x is expressed as:

$$\begin{aligned}\Delta \log c^i_{x,t+1} =& \frac{1}{\gamma_x}\Big\{\log(\beta_x) + \log(R^f_{t+1}) + \kappa_x\tilde{\tilde{\omega}}_{t+1} + \\ &\psi_x(\log R^s_{t+1} - \log R^f_{t+1} - \mu\tilde{\tilde{\omega}}_{t+1}) + \sigma_x\tilde{\varepsilon}^i_{t+1}\Big\}\,. \end{aligned}\tag{3.4}$$

where $\{\tilde{\varepsilon}^i_{t+1}\}$ are IID Student-t distributed errors with three degrees of freedom, serially independent of $\{\tilde{u}_t\}$. The computational difficulties in the estimation raised by Gaussian innovations can be explained as follows. The estimation involves computing the log-likelihood function, which relies on the conditional probability density function (pdf, henceforth) of consumption growth. We denote this pdf by $f(\Delta \log c^i_{t+1}|y^i_{t+1})$, conditional on the joint observation of equity premium and past income, $y^i_{t+1} = (\log R^s_{t+1} - \log R^f_{t+1}, \zeta^i_t)$. The pdf, unconditional on the participation status, is thus equal to the sum:

$$(1-\tilde{p}^i_t)f(\Delta \log c^i_{n,t+1}|y^i_{t+1}) + \tilde{p}^i_t(1-p^i_t)f(\Delta \log c^i_{b,t+1}|y^i_{t+1}) + \tilde{p}^i_t p^i_t f(\Delta \log c^i_{s,t+1}|y^i_{t+1}). \tag{3.5}$$

Because of the summation in equation (3.5), the log of (3.5) does not cancel out with the exponential of the normal pdf. For some outliers in the log consumption process, the pdf in equation (3.5) can be arbitrarily close to zero for exponentially decreasing tails like the normal distribution. This makes the computation of the log-likelihood function numerically challenging with normal tails, which motivates our choice of Student-t innovations in the auxiliary consumption growth process (3.4).

In a similar vein to making the consumption growth process robust to outliers, we adopt a comparable approach for the revenue process (2.1), which becomes:

$$\zeta^i_{t+1} = \xi + \rho\zeta^i_t + \sigma_\zeta \tilde{z}^i_{t+1}\,, \tag{3.6}$$

where $|\rho| < 1$, $\sigma_\zeta > 0$ and $\{\tilde{z}^i_t\}$ are IID Student-t errors with three degrees of freedom, mutually and serially independent of $\{\tilde{\varepsilon}^i_t\}$ and $\{\tilde{u}_t\}$.

Finally, the robust auxiliary model comprises equations (3.2)–(3.4) and (3.6) and can be estimated using maximum likelihood. Formally, we denote the auxiliary parameter vector by η, with $(x = n, b, s)$:

$$\eta = ((\{\beta_x\}, \{\gamma_x\}, \{\kappa_x\}, \{\psi_x\}, \{\sigma_x\})_{x=n,b,s}, \tilde{a}, a, \xi, \rho, \sigma_\zeta, \mu, a_\omega, b_\omega, \nu)'\,. \tag{3.7}$$

The weighted maximum likelihood estimator of η is defined as:

$$\tilde{\eta} = \arg\max_\eta \mathcal{L}_T(\eta|\Delta C, Y)\,, \tag{3.8}$$

where $\Delta C = \{\Delta \log c_t^i\}_{t=1,\ldots,T}^{i=1,\ldots,N}$ and $Y = \{y_t^i\}_{t=1,\ldots,T}^{i=1,\ldots,N}$ are empirical data, and

$$\mathcal{L}_T(\eta|\Delta C, Y) = \sum_{t=2}^{T}\sum_{i=1}^{N} w_t^i \left[\log f(\Delta \log c_t^i|y_t^i) + \log f(\zeta_t^i|\zeta_{t-1}^i)\right] + \sum_{t=2}^{T} \log f(\log R_t^s - \log R_t^f \,|\, \log R_{t-1}^s - \log R_{t-1}^f)\,. \tag{3.9}$$

Because of the robust specification of the auxiliary model, the computation of the maximum likelihood in estimation (3.8) is computationally reliable and accurate.

3.2.2 The indirect inference methodology

The second step of indirect inference (henceforth, II) involves estimating the structural model parameters. These parameters are chosen such that they generate pseudo-data, which are the closest to empirical data, according to the metric implied by the auxiliary model of Section 3.2.1. More precisely, given the full structural model specified in Section 2, we collect the structural parameters in the following vector:

$$\theta = ((\{\beta_x\}, \{\gamma_x\}, \{\kappa_x\}, \{\psi_x\}, \{\sigma_x\})_{x=n,b,s}, \tilde{a}, a, \xi, \rho, \sigma_\zeta, \mu, \omega, \phi, \sigma)'\,. \tag{3.10}$$

Then, for a given parameter vector θ, pseudo-data are generated at a quarterly frequency from the structural model of Sections 2.1–2.3. These pseudo-data are then used to compute the components of the auxiliary estimator $\tilde{\eta}$ defined in equation (3.8). Finally, the II method selects the parameter vector $\hat{\theta}_{II}$ that minimizes the distance between the auxiliary estimates calculated with simulated and empirical data.

The first step is the simulation of $M \geq 1$ pseudo-data samples. Each simulated sample consists of individual consumption growth rates and revenue data for the whole population, as well as of equity premia at a quarterly frequency. Using the notation ΔC and Y of equation (3.8), the pseudo-data of sample $m = 1, \ldots, M$ are denoted by $[\Delta C^m(\theta), Y^m(\theta)]$, where we make explicit the dependence in the structural parameter vector θ used to simulate data.

The second step involves computing, for each sample m, the simulated score function:

$$H^m(\theta; \tilde{\eta}) = \frac{\partial \mathcal{L}_T}{\partial \eta}(\tilde{\eta}|\Delta C^m(\theta), Y^m(\theta)). \tag{3.11}$$

This score function is the derivative of the log-likelihood of equation (3.9) and is computed using: (i) the auxiliary parameter vector $\tilde{\eta}$ of equation (3.8) (which is estimated on actual empirical data ΔC and Y); and (ii) the pseudo-data sample m generated with the structural parameter vector θ. Using these sample score functions, we can then deduce the aggregate simulated score function, defined as the average of sample scores: $H(\theta) = M^{-1}\sum_{m=1}^{M} H^m(\theta;\tilde{\eta})$.

The last step of the II method is to compute a just-identified score-based II estimator (Gallant and Tauchen, 1996; Gouriéroux, Monfort, and Renault, 1993; Smith, 1993), which minimizes the aggregate score. Formally, the II estimator $\hat{\theta}_{II}$ is defined as:

$$\hat{\theta}_{II} = \arg\min_{\theta} H(\theta)' H(\theta)\,. \tag{3.12}$$

Under the regularity conditions given in Gouriéroux, Monfort, and Renault (1993) and Gouriéroux, and Monfort (1996), an II estimator is consistent with the structural parameter θ and asymptotically normally distributed. In empirical and Monte Carlo applications, we use $M = 5$ pseudo-data samples.[22] According to Gouriéroux, Monfort, and Renault (1993), asymptotic standard errors are multiplied by the factor $\sqrt{1+M^{-1}}$ (as asymptotic variance is multiplied by $(1+M^{-1})$). The choice of $M = 5$ ensures that the loss in standard deviation due to simulations is less than 10% (since $\sqrt{1+5^{-1}} = 1.0954 < 1.1$). Using $M = 10$ would ensure a loss smaller than 5% (since $\sqrt{1+10^{-1}} = 1.0488 < 1.05$); however we would need another year of computational time.[23]

4 Empirical results

In this section, we discuss the results of the II estimation of the full model with limited participation and stochastic volatility. To highlight the role of these two model features, we compare our estimates to those of two other models. The first one is an unlimited participation model, in which all households participate in bond and stock markets at all dates. This model is nested in our general model and corresponds to a constrained specification where

[22] Programs were implemented in C and computation was performed on an Intel® Xeon® CPU E7-4830 v3 @ 2.10GHz (using the Intel icc compiler). The computation time for an II estimation (with $M = 5$) using parallel computing with 45 CPUs was approximately four days. For Monte-Carlo simulations requiring 100 such estimations, we used 96 cores, which reduced the computational time to approximately 6 months.

[23] In the Online Appendix, we assess the finite sample accuracy of our estimation method by Monte Carlo simulation. We check that the parameter estimates and the proportion of correctly identified stockholders are essentially the same for $M = 4$ and $M = 6$ compared to our benchmark case $M = 5$.

we impose $\tilde{h}_t^i = h_t^i = 1$ for all i and t, thereby implying that all households are stockholders. The second model features limited participation but constant volatility (corresponding to the restrictions $\kappa_x = \phi = \sigma = 0$ for $x = n, b, s$). We estimate both constrained models using the same II methodology from Section 3.2.[24] The results of these three estimations are presented in Section 4.1. We then compare the forecasting performances of the three models in Section 4.2. The forecasts are computed for individual consumption growth rates and stock returns. This exercise leads to a unanimous conclusion: both limited participation and stochastic volatility are key to capturing consumption and asset price dynamics.

