Entry and slow-moving capital: using asset markets to infer the costs of risk concentration*

Paymon Khorrami†

June 29, 2018

Abstract

Risk concentration is a major outstanding explanation for crisis dynamics of asset prices and macroeconomic quantities. Apparently, capital flows are slow to correct these crises. By considering costly entry in a canonical limited participation model, I illustrate how asset prices encode costs of risk concentration. These costs must be enormous to match risk premia levels and variability. This finding is robust: auxiliary features that increase risk premia levels mitigate their dynamics, through endogenous entry. In short, either entry costs are large, or limited risk-sharing arises for other reasons. One appealing possibility is extrapolative expectations, which complements entry well.

JEL Codes: D14, G11, G12.

Keywords: Asset Pricing Theory, Limited Participation, Intermediary Asset Pricing, Risk Concentration, Market Segmentation, Entry.

*I would like to acknowledge the gracious financial support of the Macro-Financial Modeling group and the Stevanovich Center for Financial Mathematics. I appreciate very helpful feedback from Ben Charoenwong, Stefano Giglio, Lars Hansen, Zhiguo He, John Heaton, Yunzhi Hu, Stavros Panageas, Jung Sakong, Lawrence Schmidt, Fabrice Tourre, Winnie van Dijk, Pietro Veronesi, Alex Zentefis, and participants in the Economic Dynamics Working Group at University of Chicago.

†University of Chicago Department of Economics and Booth School of Business. Contact: `paymon@uchicago.edu`.

1 Introduction

Does risk concentration explain observed asset price behavior? The standard limited participation economy operates through a risk concentration channel: exogenously designated "participants" hold risky assets, while "non-participants" do not. In equilibrium, all risky assets must be held by a subset of the population, so participants demand a large risk premium for their holdings. Moreover, as negative cash flow shocks accumulate and participants' risk exposures rise, risk premia must also rise to induce additional leverage in participants' portfolios. This channel is powerful – it has the potential to resolve asset pricing puzzles – and it is motivated by empirical regularities such as market segmentation in many risky asset markets.

In the real world, the decision to participate and hold risky assets is endogenous. In this paper, I show that this endogenous participation decision has a substantial impact on equilibrium properties of asset prices. More specifically, I study a limited participation economy with costly entry, whereby non-participants may pay a one-time cost to begin trading in risky asset markets forever after. In doing so, I relax the notion of rigid investor "types" that pervades the limited participation literature. These investor types might be termed "participants" or "experts" or "specialists", depending on the context to which the limited participation framework applies. Agents in my model are ex-ante identical, and ex-post heterogeneity in risk exposures arises endogenously through entry. Because non-participants become increasingly willing to participate in asset markets as risk becomes more concentrated and risk premia become more attractive, entry endogenously limits risk concentration and subdues risk price dynamics.

For example, suppose risky assets perform poorly for a long stretch of time. Because participants hold a leveraged position in the risky asset, their wealth falls relative to non-participants. The same amount of risk becomes more concentrated in the hands of less wealth, and risk premia must rise. Eventually, expected returns rise so far that non-participants find it optimal to pay a cost to enter the market. Risk concentration and expected returns are prevented from rising further. In this way, capital flows prevent asset prices in this economy from exhibiting extreme dynamics.

Consistent with this discussion, my main result is that entry greatly weakens the risk concentration mechanism, and therefore attenuates expected return levels and dynamics. Three findings in this paper support this main result. First, I prove that if entry is not too costly, the risk concentration channel is completely severed as markets become fully integrated and agents share aggregate risk equally. Second, entry costs need to be on the order of 90% of wealth in order to induce enough segmentation to generate empirically realistic asset prices. Third, I show, through a series of robustness exercises, that auxiliary features which raise unconditional risk premia tend to dampen risk premia dynamics, illustrating a trade-off in these models between the level and variability of risk premia.

The basic intuition for these three findings comes from the fact that participation in risky asset markets cannot be too profitable if agents are allowed to choose to enter these markets. Participation provides an extra average return on wealth, which translates into a large present discounted utility gain that outweighs small entry costs. For example, a standard log utility Merton investor attains a portfolio expected return equal to the squared market Sharpe ratio. If she has a $\rho = 2\%$ discount rate and faces a Sharpe ratio of $\eta = 0.1$, her risk-adjusted lifetime gains from holding risky assets are equal to $\frac{1}{2}\eta^2 \int_0^\infty e^{-\rho t} dt = 25\%$ of her wealth (the $\frac{1}{2}$ does a risk-adjustment). Small participation costs cannot dissuade investors from taking these benefits.

To put it slightly differently, if limited participation is the mechanism generating large risk premia, then implied participation costs must be enormous. It is revealing to consider a simple back-of-the-envelope calculation using the Gordon growth formula. With a 2% growth rate and 2% riskless rate, switching from a 0.5% risk premium (frictionless economy) to a 5% risk premium (economy with frictions) suggests a $1 - \frac{1/5\%}{1/0.5\%} = 90\%$ drop in aggregate wealth, consistent with the enormous implied entry costs from my model. One can think of the model as formalizing these back-of-the-envelope calculations in a fully non-linear general equilibrium environment.

Finally, any parameters that make participation more profitable necessarily reduce risk premia dynamics. With higher unconditional levels of expected returns, incentives to participate are greater, which mitigates conditional asset price dynamics arising from the traditional risk concentration channel. This reveals a trade-off between unconditional and conditional asset prices in limited participation models, a trade-off which uniquely arises when entry is endogenous. For example, higher risk aversion increases unconditional risk premia and thereby attenuates risk premia dynamics. Conversely, partial equity-issuance by participants lowers the typical levels of risk premia but increases the chances of extreme risk prices. The conclusion that financial frictions' costs are large is thus robust to certain alternative parameter choices.

I conclude by showing that a limited participation economy augmented with extrapolative expectations is immune to this critique. With extrapolative expectations, agents believe risk premia are high when they are actually low, and vice versa. As a result, entry becomes more procyclical relative to the standard economy, consistent with the data. This absence of entry in bad times also allows for more volatile risk premia. Thus, extending the model in a single direction for which there is growing empirical support (survey data on expectations) brings the model closer to the data in two important ways (procyclical entry, high volatility of risk premia).

The framework analyzed in this paper is most similar to the restricted participation model of Basak and Cuoco (1998), where markets are "segmented" in the sense that some agents are barred from trading in risky asset markets. To that framework, I add an entry margin that prevents markets from being completely segmented.

Technically speaking, the model is quite tractable. One key assumption I make is that entry costs are homogeneous in wealth. As a result, the relative consumption of participants versus non-participants fully characterizes equilibrium dynamics, and solving the model only requires solving a single free boundary problem. In the case of logarithmic utility, I establish existence/uniqueness results and analyze the model and several extensions analytically. For example, I verify the robustness of my main results to exogenous entry, equity-issuance, idiosyncratic shocks, and I cleanly characterize the interactions between extrapolative expectations and endogenous entry.

Related Literature. Recently, the limited participation mechanism has been applied to a variety of contexts, for instance asset markets that rely on arbitrageurs (e.g., futures, commodity, and options markets) or function primarily through intermediaries (e.g., credit markets and asset-backed securities). In models of such markets, because of financial frictions, arbitrageur or intermediary wealth matters for equilibrium dynamics. Asset price dynamics resemble those in the canonical limited participation model, in that risk premia increase dramatically as arbitrageur/intermediary wealth is low. My paper provides a way to understand how large financial frictions or how slow capital flows must be in these markets to generate significant risk concentration (e.g., Duffie (2010b)). See Kyle and Xiong (2001) and Gromb and Vayanos (2002) for models with arbitrage capital; see He and Krishnamurthy (2012) and He and Krishnamurthy (2013) for models with intermediated investment in risky assets; and see Adrian, Etula, and Muir (2014) and He, Kelly, and Manela (2017) for some empirical evidence that intermediary balance sheets are important for asset pricing.

In the macroeconomics literature, limited participation style models have been shown to amplify real dynamics as well. For instance, in Brunnermeier and Sannikov (2014), experts have a more efficient production technology than households. The experts thus produce in equilibrium and they are levered, borrowing from the less productive households to finance capital holdings. If hit by enough negative shocks, experts' net worth is sufficiently impaired that they begin to sell assets to less productive households, albeit slowly, which amplifies fundamental economic shocks. In this way, experts are analogous to the participants in the canonical limited participation model, while households are analogous to non-participants.

The common thread in these literatures is the presence of a risk concentration channel through assumed heterogeneity in investment opportunity sets or ex-ante differences in investor's types. On the other end of the spectrum, the model of Haddad (2014) features ex-ante identical investors and focuses on entry/exit incentives between different investment sectors. In contrast to the no-entry equilibria common in the limited participation literature, he considers a free-entry equilibrium in which investors split between sectors with active and passive investment. He finds that free entry eliminates any asset price dynamics when there are no other dynamic considerations. But neither absence of entry nor free entry are realistic; the costly entry device introduced in this paper provides a natural way to span these extremes and infer costs of financial frictions from asset markets.

My paper is also related to an older and larger literature on the effects of stock market non-participation of stock prices. Beginning with Mankiw and Zeldes (1991), the literature noticed the presence of risk concentration: only a subset of households hold stocks. Furthermore, household-level data shows that stock market investors' consumption is more volatile, and covaries more with stock returns, than non-stockholders', e.g., Parker and Vissing-Jørgensen (2009). Basak and Cuoco (1998) incorporate these features in a theoretical model, finding that non-participation and volatile stockholder consumption dynamics are a powerful mechanism to possibly resolve the equity premium puzzle. Pushing this story further, Guvenen (2009) builds a quantitative model incorporating limited stock market participation (without entry), showing that this channel can actually generate large equity premia and volatility, as long as stockholders have higher EIS than non-stockholders, consistent with the empirical findings of Vissing-Jørgensen (2002). Malloy, Moskowitz, and Vissing-

Jørgensen (2009) provide direct tests of stockholders' Euler equations under recursive preferences, finding that their long-run consumption dynamics observed in the micro-data can rationalize a variety of risk premia, using an EIS around 1 and risk aversion around 10.

That said, applying the limited participation model to households and stock markets is somewhat problematic, as risk concentration appears modest in that context. In 2007, approximately 50% of U.S. households by number and over 80% by wealth participate in stock markets, as documented in the Survey of Consumer Finances by Ackerman, Fries, and Windle (2012). The asset pricing consequences of this have been floating around in the literature. For instance, Gomes and Michaelides (2008) find that fixed entry costs help match basic facts on stock market non-participation, but has an insignificant effect on asset price moments. Indeed, when they reduce the participation cost to zero, they find minimal changes to their quantitative results on the equity premium. Empirically, stockholders are too wealthy to demand substantial risk compensation simply from a limited participation mechanism.

This is consistent with my results, which show that costly entry alone is unlikely to reproduce asset pricing regularities with the canonical limited participation model. The point of my paper, relative to Gomes and Michaelides (2008), is twofold. First, because of the tractability and simplicity of my model, I hope the reader can gain a clearer understanding of the entry channel in isolation. A full characterization of the joint (nonlinear) dynamics of risk prices and risk concentration is straightforward in my model. This allows me to uncover new findings such as the trade-off between matching the level and time-variability of risk prices in models with costly participation.

Second, as discussed above, limited participation is present in many other markets, and a growing theoretical literature on financial frictions leverages this same mechanism in understanding crises and price dynamics. As this mechanism becomes more pervasive, the literature needs a better understanding of the implied costs of financial frictions. Furthermore, my exploration of the interactions between costly participation and features like equity-issuance, idiosyncratic risk, and extrapolative expectations are all new to the literature.

The paper is organized as follows. Section 2 describes a limited participation model with entry. Section 3 solves for the equilibrium, illustrates the risk concentration channel, and argues that this channel implies very large entry costs. Section 4 shows, through several model extensions, that mechanisms increasing the level of risk premia sacrifice risk premia dynamics. Section 5 proposes extrapolative expectations as a partial remedy to some of the problems in the baseline model. Section 6 concludes. All proofs can be found in Appendix A and Online Appendix B.

2 Model

In this economy, time is continuous and spans the infinite past and future, $t \in \mathbb{R}$.

Endowment and financial markets. The aggregate endowment is given by the geometric

Brownian motion

$$\frac{dY_t}{Y_t} = \mu_Y dt + \sigma_Y dZ_t, \tag{1}$$

where $\{Z_t\}_{t\in\mathbb{R}}$ is a standard Brownian motion. Financial markets consist of a risky asset in unit supply and a locally riskless bond in zero net supply. The risky asset is a claim on $\{Y_t\}_{t\in\mathbb{R}}$ and can be thought of as productive capital, stocks, corporate debt, mortgage-backed securities, or indeed any positive-supply asset in which market segmentation might play a role. It has return dynamics

$$dR_t = \mu_{R,t}dt + \sigma_{R,t}dZ_t.$$

The bond pays an instantaneous return of $r_t dt$. Since this model will always feature a positive measure of agents trading in dynamically complete markets, we may define the *state-price density* process

$$\xi_t := \exp\Big\{ - \int_{-\infty}^{t} \Big(r_s + \frac{1}{2}\eta_s^2\Big) ds - \int_{-\infty}^{t} \eta_s dZ_s \Big\}, \tag{2}$$

where η is the *market price of risk* process. For the financial market to be absent of arbitrage opportunities, η must satisfy[1]

$$\eta_t = \frac{\mu_{R,t} - r_t}{\sigma_{R,t}}.$$

Agents and preferences. Births and deaths occur at rate π. Let b designate the birthdate of a cohort, within which there is a mass $\pi e^{-\pi(t-b)}$ of agents at time t. Agents have identical logarithmic preferences over consumption:

$$V_{t,b} := \mathbb{E}\Big[\int_{t}^{\infty} e^{-(\rho+\pi)(s-t)} \log(c_{s,b}) ds \mid \mathcal{F}_t \Big]. \tag{3}$$

In section 4, I offer an extension (3) to recursive utility of Duffie and Epstein (1992). Note that the death rate π simply augments the subjective discount rate. The only purpose of births and deaths is to help make the model stationary, nothing more.

Participants and non-participants. Let $\mathcal{P}_t$ denote the set of participants, and let $\mathcal{N}_t = \mathcal{P}_t^{\mathsf{c}}$ be the set of non-participants, who are barred from risky asset markets. Agents are born as non-participants and may begin participating at any time after birth, by paying a cost (see below). After that point, they remain participants until death. A participant does not face an entry decision, because she would never want to pay a cost to become a non-participant and consequently be constrained. As a result of this one-directional entry, the inflow of agents into $\mathcal{N}_t$ comes entirely from birth, and the only inflow of agents into $\mathcal{P}_t$ comes entirely from entry.

For notational simplicity, suppose all members of a cohort are either participants or non-

[1] Existence of a state-price density process ξ is guaranteed if η satisfies Novikov's condition $\mathbb{E}[\exp(\frac{1}{2}\int_{-\infty}^{\infty} \eta_t^2 dt)] < +\infty$ and if σ_R satisfies $\mathbb{E}[\int_{-\infty}^{\infty} \sigma_{R,t}^2 dt] < +\infty$. These must be verified in equilibrium. See Duffie (2010a), Chapter 6, for more details.

participants, so that there is no heterogeneity in decisions among members in the same cohort.[2] If $b \in \mathcal{P}_t$, this means that members of cohort b are participating in risky asset markets, and conversely for $b \in \mathcal{N}_t$.

Let $\tau_b \geq b$ denote the time cohort b begins participating in risky asset markets:

$$\tau_b := \inf\{t \geq b : b \in \mathcal{P}_t\}.$$

Since newborn agents are non-participants, their wealth dynamics are given by

$$\begin{aligned} dW_{t,b} &= (r_t W_{t,b} + \alpha\pi W_{t,b} - c_{t,b})dt, \quad t < \tau_b \\ W_{b,b} &> 0 \quad \text{given.} \end{aligned} \tag{4}$$

Upon participation, wealth dynamics are given by

$$dW_{t,b} = (r_t W_{t,b} + \theta_{t,b} W_{t,b}(\mu_{R,t} - r_t) + \alpha\pi W_{t,b} - c_{t,b})dt + \theta_{t,b} W_{t,b} \sigma_{R,t} dZ_t, \quad t \geq \tau_b. \tag{5}$$

Note that (4) resembles (5), but with the non-participation constraint $\theta_{t,b} \equiv 0$. Terms involving π represent annuity contracts: agents insure an exogenous fraction α of their wealth to death shocks by purchasing annuity contracts on competitive insurance markets, which results in flow income of $\alpha\pi W_{t,b}dt$. The insurance company takes the insured portion dying agents' wealth, which totals $\alpha \int_{-\infty}^{t} \pi e^{-\pi(t-b)} W_{t,b} db$, since dying agents are representative sample of the population. Notice this equals total payouts by insurance companies, i.e., insurance is priced fairly. The remaining fraction $1-\alpha$ of dying wealth is distributed to newborn generations ("unintended bequests"). This is similar to the insurance in the perpetual youth model of Blanchard (1985), although I assume $\alpha < 1$ to ensure that newborns have some financial wealth. This model only features financial wealth, but we may equivalently think of these unintended bequests as human wealth embodied in newborn generations, as long as labor income is pledgeable and perfectly correlated with capital income (Y_t). See Online Appendix B.5 for a discussion of this point.