The two last subsections explore the properties of the full model. Section 4.3 tests the Euler conditions and shows that our estimation is consistent with the interpretation of the three types $x = s, b, n$ as stockholders, bondholders, and nonparticipants, respectively. Finally, Section 4.4 focuses on the asset market participation uncovered by our estimation. The conclusion is that the limited participation and the stock market participation cost are both consistent with their empirical counterparts. The benefits of correctly uncovering individual asset market participation are discussed in Section 5.

4.1 Parameter estimates

We report the results of the empirical II estimation of the limited participation model in the first column of Table 1. The II estimates of the unlimited participation model and of the constant-variance limited participation model are reported in the second and third columns. Finite standard errors computed with the 100 Monte Carlo replicates reported in the Online Appendix, are in parentheses below the estimates. Significance at the 5% level is indicated with stars. We also report in the Online Appendix box plots, as well as the Monte Carlo 95%-coverages associated to Table 1.

We start by discussing the estimates of the discount factor for the three groups $\beta_n, \beta_b, \beta_s$. In the unlimited participation model with stochastic variance, the respective values of the betas are 0.52, 0.97, and 0.89 and they are significant. However, looking at equation (2.9), the effective discount rate is time-varying, equal to $\log(\beta_x) + \kappa_x \omega_{t+1}$, and dependent on the

[24] Maximum likelihood estimation is inapplicable, even in the constant volatility model, as the likelihood function is numerically challenging to evaluate for the reasons presented in the first paragraph of Section 3.2. The auxiliary model for the unlimited participation model is the model from Section 3.2.1 but with $x = b = n = s$. The auxiliary model for the constant volatility model is the model from Section 3.2.1 but with $\tilde{\tilde{\omega}}_t = a_s$, since equity premium volatility in the structural model is constant, $\tilde{u}_t$ is standard normal, and $b_s = \kappa_x = 0$ for $x = n, b, s$.

Table 1: Empirical estimates

Parameter	Symbol	Limited participation	Unlimited participation	Constant volatility
Time discount factor for non-particip. in fin. markets	β_n	0.5181* (0.1053)	–	1.0217* (0.0062)
Time discount factor for bond market particip.	β_b	0.9701* (0.0175)	–	0.9998* (0.0019)
Time discount factor for stock market particip.	β_s	0.8917* (0.0380)	0.7736* (0.0761)	0.9929* (0.0028)
Inverse of EIS for non-particip. in fin. markets	γ_n	2.3552* (1.0759)	–	0.7370* (0.2267)
Inverse of EIS for bond market particip.	γ_b	0.4614* (0.2230)	–	0.4077 (0.2930)
Inverse of EIS for stock market particip.	γ_s	1.5641* (0.5485)	0.5346* (0.1997)	2.4809* (0.7026)
Log cons. increase for non-particip. in fin. markets	κ_n	201.5889* (35.7051)	–	–
Log cons. increase for bond market particip.	κ_b	6.3829 (4.1380)	–	–
Log cons. increase for stock market particip.	κ_s	27.8652* (9.9495)	74.6469* (5.6781)	–
Volat. of indiv. shock for non-particip. in fin. markets	σ_n	0.7255 (0.5340)	–	0.3194* (0.0991)
Volat. of indiv. shock for bond market particip.	σ_b	0.3141* (0.1519)	–	0.2787 (0.2003)
Volat. of indiv. shock for stock market particip.	σ_s	0.3125* (0.1151)	0.0245 0.0767	0.5226* (0.1480)
Income multiplier in $P(\tilde{h}_t = 1)$	$\tilde{a}$	6.6840* (0.7117)	–	7.3250* (0.4169)
Risk premium multiplier in $P(h_t = 1)$	a	1.4272* (0.0155)	–	1.4324* (0.0092)
Constant in income mean	ξ	0.0114* 0.0005	0.0113* (0.0005)	0.0114* 0.0005
Persistence in income mean	ρ	0.9909* 0.0003	0.9909* (0.0002)	0.9909* 0.0003
Income volatility	σ_ζ	0.1798* 0.0003	0.1797* (0.0002)	0.1798* 0.0003
Aggregate shock for non-particip. in fin. markets	ψ_n	−5.2972* (2.3190)	–	0.0268 (0.0253)
Aggregate shock for bond market particip.	ψ_b	−0.1482 (0.1067)	–	0.0068 (0.0239)
Aggregate shock for stock market particip.	ψ_s	−0.6872* (0.3452)	−1.0459 (1.5050)	−0.0091 (0.0424)
Constant in mean equity premium	μ	6.9049* (2.3790)	8.0674 (7.4938)	6.2018* (1.9228)
Log stochastic volatility mean	ω	0.0030* (0.0007)	0.0031* (0.0013)	0.0037* (0.0005)
Persistence in log stochastic volatility	ϕ	0.0444 (0.1124)	−0.2263 (0.2078)	-
Volatility of log stochastic volatility	σ	0.0050* (0.0014)	0.0037* (0.0013)	-

The table reports empirical estimates of the limited participation (first column), unlimited participation (second column), and constant volatility (third column) models. Standard errors are in parentheses. Stars indicate significance at the 5% level.

variance of shocks in stock returns. The value of κ_n is estimated at a very high value of 201.59, showing the high sensitivity of nonparticipants in the asset markets to the aggregate variance since they have no asset hedges. This explains the value of 0.52 for β_n with stochastic variance and the value of 1.02 with constant variance. The former value is consistent with the fact that nonparticipants are very impatient and prefer consuming today to consuming tomorrow.[25] The parameters related to time discounting have higher standard errors for non-participants. Indeed, since they save very little, this makes their intertemporal behavior difficult to estimate, which is reflected in these higher standard errors. The values of κ_x are much smaller for bond- and stockholders, while the values of β_x for bond- and stockholders are closer to each other in the limited participation and constant volatility models. For bondholders, the estimate of β_b with or without variance risk is close to the estimated values obtained in studies with a representative household model and implies a discount rate between 3% and 0%. For stockholders, who directly bear the time-varying variance risk, the discount rate is close to 11% in the model with stochastic variance, falling to 0% when the variance is constant.

Our limited participation model is an expected-utility setting in which the interpretation of parameter γ is a measure of the inverse of the EIS, as is standard in the literature on limited asset market participation. Our estimated values are 0.42, 2.17, and 0.64 for nonparticipants, bondholders, and stockholders, respectively. These values are consistent with the results obtained in the seminal paper of Vissing-Jorgensen (2002), who provides estimates of the EIS based on log-linearized Euler conditions of asset holders and non-asset holders. She finds estimates of 0.3-0.4 for stockholders, 0.8-1.0 for bondholders, and small and insignificantly different from zero for non-stockholders and non-bondholders.[26] The high value of the EIS for bondholders is consistent with a low average interest rate over the estimation period.[27] Overall, the estimated values are consistent with those used and found in the literature on limited asset market participation.[28]

Our model provides interesting information about the idiosyncratic and aggregate shocks

[25] On average, the difference corresponds roughly to $\kappa_n \times \omega = 201.59 \times 0.003 = 0.60$.

[26] The standard deviation of $1/\gamma_n$ is 0.6247, which also makes the EIS non-significantly different from zero.

[27] In Vissing-Jorgensen (2002), the EIS is estimated to be larger, around 1.6, for the top layer of bondholders.

[28] Another interpretation of γ is a parameter of risk aversion, as is standard in the consumption-based asset pricing literature. Our estimated values are consistent with those used in the literature featuring household heterogeneity and endogenous stock market participation, as in this paper (see Krusell et al., 2011 and LeGrand and Ragot, 2018 among others).

for asset market participants and nonparticipants. First, it provides separate estimates of the volatility of the idiosyncratic shocks for the three groups. The volatility σ_x for nonparticipants is higher than for asset holders since they are exposed to more important life disruptions given their precarious economic status, but σ_n is not significant. Bondholders and stockholders bear individual shocks of similar magnitude. Second, it measures the impact of the aggregate shock on the consumption of each category of households. The aggregate shock is measured by the volatility of the stock market, which, in our model, is a latent stochastic process ω_t. To extract the distribution of the hidden ω_t, knowing the equity premium over the entire sample until date T, we use a particle smoother as described in Section 6 of the Online Appendix with $J = 10^4$ and $K = 10^4$. In Figure 1, we plot the sample means of the particle smoothed $\omega_{1:T}$, i.e., the sample means $K^{-1}\sum_{k=1}^{K}\omega_{1:T}^{(k,*)}$. It is clear that the time series of ω_t is counter-cyclical since it is high during the recessions in our estimation period (early 1990s, early 2000s, and 2009 and following years). The series also naturally features the peaks of the 1987 crash and the end of the 90s with the Russian crisis. The nonparticipants in asset markets are the most affected by this aggregate shock, with a coefficient ψ_n of -5.30. We presume that they are affected not through asset returns but through the related economic effects (e.g., unemployment). The stockholders come next with a ψ_s of -0.69 since they directly bear the stochastic volatility shock through stocks. The effect for bondholders ψ_b is not significantly different from zero since bond returns increase when the interest rate decreases in recessions.

The remaining parameters $\tilde{a}$ and a are at the core of the model, determining the probabilities of participation in the bond and stock markets. The parameter $\tilde{a}$ is a multiplier of individual income in the probability of participating in the bond market. The parameter a multiplies the ratio R_t^s/R_t^f in the probability of participating in the stock market. To gauge the reasonableness of the parameters $\tilde{a}$ and a entering the probability functions, we plot the probabilities of $\tilde{h}_t^i = 1$ and $\tilde{h}_t^i h_t^i = 1$ as functions of per capita household income in Figure 2. The graph shows that the probability of being a bondholder reaches approximately one for a quarterly per capita income of $8,000$ USD. Since the probability of stock participation also depends on the ratio R_t^s/R_t^f, we consider three situations in which we consider the ratio to be constant: $R_t^s/R_t^f = 1.5$ (trading stocks is *much* more profitable than trading bonds), $R_t^s/R_t^f = 1$ (trading bonds and stocks are equally profitable), and $R_t^s/R_t^f = 0.5$ (trading bonds is *much* more profitable). For a quarterly per capita income of $10,000$ USD, when

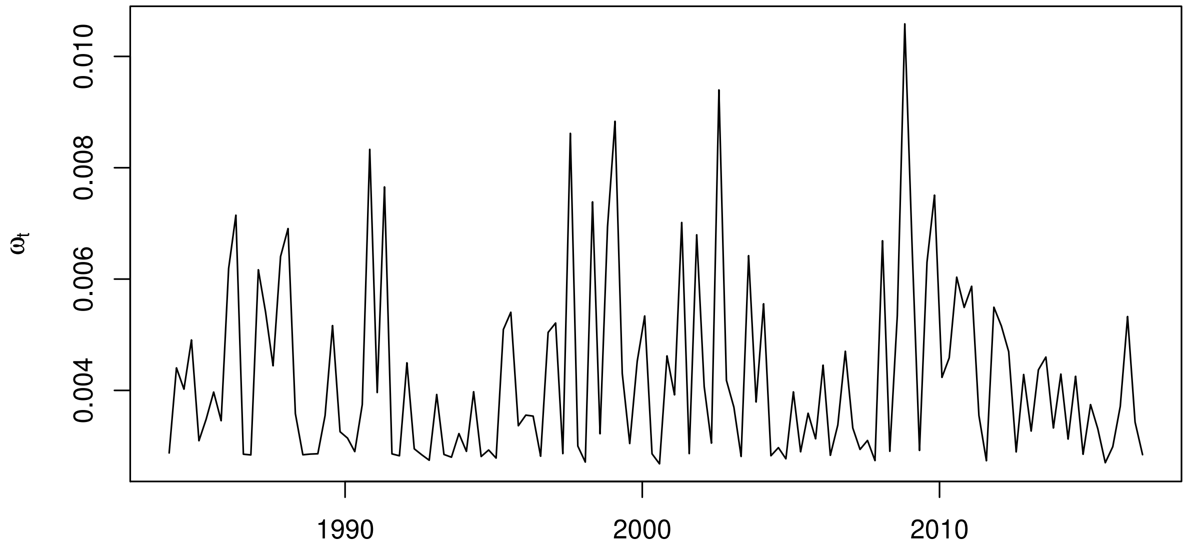

Figure 1: Mean ω_t imputed using a particle smoother.

$R_t^s/R_t^f = 1.5$, the probability of stock participation is close to 0.8, and this probability drops to 0.55 when $R_t^s/R_t^f = 1$ and to 0.25 when $R_t^s/R_t^f = 0.5$. These figures appear to be qualitatively consistent with empirical evidence. At the aggregate level, stock market participation rose during the dot bubble before declining afterwards and a similar movement occurred during the 2008 financial crisis (see Kaustia et al., 2020 for recent evidence). At the individual level, empirical studies also report a positive correlation between inflows in mutual funds and short-term performance of the funds (see Frazzini and Lamont, 2008 for instance). Perhaps more convincingly, we will later see that these participation probabilities imply accurate participation percentages over time. In Section 4.4.2, we examine in detail the time-varying participation in the bond and stock markets as well as the evolution of the stock market participation cost over our sample.

The estimates of the remaining parameters for the income process and the equity premium are as expected. The individual income process is very persistent and quite volatile, while the estimates for the equity premium process in equations (2.3) and (2.4) imply an annualized equity premium of 8% with a volatility of 22%.

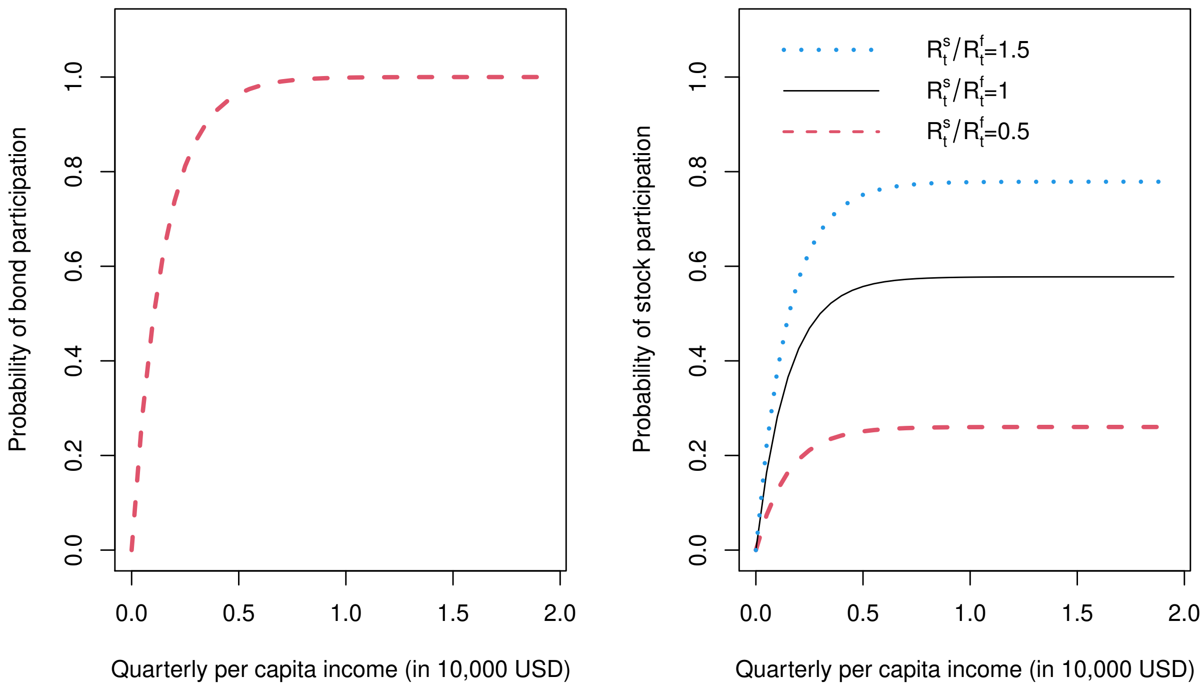

Figure 2: Probability of bond and stock participation as a function of quarterly per capita household income (in 10,000 USD).

4.2 Equity premium and individual consumption growth forecasts

In this section, we compare the three models (limited participation, unlimited participation, and constant volatility) and assess their ability to replicate the conditional distributions of quarterly consumption growth rates. To perform the one-quarter-ahead forecasts at each date t, we use a particle filter as described in Section 5 of the Online Appendix, with $J = 10^4$ particles, and estimate the distribution of the hidden state ω_{t+1} given equity premium observations until date t. We then construct 90% and 95% prediction intervals for the equity premium and for each individual consumption growth rate at date $t+1$. Once we have the prediction intervals for all dates, we count the proportion of times that the actual equity premium and individual consumption growth rates were outside the prediction intervals. We replicate the exercise for the unlimited participation and constant volatility models. The failure rates of conditional 90% and 95% prediction intervals are shown in Table 2. We note that the unlimited participation model predicts the one-quarter-ahead equity premium distribution fairly well but is unable to predict the conditional distributions of the individual quarterly consumption growths. The constant volatility model accurately predicts the conditional distributions of the individual consumption growths but very poorly predicts the distribution of the equity premium. Finally, the limited participation model with stochastic volatility is the only model that is able to accurately and simultaneously predict the conditional distributions of both equity premium and individual consumption growths.

Table 2: Failures of consumption and equity premium predictions

	Failure rates of prediction intervals		
Prediction intervals	Limited part.	Unlimited part.	Constant volatility
90%, Cons. growth	11.94%	70.87%	10.25%
95%, Cons. growth	7.01%	64.32%	6.04%
90%, Equity prem.	9.85%	11.36%	14.39%
95%, Equity prem.	6.82%	8.33%	10.61%

The table reports failure rates of conditional 90% and 95% prediction intervals for individual consumption growth and equity premium processes.