Entry cost. To begin participating in risky asset markets, a non-participant must pay a non-pecuniary (utility) entry cost of

$$\Phi := -(\rho + \pi)^{-1} \log(1 - \phi), \quad \phi \in (0, 1). \tag{6}$$

With log utility, a constant entry cost has the desired homogeneity properties. The parameter ϕ represents the degree of entry costs: for an individual agent, (6) leads to equivalent entry incentives as if she were required to pay a fixed fraction ϕ of her wealth. Letting the cost be non-pecuniary has substantial benefits, however, because there is no need to account for deadweight losses from

[2] In equilibrium, it will turn out that the exact identities of participants and non-participants are not pinned down uniquely, although the wealth and consumption shares of participants will be determined. Having all members of a cohort participate (or not) together is among the possible equilibria, and all other quantities and prices are identical with and without this assumption, so this is without loss of generality.

entry. These assumptions are unlike papers on stock market participation and equity premia, e.g., Gomes and Michaelides (2008), which typically have non-homogeneous entry costs to address the wealth-participation gradient. My homogeneous specification adds significant tractability. See Online Appendix B.5 for discussion of this issue and about other types of fixed entry costs.

Basic properties due to homogeneity. The problem of participants is to maximize (3) subject to (5). Define their continuation utility by $V^P_{t,b}$, i.e.,

$$V^P_{t,b} := \sup_{c,\theta} \mathbb{E}\Big[\int_t^\infty e^{-(\rho+\pi)(s-t)} \log(c_{s,b}) ds \mid \mathcal{F}_t\Big], \tag{7}$$

subject to (5). By analogy, let $V^N_{t,b}$ be the continuation value of non-participants who were born at b,

$$V^N_{t,b} := \sup_{c,\tau} \mathbb{E}\Big[\int_t^\tau e^{-(\rho+\pi)(s-t)} \log(c_{s,b}) ds + e^{-(\rho+\pi)(\tau-t)}(V^P_{\tau,b} - \Phi) \mid \mathcal{F}_t\Big], \tag{8}$$

subject to the wealth dynamics given in (4) and the entry cost in (6).

Several homogeneity assumptions simplify the analysis of the model. The preferences in (3) are homothetic, so coupled with the linearity of the wealth dynamics in (5), there exists a process G^P such that participants' value in (7) is given by

$$V^P_{t,b} = (\rho+\pi)^{-1}[\log(W_{t,b}) + G^P_t].$$

Coupling this result with the formulation of entry costs in (6), the payoff to an entrant at time $t \geq b$ is $V^P_{t,b} - \Phi = (\rho+\pi)^{-1}[\log((1-\phi)W_{t,b}) + G^P_t]$, which confirms that the cost Φ is perceived as a fraction ϕ of wealth.

Because of the convenient functional form of the entry payoff $V^P_{t,b} - \Phi$, the non-participants' problem (8) is also homogeneous, and it is easy to show that

$$V^N_{t,b} = (\rho+\pi)^{-1}[\log(W_{t,b}) + G^N_t],$$

for some process G^N. The endogenous objects G^P and G^N, which are identical for all participants and non-participants, proxy for agents' investment opportunity sets.

As is standard in the theory of optimal stopping, non-participants compare the current payoff, $V^N_{t,b}$, against the best possible future entry payoff, $V^P_{t,b} - \Phi$, to decide when to enter. Non-participants optimally enter when the latter dominates, or

$$\tau_b = \inf\big\{t \geq b : \log(1-\phi) + G^P_t \geq G^N_t\big\}, \tag{9}$$

which is independent of wealth and cohort b. Thus, all agents have identical entry incentives at all times, an important property in solving for equilibrium. The intuition of (9) is that non-participants enter when risky assets are sufficiently attractive, as measured by the investment opportunity value G^P_t relative to G^N_t.

3 Equilibrium

An equilibrium is a set of price and allocation processes such that all agents maximize utility through consumption, portfolio, and entry decisions, and all markets clear. The entry decisions are slightly non-standard and merit some discussion. Define the equilibrium set of time points when entry occurs by $\mathcal{T}^*$. Entry incentives are the same for all non-participants regardless of their birthdates or their accumulated wealth, as τ_b in (9) is independent of b and $W_{t,b}$. Instead, entry incentives only depend on G_t^N and G_t^P, which only depend on the history of aggregate shocks $\{Z_s\}_{s\leq t}$. This implies that the identity of entrants is not uniquely determined in equilibrium, i.e., $\mathcal{P}_t$ is not uniquely determined. It also implies that at any time point of entry, $t^* \in \mathcal{T}^*$, all non-participants must be indifferent between inaction and entry.[3] As a result, entry incentives can be written as follows:

$$t \in \mathcal{T}^* : \ \log(1-\phi) + G_t^P = G_t^N; \tag{10}$$

$$t \notin \mathcal{T}^* : \ \log(1-\phi) + G_t^P < G_t^N. \tag{11}$$

Besides entry, the equilibrium definition is standard in securities market models. The market clearing equations are as follows. Note that P_t is defined as the aggregate value of the stock market.

- Goods market clearing:

$$Y_t = \int_{-\infty}^{t} \pi e^{-\pi(t-b)} c_{t,b} db. \tag{12}$$

- Stock market clearing:

$$P_t = \int_{-\infty}^{t} \pi e^{-\pi(t-b)} \theta_{t,b} W_{t,b} db. \tag{13}$$

- Bond market clearing:

$$0 = \int_{-\infty}^{t} \pi e^{-\pi(t-b)} (1-\theta_{t,b}) W_{t,b} db. \tag{14}$$

- Newborn transfers equal collections from deceased:

$$\pi W_{t,t} = \pi(1-\alpha) \int_{-\infty}^{t} \pi e^{-\pi(t-b)} W_{t,b} db. \tag{15}$$

Markov equilibrium. We seek a *stationary Markov equilibrium* in the state variable

$$X_t := Y_t^{-1} \int_{\mathcal{P}_t} \pi e^{-\pi(t-b)} c_{t,b} db, \tag{16}$$

[3] If instead, all non-participants strictly preferred entry at t^*, the economy would become populated solely by participants after t^*. This implies that, immediately after t^*, the new entrants would regret their decision to enter, contradicting the optimality of τ_b. Because of the required indifference between inaction and entry, we also note that non-participants cannot all be using stopping-time strategies in deciding to enter. The stopping time τ_b tracks incentives to enter but cannot determine entry decisions for non-participants.

which represents the consumption share of the participants. This single endogenous variable is sufficient to characterize the entire equilibrium. Conjecture the following dynamics for X,

$$dX_t = \mu_X(X_t)dt + \sigma_X(X_t)dZ_t + dA_t^{x^*}, \tag{17}$$

where A^{x^*} is a singular process that will correspond to time-points of entry $\mathcal{T}^*$. In Markov equilibrium, we may assume entry occurs when $X_t \leq x^*$ for some point $x^* \in [0,1]$.[4] In this case, the set of entry times is

$$\mathcal{T}^* = \{t : X_t = x^*\}, \tag{18}$$

and A^{x^*} is the barrier process at x^*. It is a non-decreasing, continuous process keeping $X_t \geq x^*$ almost-surely by increasing when $X_t = x^*$. The *reflecting boundary* x^* is a key equilibrium object.

Definition 1. *A* stationary Markov equilibrium *in X_t, defined in (16)-(17), consists of an entry point x^* and a set of functions characterizing agents' optimal policies, agents' value processes, asset prices, and state dynamics such that individual agents solve (7) and (8), and such that markets clear as in (12)-(15). Value processes are characterized by*

$$V^P(W_{t,b}, X_t) := V_{t,b}^P \quad \textit{and} \quad V^N(W_{t,b}, X_t) := V_{t,b}^N \tag{19}$$

$$g_P(X_t) := (\rho+\pi)^{-1}G_t^P \quad \textit{and} \quad g_N(X_t) := (\rho+\pi)^{-1}G_t^N. \tag{20}$$

Asset prices are characterized by

$$\eta_t = \eta(X_t),\ r_t = r(X_t),\ p_t := \frac{P_t}{Y_t} = p(X_t),\ \sigma_{R,t} = \sigma_R(X_t),\ \mu_{R,t} = \mu_R(X_t).$$

Equilibrium Asset Prices. Assume that the equilibrium entry point x^* is given. Then, all asset prices and state dynamics from Definition 1 can be determined in closed form. We have the following proposition. The basic steps in determining η, r, μ_X, and σ_X are to apply Itô's formula to the goods market clearing equation and the definition of the state variable, for $X_t \in [x^*, 1)$. The proof is in Appendix A.4.

Proposition 1. *Suppose entry point $x^* \in (0,1)$ is given. Then, the state-price density process ξ_t exists uniquely, and is characterized by*

$$\eta(x) = \frac{\sigma_Y}{x} \quad \textit{and} \quad r(x) = \rho + \pi + \mu_Y - \frac{\sigma_Y^2}{x}, \quad x \in [x^*, 1).$$

[4] This is essentially without loss of generality in a stationary equilibrium. Suppose the diffusive part of X, (μ_X, σ_X), is "regular" in the following sense: without any entry, X would eventually visit all states in $(0,1)$ in finite time, a.s. Lemma B.6 shows that a sufficient condition for this "regularity" is $\gamma > \frac{\psi}{\psi+1}$. Now assume there were an entire family of entry points x_i^*, with minimum and maximum points $0 < \underline{x}^* \leq \overline{x}^* < 1$. Since entry only increases X, it follows that X_t eventually exceeds $\overline{x}^*$, so we may take $x^* = \overline{x}^*$ to be our entry point.

The state variable X is the unique strong solution to $dX_t = \mu_X(X_t)dt + \sigma_X(X_t)dZ_t + dA_t^{x^*}$, where

$$\mu_X(x) = -\pi(1-\alpha)x + \sigma_Y^2 \frac{(1-x)^2}{x} \quad \text{and} \quad \sigma_X(x) = (1-x)\sigma_Y, \quad x \in [x^*, 1).$$

Finally, the non-degenerate stationary density of X_t is given by

$$h(x) = K_0 \Big(\frac{x}{1-x}\Big)^2 (1-x)^{-\frac{2\pi(1-\alpha)}{\sigma_Y^2}} \exp\Big(-\frac{2\pi(1-\alpha)}{\sigma_Y^2(1-x)}\Big), \quad x \in [x^*, 1), \tag{21}$$

where K_0 is a constant to ensure that h integrates to 1.

Proposition 1 illustrates several key features of the model. First, we can see the generic asset pricing properties of limited participation economies. The fact that $X_t < 1$ raises the market Sharpe ratio, given by η_t, and lowers the market interest rate, given by r_t. With limited participation, all risk must be borne by a subset of individuals, meaning they must be levered. These levered participants are marginal in all asset markets. But voluntarily taking a leveraged position requires lower borrowing costs and higher expected rates of return, compared to a similar economy without leverage. Furthermore, as participants' relative wealth falls due to negative shocks (notice that $\sigma_X > 0$), these forces increase, so the effects on η and r strengthen. This is the so-called risk concentration channel referenced in the introduction.

Second, as long as entry is such that $x^* > 0$ (which corresponds to $\phi < 1$ as we will see below), the economy is well-behaved in the sense that an equivalent martingale measure exists. Without entry, the equilibrium tends to explode in bad times: if $x^* \to 0$, we have $r(x^*) \to -\infty$ and $\eta(x^*) \to +\infty$. One downside is that this can lead limited participation equilibria to have bubbles and arbitrage opportunities, as shown in related work by Hugonnier (2012). With $\phi < 1$ such that $x^* > 0$, this explosion is prevented. Proposition 19 shows that the equilibrium with $\phi = 1$ always has bubbles, while $\phi < 1$ never has bubbles. In that sense, there is a discontinuity in the equilibria as $\phi \to 1$. See Appendix A.5 for an extended discussion of this point.

Finally, unlike models such as Basak and Cuoco (1998), the entry device included here, along with the OLG environment, ensures long-run stationarity. To see this intuitively, note that at $X_t = x^*$, enough entry occurs so that $dX_t \geq 0$. Thus, entry keeps X_t above x^*. As $X_t \to 1$, we have $\mu_X(1) < 0$ while $\sigma_X(1) = 0$. Thus, birth and death keeps X_t away from 1. Formally, a non-degenerate stationary distribution exists, given by (21).

Equilibrium Entry. The equilibrium entry point x^* is determined as follows. I apply dynamic programming to the participants' and non-participants' problems, leading to two ODEs (the HJB equations) for g_P and g_N:

$$0 = \log(\rho+\pi) - 1 + (\rho+\pi)^{-1}\Big(\alpha\pi + r + \frac{1}{2}\eta^2\Big) - (\rho+\pi)g_P + \mu_X g_P' + \frac{1}{2}\sigma_X^2 g_P'' \tag{22}$$

$$0 = \log(\rho+\pi) - 1 + (\rho+\pi)^{-1}(\alpha\pi + r) - (\rho+\pi)g_N + \mu_X g_N' + \frac{1}{2}\sigma_X^2 g_N'', \tag{23}$$

which hold on $(x^*, 1)$. Boundary conditions for these ODEs are the following. First, equation (10) implies the *value-matching* condition

$$g_P(x^*) - \Phi = g_N(x^*). \tag{24}$$

Next, as shown in Lemma A.3, the entry inequality (11) is equivalent to the following *smooth-pasting* conditions:

$$g_P'(x^*) = g_N'(x^*) = 0. \tag{25}$$

Finally, given $\sigma_X \to 0$ as $x \to 1$, we may take limits of (22)-(23) as $x \to 1$ to obtain two additional boundary conditions. These five boundary conditions are sufficient to solve ODEs (22)-(23) and the entry point x^*.

However, some simplifications can be made by noticing that the coefficients on g_P and g_N are identical in (22)-(25). Putting $\Delta g := g_P - g_N$ and taking differences between the HJB equations yields one linear ODE

$$0 = \frac{1}{2}(\rho + \pi)^{-1}\eta^2 - (\rho + \pi)\Delta g + \mu_X \Delta g' + \frac{1}{2}\sigma_X^2 \Delta g'', \quad x \in (x^*, 1), \tag{26}$$

with boundary conditions

$$0 = \Delta g(x^*) - \Phi \tag{27}$$

$$0 = \Delta g'(x^*). \tag{28}$$

Conditions (27)-(28), along with the limiting condition arising from taking $x \to 1$ in (26), are sufficient to solve for x^*. In fact, the entry point is determined uniquely, as the following proposition claims. This establishes existence and uniqueness of the equilibrium.

Proposition 2. *There is exactly one pair* $(\Delta g, x^*)$ *that satisfies the ODE (26) and boundary conditions in (27)-(28). Consequently, the equilibrium of Definition 1 exists uniquely.*

There are likely many ways to prove Proposition 2. I choose to demonstrate the equivalence of the model to a relatively standard variational inequality, which is portable to higher dimensions and potentially useful in other models. See Appendix A.4 for the proof. In Proposition 18, we obtain Δg as a power series solution to (26), which must be the only solution by uniqueness.

To characterize the entry decisions, I find it more informative to examine the Feynman-Kac representation of the value function.

Proposition 3. *The function Δg can be equivalently represented in the following three ways:*

$$\Delta g(x) = \mathbb{E}^x\Big[\frac{1}{2}(\rho+\pi)^{-1}\int_0^\infty e^{-(\rho+\pi)t}\eta^2(X_t)dt\Big] \tag{29}$$

$$= \mathbb{E}^x\Big[\frac{1}{2}(\rho+\pi)^{-1}\int_0^{\tau_{x^*}} e^{-(\rho+\pi)t}\eta^2(X_t)dt + e^{-(\rho+\pi)\tau_{x^*}}\Phi\Big] \tag{30}$$

$$= \inf_\tau \mathbb{E}^x\Big[\frac{1}{2}(\rho+\pi)^{-1}\int_0^{\tau} e^{-(\rho+\pi)t}\eta^2(X_t)dt + e^{-(\rho+\pi)\tau}\Phi\Big], \tag{31}$$

where $\tau_{x^*} := \inf\{t \geq 0 : X_t = x^*\}$, the minimization in (31) is over the set of stopping times, and (X, A^{x^*}) is the unique strong solution to $dX_t = \mu_X(X_t)dt + \sigma_X(X_t)dZ_t + dA_t^{x^*}$ with $X_0 = x$.

Given (29), we can now interpret Δg, which encodes the difference between the investment opportunities of participants and non-participants, as the foregone costs of non-participation. Indeed, every instant, a participant expects to earn $\eta_t\sigma_Y dt$ in excess returns per unit of investment, and optimally invests η_t/σ_Y, resulting in η_t^2 in gains per unit of time. The scaling by $\frac{1}{2}$ is a Jensen risk-adjustment. These gains are discounted by $e^{-(\rho+\pi)t}$ and then cumulated to produce lifetime gains. Finally, the scaling by $(\rho+\pi)^{-1}$ translates from monetary to utility gains.

Because of (27), we can re-write (29) to "back out" the implied entry cost ϕ:

$$\phi = 1 - \exp\Big(-\frac{1}{2}\mathbb{E}^{x^*}\Big[\int_0^\infty e^{-(\rho+\pi)t}\eta^2(X_t)dt\Big]\Big). \tag{32}$$

We can read (32) as the fraction of wealth a typical investor is willing to pay to participate in risky asset markets. The willingness-to-pay is related to the expected present discounted value of squared Sharpe ratios, starting from the worst state of the world (i.e., $x = x^*$). Thus, in computing implied entry costs, we need to account for extreme Sharpe ratios, but also their speed of transition back to normal levels, which is embedded in the dynamics of X_t.

Representations (30) and (31) show that entry solves a relatively standard optimal stopping problem. One could think of a fictitious planner controlling the wealth distribution in a way that balances non-participants' opportunity costs from participation with their direct participation costs.[5]

Small Entry Costs. When entry is free, i.e., $\phi = 0$, all agents participate in risky asset markets, and the economy features full market integration. This implies that $x^* = 1$ so that continuous entry keeps $X_t \equiv 1$. Asset prices behave as in Proposition 1 with $x = 1$. This economy is equivalent to an unconstrained OLG economy, i.e., a homogeneous economy comprised of participants.

For small enough entry costs, it turns out that the same full-integration equilibrium prevails. Writing the equilibrium outcome x^* as a function of ϕ, $x^*(\phi)$, define

$$\phi^* := \sup\{\phi : x^*(\phi) = 1\} \tag{33}$$

[5]These representations are not specific to the one-dimensional model presented here. In Appendix A.8, I show that similar formulas hold in many dimensions, after introducing an arbitrary number of exogenous state variables.

to be the largest fixed cost such that the free entry equilibrium prevails. In the following proposition, proven in Appendix A.4, I show that $\phi^* > 0$, a stark result implying that the results of traditional limited participation models are, in some limiting sense, not robust to entry. Indeed, the notion that agents have rigid types ("experts" versus "non-experts" or "investors" and "households") cannot be justified by small participation costs.

Proposition 4. *Define* $\phi^* := \sup\{\phi : x^*(\phi) = 1\}$. *Then,* $\phi^* > 0$ *and in particular,*

$$\phi^* = 1 - \exp\Big(- \frac{1}{2}(\rho+\pi)^{-1}(\eta^*)^2 \Big) = \frac{1}{2}(\rho+\pi)^{-1}(\eta^*)^2 + O(\sigma_Y^4), \tag{34}$$

where $\eta^* := \eta(1)$ *is the full-integration risk price.*

This result is somewhat counterintuitive at first. The costs ϕ are *fixed costs*, which typically implies some inaction (hesitation to enter) for any positive cost. Here, it is the opposite. The reason for the result of Proposition 4 lies in the fact that participation strictly dominates non-participation as an investment technology. This is because the risk in the economy is aggregate risk, which does not dissipate even if shared maximally among agents, yielding positive risk prices, $\eta^* > 0$. Hence, a discrete gain in lifetime utility is possible from participation, which justifies immediate entry (at birth) despite a fixed cost.

To get an estimate of the size of ϕ^* in Proposition 4, we suppose $X_t = 1$ for all t and substitute this into (32). This estimate is large: using a small Sharpe ratio of $\eta^* = 0.10$, and a discount plus birth/death rate of $\rho + \pi = 0.02$, we find that $\phi^* \approx 0.25$. This is the 25% of wealth estimate quoted in the introduction. This gives us a preliminary understanding that the risk concentration mechanism, if it is behind large risk premia, must imply very large entry costs.

Larger Entry Costs. Now, we consider a different thought experiment. If we allow entry costs ϕ to be larger, such that $x^* < 1$, what values of ϕ are consistent with realistic levels and dynamics of asset prices? Intuitively, the primary effect of entry is to prevent the economy from reaching high return states, mitigating asset price dynamics. As entry costs increase, the economy is more likely to access those states.

Table 1: Baseline calibration of model parameters.

Parameter:	ρ	π	α	μ_Y	σ_Y
Value:	0.01	0.02	0.50	0.02	0.04

Figure 1 plots Sharpe ratios η and the stationary density h for four different entry costs ϕ (other parameters for this section are given in Table 1). For lower entry costs, the stationary distribution is truncated at the entry point x^*, and high-η states are averted. For higher entry costs, the economy is more likely to visit high-η states. It is worth noting that figure 1 examines entry costs of 20%-35%, which are already quite large, and finds little evidence for extreme risk concentration or extremely high risk prices. That said, the pattern in figure 1 suggests that choosing ϕ high enough could theoretically lead to significant risk concentration and more realistic asset price behavior.

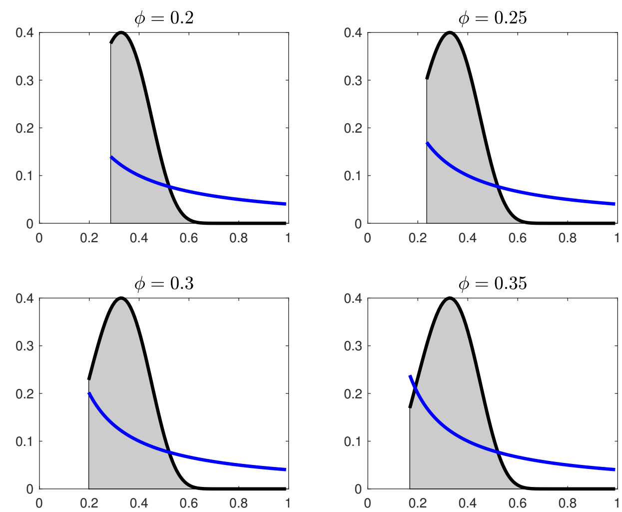

Figure 1: Market price of risk η (blue line) and stationary density for X (gray area), for four different entry costs ϕ. The horizontal axis is the participants' consumption share x. Parameters are in Table 1.

It turns out that this is not the case. Figure 2 depicts Sharpe ratios for nine higher entry cost parameters. Although the entry point x^* does fall as ϕ rises, most of the mass in the stationary density is relatively stable in ϕ. While $\eta(x)$ can technically be very high in low-x states, there is essentially zero probability of X_t reaching those states.

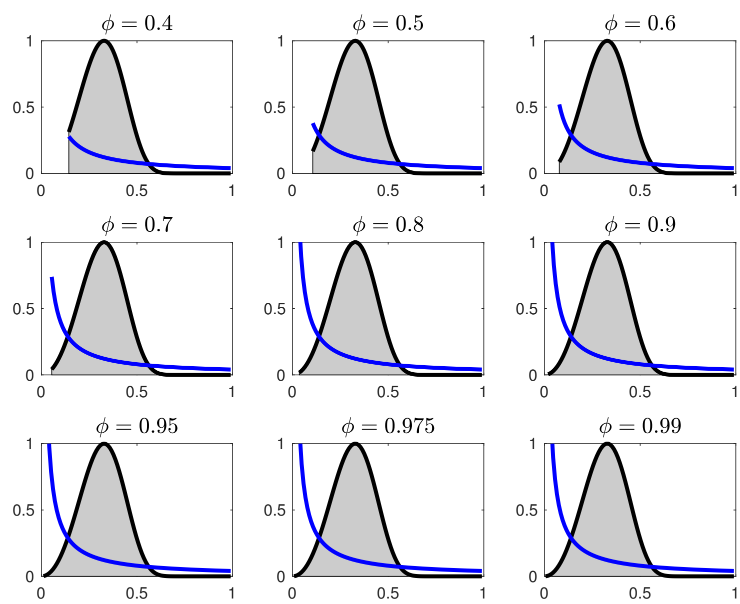

Figure 2: Market price of risk η (blue line) and stationary density for X (gray area), for nine different entry costs ϕ. The horizontal axis is the participants' consumption share x. Parameters are in Table 1.

This story is confirmed by Table 2, which considers the same entry costs from figures 1-2. As entry costs ϕ increase, all of the following increase: average Sharpe ratios, Sharpe ratio volatility, and maximal Sharpe ratios. However, once entry costs reach 70%, $\mathbb{E}[\eta(X_t)]$ and $\text{std}[\eta(X_t)]$ stabilize, even as $\sup[\eta(X_t)]$ rises even more dramatically.

Table 2: Entry cost ϕ and different measures of market Sharpe ratios (stationary average SR, standard deviation of SR, and maximal SR). Parameters are in Table 1.

ϕ	0.20	0.25	0.30	0.35	0.40	0.50	0.60	0.70	0.80	0.90	0.95	0.97	0.99
$\mathbb{E}[\eta(X_t)]$	0.10	0.11	0.12	0.13	0.13	0.14	0.14	0.14	0.15	0.15	0.15	0.15	0.15
$\mathrm{std}[\eta(X_t)]$	0.02	0.03	0.03	0.04	0.04	0.05	0.06	0.07	0.08	0.08	0.09	0.09	0.09
$\sup[\eta(X_t)]$	0.14	0.17	0.20	0.24	0.28	0.38	0.52	0.74	1.14	2.11	3.51	5.31	8.12

Intuitively, as expected future returns rise, participants' wealth rebounds very quickly from a series of poor returns, creating a "buoying effect" on participant wealth. This force, that high risk premia create very fast recovery, is present in any limited participation economy. This buoying effect acts as a kind of natural entry, substituting for entry when costs are high, in that it helps the economy avoid crisis states.

Exogenous Entry. For comparison, I modify the entry assumptions so that a fraction ν of the newborns exogenously become designated participants, while $1 - \nu$ fraction become non-participants. No endogenous entry is possible. Appendix A.6 discusses this model and derives the entire equilibrium in closed form (Proposition 20). The difference between the two economies is exclusively the dynamics of the state variable:

$$\begin{aligned} \text{endogenous entry: } \mu_X(x) &= -\pi(1-\alpha)x + \sigma_Y^2 \frac{(1-x)^2}{x} \\ \text{exogenous entry: } \mu_X(x) &= -\pi(1-\alpha)x + \sigma_Y^2 \frac{(1-x)^2}{x} + \pi(1-\alpha)\nu. \end{aligned}$$

With exogenous entry, there is continuous entry at the rate $\pi(1-\alpha)\nu dt$, rather than the discrete entry at $dA_t^{x^*}$. Figure 3 compares these two economies.

The top four panels show four endogenous entry economies, indexed by their entry costs ϕ. In each plot, risk prices η are displayed along with the ergodic distribution of X_t. Notice that the distribution of X_t is truncated by entry, with less truncation occurring as the entry cost rises. The bottom four panels show four comparable exogenous entry economies, indexed by their participant fraction ν. The parameter ν is chosen so that the stationary mean $\overline{x}$ matches that of the endogenous entry economy plotted directly above.

For relatively small costs (e.g., $\phi = 0.10, 0.25$), figure 3 shows that endogenous entry constrains the dynamics of X_t and η_t much more than a comparable amount of exogenous entry ($\nu = 0.38, 0.10$). Despite having the same stationary mean, economies with endogenous entry spend significantly less time in low-x and high-η states, the sense in which endogenous entry limits market segmentation.

For larger entry costs (e.g., $\phi = 0.40, 0.60$), there is less of a distinction between endogenous and exogenous entry. Equilibrium asset prices become increasingly similar between the models as entry is eliminated. In fact, we have the following result.

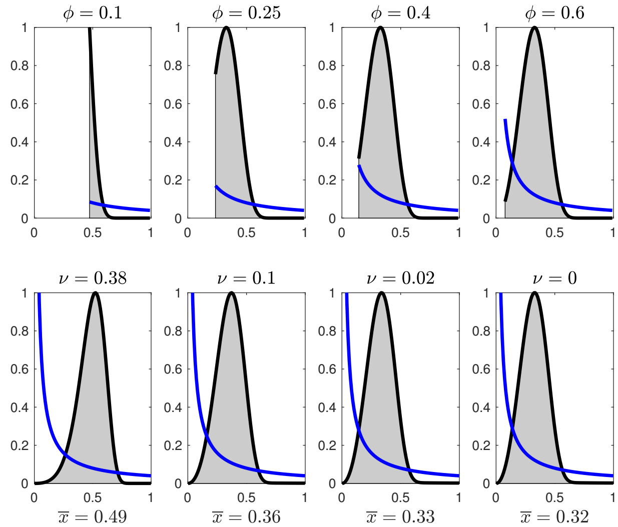

Figure 3: Each plot features the market price of risk η (blue line) and stationary density for X (gray area). The horizontal axis is the participants' consumption share x. Top four panels: Each plot corresponds to a different entry cost ϕ. Bottom four panels: Each plot corresponds to a different participant population share ν. The share of participants ν is chosen to match the stationary mean $\overline{x} := \mathbb{E}X_t$ in the endogenous entry economy plotted directly above. For example, the endogenous entry model with cost $\phi = 10\%$ and the exogenous entry model with $\nu = 38\%$ both have the same average participant consumption share $\overline{x} = 49\%$. Parameters are in Table 1.

Proposition 5. *As entry disappears (i.e., $\nu \to 0$ and $\phi \to 1$), the exogenous and endogenous entry economies coincide.*

More General Preferences. The previous model is restrictive in its preference specification. In this section, I generalize preferences to the recursive preference specification of Duffie and Epstein (1992). First of all, more general preferences allow for hedging demands that may potentially dissuade entry, even with moderate participation costs. Second, with log utility, price-dividend ratios are constant, so all variation in the risk premium is counterfactually due to the interest rate. Third, recursive preferences allow me to experiment with risk aversion in isolation, through which I uncover a trade-off between conditional and unconditional risk premia. Here, I introduce and analyze the recursive utility model, but leave a more detailed discussion to Appendix A.1.

Mathematically, the continuation value now satisfies

$$V_{t,b} := \mathbb{E}\Big[\int_t^\infty f(c_{s,b}, V_{s,b})ds \mid \mathcal{F}_t\Big], \tag{35}$$

where the felicity function f is defined by

$$f(c,V) := \frac{1}{1-\psi}\Big(c^{1-\psi}[V(1-\gamma)]^{\frac{\psi-\gamma}{1-\gamma}} - (\rho+\pi)V(1-\gamma)\Big). \tag{36}$$

In (36), parameter γ is the coefficient of relative risk aversion (RRA), and ψ^{-1} is the elasticity of intertemporal substitution (EIS). Assume $\gamma, \psi \neq 1$. Again, the death rate π simply augments the subjective discount rate, as shown by Gârleanu and Panageas (2015) for these preferences.

To maintain tractability, I modify the participation cost. The cost $\Phi_{t,b}$ now has a time and cohort dimension and is given by

$$\Phi_{t,b} := [1 - (1-\phi)^{1-\gamma}]V^P_{t,b}, \tag{37}$$

where $V^P_{t,b}$ is the participant value function. With this specification, parameter $\phi \in (0,1)$ still denotes the perceived fraction of wealth a non-participant must pay to begin participation.

Proposition 14 in Appendix A.1 derives the equilibrium under these assumptions. The equilibrium is more complicated than the log utility model. Now we must solve for two value functions, one each for participants and participants, as part of a free-boundary problem for x^*. Furthermore, asset prices are no longer independent of agents' value functions. In terms of parameters, I choose both $\gamma, \psi^{-1} > 1$, to bring the model slightly closer to the asset pricing literature. By increasing the RRA, the model generates higher levels of risk prices. By choosing EIS larger than 1, the model generates procyclical price-dividend ratios, thus potentially more volatile risk prices.

That said, the economy behaves qualitatively the same as the log utility economy. Figures 4-5 are the recursive utility versions of figures 1-2. Entry continues to limit extreme asset price dynamics, by truncating the distribution of risk prices η.