One conclusion is that time-varying participation and stochastic volatility are two key ingredients in a consumption-based asset pricing model. While dropping one of them comes at the expense of a poorer empirical fit, including both in an otherwise standard model (standard additive preferences, expected utility model, standard market arrangement) turns out to be sufficient to capture consumption and asset price dynamics. The remainder of the paper will therefore focus solely on the limited participation model.

4.3 Time series tests of Euler conditions

In this section, we check that our estimation yields a limited participation that is internally consistent – in other words that is consistent with Euler conditions. Our estimation involves jointly estimating individual consumption growths and asset prices at a quarterly frequency without imposing Euler equations for either bonds or stocks. Propositions 1 and 2 explain that Euler conditions can be tested in a straightforward way by comparing the expected discounted intertemporal marginal rate of substitution, for bonds and stocks, to 1. We use the 5% significance level throughout this section.

We first test the Euler conditions (2.11)–(2.19) dynamically, that is at each date t, we test whether:

(T1) $E^B_{s,t}$ is not significantly different from 1 (i.e., stockholders trade bonds);

(T2) $E^B_{b,t}$ is not significantly different from 1 (i.e., bondholders trade bonds);

(T3) $E^B_{n,t}$ is significantly less than 1 (i.e., nonparticipants do not trade bonds);

(T4) $E^S_{s,t}$ is not significantly different from 1 (i.e., stockholders trade stocks);

(T5) $E^S_{b,t}$ is significantly greater than 1 (i.e., bondholders do not trade stocks because of a stock market participation cost);

(T6) $E^S_{n,t}$ is significantly less than 1 (i.e., nonparticipants trade stocks).

Since the above expectations are conditional on observations up to date t, we use the particle filter described in Section 4 of the Online Appendix with $J = 10^6$ particles $\{\omega_t^j\}_{j=1}^J$ to impute the distribution of ω_t at each date t. In Section 8 of the Online Appendix, we report the sample means $\{\overline{E}^y_{x,t}\}_{t=1}^T\}$ obtained using the J particles $\{\omega_t^j\}_{j=1}^J$ for $x = n, b, s$ and $y = B, S$. In addition, we also plot the 5% critical values for $\{\overline{E}^y_{x,t}\}_{t=1}^T\}$ associated with the tests (T1)–(T6).

For bonds, the quantities $\{E^B_{s,t}\}_{t=1,\ldots,T}$ and $\{E^B_{b,t}\}_{t=1,\ldots,T}$ lie within the Euler acceptance regions and (T1) and (T2) hold for all t. The quantities $E^B_{n,t}$ are less than one but not significantly for all t and we conclude that (T3) holds – though not significantly – at all dates t.

For stocks, $E^S_{s,t}$ is not significantly different from 1 and (T4) holds for all t. The quantities $E^S_{b,t}$ are significantly greater than one and (T5) holds for all t, thereby confirming the presence of a stock market participation cost. The quantities $E^S_{n,t}$ are less than one, but not significantly, and (T6) holds, but not significantly, for all t. We conclude that our estimation is consistent with our initial interpretation of household types: stockholders (who also hold bonds) are of type s, bondholders (who do not hold stocks) are of type b, and nonparticipants are of type n.

Although these conditions hold for the stock index we use in the estimation, it is a weak test in the sense that the preference and participation parameters have been estimated jointly with the stock dynamics parameters to optimize the II objective function. A more convincing test is to maintain the preference and participation parameters at their optimal values and to test whether the Euler conditions are satisfied for other stock portfolios. We use equity return data available from Kenneth French's Dartmouth website on 25 size/book-to-market sorted portfolios, 10 long-run reversal portfolios, 25 size/operating profitability portfolios (Size/OP), 25 size/investment portfolios, and 10 industry portfolios. For each of these portfolios, we estimate the four equity premium parameters μ, ω, ϕ, and σ of equations (2.2)–(2.4) driving the dynamics of the equity premium and of its stochastic volatility using the simulated maximum likelihood described in Section 7 of the Online Appendix. We report in Section 9 of the Online Appendix, for each portfolio, the sample means and associated standard errors of $\{E^B_{x,t}\}_{t=1,\ldots,T}$ for bonds and $\{E^S_{x,t}\}_{t=1,\ldots,T}$ for stocks. This allows us to conduct t-tests on Euler conditions. The difference compared to the previous exercise is that the test is conducted unconditionally through time and not per period. The results are clear and strongly support the limited participation model at the 5% significance level. For bond holding, the three Euler conditions hold. For stock holding, the Euler condition for nonparticipants is significantly less than one most of the time. For stockholders, E^S_s is not significantly different from one most of the time. For bondholders, E^S_b is greater than one for most portfolios but not significantly. Overall, the tests for Euler conditions are consistent with the model and confirm that, in the time series, our estimated model properly isolates three categories of households: stockholders, bondholders, and nonparticipants.

4.4 Financial market participation and stock market participation cost

As we saw in Sections 4.2 and 4.3, limited asset market participation, which is a key component of our model, is consistent with the Euler conditions and is thus theoretically sound. In this section, we investigate the empirical relevance of the limited participation implied by our estimation. We proceed in two steps. First, in Section 4.4.1, we compare the participation rates predicted by our model to their external empirical counterparts. Second, in Section 4.4.2, we reveal the stock market participation cost and show its time series evolution.

4.4.1 Financial market participation

At the aggregate level. Using our empirical parameter estimates, for each date t and household i, we calculate the probabilities of belonging to the participation categories n, b, s: $1 - \tilde{p}_t^i$, $\tilde{p}_t^i(1 - p_t^i)$, and $\tilde{p}_t^i p_t^i$, respectively. The sums:

$$NP_t = \sum_i (1 - \tilde{p}_t^i), \;\; B_t = \sum_i \tilde{p}_t^i (1 - p_t^i), \;\; S_t = \sum_i \tilde{p}_t^i p_t^i \,,$$

provide a date-t estimate of the number of nonparticipants in financial markets, bond (only) market participants, and stock market participants, respectively. By normalizing the above sums, we obtain the proportion of participants in the three categories at date t:

$$\frac{NP_t}{NP_t + B_t + S_t}, \;\; \frac{B_t}{NP_t + B_t + S_t}, \;\; \frac{S_t}{NP_t + B_t + S_t} \,,$$

Figure 3 plots the yearly (left panels) and three-year (right panels) proportions of participants in each category (top panels) and the total number of households available (bottom panels) at each date from 1984 to 2017. In addition, we report the stock holdings proportion as provided by the Survey of Consumer Finances (SCF). It is shown as a continuous line whose data is only available at the three-year intervals. The top right panel shows that our model very accurately uncovers the hidden distribution of stock market participants in our sample, in particular for the years after 1995. It should be recalled that our estimation only involves asset prices and consumption data. In particular, it involves no data relating to asset market participation, be it at the individual level through data regarding wealth or asset holdings, or at the aggregate level through asset participation shares. Furthermore, the source of the data we use for the estimation (CEX survey) is distinct from the source for the stock holding shares (SCF). The asset market participation rates are therefore genuinely "uncovered" by the model and its internal structure from individual consumption data and asset prices. Here, we show that implied limited participation is perfectly consistent with empirical data, which is another argument supplementing those of Sections 4.2 and 4.3 and highlighting the relevance of our modeling strategy.

A caveat about Figure 3 is that the fit between model predictions and empirical values is not as good for the pre-1995 period as for the post-1995 period. This is mostly due to the small number of participants in the CEX survey prior to 1995. As can be seen from the

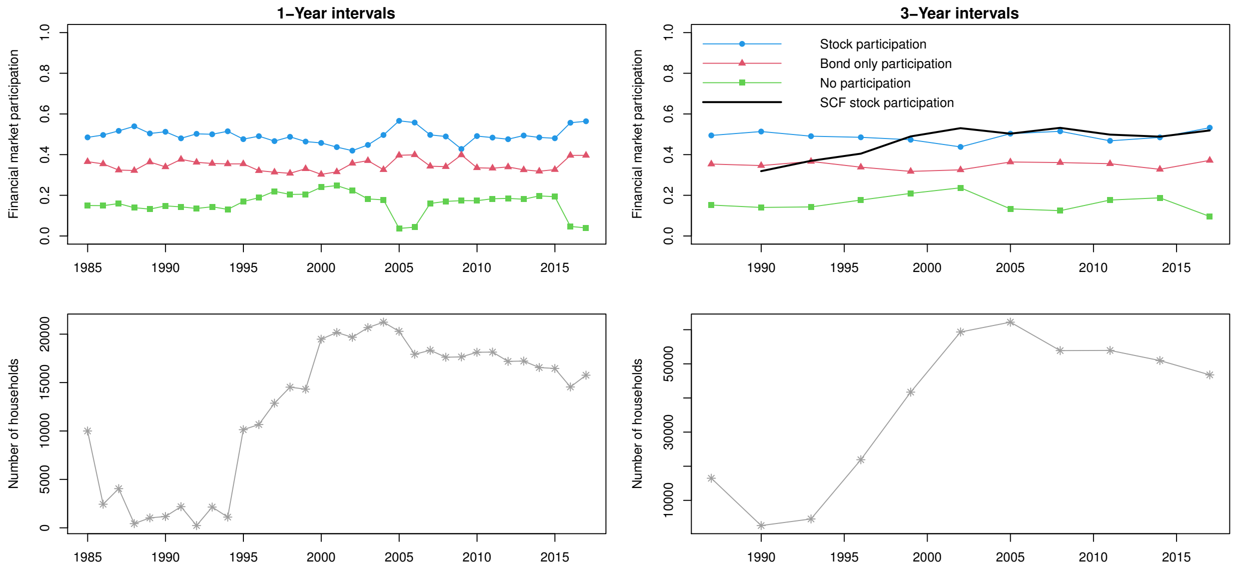

Figure 3: Yearly (left panels) and three-year (right panels) proportions of participants in financial markets (top panels) and the numbers of households available in the survey (bottom panels). In the top panels, participants in stock markets are plotted with filled circle lines, in bond (only) markets with filled triangle lines, and nonparticipants with filled square lines. The continuous line corresponds to the data provided by the SCF.

bottom panels, the number of households participating in the CEX is extremely low in the years prior to 1995. More precisely, the yearly average number of households in the CEX is 3,174 in the period pre-1995, while it amounts to 17,247 in the period post-1995. This implies that our sample is very unbalanced and that approximately 92% of the households in our sample correspond to the period post-1995.