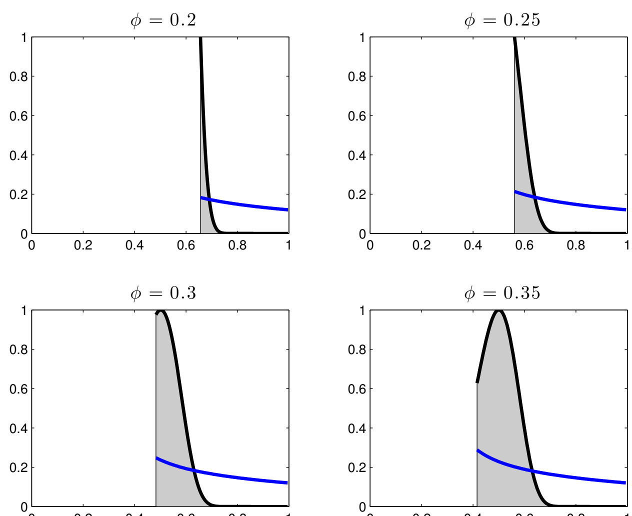

Figure 4: Market price of risk η (blue line) and stationary density for X (gray area), for four different entry costs ϕ. The horizontal axis is the participants' consumption share x. I set $\gamma = 3$ and $\psi = 3/4$. All other parameters are as in Table 1.

A new property can be seen by comparing the first 5 panels of figure 5 with the final 4 panels. The first 5 panels show that Sharpe ratios are almost always in the 0.15-0.3 range, for any entry cost between 40% and 80% of wealth. In panels 6-9, the stationary distribution does begin allowing risk prices to attain crisis dynamics.

These results arise because of the hedging motives brought about by $\gamma \neq 1$. Suppose x_0 is a state with very high and volatile risk prices, $\eta(x_0)$ and $|\eta'(x_0)|$ very large. At x_1 slightly above x_0, a

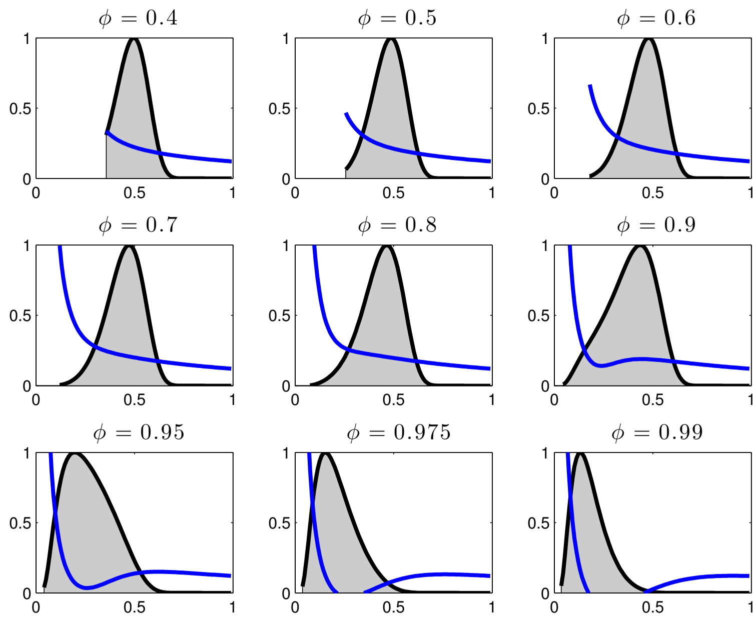

Figure 5: Market price of risk η (blue line) and stationary density for X (gray area), for nine different entry costs ϕ. The horizontal axis is the participants' consumption share x. I set $\gamma = 3$ and $\psi = 3/4$. All other parameters are as in Table 1.

negative shock improves the investment opportunity set, which generates hedging demand for risky assets. This means that $\eta(x_1)$ should be much smaller than $\eta(x_0)$. The intuition is "heads, I win; tails, I will win soon." That these features only emerge when entry costs are 90% or larger. But at least there exist entry costs such that the recursive utility model, unlike the log utility model, can match the high level and variability of risk prices. This begs the question as to whether other utility parameters can match empirical asset prices for more moderate entry costs, a subject we turn to immediately in section 4.

4 Trade-off between conditional and unconditional risk premia

The previous section showed that limited participation economies imply large entry costs, given large and time-varying risk premia. In this section, I show that this conclusion is general in the sense that model extensions which raise unconditional risk premia necessarily attenuate conditional asset price dynamics, due to the entry channel. As a result, such model extensions require even larger entry costs to justify crisis dynamics of risk premia, if such dynamics are to be attributed to limited participation. The model extensions I consider here are increasing agents' risk aversions, allowing equity-issuance, and introducing idiosyncratic risk.

Higher Risk Aversion. By comparing the model equilibrium for different values of γ, we uncover a trade-off between unconditional and conditional risk premia in limited participation models. Intuitively, higher risk aversion works to increase the level of risk premia, which incentivize entry, thus mitigating the time-variability of risk premia. Conversely, lower risk aversion leads to more substantial risk premia dynamics, but lower average risk premia. In partial equilibrium, we would expect more frequent entry when agents are more risk-tolerant or when the economy is less risky. In general equilibrium, this effect is reversed, revealing the conclusions of this section.

This result is important because the asset pricing literature frequently chooses γ structurally to match empirical asset prices. Having little direct evidence on investors' risk aversions, values of γ up to 10 are not considered unusual in this literature. Here, I show that such calibrations of a limited participation model are not a panacea for asset pricing puzzles.

To start, we have the following generalization of Proposition 4, which shows how to calculate the implied entry cost such that the economy is fully integrated. The proof is in Appendix A.3.

Proposition 6. *Define $\phi^* := \sup\{\phi : x^*(\phi) = 1\}$. Then, $\phi^* > 0$ and in particular,*

$$\phi^* \geq 1 - \Big(\frac{\rho + \pi + (\psi - 1)(r^* + \alpha\pi + \frac{1}{2\gamma}(\eta^*)^2)}{\rho + \pi + (\psi - 1)(r^* + \alpha\pi)}\Big)^{\frac{\psi}{1-\psi}} = \frac{\frac{1}{2}\sigma_Y \eta^*}{\rho + \pi(1-\alpha) + \psi(\mu_Y + \pi\alpha)} + O(\sigma_Y^4), \quad (38)$$

where $\eta^* := \gamma\sigma_Y$ is the full-integration risk price.

The positive relationship between risk aversion and entry incentives is depicted in figure 6. For instance, with $\gamma = 3$, the full participation equilibrium is attained for $\phi \leq \phi^* \approx 10\%$, while for $\gamma = 10$, the full participation equilibrium is attained for $\phi \leq \phi^* \approx 27\%$ (left panel of figure 6). This occurs because higher γ raises equilibrium risk prices $\eta^* = \gamma\sigma_Y$, which appear directly in (38). Non-participants will want to enter to claim these benefits. Thus, for moderate entry costs, risk price dynamics are completely eliminated with higher risk aversion. More generally, we find numerically that increasing γ increases both the entry point, x^*, and the stationary mean $\overline{x} := \mathbb{E}X_t$ (middle panel of figure 6).

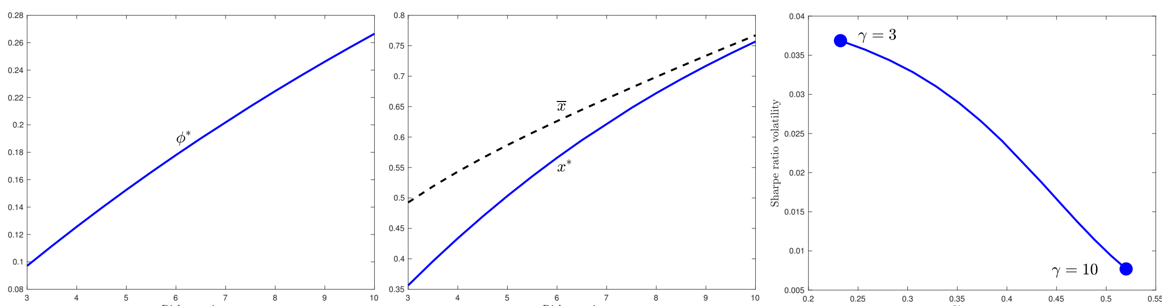

Figure 6: Left panel: full-integration cost $\phi^* := \sup\{\phi : x^*(\phi) = 1\}$, as a function of risk aversion γ. Middle panel: entry boundary, x^*, and stationary mean, $\overline{x} := \mathbb{E}X_t$, as a function of risk aversion γ. Right panel: trade-off between $\mathbb{E}[\eta(X_t)]$ and std$[\eta(X_t)]$ as a function of risk aversion γ. I set $\psi = 3/4$ and $\phi = 0.4$. All other parameters are as in Table 1.

The increase of the entry point x^* is very informative about risk price variability. Indeed, we have the following proposition.

Proposition 7. *In the recursive utility model,*

$$\frac{\eta(x^*)}{\eta(1)} = \frac{1}{x^*}. \quad (39)$$

Recall that higher γ tends to lead to increase x^* (middle panel of figure 6). If we consider $\eta(x^*)$ a proxy for the maximal risk price and $\eta(1)$ a proxy for the minimal risk price,[6] then $\eta(x^*)/\eta(1)$ proxies for risk price variability. Thus, equation (39) shows that higher risk aversion tends to decrease risk price variability.

The same trade-off is visible if we measure "level" and "variability" by mean and standard deviation of risk prices (right panel of figure 6): as γ increases from 3 to 10, average risk prices increase, but risk price volatility falls. Putting these results all together, it seems that higher risk aversion increases the level of risk prices, which raises entry incentives, and thereby attenuates risk price dynamics.[7]

Equity-Issuance. In the baseline model, non-participants can only share risky asset returns through the bond market. This type of setup is naturally interpreted as a model of household participation in stock markets. However, the limited participation framework is more general, and the results of this paper extend to other markets. In fact, stock market participation is perhaps the least appealing application of the limited participation mechanism, as discussed in the literature review. A model featuring financial intermediaries, whereby insiders in the intermediary must keep concentrated positions or "skin-in-the-game" for incentive reasons, is also a type of limited participation model.[8] Such a model may be applied to corporate bond markets, mortgage-backed securities, commodity markets, and others.

Here, I extend the model to allow partial equity issuance by participants, to facilitate a financial intermediary interpretation. Participants' ability to partially share risks through public equity markets lowers typical risk premia levels, but raises the chances of extreme risk prices. Again, we see the trade-off between conditional and unconditional risk premia.

Participants keep a fraction $\chi_{t,b} \geq \chi^*$ of their equity risk on their books, with $\chi^* \in (0,1)$, capturing partial equity issuance. They offload the remaining $1 - \chi_{t,b}$ of risk to financial markets. Here, $\chi_{t,b}$ is a choice variable. Non-participants' position in participants' outside equity is given by $\tilde{\theta}_{t,b}$. Finally, I allow participants to purchase long-only diversified positions in other participants' outside equity ($\tilde{\theta}_{t,b} \geq 0$ for $b \in \mathcal{P}_t$), which they might want to do if their aggregate risk exposure is too low after their equity issuance.

Participants are compensated for their equity issuance constraints by additional returns, which are captured mathematically by two different risk prices: one for inside equity (η_t) and one for

[6] These proxies are exact under $\gamma, \psi = 1$ (log utility), but $\eta(1)$ may not be the minimal risk price when $\gamma, \psi \neq 1$, as the last three panels of figure 5 show.

[7] Similarly, higher fundamental volatility σ_Y will increase unconditional risk prices and simultaneously lead to more entry. A simple way to see this is to re-examine formula (38) for ϕ^*, in which $\frac{d\phi^*}{d\sigma_Y} > 0$. Since $\sigma_Y \eta^* = \gamma \sigma_Y^2$ appears in the numerator, increasing either risk quantity or risk aversion will increase entry incentives and reduce risk concentration. This will clearly reduce risk price variability, as discussed above. In numerical calculations, I have found that increasing σ_Y and increasing γ produce analogous results with multiple calibrations. In particular, figure 6 looks very similar if we were to vary σ_Y, rather than γ.

[8] A similar model with equity issuance is that considered by He and Krishnamurthy (2012) and He and Krishnamurthy (2013). In those models, "specialists" manage intermediaries, and "households" can only invest in risky assets through intermediaries. Intermediaries issue equity to households, and aggregate risk is shared. However, for incentive reasons, specialists must keep sufficient "skin in the game," so their equity issuance is only partial.

outside equity ($\tilde{\eta}_t$). We have $\mu_{R,t} - r_t - (1-\chi_{t,b})\sigma_{R,t}\tilde{\eta}_t$ of returns available to participants after equity issuance. Define η_t as the risk price on these insider returns, i.e., $\chi_{t,b}\sigma_{R,t}\eta_t := \mu_{R,t} - r_t - (1-\chi_{t,b})\sigma_{R,t}\tilde{\eta}_t$. Finally, budget constraints must be modified to (81)-(82), and the outside equity market must clear, accounted for by (79). Complete details for this section are in Appendix A.7.

Proposition 8. *There exists a unique equilibrium with equity issuance, which is governed by the state variable X_t. When $X_t \geq \chi^*$, aggregate risk is shared perfectly with $\sigma_{X,t} = 0$. When $X_t \in (x^*, \chi^*)$, participants are constrained in the sense that $\chi_t = \chi^*$ and $\tilde{\theta}_t = 0$. When $X_t \leq x^*$, entry occurs until $X_t \geq x^*$, where x^* is determined by solving the ODE (83). Equilibrium objects are given by the following set of functions of x which hold for $x \in [x^*, 1]$:*

$$
\begin{aligned}
\eta(x) &= \frac{\max(x,\chi^*)}{x}\sigma_Y \quad \text{and} \quad \tilde{\eta}(x) = \frac{1-\max(x,\chi^*)}{1-x}\sigma_Y \\
r(x) &= \rho + \pi + \mu_Y - \Big[x\eta^2(x) + (1-x)\tilde{\eta}^2(x)\Big] \\
\sigma_X(x) &= x(1-x)\Big[\eta(x) - \tilde{\eta}(x)\Big] \\
\mu_X(x) &= -\pi(1-\alpha)x + x(1-x)\Big[\eta^2(x) - \tilde{\eta}^2(x)\Big] - x(1-x)[x\eta(x) + (1-x)\tilde{\eta}(x)][\eta(x) - \tilde{\eta}(x)].
\end{aligned}
$$

The equilibrium in Proposition 8 features a "safe" risk-sharing region and a "vulnerable" constrained region. In particular, when $X_t \geq \chi^*$, observe that $\eta_t = \tilde{\eta}_t$ and $\sigma_{X,t} = 0$. In this safe region, participants and non-participants perfectly share aggregate risk, because they obtain the same risk compensation and have the same risk preferences. The resulting risk prices equal the representative agent risk price σ_Y.

However, this region is transient. Indeed, notice $\mu_{X,t} < 0$ when $X_t \geq \chi^*$, so that the economy slowly progresses toward the constrained region. Eventually, $X_t < \chi^*$, and perfect risk-sharing is no longer possible, so participants concentrate risk on their balance sheets. In the long run, the economy never leaves this region: the stationary density places zero mass on $\{x \geq \chi^*\}$.

Since risk is concentrated on participants, negative fundamental shocks translate lead to even more concentrated risk ($x \downarrow$), which leads to higher risk prices ($\eta \uparrow$). As such, the equity-issuance economy on $(x^*(\chi^*), \chi^*)$ behaves qualitatively similarly for any χ^*. Despite this qualitative similarity, χ^* has opposing effects on the level of risk prices and their extreme values.

Proposition 9. *Consider a set of alternative economies $\mathcal{E}$, which is parameterized by various values of $\chi^* \in \mathcal{E}$. Let $\eta_t^{\chi^*}, \tilde{\eta}_t^{\chi^*}$ be the equilibrium risk prices in the χ^*-economy. Let $\tau_{x^*}^{\chi^*} := \inf\{t \geq 0 : X_t \leq x^*(\chi^*)\}$ be the first entry time in the χ^*-economy. Then, the following hold:*

(i) For $T \leq \inf_{\mathcal{E}}(\tau_{x^*}^{\chi^*})$, the path $\{\eta_t^{\chi^*} : t \leq T\}$ is uniformly increasing in χ^*, almost-surely.

(ii) Risk price variability $\sup_t(\eta_t^{\chi^*})/\inf_t(\eta_t^{\chi^*})$ is decreasing in χ^*.

(iii) Entry occurs earlier in the sense that $\tau_{x^*}^{\chi^*}$ is decreasing in χ^*.

Proposition 9 shows that more skin-in-the-game implies (i) higher typical risk prices but (ii) less extreme risk price dynamics. This operates through the entry channel, as (iii) suggests. Intuitively, a higher level of risk prices in good times means that non-participants will enter the participant sector earlier, and risk prices at entry will be more moderate. Once again, this illustrates a trade-off between the level and variability of risk compensation in the presence of entry.

Idiosyncratic Risk. One possible reason for slow-moving capital into complex risky asset markets is the presence of idiosyncratic risk embedded in the assets.[9] If such risk in non-diversifiable for participants, entry may be dissuaded even with moderate entry costs.

To study this possibility, modify the economy as follows, with further details in Appendix A.7. Participants' risky asset position is now a claim to $\{\hat{Y}_t\}$, which follows

$$d\hat{Y}_t = \hat{Y}_t[\mu_Y dt + \sigma_Y dZ_t + \hat{\sigma}_Y d\hat{Z}_t],$$

where $\hat{Z}$ is an idiosyncratic Brownian motion, independent of Z. Each participant draws an independent copy of $\hat{Z}$, so that the total risky asset claims in the participant sector will be equal to Y_t, due to the Law of Large Numbers. With these cash flows, participants' risky asset return is

$$dR_t = \mu_{R,t} dt + \sigma_{R,t} dZ_t + \hat{\sigma}_Y d\hat{Z}_t.$$

Participants lever up this asset by the choice variable $\theta_{t,b}$, giving them a total risk exposure of $\theta_{t,b}(\sigma_{R,t} dZ_t + \hat{\sigma}_Y d\hat{Z}_t)$.

In equilibrium, participants will be compensated for their idiosyncratic risk exposure by additional returns. This is captured mathematically by a new idiosyncratic risk price $\hat{\eta}_t$, which is a fictitious construct to capture the residual returns available to participants after they are fairly compensated for aggregate risk. We define η_t and $\hat{\eta}_t$ such that

$$\sigma_{R,t}\eta_t + \hat{\sigma}_Y \hat{\eta}_t := \mu_{R,t} - r_t.$$

The equilibrium and the new idiosyncratic risk price are given as follows. Note that the independence of the idiosyncratic shocks, plus the scale invariance (in wealth) of participants' optimization problems, allows us to continue to study a Markov equilibrium in the single state variable X_t.

Proposition 10. *There exists a unique equilibrium with idiosyncratic risk $\hat{\sigma}_Y > 0$, which is governed by the state variable X_t. In this equilibrium, the idiosyncratic risk price is given by*

$$\hat{\eta}(x) = \hat{\sigma}_Y / x. \tag{40}$$

Since participants earn $\hat{\eta}$, which is increasing in $\hat{\sigma}_Y$ by (40), the presence of idiosyncratic risk makes participants wealthier in the long-run, thus leading to lower aggregate risk prices. Formally, we have the following proposition, which is the analog of claim (i) of Proposition 9.

[9]See Eisfeldt, Lustig, and Zhang (2017) for example. Similarly, the "experts" in Di Tella (2017) are subject to idiosyncratic risk, motivating my choice to include it in this section. The contexts where idiosyncratic risks might be most prevalent include real investment projects by firms' insiders and complex financial markets.

Proposition 11. *Consider a set of alternative economies $\mathcal{E}$, which is parameterized by various values of $\hat{\sigma}_Y \in \mathcal{E}$. Let $\eta_t^{\hat{\sigma}_Y}$ be the equilibrium aggregate risk price in the $\hat{\sigma}_Y$-economy. Let $\tau_{x^*}^{\hat{\sigma}_Y} := \inf\{t \geq 0 : X_t \leq x^*(\hat{\sigma}_Y)\}$ be the first entry time in the $\hat{\sigma}_Y$-economy. Then, for $T \leq \inf_{\mathcal{E}}(\tau_{x^*}^{\hat{\sigma}_Y})$, the path $\{\eta_t^{\hat{\sigma}_Y} : t \leq T\}$ is uniformly decreasing in $\hat{\sigma}_Y$, almost-surely.*

What happens to entry incentives? In this model, participants earn both aggregate and idiosyncratic risk premia, and their entry incentives take both into account. Similar to equation (32), we can write the implied entry costs of this economy as

$$\phi = 1 - \exp\Big(- \frac{1}{2}\mathbb{E}^{x^*}\Big[\int_0^\infty e^{-(\rho+\pi)t}[\eta^2(X_t) + \hat{\eta}^2(X_t)]dt\Big]\Big). \tag{41}$$

With larger $\hat{\sigma}_Y$, $\hat{\eta}_t$ tends to be larger (Proposition 10), but η_t tends to be smaller (Proposition 11). Thus, there is an ambiguous effect on entry incentives. This ambiguity disappears if we study the full-integration cost ϕ^*, analogously to Proposition 4.

Proposition 12. *Define $\phi^* := \sup\{\phi : x^*(\phi) = 1\}$. Then,*

$$\phi^* = 1 - \exp\Big(- \frac{1}{2}(\rho + \pi)^{-1}[\sigma_Y^2 + \hat{\sigma}_Y^2]\Big). \tag{42}$$

From (42), we see that ϕ^*, a measure of participation incentives, is increasing in $\hat{\sigma}_Y$. Since idiosyncratic risk is compensated, entry can become more attractive, not less.[10] Combining this result with the result of Proposition 11, we find that the presence of idiosyncratic risk can reduce both the level and variability of aggregate risk prices.

5 Extrapolative expectations

The results so far, that asset markets imply unreasonably high entry costs, can be rephrased in the form of a question: why is capital so slow-moving, especially in crises when high risk premia prevail (e.g., Duffie (2010b))? In this section, I introduce extrapolative expectations into the model to help explain why entry may not occur when risk premia are high.

To do this, I build on Barberis, Greenwood, Jin, and Shleifer (2015), who present a model in which some investors extrapolate past price movements rather than rationally computing expected price changes. In equilibrium, agents tend to believe expected returns are high in "good times," which have resulted from a run-up in prices, and low in "bad times," in accordance with survey evidence (e.g., Greenwood and Shleifer (2014)), but at odds with reality.

[10] Combining idiosyncratic risk, equity-issuance, and recursive preferences all together, corresponding loosely to a model like Di Tella (2017), we find

$$\phi^* \approx \frac{\frac{1}{2}\chi^*\hat{\sigma}_Y\hat{\eta}^*}{\rho + \pi(1-\alpha) + \psi(\mu_Y + \pi\alpha)} + O(\sigma_Y^4) + O(\hat{\sigma}_Y^4).$$

Under parameters of Table 1, and also with $\gamma = 5$, $\psi = 0.5$, $\chi^* = 0.2$, and $\hat{\sigma}_Y = 0.25$ (all exactly as in Di Tella (2017)), this approximation delivers $\phi^* \approx 89\%$. Thus, this class of models requires a huge implied entry cost to deliver any dynamics at all.

Define agents' *sentiment* about financial markets by S_t, which is an exponentially-weighted average of previous returns, with higher weights on more recent returns:

$$S_t := \beta \int_{-\infty}^{t} e^{-\beta(t-s)} dR_s. \tag{43}$$

In changes, sentiment follows

$$dS_t = \beta(dR_t - S_t dt). \tag{44}$$

Equations (43) and (44) capture the idea that a string of positive (negative) returns increases (decreases) sentiment, while sentiment mean-reverts in absence of trends.[11]

Sentiment is the basis of extrapolative expectations: all agents have biased expectations about returns, leaning in the direction of their sentiment. Mathematically, I assume perceived expected returns are a weighted average of actual expected returns and the level of sentiment:

$$\tilde{\mu}_{R,t} := \mu_{R,t} + \lambda(S_t - \mu_{R,t}). \tag{45}$$

Equation (45) is the key assumption of this section. Note that λ controls the degree of bias in agents' expectation-formation: $\lambda = 1$ implies agents are fully extrapolative, while $\lambda = 0$ implies agents are fully rational.[12]

The equilibrium is given by the following proposition. The proof is in Appendix A.9.

Proposition 13. *There exists a unique Markov equilibrium with sentiments, which is governed by the state variable (X_t, S_t). Entry occurs whenever $X_t \leq x^*$, where x^* is identical to the entry point without sentiments. Equilibrium objects are characterized by the following functions of (x, s) which hold on $[x^*, 1) \times (-\infty, +\infty)$. For asset prices, we have*

$$\eta(x, s) = \left[1 + \frac{\lambda}{\sigma_Y^2}(s - \bar{s})\right]^{-1} \frac{\sigma_Y}{x}$$

$$r(x, s) = \bar{s} + \lambda(s - \bar{s}) - \frac{\sigma_Y^2}{x},$$

[11] Notice that sentiment is based on returns rather than fundamentals. Fundamentals extrapolation appears similar on the surface but ultimately would not deliver the appropriate dynamics, because prices adjust immediately to biased beliefs about dividend growth.

[12] This assumption allows for flexibility in the degree of rationality in the economy, while maintaining tractability. In tying the expectation function to the true expected returns, we might also interpret λ as a reduced-form for heterogeneity in investor sophistication and rationality. Barberis et al. (2015) include both rational and irrational traders, assuming the irrational traders have expectations that only depend on sentiment and not the actual expected returns. In such a model, we would additionally have to keep track of the wealth distribution between extrapolators and non-extrapolators, making it a less tractable alternative.

and for state dynamics

$$\begin{aligned}
\sigma_X(x,s) &= (1-x)\sigma_Y \\
\mu_X(x,s) &= -\pi(1-\alpha)x + \sigma_Y^2\frac{(1-x)^2}{x} - \lambda(s-\overline{s})(1-x) \\
\sigma_S(x,s) &= \beta\sigma_Y \\
\mu_S(x,s) &= -\beta(s-\overline{s}),
\end{aligned}$$

where $\overline{s} := \mathbb{E}S_t = \rho + \pi + \mu_Y$ *is the average sentiment level.*

In the model with sentiment, entry occurs when the perceived, rather than actual, participation benefits are high. Perceived benefits are given by

$$\frac{1}{2}(\rho+\pi)^{-1}\tilde{\mathbb{E}}\Big[\int_0^\infty e^{-(\rho+\pi)t}\tilde{\eta}_t^2 dt \mid \mathcal{F}_0\Big],$$

where $\tilde{\mathbb{E}}$ denotes expectations under the extrapolative beliefs and $\tilde{\eta}_t$ denotes the perceived Sharpe ratio (this formula is the irrational beliefs generalization of (29)). The perceived Sharpe ratio is given by $\tilde{\eta}(x,s) = \frac{\sigma_Y}{x}$, which is independent of s and identical to the actual Sharpe ratio in the rational model.[13] Similarly, extrapolative agents with log utility perceive the dynamics of (X_t, S_t) to be independent of S_t, so the subjective expectation $\tilde{\mathbb{E}}$ is unaffected by sentiments. The independence of $\tilde{\eta}_t$ and $\tilde{\mathbb{E}}$ from S_t is why the computation of the entry boundary is simple and unaffected by sentiments.

As anticipated, market risk premia are more volatile with sentiments. The actual Sharpe ratio is given by $\frac{\mu_R(x,s)-r(x,s)}{\sigma_R(x,s)} = \frac{\sigma_Y}{x} - \frac{\lambda}{\sigma_Y}(s-\overline{s})$, which is equal to the perceived Sharpe ratio $\tilde{\eta}$, plus a term capturing sentiments. Thus, the Sharpe ratio is decreasing in both X_t and S_t, which are positively correlated (i.e., $\sigma_X, \sigma_S > 0$). Mathematically, the local volatility of the Sharpe ratio is

$$\operatorname{std}\Big[\frac{\mu_{R,t}-r_t}{\sigma_{R,t}}\Big] = \lambda\beta + \sigma_Y^2\frac{1-X_t}{X_t^2}, \tag{46}$$

which is increasing in λ, the degree of extrapolation in this economy. Thus, Sharpe ratios are more volatile with sentiments.[14]

These results suggest that a model with sentiments can generate volatile risk prices and slow-moving capital in crises. Because $\frac{\mu_R - r}{\sigma_R}$ is decreasing in s, while entry decisions are independent of s, sentiments provide a channel in which risk prices can rise in bad times without entry. With rational beliefs, entry occurs when risk prices are high, preventing even higher risk prices. With irrational beliefs, entry is less likely, as agents may not perceive high risk prices in crises.

[13] The reasoning is as follows. Asset market clearing pins down participants' risky portfolios, i.e., $\theta_P = 1/x$. With log utility, participants have no hedging demand against changes in sentiments, and $\theta_P = \tilde{\eta}/\sigma_Y$. Hence, $\tilde{\eta} = \sigma_Y/x$.

[14] Importantly, this result does not rely on increasing agents' consumption growth volatility. Indeed, $\tilde{\eta}_t = \sigma_Y/X_t$ represents participants' local consumption growth volatility, which is independent of λ.

This story is confirmed in the left panel of figure 7, which plots average Sharpe ratios against Sharpe ratio volatility, for different values of λ. Clearly, Sharpe ratio volatilities are increasing in λ, as suggested by the local equation (46) above. At the same time, average Sharpe ratios are not decreasing in λ. Intuitively, extrapolative agents do not perceive the portion of Sharpe ratio volatility that is due to their sentiments. Thus, an increase in λ can increase Sharpe ratio volatility without increasing entry incentives, thereby sidestepping the conditional-unconditional trade-off documented in section 4. With 25% entry costs ($\phi = 0.25$) and full extrapolation ($\lambda = 1$), this economy can generate an average Sharpe ratio of 0.3 and a Sharpe ratio volatility of 0.38.

This model also adds an element of procyclicality to participation, consistent with the stock market evidence in Kaustia and Knüpfer (2012). This is confirmed by calculating the probability that entry occurs in good times. Specifically, in the right panel of figure 7, I plot the probability that, when entry occurs, sentiments are high ($S_t \geq \overline{s}$), for different values of λ. This probability, which I interpret as the likelihood of procyclical entry, increases in the degree of extrapolation λ.

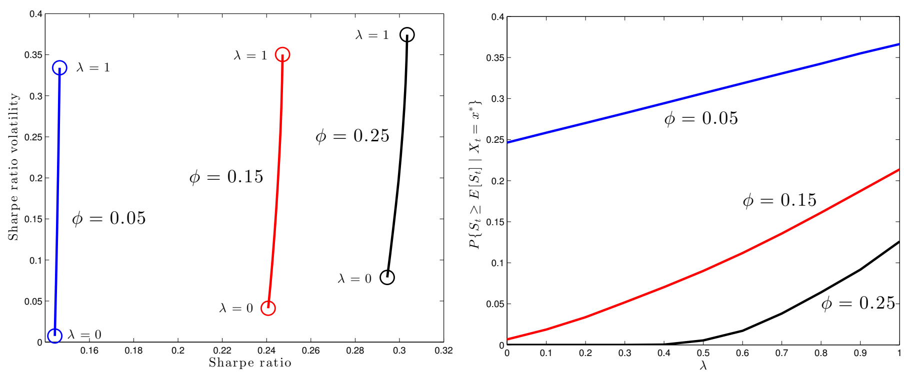

Figure 7: Left panel: unconditional average Sharpe ratios (horizontal axis) against unconditional Sharpe ratio volatility (vertical axis). Right panel: probability of entry in high-sentiment states (defined by $S_t \geq \mathbb{E}S_t = \rho + \pi + \mu_Y$) as a function of λ. Averages are computed by Monte Carlo simulations. Fixed parameters are $\rho = 0.01$, $\alpha = 0.5$, $\mu_Y = 0.02$, $\sigma_Y = 0.04$, $\pi = 0.2$, $\gamma = \psi = 1$, $\beta = 0.2$. Variable parameters are ϕ and λ.

The logic for procyclical participation can be seen most clearly in the phase diagram in figure 8. The arrows represent the drifts of the state variable (X_t, S_t). The entry point is given by x^*, which is the dotted vertical line. Suppose X_t is low so that entry is possible in the near future. If S_t is also low (bottom shaded region, in blue), Sharpe ratios are high, so participants accumulate wealth quickly (i.e., μ_X is high). This force buoys the economy, and entry may not be needed. Conversely, if S_t is high (top shaded region, in gray), Sharpe ratios are low, so participants are losing wealth share (i.e., μ_X is low). As X_t drifts towards x^*, a small number of negative fundamental shocks will induce entry. Thus, relative to an economy without sentiments, entry occurs more often in good times when Sharpe ratios are low, consistent with the data. And since higher λ increases Sharpe ratio volatility, the degree of entry procyclicality strengthens with stronger extrapolation.

In summary, extrapolative expectations can bring models of market segmentation closer to the data, despite the presence of entry: sentiments increase risk price volatilities, provide some procyclical motives for entry, and do all this without compromising risk price levels.

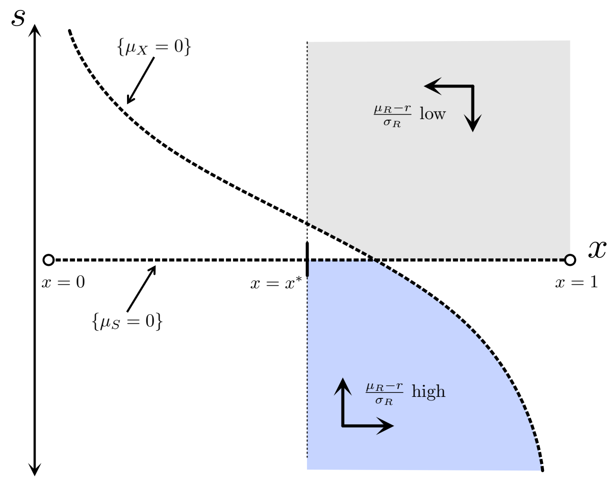

Figure 8: Phase diagram. The thick dashed lines show denote the loci of $\mu_X = 0$ and $\mu_S = 0$. The intersection of these dashed lines separate the state space into quadrants. In the shaded regions, the arrows point in the direction of the state variable drifts. The entry boundary is denoted by the thin dotted line at the point $x = x^*$. All elements of the figure are purely illustrative, although the shapes of the loci, state drifts, and entry boundary are general.