At the individual level. The previous exercise verifies that our model correctly predicts stock market participation at the aggregate level. However, these correct aggregate participation rates could result from the aggregation of incorrect household-level participation predictions. To alleviate this concern, we also assess the accuracy of the estimated stock market participation using the CEX dataset, which provides stock market participation decision at the household level.

However, it should be noted that CEX provides a rather low-quality information regarding stock market participation. Indeed, in addition to under-reporting, the participation question is only asked once in the last quarter of the sample. Furthermore, the survey question is also not completely consistent across our sample and does not provide a very precise information.

Indeed, before 2013, the variable 'SECESTX' provides information on holdings of "stocks, bonds, mutual funds and other such securities", while after 2013, the variable 'STOCKX' answers the question "As of today, what is the total value of all directly-held stocks, bonds, and mutual funds?".

Despite these limitations, we measure the share of households that our model classifies as stockholders with respect to the identified stockholders in the CEX. Concretely, a household i will be said to be participating in the stock market at date t if $\text{SECESTX}_t^i > 0$ or $\text{STOCKX}_t^i > 0$ (depending on the year under consideration). Then, for each of these households, we calculate the estimated probabilities $\tilde{p}_t^i$ and p_t^i using the parameter estimates. If the estimated stock participation probability $\tilde{p}_t^i p_t^i$ is above 0.5 we infer that household i participates in the stock market at date t.[29] Overall, in our sample, the proportion of correctly predicted stock market participants is $\hat{\pi}_S = 93.39\%$. In other words, 93.39% of stock market participants in the CEX data are predicted to be stock participants by our model.[30]

The model is therefore capable of correctly assessing stock market participation. The fact that the model provides participation rates that are consistent with the SCF is the result of the aggregation of empirically-sound household-level participation predictions.

4.4.2 Stock market participation cost

As discussed in Section 4.3, an implication of Proposition 2 and of the estimation results of Table 1 is that the stock market involves a participation cost. The participation shares uncovered by the model are consistent with their empirical counterpart, and we now assess the empirical relevance of the participation cost.

This participation cost can cover various costs, either monetary or non-monetary. For instance, on the monetary side, it may include trading fees or financial intermediation fees. On the non-monetary side, the cost may reflect the effort required to acquire and maintain financial literacy and knowledge about stock markets, as well as monitoring financial news on a regular basis. Participation costs have already been used in a number of papers, such

[29] We run some sensitivity tests and changing the zero threshold for SECESTX and STOCKX to \$ 100 or \$ 1000 or varying the probability threshold of 0.5 have very limited impact.

[30] If we interpret the CEX variables 'SECESTX' and 'STOCKX' as indicating bond holdings instead of stock holdings (which is possible given the loose formulation of the question), our model accurately predicts 97.36% of bondholders. However, it seems that the literature has mostly used this variable as an indicator of stock holdings, hence our choice.

as Allen and Gale (1990, 1994), or more recently Favilukis (2013).[31] There are two types of participation costs. First, households can pay a once-in-a-lifetime participation cost, implying that the main hurdle restricting participation in financial markets is informational. This cost would correspond to acquiring financial literacy. Second, the cost can be recurring, corresponding to trading-related fees. Although our setup is compatible with the presence of one or both of these costs, our estimation focuses on a per-period cost.

We interpret the stock market participation cost as a transfer, in consumption units, that would make a bondholder (not trading stocks) indifferent between participating in stock markets or not participating in them. We take advantage of the structure of our model to obtain a closed-form expression for our cost. As proved in Section 10 of the Online Appendix, the average individual cost at date t, denoted by $\overline{\tau}_t$, can be expressed as follows:

$$\overline{\tau}_t = \frac{1}{\#\{j : \tilde{h}_t^j = 1, h_t^j = 0\}} \sum_{i \in \{j : \tilde{h}_t^j = 1, h_t^j = 0\}} c_t^i \times \left(\frac{\left(1 + \beta_b^{\frac{1}{\gamma_b}} \left(R_{t+1}^f\right)^{\frac{1-\gamma_b}{\gamma_b}} e^{\left(\frac{1-\gamma_b}{\gamma_b}\right)^2 \frac{\sigma_b^2}{2}} \mathbb{E}_t e^{\left(\kappa_b + \frac{1-\gamma_b}{\gamma_b}\frac{\psi_b^2}{2}\right)\frac{1-\gamma_b}{\gamma_b}\omega_{t+1}}\right)^{\frac{1}{1-\gamma_b}}}{\left(1 + \beta_b \beta_s^{\frac{1-\gamma_b}{\gamma_s}} \left(R_{t+1}^f\right)^{\frac{1-\gamma_b}{\gamma_s}} e^{\left(\frac{1-\gamma_b}{\gamma_s}\right)^2 \frac{\sigma_s^2}{2}} \mathbb{E}_t e^{\left(\kappa_s + \frac{1-\gamma_b}{\gamma_s}\frac{\psi_s^2}{2}\right)\frac{1-\gamma_b}{\gamma_s}\omega_{t+1}}\right)^{\frac{1}{1-\gamma_b}}} - 1 \right),$$

where $\{j : \tilde{h}_t^j = 1, h_t^j = 0\}$ denotes the set of bondholders, featuring $\tilde{h}_t^j = 1$ and $h_t^j = 0$. The dynamics of the stochastic volatility ω_{t+1} are defined in equations (2.3)–(2.4).

We then deduce a per-period stock market participation cost $\tilde{\tau}_{c,t}$, expressed as a percentage of the average consumption of bondholders:

$$\tau_{c,t} = \frac{\overline{\tau}_t}{\frac{1}{\#\{j:\tilde{h}_t^j=1,h_t^j=0\}} \sum_{i\in\{j:\tilde{h}_t^j=1,h_t^j=0\}} c_t^i}. \tag{4.1}$$

Figure 4 plots the temporal evolution of the participation cost $\tau_{c,t}$, where $\overline{\tau}_t$ is estimated using 100 Monte Carlo replicates of participation panels of the same size as the empirical dataset. The left panel of this figure considers yearly intervals t, while the right panel considers three-year intervals. Over the whole period, the cost amounts to approximately

[31] An alternative explanation for the presence of stock market participation costs would be to state that stock holdings are forbidden for some households. This is the case in Basak and Cuoco (1998), or Guvenen (2009), for instance.

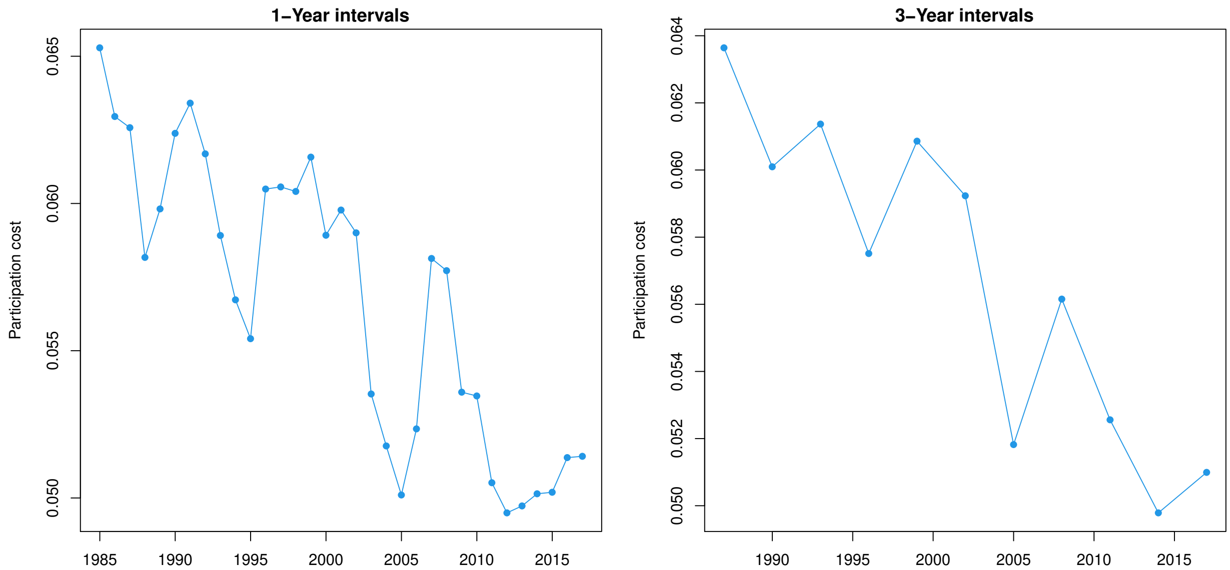

Figure 4: Temporal evolution of the stock participation cost (as a percentage of total consumption) at the yearly level (left panel) and three-year level (right panel).

5.7% of the average quarterly consumption, or approximately 350 USD (in 2000 values) per year. This empirical estimate is in line with the literature. Vissing-Jorgensen (2002b) estimates an annual participation cost of between 150 and 350 USD. Our value is also comparable to the values chosen by Gomes and Michaelides (2008) and Favilukis (2013) in their respective calibrations. Finally, Figure 4 shows that our estimation of the per-period cost exhibits a decreasing temporal pattern, falling from 6.4% in 1984 to 5.0% in 2017. This decrease can be explained by several factors. First, financial innovation, through the development of mutual funds, new securities such as ETFs, or other intermediation vehicles with low transaction fees, may foster stock market participation. Using Swedish data, Calvet et al. (2016b) report the positive impact of financial innovation on stock market participation. Second, the development of the internet, lowering both transaction costs and informational frictions, may also boost stock market participation. Barber and Odean (2002), for instance, report a decrease in financial transaction costs due to the development of online trading. A similar effect is also documented in Bogan (2008).

The stock market participation in our model can be explained by a stock market participation cost that decreases over time, similar to the pattern observed in the data. This is an additional external validation check for our model.

5 The implications of limited participation

The key strength of our model is its ability to successfully uncover the asset market participation status of households using asset prices and consumption data. We explore how the information regarding individual participation status can be used, either in macroeconomics to quantify marginal propensities-to-consume (Section 5.1) or in asset pricing to isolate marginal investors (Section 5.2).

5.1 Implications for macroeconomics

As shown in Sections 4.3 and 4.4, our model generates empirically relevant limited participation at the individual and aggregate levels. We now investigate how this information regarding household market participation can be used in macroeconomics when estimating propensities to consume, which measure how household consumption reacts to income variations. Propensities to consume play an important role in macroeconomics and public economics as they capture the effect of a fiscal stimulus on household consumption. They influence the design of the stimulus (e.g., size of the stimulus and households targeted), which optimizes the aggregate response and avoids the stimulus being saved rather than spent. A robust lesson from the macroeconomic literature is that there is a pervasive and large heterogeneity in marginal propensities-to-consume (see Lewis et al. 2021 for a recent discussion). Here, we propose using the individual asset market participation statuses to accurately estimate marginal propensities-to-consume.

More precisely, we compute the marginal propensities-to-consume in each of the three categories of households we consider: nonparticipants, bondholders, and stockholders. To do so, we collect, for each period, the empirical differences:

$$\left\{\Delta C_t^i\right\} = \left\{C_{t+1}^i - C_t^i\right\} \quad \text{and} \quad \left\{\Delta \zeta_t^i\right\} = \left\{\zeta_{t+1}^i - \zeta_t^i\right\},$$

where C_t^i is the per capita consumption and ζ_t^i is the per capita income at date t of household i (in US dollars). For this exercise, we only consider households with nonzero revenues and with relative revenue increases smaller than 10%. We calculate three groups of weights:

$$\mathcal{NP} = \{1 - \tilde{p}_t^i\},\, \mathcal{B} = \{\tilde{p}_t^i(1 - p_t^i)\},\, \mathcal{S} = \{\tilde{p}_t^i p_t^i\}\,.$$

Using each of the three normalized weights, we then run three distinct regressions in which we calculate the weighted least squares slope estimates of $\{\Delta C_t^i\}$ regressed on $\{\Delta\zeta_t^i\}$. The slope of the regression can be interpreted as the marginal propensity-to-consume out of income and will be denoted by MPCY. We report the MPCYs with t-statistics in parentheses along with the proportions and numbers of households in each of the three categories in Table 3. This table also shows the original number of households in each category before and after truncating the households with zero revenues or those whose relative revenue increases are larger than or equal to 10%.

Table 3: Marginal propensity-to-consume out of income

	Nonparticip.	Bond (only) particip.	Stock particip.
MPCY	0.1383 (2.0567)	0.0723 (9.3321)	0.0715 (9.1745)
Prop. of households	6.5878%	38.8475%	54.5647%
Total nb. of households	16,397	96,692	135,813
Total nb. before truncation	66,826	144,904	203,594

The table reports MPCY values obtained by regressing $\{\Delta C_t^i\}$ on $\{\Delta\zeta_t^i\}$.

Our regression results show that the MPCYs are overall quite low among all groups, which is consistent with the existing literature (see Aguiar et al. 2020, for instance). As explained in Blundell et al. (2008), the MPCY can be interpreted as the marginal propensity to consume after a permanent shock in income. It should be distinguished from the marginal propensity to consume (MPC) after transitory income shocks, which are usually much higher. The main outcome of our regression results in Table 3 is that the MPCYs of non-participating households are twice as large as the MPCYs of participating ones, while no significant difference between the MPCYs of bondholders and stockholders is observed. These results are consistent with the fact that non-participating households are credit constrained and hence unlikely to be able to smooth out their consumption. Their consumption is thus prone to react to income variations. Conversely, participating households save through bonds or stocks and can hence smooth out their consumption. Unsurprisingly, their MPCYs are thus smaller than those of non-participating households. Their MPCYs are also close to each other, reflecting that what matters most for consumption smoothing is the ability to save rather than the composition of saving portfolios.

We conclude with a remark regarding sample selection. We opted for limited and simple criteria and only excluded zero revenues, as well as those whose variation in revenues is larger than or equal to 10%. First, our results remain similar if we change the threshold moderately. Second, Aguiar et al. (2020), who use PSID data (from 1999 to 2015) instead of CEX data, focus on a more restricted population of households. They only consider households whose head age is between 25 and 64, and discard households whose revenues and consumption are below \$ 2,000 per year, as well as those who do not appear at least three times in the sample. Their main restriction criterion is based on the level of consumption and revenues, while ours is based on the variation in revenues. Applied to our sample, their restriction implies a stricter selection since large income variations mostly occur in households experiencing low income in one period. Despite this difference in selection criteria, we obtain comparable MPCY results to those obtained by Aguiar et al. (2020).

The main take-away of this section is that the participation process uncovered by our estimated model is a useful piece of information for estimating marginal propensities to consume. Indeed, the participation process enables us to distinguish between constrained and unconstrained households, which is a key parameter for computing marginal propensities-to-consume.

5.2 Implications for asset pricing

We now use the information regarding participation to characterize marginal investors in stock markets. The intuition behind this is that the consumption and the marginal utility of these investors are likely to be relevant for pricing securities. We consider the same set of Fama-French equity characteristic portfolios as in Section 4.3: 25 size/book-to-market sorted portfolios (Size/BM), 10 long-run reversal portfolios (REV), 25 size/ operating profitability portfolios (Size/OP), and 25 size/investment portfolios (Size/INV). We test whether a factor model similar to Lettau et al. (2019) prices them well on average. Lettau et al. (2019) show that a single macroeconomic factor based on growth in the capital share of aggregate income exhibits significant explanatory power for expected returns across a range of equity characteristic portfolios, with risk price estimates that are of the same sign and similar in magnitude. We follow their methodology with our own consumption and income data in the CEX. We use our estimated model to identify marginal investors, defined as the top 10% and top 5% of households participating in the bond and stock markets, respectively.

The inequality-based asset pricing literature predicts that the income shares of wealthy capital owners (who we associate with the highest-income percentiles of the stock holding households) should vary positively with the national capital share. Therefore, a first step in building the macroeconomic factor is to regress the income ratio of these households on the capital share aggregate variable.[32] We identify the stockholders at each date using our parameter estimates and households' probabilities of participating in the stock market. We denote by N_t the number of all households at date t. We denote by $Y_t = \sum_{i \in N_t} \tilde{p}_t^i p_t^i \zeta_t^i$ the sum of all per capita incomes of stockholders at date t. We subdivide all stockholders into three categories according to their income: bottom 90%, top 10%, and top 5%. We compute $Y_t^J = \sum_{i \in N_t}^{i \in J} \tilde{p}_t^i p_t^i \zeta_t^i$ i.e., the sum of all per capita incomes in the stockholders category J at date t. The right panel in Table 4 reports, for each group, the coefficients of the OLS regression of the income share $\frac{Y_t^J}{Y_t}$ on the capital share KS_t. We provide similar statistics for all households in the left panel of the table.