6 Conclusion

Asset market data suggest that, in the context of limited participation models, entry costs must be unreasonably high. Because the compensation to participating in financial markets is large, entry costs must be very large to generate significant market segmentation.

First, because the compensation to participating in financial markets is large, participation costs must be at least as large to generate any market segmentation. Second, unless participation costs are very large, the economy only displays subdued risk price dynamics due to a combination of entry and the buoying effect of higher returns on participant wealth. Third, I illustrate a trade-off between risk price levels and their dynamics in limited participation models.

What types of features can interact productively with endogenous entry and avoid the critiques outlined in this paper? I propose one possibility: extrapolative expectations. Whereas the limited participation model with rational agents generates countercyclical entry, extrapolative expectations add a procyclical motive, namely that perceived entry benefits are high when actual expected returns are low. Among the mechanisms considered in this paper, this is the only one which can generate both high Sharpe ratio levels and high Sharpe ratio volatility. These features help bring the limited participation model closer to the participation and asset price data, while remaining disciplined by the survey data which motivates extrapolative expectations in the first place.

A Appendix

In section A.1, I generalize the model by introducing recursive preferences and derive the equilibrium. In section A.2, I verify that the HJB equations and associated boundary conditions are sufficient for optimality in individual agents' control problems. Section A.3 provides details on the full-integration equilibrium (i.e., for ϕ small enough such that $x^* = 1$). Section A.4 provides proofs for the model with log utility. Section A.5 discusses the fact that endogenous entry rules out bubbles and arbitrages in the economy, which are present in many limited participation economies. Section A.6 presents an economy with log utility and exogenous segmentation. Section A.7 adds both equity issuance and idiosyncratic risks to the log utility economy, and derives the equilibrium. Section A.8 adds exogenous state variables and shows that the entry decisions continue to be made according to the level and dynamics of risk prices. Finally, section A.9 adds sentiments and extrapolative expectations to the log utility economy, and derives the equilibrium. The Online Appendix B contains further results.

A.1 Generalization to recursive utility

This section presents more details on the model environment under the recursive utility generalization. The utility function is defined by (35)-(36) in the text. The returns on the risky asset need to be modified to

$$dR_t = \mu_{R,t}dt + \sigma_{R,t}dZ_t + dA_t^R,$$

where A^R is a non-decreasing, singularly continuous process.[15] The bond pays an instantaneous return of $r_t dt + dA_t^R$.[16] The presence of the singular term dA_t^R, though unusual, is due to the equilibrium entry in the model in conjunction with these more general preferences. The state-price density process is modified to

$$\xi_t := \exp\Big\{ -\int_{-\infty}^{t} \Big(r_s + dA_s^R + \frac{1}{2}\eta_s^2 \Big) ds - \int_{-\infty}^{t} \eta_s dZ_s \Big\}. \tag{47}$$

Participants' wealth dynamics are now given by

$$dW_{t,b} = (r_t W_{t,b} + \theta_{t,b} W_{t,b}(\mu_{R,t} - r_t) + \alpha\pi W_{t,b} - c_{t,b})dt + W_{t,b}dA_t^R + \theta_{t,b}W_{t,b}\sigma_{R,t}dZ_t, \quad t \geq \tau_b. \tag{48}$$

Given this new utility and budget constraint, participants' optimization problems are now given by

$$V_{t,b}^P = \sup_{c,\theta} \mathbb{E}\Big[\int_t^{\infty} f(c_{s,b}, V_{s,b}^P) ds \mid \mathcal{F}_t \Big], \tag{49}$$

subject to (48). Non-participants solve

$$V_{t,b}^N = \sup_{c,\tau} \mathbb{E}\Big[\int_t^{\tau} f(c_{s,b}, V_{s,b}^N) ds + V_{\tau,b}^P - \Phi_{\tau,b} \mid \mathcal{F}_t \Big], \tag{50}$$

where entry costs $\Phi_{t,b}$ are given by (37), and the optimization problem is subject to the wealth dynamics

$$dW_{t,b} = (r_t W_{t,b} + \alpha\pi W_{t,b} - c_{t,b})dt + W_{t,b}dA_t^R, \quad t < \tau_b, \quad W_{b,b} > 0 \quad \text{given.} \tag{51}$$

[15] So A^R is of bounded variation but not absolutely continuous with respect to Lebesgue measure.

[16] No arbitrage requires that the singular component of the bond process be identical to that of the stock. See Karatzas and Shreve (1998), appendix B, for a proof.

Homogeneity properties. As before, scalability properties of the model allow for a convenient representation of value functions. We have

$$V^P_{t,b} = \frac{W^{1-\gamma}_{t,b}}{1-\gamma} G^P_t \quad \text{and} \quad V^N_{t,b} = \frac{W^{1-\gamma}_{t,b}}{1-\gamma} G^N_t,$$

where G^P and G^N are processes independent of agents' wealth. The key to achieving this is the homogeneity of the entry cost (37), so that the payoff to an entrant at time $t \geq b$ is $V^P_{t,b} - \Phi_{t,b} = (1-\phi)^{1-\gamma} V^P_{t,b} = \frac{((1-\phi)W_{t,b})^{1-\gamma}}{1-\gamma} G^P_t$. Thus, the cost $\Phi_{t,b}$ is perceived as a fraction ϕ of wealth, as in the log utility model. Consequently, entry incentives are summarized by

$$t \in \mathcal{T}^* : \ (1-\phi)^{1-\gamma} G^P_t = G^N_t; \tag{52}$$

$$t \notin \mathcal{T}^* : \ (1-\phi)^{1-\gamma} G^P_t (1-\gamma)^{-1} < G^N_t (1-\gamma)^{-1}, \tag{53}$$

where $\mathcal{T}^*$ denotes the set of entry times.

Solving for Markov equilibrium. In a Markov equilibrium, with state variable X_t, there are functions g_P and g_N such that $G^P_t = g_P(X_t)^{\frac{\psi(1-\gamma)}{1-\psi}}$ and $G^N_t = g_N(X_t)^{\frac{\psi(1-\gamma)}{1-\psi}}$. As before, apply dynamic programming to the participants' and non-participants' problems, leading to two ODEs (the HJB equations) for the wealth-consumption ratios g_P and g_N:

$$\begin{aligned} 0 = \psi + \Big[-\rho - \pi + (1-\psi)\Big(r + \alpha\pi + \frac{1}{2\gamma}\eta^2\Big)\Big] g_P + \Big[\psi\mu_X + \frac{\psi}{\gamma}(1-\gamma)\eta\sigma_X\Big] g_P' + \frac{1}{2}\psi\sigma_X^2 g_P'' \\ + \frac{1}{2}\frac{\psi(\psi-\gamma)}{\gamma(1-\psi)}\sigma_X^2 \frac{(g_P')^2}{g_P} \end{aligned} \tag{54}$$

$$\begin{aligned} 0 = \psi + \Big[-\rho - \pi + (1-\psi)(r + \alpha\pi)\Big] g_N + \psi\mu_X g_N' + \frac{1}{2}\psi\sigma_X^2 g_N'' \\ + \frac{1}{2}\Big(\frac{\psi(\psi-\gamma)}{\gamma(1-\psi)} - \frac{\psi^2(1-\gamma)^2}{\gamma(1-\psi)}\Big)\sigma_X^2 \frac{(g_N')^2}{g_N}. \end{aligned} \tag{55}$$

These ODEs are solved on $(x^*, 1)$ with an endogenous boundary x^*. Boundary conditions for these ODEs are the following. First, the value functions satisfy (52) at entry times, i.e., times where $X_t = x^*$, implying the *value-matching* condition

$$(1-\phi)^{\frac{1-\psi}{\psi}} g_P(x^*) = g_N(x^*). \tag{56}$$

Next, the *smooth-pasting* conditions

$$g_P'(x^*) = g_N'(x^*) = 0 \tag{57}$$

also hold at the entry point x^*. These are three boundary conditions at $x = x^*$. The other two conditions are derived by taking the limits of (54)-(55) as $x \to 1$. In Appendix A.2, I derive the HJB equations, discuss boundary conditions, and finally prove that the HJB equations and associated boundary conditions are sufficient for individual optimality (Proposition 16).

In Proposition 14, I solve for all equilibrium objects, up to the solutions g_P and g_N to (54)-(55). This demonstrates the tractability of the setup, although it is more complicated than the log utility model. The basic steps in determining η, r, μ_X, and σ_X are to apply Itô's formula to the goods market clearing equation and the definition of the state variable, for $X_t \in [x^*, 1)$. The proof is at the end of this section.

Proposition 14. *There exists a stationary Markov equilibrium defined by asset prices*

$$
\begin{aligned}
\eta(x) &:= \Big[1 + \frac{1-x}{x}\omega(x)\Big]\gamma\sigma_Y + \Big(\frac{\gamma-\psi}{1-\psi}\Big)\sigma_Y(1-x)\frac{g_P'(x)}{g_P(x)} \\
r(x) &:= \rho + \psi\mu_Y - \frac{1}{2}\gamma(\psi+1)\sigma_Y^2 + \pi(1-\alpha) + \psi\pi - \psi\pi(1-\alpha)\frac{p(x)}{g_N(x)} - \frac{1}{2}\gamma(\psi+1)\sigma_Y^2\Big(\frac{1-x}{x}\Big)\omega^2(x) \\
&\quad + \frac{1}{2}\Big(\frac{\gamma-\psi}{1-\psi}\Big)\sigma_Y^2(1-x)(1-\omega(x))^2 - \Big(\frac{\gamma-\psi}{1-\psi}\Big)\sigma_Y^2\Big[x\zeta(x) + \frac{x}{2}\zeta^2(x) + (1-x)\omega(x)\zeta(x)\Big] \\
p(x) &:= xg_P(x) + (1-x)g_N(x) \\
\sigma_R(x) &:= \sigma_Y\Big[1 + (1-x)\omega(x)\frac{p'(x)}{p(x)}\Big] \\
\mu_R(x) &:= r(x) + \sigma_R(x)\eta(x)
\end{aligned}
$$

and state dynamics

$$
\begin{aligned}
\mu_X(x) &:= -\pi(1-\alpha)x\frac{p(x)}{g_N(x)} + \Big(\frac{\gamma(\psi+1)}{\psi} - 1\Big)\sigma_Y^2(1-x)\omega(x) \\
&\quad + \frac{1}{2}\frac{\gamma(\psi+1)}{\psi}\sigma_Y^2\Big(\frac{1-x}{x}\Big)(1-2x)\omega^2(x) + \frac{1}{2}\Big(\frac{\gamma-\psi}{\psi(1-\psi)}\Big)\sigma_Y^2 x(1-x)(1-\omega(x))^2 \\
&\quad + \Big(\frac{\gamma-\psi}{\psi(1-\psi)}\Big)\sigma_Y^2(1-x)\Big[x\zeta(x) + \frac{x}{2}\zeta^2(x) + (1-x)\omega(x)\zeta(x)\Big] \\
\sigma_X(x) &:= (1-x)\omega(x)\sigma_Y
\end{aligned}
$$

on $[x^*, 1)$*, where*

$$
\begin{aligned}
\omega(x) &:= \Big(1 - (1-x)\frac{g_N'(x)}{g_N(x)}\Big)^{-1} \\
\zeta(x) &:= (1-x)\frac{g_P'(x)}{g_P(x)}\omega(x),
\end{aligned}
$$

and functions g_P *and* g_N*, with endogenous entry point* x^**, satisfy the ordinary differential equations (22) and (23) subject to boundary conditions given by (56), (57), (65), and (66), assuming these ODEs have a solution. In that case,* (X, A^{x^*}) *is the unique strong solution to (17). Finally, the non-degenerate stationary density of* X_t *is given by*

$$
h(x) = \frac{K_0}{\sigma_X^2(x)}\exp\Big(\int_{x^*}^{x}\frac{2\mu_X(y)}{\sigma_X^2(y)}dy\Big),
$$

for $x \in [x^*, 1)$*, where* K_0 *is a constant chosen to ensure* h *integrates to 1, i.e.,* $\int_{x^*}^1 h(x)dx = 1$.

Canonical limited participation dynamics. To see that the equilibrium in Proposition 14 corresponds closely to the canonical limited participation model, I plot the key equilibrium asset prices in figure 9 (table 3 lists the benchmark parameters). I compare asset prices with costly entry (solid blue) to asset prices in an economy with costless entry and full participation (dashed black). Notice that risk prices (η) are higher and risk-free rates (r) lower with costly entry, and these effects strengthen as participants' wealth falls. This is qualitatively the same as the log utility model.

The results of figure 9 can be verified analytically. Notice that equilibrium objects can be decomposed into terms from a frictionless economy, terms arising due to limited participation ("LP effects"), and terms

Table 3: Recursive utility calibration of model parameters.

Parameter:	ρ	π	α	RRA (γ)	EIS (ψ^{-1})	μ_Y	σ_Y	ϕ
Value:	0.01	0.02	0.50	3	1.33	0.02	0.04	0.4

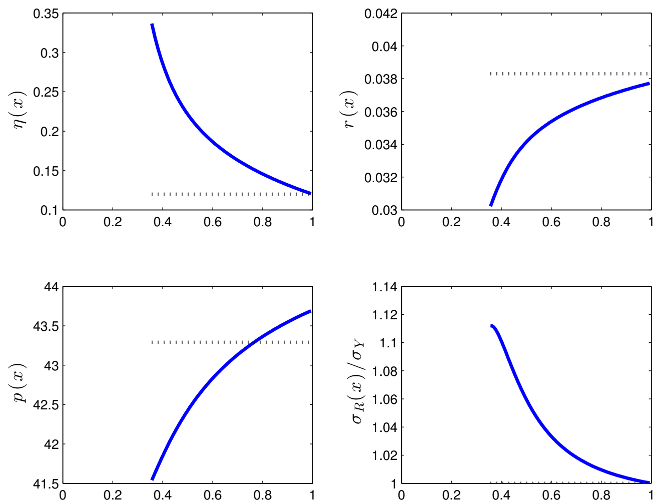

Figure 9: Asset prices in the benchmark limited participation economy with recursive preferences and costly entry (blue) versus complete integration (dashed black). The horizontal axis is the participants' consumption share x.

due to recursive preferences. For example, the market price of risk can be understood this way:

$$\eta(x) = \underbrace{\gamma\sigma_Y}_{\text{frictionless}} + \underbrace{\gamma\sigma_Y\Big(\frac{1-x}{x}\Big)\omega(x)}_{\text{LP effects}} + \underbrace{\Big(\frac{\gamma-\psi}{1-\psi}\Big)\sigma_Y(1-x)\frac{g_P'(x)}{g_P(x)}}_{\text{recursive preferences}}. \tag{58}$$

When $\psi = \gamma$, corresponding to power utility, the recursive preference terms disappear. The risk-free rate has additional terms arising from the OLG environment, and can be decomposed as follows:

$$r(x) = \underbrace{\rho + \psi\mu_Y - \frac{1}{2}\gamma(\psi+1)\sigma_Y^2}_{\text{frictionless}} + \underbrace{\pi(1-\alpha) + \psi\pi - \psi\pi(1-\alpha)\frac{p(x)}{g_N(x)}}_{\text{OLG effects}} \underbrace{- \frac{1}{2}\gamma(\psi+1)\sigma_Y^2\Big(\frac{1-x}{x}\Big)\omega^2(x)}_{\text{LP effects}}$$
$$\underbrace{+ \frac{1}{2}\Big(\frac{\gamma-\psi}{1-\psi}\Big)\sigma_Y^2(1-x)(1-\omega(x))^2 - \Big(\frac{\gamma-\psi}{1-\psi}\Big)\sigma_Y^2\Big[x\zeta(x) + \frac{x}{2}\zeta^2(x) + (1-x)\omega(x)\zeta(x)\Big]}_{\text{recursive preferences}}. \tag{59}$$

Next, I show that the contributions of the "LP effects" terms are generally to lower r and increase η. The proof is at the end of this section.