Table 4: OLS regression of income share on capital share

	All households			Stockholders		
Group J	$\hat{\alpha}_{KS}^J$	$\hat{\beta}_{KS}^J$	R^2	$\hat{\alpha}_{KS}^J$	$\hat{\beta}_{KS}^J$	R^2
$< 90\%$	0.7554 (21.6449)	-0.2447 (-2.7883)	0.0577 (0.0061)	0.8451 (21.7230)	-0.3816 (-3.9006)	0.1070 (0.0002)
90% to 100%	0.2446 (7.0091)	0.2447 (2.7883)	0.0577 (0.0061)	0.1549 (3.9806)	0.3816 (3.9006)	0.1070 (0.0002)
95% to 100%	0.1270 (4.5085)	0.2275 (3.2118)	0.0751 (0.0017)	0.0674 (2.1848)	0.3223 (4.1560)	0.1197 ($6{\cdot}10^{-5}$)

The table reports regression intercepts $\hat{\alpha}_{KS}^J$, slopes $\hat{\beta}_{KS}^J$, and R^2 when regressing income shares $\frac{Y_t^J}{Y_t}$ on capital shares KS_t for three groups: bottom 90%, top 10%, and top 5% income households. For convenience, t-statistics are shown below the regression coefficients and p-values are shown below the R^2s.

The $\hat{\beta}_{KS}^J$ coefficients are positive and significant in the top 10% and 5% categories, sizable (between 0.2 and 0.4), and higher for stockholders than for all households. The predicted values of the income ratios from the regression of the top 10% of shareholders ($\widehat{Y_t^{>90\%}/Y_t} = 0.1549 + 0.3816\, KS_t$) can be used to compute the consumption-based macroeconomic factor

[32] The quarterly labor share level can be found from the data set at `https://www.bls.gov/lpc/special_requests/msp_dataset.zip`. We use non-farm business sector labor share. For details on this measure, see the data description in the online appendix of Lettau et al. (2019).

$F_t^H = \frac{C_t}{C_{t-H}} \left[\frac{\widehat{Y_t^{>90\%}/Y_t}}{\widehat{Y_{t-H}^{>90\%}/Y_{t-H}}} \right]$, where C_t is the total per capita consumption and H is set to four and eight quarters to smooth out quarterly noise.[33]

We can now run our Fama-McBeth analysis for the characteristic-based portfolios. First, for each set of portfolios j (Size/BM, REV, Size/OP and Size/INV), we regress the excess returns $R_{j,t}^{e,H} = R_{j,t}^{H} - R_t^{f,H}$ on the macroeconomic factor F_t^H:

$$R_{j,t}^{e,H} = \alpha_F^j + \beta_F^j F_t^H + \epsilon_t \,, \tag{5.1}$$

to obtain the corresponding loading $\hat{\beta}_F^j$. We then estimate the cross-sectional regression of the average excess returns of characteristic-based portfolios on their estimated loadings:

$$\mathbb{E}[R_t^j - R_t^f] = \alpha_R^H + \beta_R^H \hat{\beta}_F^j + \upsilon_j. \tag{5.2}$$

Table 5 provides the regression results for the four sets of Fama-French portfolios (Size/BM, REV, Size/OP, and Size/INV). The prices of risk obtained for the four sets of portfolios are similar to those obtained in Lettau et al. (2019) and are close to each other. The $\frac{RMSE}{RMSR}$ ratios are also of the same magnitude as those reported in their paper. We infer from these results that the rich stockholders identified by our model qualify as marginal investors for these equity portfolios. It is noteworthy that stock holding is uncovered by our model rather than being directly observed in the data. We use the CEX survey, while Lettau et al. (2019) rely on the data set of Saez and Zucman (2016), which includes detailed income tax return data from 1963 to 2012. This enables them to compute a joint distribution of income and wealth across households, as well as to directly observe stock holding from dividends and capital gains. By way of comparison, the CEX data set spans a shorter period (from 1984 to 2016 in our case) and is known to report relatively inaccurate data for wealthy households, for several reasons: it is not targeted toward wealthy households unlike the SCF, wealth data is missing, and there is under-reporting of wealthy households (see the discussions in Lettau et al., 2019 and Sabelhaus et al., 2014).[34]

[33]Lettau et al. (2019) explain that using the observed income ratio for a percentile group may be too noisy since some of the variation in the ratio across percentile groups is likely to be idiosyncratic and therefore not priced. The predicted ratio based on capital share provides a better measure that isolates the systematic risk component of the income share variation that is priced.

[34]These elements may also explain why our R^2 values are not as high as in Lettau et al. (2019), despite the regression results being otherwise very similar.

Following the example of marginal propensities-to-consume in Section 5.1, this regression illustrates the benefits for asset pricing of uncovering asset market participation.[35]

Table 5: OLS regressions for the Fama-French portfolios

	Equity portfolios			
	Panel A: Size/BM		Panel B: REV	
H	4	8	4	8
$\hat{\alpha}_R^H$	3.24 [2.93,3.55]	2.94 [2.68,3.20]	3.08 [2.86,3.29]	2.92 [2.55,3.30]
$\hat{\beta}_R^H$	2.61 [0.91,4.32]	1.19 [0.40,1.99]	1.93 [0.70,3.16]	1.03 [−0.27,2.33]
R^2	0.28 [0.04,0.57]	0.27 [0.04,0.57]	0.54 [0.12,0.86]	0.23 [0.00,0.72]
$\frac{RMSE}{RMSR}$	0.16	0.19	0.09	0.16
	Panel C: Size/INV		Panel D: Size/OP	
H	4	8	4	8
$\hat{\alpha}_R^H$	3.25 [2.96,3.54]	3.10 [2.84,3.36]	3.05 [2.63,3.47]	2.96 [2.63,3.29]
$\hat{\beta}_R^H$	2.41 [0.61,4.21]	1.59 [0.66,2.53]	1.31 [−1.00,3.61]	1.34 [0.19,2.48]
R^2	0.23 [0.02,0.53]	0.33 [0.07,0.61]	0.05 [0.00,0.32]	0.18 [0.01,0.49]
$\frac{RMSE}{RMSR}$	0.15	0.16	0.19	0.19

This table reports OLS regression estimates of the models $\mathbb{E}[R_t^j - R_t^f] = \alpha_R^H + \beta_R^H \hat{\beta}_F^j + \upsilon_j$, where $\{\hat{\beta}_F^j\}$ are obtained from OLS regressions of the models $R_{j,t}^{e,H} = \alpha_F^j + \beta_F^j F_t^H + \epsilon_t$ for Reversal Fama stocks, with $F_t^H = \frac{C_t}{C_{t-H}} \left[\frac{\widehat{Y_t^{>90\%}/Y_t}}{\widehat{Y_{t-H}^{>90\%}/Y_{t-H}}} \right]$ and $\widehat{Y_t^{>90\%}/Y_t} = 0.1549 + 0.3816\, KS_t$ and $R_{j,t}^{e,H} = R_{j,t}^H - R_t^{f,H}$. Bootstrapped 95% confidence intervals are reported under regression values. Intercept and slope coefficients are multiplied by 100.

6 Conclusion

This paper proposes an asset pricing model based on limited participation and heterogeneity that is able to accurately replicate asset pricing properties and the distribution of individual consumption at a quarterly frequency. Our model is based on Euler conditions for bonds

[35]We also illustrate the implications for asset pricing by directly considering the pricing kernel implied by the model for stockholders (but for a horizon $H = 1$). The results can be found in the Online Appendix.

and stocks and they are not merely used as moment conditions for parameter estimation. We develop an indirect inference method to estimate the parameters of the model. We can draw several conclusions from our estimation. First, limited asset market participation and stochastic volatility are two key ingredients required to jointly replicate individual consumption behaviors and asset prices. Second, testing Euler conditions on the CEX data shows evidence of endogenous bond market participation and of non-zero stock market participation costs. Third, the estimated model successfully uncovers time series participation shares and a stock market participation cost that are both empirically relevant. To conclude, our model succeeds in extracting "hidden" information regarding individual bond and stock market participation statuses from consumption data and asset prices. This is a key feature, both for macroeconomic purposes (because of the isolation of credit-constrained households) and asset pricing purposes (because of the isolation of stockholders). An open question is whether an extended variant of our model would be able to capture the so-called wealthy hand-to-mouth households who are credit-constrained because they only hold illiquid assets (see Kaplan and Violante 2014). We leave this question for future research.