Proposition 15. *Define $r_{LP}(x) := -\frac{1}{2}\gamma(\psi+1)\sigma_Y^2(\frac{1-x}{x})\omega^2(x)$ and $\eta_{LP}(x) := \gamma\sigma_Y(\frac{1-x}{x})\omega(x)$, which are the effects of limited participation in (58) and (59). Then, there exist choices for the parameters and $\varepsilon > 0$ such that the following hold for all $x \in [x^*, x^*+\varepsilon) \cup (1-\varepsilon, 1)$,*

$$r_{LP}(x) < 0 \quad \textit{and} \quad r_{LP}'(x) > 0$$
$$\eta_{LP}(x) > 0 \quad \textit{and} \quad \eta_{LP}'(x) < 0.$$

The choice to include recursive utility in this model is partially motivated by the desire to give the limited participation model its best shot quantitatively. I relegate a detailed discussion to Online Appendix B.1, as the effects of this utility specification on asset prices are relatively standard, even in this model. In short, with recursive utility, the model can deliver (a) low risk-free rates; (b) procyclical price-dividend ratios and counter-cyclical return volatility; and (c) a large component of equity volatility which is unrelated to the risk-free rate. Some of these properties I verify analytically in the Appendix (see Lemmas B.1 and B.2). Without recursive utility, none of these features would be present, which might lead us to conclude a failure of the model, but which is actually unrelated to the limited participation channel. To avoid this, I simply endow agents with recursive utility.

Entry and singular asset prices. Examining the proof of Proposition 14 shows that returns contain a singularly continuous component at points of entry, i.e.,

$$dA_t^R = \frac{p'(x^*)}{p(x^*)} dA_t^{x^*} \neq 0$$

in return dynamics $dR_t = \mu_{R,t}dt + \sigma_{R,t}dZ_t + dA_t^R$.[17] Consequently, there is no well-defined expected rate of return on assets at those times. Importantly, singular asset prices are not inconsistent with absence of arbitrage as long as the singular components of risky and risk-free assets coincide, as shown by Karatzas, Lehoczky, and Shreve (1991). The intuition is that there can be an infinite expected rate of return in the economy if and only if the opportunity cost of obtaining that return is also infinite. This model, which features micro-founded entry decisions, is arbitrage-free and yet has singular asset prices.[18]

Below, I prove Proposition 14. The proof takes as given agents' optimal controls to construct an equilibrium. These optimal controls are contained in (63), (64), and (70) in Appendix A.2. For verification of the optimality of these controls, see Proposition 16 in Appendix A.2.

Proof of Proposition 14. The proof proceeds in four steps. First, we derive coefficients for the state-price density (ξ) and consumption distribution (X), in terms of the participants' and non-participants' wealth-consumption ratios (g_P and g_N). Second, we solve for the price-dividend ratio, stock volatility, expected stock returns, and the singular component of returns. Third, we verify some technical conditions, required for existence of a state-price density and a solution to the SDE for X. Finally, we compute the stationary density of X.

Step 1: State-price density and consumption distribution.

[17] The fact that this singular component is non-zero can be attributed to the non-zero derivative of the price-dividend ratio at x^*. Using the value-matching and smooth-pasting conditions (56)-(57),

$$p'(x^*) = x^* g_P'(x^*) + (1-x^*)g_N'(x^*) + g_P(x^*) - g_N(x^*) = \left[1-(1-\phi)^{\frac{1-\psi}{\psi}}\right]g_P(x^*) \gtrless 0 \quad \text{as} \quad \psi \lessgtr 1.$$

The intuition for the sign of $p'(x^*)$ is as follows. When $\psi < 1$, the EIS is high, and agents are willing to tolerate less consumption-smoothing than they would with log preferences. The standard result is that the price-dividend ratio is procyclical, and it is sensible by extension that $p(x)$ would rise as entry occurs.

[18] Conversely, in many models with reflecting boundaries, an equilibrium is determined in which returns contain no singular components, in order to prevent arbitrage. See Brunnermeier and Sannikov (2014), the version of the model in which utility is logarithmic, or He and Krishnamurthy (2014). The presence of singularities in returns depends on the derivative of the price-dividend ratio at x^*, which is zero in those papers but non-vanishing in this model (see previous footnote).

First, write down the consumption dynamics for participants and non-participants by applying Itô's formula to $c_{t,b}$. For $i \in \{P, N\}$ according to whether $b \in \mathcal{P}_t$ or $\mathcal{N}_t$, we have

$$\begin{aligned}
dc_{t,b} = d\Big(\frac{W_{t,b}}{g_i(X_t)}\Big) &= \Big[\frac{W_{t,b}}{g_i(X_t)}\Big(r_t + \alpha\pi + \theta_{t,b}(\mu_{R,t} - r_t) - g_i(X_t)^{-1}\Big) - \frac{W_{t,b}}{g_i(X_t)}\frac{g_i'(X_t)}{g_i(X_t)}\mu_{X,t} \\
&\quad - \frac{1}{2}\frac{W_{t,b}}{g_i(X_t)}\Big(\frac{g_i''(X_t)}{g_i(X_t)} - \Big(\frac{g_i'(X_t)}{g_i(X_t)}\Big)^2\Big)\sigma_{X,t}^2 - \frac{W_{t,b}}{g_i(X_t)}\frac{g_i'(X_t)}{g_i(X_t)}\theta_{t,b}\sigma_{R,t}\sigma_{X,t}\Big]dt \\
&\quad + \Big[\frac{W_{t,b}}{g_i(X_t)}\theta_{t,b}\sigma_{R,t} - \frac{W_{t,b}}{g_i(X_t)}\frac{g_i'(X_t)}{g_i(X_t)}\sigma_{X,t}\Big]dZ_t \\
&= c_{t,b}\Big[r_t + \alpha\pi + \theta_{t,b}(\mu_{R,t} - r_t) - g_i(X_t)^{-1} - \frac{g_i'(X_t)}{g_i(X_t)}\mu_{X,t} + \Big(\frac{g_i'(X_t)}{g_i'(X_t)}\Big)^2\sigma_{X,t}^2 \\
&\quad - \frac{1}{2}\frac{g_i''(X_t)}{g_i(X_t)}\sigma_{X,t}^2 - \frac{g_i'(X_t)}{g_i(X_t)}\theta_{t,b}\sigma_{R,t}\sigma_{X,t}\Big]dt + c_{t,b}\Big[\theta_{t,b}\sigma_{R,t} - \frac{g_i'(X_t)}{g_i(X_t)}\sigma_{X,t}\Big]dZ_t.
\end{aligned}$$

Using the fact that $\theta_{t,b} = \frac{\eta_t}{\gamma\sigma_{R,t}} + \frac{\sigma_{X,t}}{\sigma_{R,t}}\frac{\psi(1-\gamma)}{\gamma(1-\psi)}\frac{g_P'(X_t)}{g_P(X_t)}$ for $b \in \mathcal{P}_t$ and $\theta_{t,b} \equiv 0$ for $b \in \mathcal{N}_t$, we have

$$\begin{aligned}
\frac{dc_{t,b}}{c_{t,b}} &= \Big[r_t + \alpha\pi + \frac{\eta_t^2}{\gamma} - g_P(X_t)^{-1} + \Big(\Big(\frac{\psi(1-\gamma)}{\gamma(1-\psi)} - \frac{1}{\gamma}\Big)\eta_t\sigma_{X,t} - \mu_{X,t}\Big)\frac{g_P'(X_t)}{g_P(X_t)} \\
&\quad - \frac{\psi-\gamma}{\gamma(1-\psi)}\sigma_{X,t}^2\Big(\frac{g_P'(X_t)}{g_P(X_t)}\Big)^2 - \frac{1}{2}\sigma_{X,t}^2\frac{g_P''(X_t)}{g_P(X_t)}\Big]dt \\
&\quad + \Big(\frac{\eta_t}{\gamma} + \frac{\psi-\gamma}{\gamma(1-\psi)}\sigma_{X,t}\frac{g_P'(X_t)}{g_P(X_t)}\Big)dZ_t, \quad b \in \mathcal{P}_t \\
\frac{dc_{t,b}}{c_{t,b}} &= \Big[r_t + \alpha\pi - g_N(X_t)^{-1} - \mu_{X,t}\frac{g_N'(X_t)}{g_N(X_t)} + \sigma_{X,t}^2\Big(\frac{g_N'(X_t)}{g_N(X_t)}\Big)^2 - \frac{1}{2}\sigma_{X,t}^2\frac{g_N''(X_t)}{g_N(X_t)}\Big]dt \\
&\quad - \sigma_{X,t}\frac{g_N'(X_t)}{g_N(X_t)}dZ_t, \quad b \in \mathcal{N}_t.
\end{aligned}$$

Substituting the second derivatives from the HJB equations (22)-(23), we have

$$\begin{aligned}
\frac{dc_{t,b}}{c_{t,b}} &= \frac{1}{\psi}\Big[r_t + \alpha\pi - \rho - \pi + \frac{\psi+1}{2\gamma}\eta_t^2 + \frac{\psi(\psi-\gamma)}{\gamma(1-\psi)}\eta_t\sigma_{X,t}\frac{g_P'(X_t)}{g_P(X_t)} - \frac{1}{2}\frac{\psi(\psi-\gamma)}{\gamma(1-\psi)}\sigma_{X,t}^2\Big(\frac{g_P'(X_t)}{g_P(X_t)}\Big)^2\Big]dt \\
&\quad + \Big(\frac{\eta_t}{\gamma} + \frac{\psi-\gamma}{\gamma(1-\psi)}\sigma_{X,t}\frac{g_P'(X_t)}{g_P(X_t)}\Big)dZ_t, \quad b \in \mathcal{P}_t \\
\frac{dc_{t,b}}{c_{t,b}} &= \frac{1}{\psi}\Big[r_t + \alpha\pi - \rho - \pi + \frac{1}{2}\frac{\psi}{1-\psi}(1-\gamma\psi)\sigma_{X,t}^2\Big(\frac{g_N'(X_t)}{g_N(X_t)}\Big)^2\Big]dt - \sigma_{X,t}\frac{g_N'(X_t)}{g_N(X_t)}dZ_t, \quad b \in \mathcal{N}_t.
\end{aligned}$$

Apply Itô's formula to the goods market clearing equation (12) and match drifts and diffusions,

$$\begin{aligned}
\mu_Y &= \pi\Big(\frac{c_{t,t}}{Y_t}\Big) - \pi + \frac{X_t}{\psi}\Big[r_t + \alpha\pi - \rho - \pi + \frac{\psi+1}{2\gamma}\eta_t^2 + \frac{\psi(\psi-\gamma)}{\gamma(1-\psi)}\eta_t\sigma_{X,t}\frac{g_P'(X_t)}{g_P(X_t)} - \frac{1}{2}\frac{\psi(\psi-\gamma)}{\gamma(1-\psi)}\sigma_{X,t}^2\Big(\frac{g_P'(X_t)}{g_P(X_t)}\Big)^2\Big] \\
&\quad + \frac{1-X_t}{\psi}\Big[r_t + \alpha\pi - \rho - \pi + \frac{1}{2}\frac{\psi}{1-\psi}(1-\gamma\psi)\sigma_{X,t}^2\Big(\frac{g_N'(X_t)}{g_N(X_t)}\Big)^2\Big] \\
\sigma_Y &= X_t\Big[\frac{\eta_t}{\gamma} + \frac{\psi-\gamma}{\gamma(1-\psi)}\sigma_{X,t}\frac{g_P'(X_t)}{g_P(X_t)}\Big] - (1-X_t)\sigma_{X,t}\frac{g_N'(X_t)}{g_N(X_t)}
\end{aligned}$$

Note that $\frac{c_{t,t}}{Y_t} = \frac{c_{t,t}}{W_{t,t}}\frac{W_{t,t}}{P_t}\frac{P_t}{Y_t} = (1-\alpha)g_N(X_t)^{-1}p(X_t)$ by the newborn transfer equation (15). Do the same to the state equation (16) to obtain

$$
\begin{aligned}
X_t\mu_Y + \mu_{X,t} + \sigma_Y\sigma_{X,t} &= \frac{X_t}{\psi}\Big[r_t + \alpha\pi - \rho - \pi - \psi\pi + \frac{\psi+1}{2\gamma}\eta_t^2 + \frac{\psi(\psi-\gamma)}{\gamma(1-\psi)}\eta_t\sigma_{X,t}\frac{g_P'(X_t)}{g_P(X_t)} \\
&\qquad - \frac{1}{2}\frac{\psi(\psi-\gamma)}{\gamma(1-\psi)}\sigma_{X,t}^2\Big(\frac{g_P'(X_t)}{g_P(X_t)}\Big)^2\Big] \\
X_t\sigma_Y + \sigma_{X,t} &= X_t\Big[\frac{\eta_t}{\gamma} + \frac{\psi-\gamma}{\gamma(1-\psi)}\sigma_{X,t}\frac{g_P'(X_t)}{g_P(X_t)}\Big]
\end{aligned}
$$

Solving these four equations for r, η, μ_X, and σ_X gives the expressions in the text.

Step 2: Solve for other asset pricing objects.

To determine the price-dividend ratio p, combine the stock and bond market clearing conditions (13)-(14) to get the asset market clearing condition $P_t = \int_{-\infty}^t \pi e^{-\pi(t-b)} W_{t,b} db$. Then,

$$
\begin{aligned}
p_t &= Y_t^{-1}\int_{-\infty}^t \pi e^{-\pi(t-b)}\frac{W_{t,b}}{c_{t,b}}c_{t,b}db \\
&= g_P(X_t)\int_{\mathcal{P}_t}\pi e^{-\pi(t-b)}\frac{c_{t,b}}{Y_t}db + g_N(X_t)\int_{\mathcal{N}_t}\pi e^{-\pi(t-b)}\frac{c_{t,b}}{Y_t}db \\
&= X_t g_P(X_t) + (1-X_t)g_N(X_t).
\end{aligned}
$$

To determine μ_R, σ_R, and dA^R, apply Itô's formula to stock prices:

$$
\begin{aligned}
dR_t &= \frac{dP_t}{P_t} + \frac{Y_t}{P_t}dt = \frac{d(Y_t p_t)}{Y_t p_t} + \frac{1}{p_t}dt \\
&= \Big(\mu_Y + \mu_{X,t}\frac{p'(X_t)}{p(X_t)} + \sigma_Y\sigma_{X,t}\frac{p'(X_t)}{p(X_t)} + \frac{1}{2}\sigma_{X,t}^2\frac{p''(X_t)}{p(X_t)} + \frac{1}{p(X_t)}\Big)dt \\
&\quad + \Big(\sigma_Y + \sigma_{X,t}\frac{p'(X_t)}{p(X_t)}\Big)dZ_t + \frac{p'(X_t)}{p(X_t)}dA_t^{x^*}.
\end{aligned}
$$

Matching coefficients on the diffusion and singularly continuous component, we obtain formulas for σ_R and dA^R. To obtain μ_R, apply the no-arbitrage relationship $\mu_R = r + \eta\sigma_R$.

Step 3: Verify technical conditions.

Suppose that $x^* \in (0,1)$. Then, $\mu_X(x^*+)$ and $\mu_X(1-)$ are finite, so $\mu_X(x)$ is bounded by continuity. Similarly, $\sigma_X(x^*+)$ and $\sigma_X(1-)$ are finite, so $\sigma_X(x)$ is bounded by continuity. In addition, one can verify that μ_X and σ_X are Lipschitz in the interior $(x^*, 1)$. Indeed, g_P and g_N are bounded away from infinity and zero, and g_P' and g_N' are both continuously differentiable on $(x^*, 1)$. As a result, for any $\delta > 0$, μ_X is bounded on $[x^*, 1-\delta]$ and σ_X is bounded away from 0 and continuously differentiable on $[x^*, 1-\delta]$. Therefore, given any point $x_0 \in (x^*, 1)$ where $X_0 = x_0$, the assumptions of Theorem 3.1 of Zhang (1994) hold, so there is a unique strong solution $(\{X_t^\delta\}_{t\in[0,\tau_{1-\delta}\wedge T]}, \{A_t^{x^*,\delta}\}_{t\in[0,\tau_{1-\delta}\wedge T]})$ to the SDE (17), where $\tau_{1-\delta} := \inf\{t \geq 0 : X_t^\delta = 1-\delta\}$ and $T > 0$. Take the limit $\delta \to 0$ in the solutions $(X^\delta, A^{x^*,\delta})$ to obtain a candidate solution $(\{X_t\}_{t\in[0,T]}, \{A_t^{x^*}\}_{t\in[0,T]})$ to the SDE (17), given functions g_P and g_N. Indeed, the limit exists almost-surely due to the following reasoning. First, on $\{\tau_{1-\delta} \geq T, \text{ some } \delta\}$, there exists δ^* such that $(X^{\delta'}, A^{x^*,\delta'}) = (X^{\delta^*}, A^{x^*,\delta^*})$ for all $\delta' < \delta^*$. Second, similar to the proof of Lemma B.6 in section B.3 of the Online Appendix, we could prove that any solution to (17) must satisfy $\mathbb{P}\{X_t < 1, \forall t \geq 0\} = 1$ (although Lemma B.6 makes some parametric assumptions, the proof that X_t never reaches 1 uses none of these assumptions). This implies that $\mathbb{P}\{\tau_{1-\delta} < T,\, \forall\delta\} = 0$. Hence, in $\mathbb{P}$-almost-every case, the limit is reached at some positive δ, i.e.,

$$\mathbb{P}\Big\{\exists\delta > 0 : (\{X_t^\delta\}_{t\in[0,\tau_{1-\delta}\wedge T]}, \{A_t^{x^*,\delta}\}_{t\in[0,\tau_{1-\delta}\wedge T]}) = (\{X_t\}_{t\in[0,T]}, \{A_t^{x^*}\}_{t\in[0,T]})\Big\} = 1.$$

This limit is clearly unique by construction. Finally, to get the solution for $t \in \mathbb{R}$, one just pieces together solutions on finite intervals.