References

Aguiar, M., M. Bils, and C. Boar (2020). Who Are the Hand-to-Mouth? *Working Paper, Princeton University.* 1, 5.1

Aiyagari, S. (1994). Uninsured Idiosyncratic Risk and Aggregate Saving. *Quarterly Journal of Economics* **109**, 659–684. 2

Allen, F., and D. Gale (1990). Incomplete Markets, and Incentives to Set Up an Options Exchange. *Geneva Papers on Risk, and Insurance Theory* **15**, 17–46. 4.4.2

Allen, F., and D. Gale (1994). Limited Market Participation, and Volatility of Asset Prices. *American Economic Review* **84**, 933–955. 4.4.2

Alperovych, Y., D. Cumming, V. Czellar and A. Groh (2021). M&A Rumors about Unlisted Firms. *Journal of Financial Economics*, **142**, 1324–1339. 4

Attanasio, O., and S. Davis (1996). Relative Wage Movements and the Distribution of Consumption. *Journal of Political Economy* **104**, 1227–1262. 2

Bansal, R., and A. Yaron (2004). Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles. *Journal of Finance* **59**, 1481–1509. 1

Barber, B., and T. Odean (2002). Online Investors: Do the Slow Die First? *Review of Financial Studies* **15**, 455–487. 4.4.2

Basak, S., and D. Cuoco (1998). An Equilibrium Model with Restricted Stock Market Participation. *Review of Financial Studies* **11**, 309–341. 31

Berger, D., and J. Vavra (2015). Consumption Dynamics During Recessions. *Econometrica* **83**, 101–154. 5, 1

Bewley, T. (1983). A Difficulty with the Optimum Quantity of Money. *Econometrica* **51**, 1485–1504. 2

Blundell R., L. Pistaferri, and I. Preston (2008). Consumption Inequality and Partial Insurance. *American Economic Review* **98**, 1887–1921. 5.1

Bogan, V. (2008). Stock Market Participation, and the Internet. *Journal of Financial, and Quantitative Analysis* **43**, 191–211. 4.4.2

Bonomo, M., R. Garcia, N. Meddahi, and R. Tedongap (2011). Generalized Disappointment Aversion, Long-Run Volatility Risk, and Asset Prices. *Review of Financial Studies* **24**, 82–122. 1

Brav, A., G. Constantinides, and C. Geczy (2002). Asset Pricing with Heterogeneous Consumers, and Limited Participation: Empirical Evidence. *Journal of Political Economy* **110**, 793–824. 1

Calvet L., C. Célérier, P. Sodini, and B. Vallée (2016). Financial Innovation, and Stock Market Participation. *IFSID working paper* **WP–16-01**. 4.4.2

Calvet L., and V. Czellar (2015). Through the Looking Glass: Indirect Inference Via Simple Equilibria. *Journal of Econometrics* **185**, 343–358. 4

Calzolari G., G. Fiorentini, and E. Sentana (2004). Constrained Indirect Estimation. *Review of Economic Studies* **71**, 945–973. 4

Campbell, J., and J. Cochrane (1999). By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior. *Journal of Political Economy* **107**, 205–251. 1

Constantinides, G., and D. Duffie (1996). Asset Pricing with Heterogeneous Consumers. *Journal of Political Economy* **104**, 219–240. 1, 2

Czellar, V., G. Karolyi, and E. Ronchetti (2007). Indirect Robust Estimation of the Short-Term Interest Process. *Journal of Empirical Finance* **14**, 546–563. 4

Epstein, L., and S. Zin (1989). Substitution, Risk Aversion, and the Temporal Behavior of Consumption, and Asset Returns: A Theoretical Framework. *Econometrica* **57**, 937–969. 1

Fagereng, A., C. Gottlieb, and L. Guiso (2017). Asset Market Participation and Portfolio Choice over the Life Cycle. *Journal of Finance* **72**, 705–750. 2.1

Favilukis, J. (2013). Wealth Inequality, Stock Market Participation, and the Equity Premium. *Journal of Financial Economics* **107**, 740–759. 3, 1, 4.4.2, 4.4.2

Frazzini A. and O. A. Lamont (2008). Dumb Money: Mutual Fund Flows and the Cross-Section of Stock Returns. *Journal of Financial Economics* **88**, 299–322. 4.1

Gallant, A., and G. Tauchen (1996). Which Moments to Match? *Econometric Theory* **12**, 657–681. 3.2.2

Garcia, R., E. Renault, and D. Veredas (2011). Estimation of Stable Distributions by Indirect Inference. *Journal of Econometrics* **161**, 325–337. 4

Genton M.G. and E. Ronchetti (2003). Robust Indirect Inference. *Journal of the American Statistical Association* **98**, 67–76. 3.2.1

Gomes, F., and A. Michaelides (2008). Asset Pricing with Limited Risk Sharing, and Heterogeneous Agents. *Review of Financial Studies* **21**, 415–448. 1, 4.4.2

Gouriéroux, C., and A. Monfort (1996). *Simulation-Based Econometric Methods.* Oxford University Press. 3.2.2

Gouriéroux, C., A. Monfort, and E. Renault (1993). Indirect Inference. *Journal of Applied Econometrics* **8**, S85–S118. 1, 3.2, 3.2.2, 3.2.2

Gourinchas, P.-O., and J. Parker (2002). Consumption over the Life Cycle. *Econometrica* **70**, 47–89. 1

Grant, C. (2007). Estimating Credit Constraints among US Households. *Oxford Economic Papers* **59**, 583–605. 2.1

Guvenen, F. (2009). A Parsimonious Macroeconomic Model for Asset Pricing. *Econometrica* **77**, 1711–1750. 3, 1, 31

Guvenen, F., and A. Smith (2014). Inferring Labor Income Risk, and Partial Insurance from Economic Choices. *Econometrica* **82**, 2085–2129. 5, 1

Huggett, M. (1993). The Risk Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies. *Journal of Economic Dynamics, and Control* **17**, 953–969. 2

Jappelli, T. (1990). Who is Credit Constrained in the U.S. Economy? *Quarterly Journal of Economics* **108**, 219–234. 2.1

Jappelli, T. and L. Pistaferri (2006). Intertemporal Choice and Consumption Mobility. *Journal of the European Economic Association* **4**, 75–115. 2

Kaplan, G., and G. Violante (2014). A Model of the Consumption Response to Fiscal Stimulus Payments. *Econometrica* **82**, 1199–1239. 14, 6

Karni, E. (1990). State-dependent Preferences. In: Eatwell, J., Milgate, M. and Newman, P. (eds) *Utility and Probability*. The New Palgrave. Palgrave Macmillan, London. 2.2

Kaustia M., Conlin A. and N. Luotonen (2023). What Drives Stock Market Participation? The Role of Institutional, Traditional, and Behavioral Factors. *Journal of Banking and Finance* **148**, 106743. 4.1

Krusell, P., Mukoyama T., and A. A. Smith Jr. (2011). Asset Prices in a Huggett Economy. *Journal of Economic Theory* **146**, 812–844. 28

LeGrand F. and X. Ragot (2018). A Class of Tractable Incomplete-Market Models for Studying Asset Returns and Risk Exposure. *European Economic Review* **103**, 39–59. 28

Lettau, M., S. C. Ludvigson, and S. Ma (2019). Capital Share Risk in U.S. Asset Pricing. *The Journal of Finance* **74**, 1753–1792. 1, 5.2, 32, 5.2, 33, 34

Lewis, D., D. Melcangi, and L. Pilossoph (2021). Latent Heterogeneity in the Marginal Propensity to Consume. *Working Paper, Federal Reserve Bank of New York*. 5.1

Mankiw, N. (1986). The Equity Premium, and the Concentration of Aggregate Shocks. *Journal of Financial Economics* **17**, 211–219. 2

Mehra, R., and E. Prescott (1985). The Equity Premium: A Puzzle. *Journal of Monetary Economics* **15**, 145–61. 1

Primiceri, G., and T. van Rens (2009). Heterogeneous Life-Cycle Profiles, Income Risk, and Consumption Inequality. *Journal of Monetary Economics* **56**, 20–39. 1

Sabelhaus, J., D. Johnson, S. Ash, D. Swanson, T. Garner, J. Greenlees, and S. Henderson (2014). Is the Consumer Expenditure Survey Representative by Income? In: Carrol, C. D., Crossley T., and Sabelhaus J. (eds) *Improving the Measurement of Consumer Expenditures*. University of Chicago Press, Chicago. 5.2

Saez, E., and G. Zucman (2016). Wealth Inequality in the United States Since 1913: Evidence from Capitalized Income Yax Data. *Quarterly Journal of Economics* **131**, 519–578. 5.2

Sentana, E., G. Calzolari, and G. Fiorentini (2008). Indirect Estimation of Large conditionally Heteroskedastic Factor Models, with an Application to the Dow 30 Stocks. *Journal of Econometrics* **146**, 10–25. 4

Smith, A. (1993). Estimating Nonlinear Time Series Models Using Simulated Vector Autoregressions. *Journal of Applied Econometrics* **8**, S63–S84. 1, 3.2, 3.2.2

Storesletten K., C. Telmer, and A. Yaron (2007). Asset Pricing with Idiosyncratic Risk, and Overlapping Generations. *Review of Economic Dynamics* **10**, 519–48. 2

Vissing-Jorgensen, A. (2002). Limited Asset Market Participation, and the Elasticity of Intertemporal Substitution. *Journal of Political Economy* **110**, 825–853. 1, 2.1, 2.2, 3.1, 4.1, 27

Vissing-Jorgensen, A. (2002). Towards an Explanation of Household Portfolio Choice Heterogeneity: Nonfinancial Income, and Participation Cost Structures. *NBER Working Paper* **No 8884**. 1, 4.4.2

Zeldes, S. (1989). Optimal Consumption with Stochastic Income: Deviations from Certainty Equivalence. *Quarterly Journal of Economics* **104**, 275–298. 2