Next, conjecture that $g_P'(1)$, $g_N'(1) < +\infty$, which in conjunction with the smooth-pasting condition (25), implies that g_P' and g_N' are bounded on $[x^*, 1)$. Then, as $p(x) = xg_P(x) + (1-x)g_N(x)$ and $p'(x) = xg_P'(x) + (1-x)g_N'(x) + g_P(x) - g_N(x)$ are both bounded on $[x^*, 1)$, we know that σ_R is bounded. Similarly, it is easily verified that η and r are bounded. Finally, the guess $g_P'(1)$, $g_N'(1) < +\infty$ may be verified by using boundary conditions (65)-(66) and the boundedness of η and r. As a result, processes $\sigma_{R,t} = \sigma_R(X_t)$ and $\eta_t = \eta(X_t)$ are uniformly bounded, and for every $T > 0$,

$$\mathbb{E}\Big[\exp\Big(\frac{1}{2}\int_{-T}^{T}\eta_t^2 dt\Big)\Big] < +\infty \quad \text{and} \quad \mathbb{E}\Big(\int_{-T}^{T}\sigma_{R,t}^2 dt\Big) < +\infty.$$

Hence, given the results in Chapter 6 of Duffie (2010a) and Appendix B of Karatzas and Shreve (1998), a state price density ξ defined in (2) exists and is consistent with no arbitrage in this economy.

Step 4: Stationary distribution.

Let h denote the stationary density of X_t. Then, as is well known, h satisfies the Kolmogorov forward equation (c.f. Karatzas and Shreve (1991), section 5.7B)

$$0 = -\frac{d}{dx}(\mu_X h) + \frac{1}{2}\frac{d^2}{dx^2}(\sigma_X^2 h)$$

subject to the reflecting boundary condition at $x = x^*$:

$$0 = -\mu_X(x^*)h(x^*) + \frac{1}{2}\frac{d}{dx}(\sigma_X^2 h)\Big|_{x=x^*}.$$

Integrating the forward equation and using the reflecting boundary condition, we obtain

$$0 = -\mu_X h + \frac{1}{2}\frac{d}{dx}(\sigma_X^2 h). \tag{60}$$

Equation (60) can be solved subject to the condition that h is in fact a probability density, i.e., $\int_{x^*}^1 h(x)dx = 1$. A convenient approach to solving (60) is to make the change-of-variables $\hat{h}(x) := \sigma_X^2(x)h(x)$, which satisfies

$$\hat{h}' = \frac{2\mu_X}{\sigma_X^2}\hat{h}.$$

Integrating from x^* to x, and then inverting the change-of-variable from $\hat{h}$ to h, we get

$$h(x) = \frac{K_0}{\sigma_X^2(x)}\exp\Big(\int_{x^*}^{x}\frac{2\mu_X(y)}{\sigma_X^2(y)}dy\Big),$$

where K_0 is a constant chosen to ensure h integrates to 1 on $[x^*, 1)$. □

Two key properties of limited participation models are that they tend to lower risk-free rates and raise risk prices, which is the content of Proposition 15, whose proof is below.

Proof of Proposition 15. Suppose $x^* < 1$. Recall the definitions of the limited participation contributions:

$$r_{LP}(x) := -\frac{1}{2}\gamma(\psi+1)\sigma_Y^2\Big(\frac{1-x}{x}\Big)\omega^2(x)$$
$$\eta_{LP}(x) := \gamma\sigma_Y\Big(\frac{1-x}{x}\Big)\omega(x),$$

where

$$\omega(x) := \Big(1-(1-x)\frac{g_N'(x)}{g_N(x)}\Big)^{-1}.$$

An important quantity for our calculations below is

$$\varphi(x) := \frac{d}{dx}\Big(\frac{1-x}{x}\Big)\omega(x) = -\frac{\omega(x)}{x^2} - \omega^2(x)\Big[\frac{g_N'(x)}{g_N(x)} - (1-x)\frac{g_N(x)g_N''(x)-g_N'(x)^2}{g_N(x)^2}\Big].$$

Note that $\varphi(x^*) = -\frac{1}{x^*} + (1-x^*)\frac{g_N''(x^*)}{g_N(x^*)}$ by the smooth-pasting condition (25), and we also have $\varphi(1) = -1 - \frac{g_N'(1)}{g_N(1)}$.

Inspection shows that $r_{LP}(x) < 0$ for all $x \in [x^*, 1)$. This immediately implies $r_{LP}'(1) \geq 0$, and since $r_{LP}'(1) = -\frac{1}{2}\gamma(\psi+1)\sigma_Y^2 a(1)$, we have $a(1) = -1 - \frac{g_N'(1)}{g_N(1)} \leq 0$. Because of boundary condition (66), which relates $g_N'(1)$ to $g_N(1)$, we can find parameters such that $\varphi(1) > 0$, and hence $r_{LP}'(1) < 0$ (for example, $\psi \to 1$ works). As a result, there exists $\varepsilon_1^r > 0$ such that $r_{LP}(x) < 0$ and $r_{LP}'(x) > 0$ for all $x \in (1-\varepsilon_1^r, 1)$. Similarly, $\eta_{LP}(1) = 0$ and $\eta_{LP}'(1) = \gamma\sigma_Y\varphi(1) < 0$. Thus, there exists $\varepsilon_1^\eta > 0$ such that $\eta_{LP}(x) > 0$ and $\eta_{LP}'(x) < 0$ for all $x \in (1-\varepsilon_1^\eta, 1)$. Set $\varepsilon_1 = \varepsilon_1^r \wedge \varepsilon_1^\eta$.

Next, $r_{LP}'(x^*) = -\frac{1}{2}\gamma(\psi+1)\sigma_Y^2 a(x^*)$. In light of HJB (23), there exist parameters such that $\varphi(x^*) < 0$ (for example, again taking $\psi \to 1$ implies $g_N''(x^*) = 0$ so that $\varphi(x^*) < 0$). We have already seen that $r_{LP}(x) < 0$ in a neighborhood of x^*, so there exists $\varepsilon_2^r > 0$ such that $r_{LP}(x) < 0$ and $r_{LP}'(x) > 0$ for all $x \in (x^*, x^* + \varepsilon_2^r)$. Likewise, $\eta_{LP}(x^*) = \gamma\sigma_Y(\frac{1-x^*}{x^*}) > 0$ and $\eta_{LP}'(x^*) = \gamma\sigma_Y\varphi(x*) < 0$. Continuity of η_{LP} implies that there exists $\varepsilon_2^\eta > 0$ such that $\eta_{LP}(x) > 0$ and $\eta_{LP}'(x) < 0$ for all $x \in [x^*, x^* + \varepsilon_2^\eta)$. To complete the proof, simply set $\varepsilon_2 = \varepsilon_2^r \wedge \varepsilon_2^\eta$, then set $\varepsilon = \varepsilon_1 \wedge \varepsilon_2$. □

Proof of Proposition 7. Take $x \to x^*$ and $x \to 1$ in Proposition 14, using the smooth-pasting condition (25) to find $\eta(x^*) = \gamma\sigma_Y/x^*$ and $\eta(1) = \gamma\sigma_Y$. □

A.2 HJBs and proof of optimality

Admissible controls.

For completeness, I state the participants' and non-participants' problems here. For the sake of generality, this is done for the model with recursive utility. An *admissible control* for participants is given by consumption and portfolio processes that satisfy the dynamic budget constraint and lead to finite utility, i.e., $(c_{t,b}, \theta_{t,b})$ such that (48) has a unique strong solution and

$$\mathbb{E}\Big[\int_t^\infty \big|f(c_{s,b}, V_{s,b}^P)\big|ds \mid \mathcal{F}_t\Big] < +\infty,$$

where f is the felicity function defined in (36). The set of admissible participant controls as of time t is denoted by $\mathcal{A}_t^P$. Similarly, non-participants must choose $(c_{t,b}, \tau_b)$ such that τ_b is a stopping time,

$$dW_{t,b} = (r_t W_{t,b} + \alpha\pi W_{t,b} - c_{t,b})dt + W_{t,b} dA_t^R$$

has a unique strong solution, and

$$\mathbb{E}\Big[\int_t^\infty \big|f(c_{s,b}, V_{s,b}^N)\big| ds \mid \mathcal{F}_t\Big] < +\infty.$$

The set of admissible non-participant controls is denoted by $\mathcal{A}_t^N$. Then, participants and non-participants solve

$$b \in \mathcal{P}_t : \; V_{t,b}^P = \sup_{c,\theta \in \mathcal{A}_t^P} \mathbb{E}\Big[\int_t^\infty f(c_{s,b}, V_{s,b}^P) ds \mid \mathcal{F}_t\Big] \tag{61}$$

$$b \in \mathcal{N}_t : \; V_{t,b}^N = \sup_{c,\tau \in \mathcal{A}_t^N} \mathbb{E}\Big[\int_t^\tau f(c_{s,b}, V_{s,b}^N) ds + V_{\tau,b}^P - \Phi_{\tau,b} \mid \mathcal{F}_t\Big], \tag{62}$$

where $\Phi_{t,b}$ is given by (37).

Heuristic derivation of HJB equations.

The HJB equations for $V = V^P$ and $V = V^N$ are as follows (where if $V = V^N$, we require $\theta \equiv 0$):

$$\begin{aligned}
0 &= \sup_{c,\theta} \; f(c, V) + V_w w\Big(r + \alpha\pi + \theta(\mu_R - r) - \frac{c}{w}\Big) + \frac{1}{2} V_{ww} w^2 \theta^2 \sigma_R^2 \\
&\quad + V_x \mu_X + \frac{1}{2} V_{xx} \sigma_X^2 + V_{xw} w \theta \sigma_R \sigma_X \\
&= \sup_{c,\theta} \; \frac{1}{1-\psi}\Big(\Big(\frac{c}{w}\Big)^{1-\psi} g(x)^{1-\psi} - (\rho + \pi) g(x)\Big) + g(x)\Big(r + \alpha\pi + \theta(\mu_R - r) - \frac{c}{w}\Big) \\
&\quad - \frac{\gamma}{2} g(x) \theta^2 \sigma_R^2 + \frac{\psi}{1-\psi} g'(x) \mu_X + \frac{\psi(1-\gamma)}{1-\psi} g'(x) \theta \sigma_R \sigma_X \\
&\quad + \frac{1}{2} \frac{\psi}{1-\psi}\Big[g''(x) + \Big(\frac{\psi(1-\gamma)}{1-\psi} - 1\Big)\frac{g'(x)^2}{g(x)}\Big]\sigma_X^2.
\end{aligned}$$

The FOCs for c and θ imply

$$c_{t,b} = \begin{cases} W_{t,b} g_P(X_t)^{-1}, & \text{if } b \in \mathcal{P}_t; \\ W_{t,b} g_N(X_t)^{-1}, & \text{if } b \in \mathcal{N}_t. \end{cases} \tag{63}$$

and

$$\theta_{t,b} = \frac{\mu_{R,t} - r_t}{\gamma \sigma_{R,t}^2} + \frac{\sigma_{X,t}}{\sigma_{R,t}} \frac{\psi(1-\gamma)}{\gamma(1-\psi)} \frac{g_P'(X_t)}{g_P(X_t)} \quad \text{if } b \in \mathcal{P}_t. \tag{64}$$

Substituting these back into the HJB equations, we get two ODEs for the wealth-consumption ratios, given by (54) and (55).

Boundary conditions.

The boundary conditions for (54)-(55) at $x = x^*$ are given by the value-matching and smooth-pasting conditions (56)-(57) in the text. The boundary conditions at $x = 1$ are derived by taking the limit $x \to 1$ the HJB equations, assuming that $(1-x)g_i'(x) \to 0$ and $(1-x)^2 g_i''(x) \to 0$ as $x \to 1$, for $i \in \{P, N\}$. Since

$\sigma_X(x) \to 0$ as $x \to 1$, all terms multiplying σ_X vanish in this limit. Then, passing to the limit $x \to 1$,

$$0 = \psi + \Big[-\rho - \pi + (1-\psi)\Big(r(1) + \alpha\pi + \frac{1}{2\gamma}\eta^2(1)\Big)\Big] g_P(1) + \psi\mu_X(1) g_P'(1) \tag{65}$$

$$0 = \psi + \Big[-\rho - \pi + (1-\psi)\Big(r(1) + \alpha\pi\Big)\Big] g_N(1) + \psi\mu_X(1) g_N'(1). \tag{66}$$

The following proposition establishes that these arguments are in fact sufficient for optimality in the investors' problems.

Proposition 16 (Verification of Optimality)**.** *Let X be the unique strong solution to the stochastic differential equation $dX_t = \mu_{X,t}dt + \sigma_{X,t}dZ_t + dA_t^{x^*}$ on $[x^*, 1)$, assuming it exists, where $A_t^{x^*} = \int_{-\infty}^t \mathbf{1}_{\{X_s \le x^*\}} dA_s^{x^*}$ is a singularly continuous non-decreasing process. Define $\mathcal{S} := \mathbb{R}_+ \times [0, x^*]$ and $\mathcal{O} := \mathbb{R}_+ \times (x^*, 1]$. Consider two functions J^P and J^N satisfying*

$$J^i(w,x) = \frac{w^{1-\gamma}}{1-\gamma} g_i(x)^{\frac{\psi(1-\gamma)}{1-\psi}}, \tag{67}$$

for strictly positive bounded functions $g_i \in C^1([0,1]) \cap C^2((0,1)\setminus\{x^*\})$, for $i \in \{P, N\}$. Additionally, suppose J^P and J^N satisfy

(i) $J^P((1-\phi)w, x) = J^N(w,x)$ *on* $\mathcal{S}$, *and* $J^P((1-\phi)w, x) < J^N(w,x)\}$ *on* $\mathcal{O}$.

(ii) $\sup_{c\in\mathcal{A}^N} \mathcal{D}^N J^N + f(c, J^N) = 0$ *on* $\mathcal{O}$, *and* $\sup_{c\in\mathcal{A}^N} \mathcal{D}^N J^N + f(c, J^N) \le 0$ *on* $\mathcal{S}$, *where*

$$\mathcal{D}^N J^N := w\big(r + \alpha\pi - \tfrac{c}{w}\big)\partial_w J^N + \mu_X \partial_x J^N + \tfrac{1}{2}\sigma_X^2 \partial_{xx} J^N. \tag{68}$$

$\sup_{c,\theta\in\mathcal{A}^P} \mathcal{D}^P J^P + f(c, J^P) = 0$ *on* $\mathcal{O}\cup\mathcal{S}\setminus\{x^*\}$, *where*

$$\begin{aligned}\mathcal{D}^P J^P := {} & w\big(r + \alpha\pi + \theta(\mu_R - r) - \tfrac{c}{w}\big)\partial_w J^P + \mu_X \partial_x J^P \\ & + \tfrac{1}{2}w^2\theta^2\sigma_R^2 \partial_{ww} J^P + \tfrac{1}{2}\sigma_X^2 \partial_{xx} J^P + w\theta\sigma_R\sigma_X \partial_{wx} J^P.\end{aligned} \tag{69}$$

(iii) $\partial_x J^N(w, x^*) = \partial_x J^P(w, x^*) = 0$.

(iv) Strategies $c_{t,b}$ and $\theta_{t,b}$ are such that $J^i(W_{t,b}, X_t)$, $\theta_{t,b}\sigma_{R,t} J^i(W_{t,b}, X_t)$, and $\sigma_{X,t}\frac{g_i'(X_t)}{g_i(X_t)} J^i(W_{t,b}, X_t)$ belong to $\mathcal{H}^2 := \{h : \mathbb{E}\int_{-T}^T |h_t|^2 dt < +\infty,$ for all $T\}$.

(v) $\lim_{T\to\infty} \mathbb{E}[J^i(W_{T,b}, X_T) \mid \mathcal{F}_t] = 0$ *for* $i \in \{P, N\}$.

Then, if $\{V_{t,b}^P\}_{t\ge b}$ and $\{V_{t,b}^N\}_{t\ge b}$ are unique solutions to (61)-(62), we have $J^i(W_{t,b}, X_t) = V_{t,b}^i$ for $i \in \{P, N\}$. In addition, optimal decisions are given by $c_{t,b}$ in (63), $\theta_{t,b}$ in (64), and

$$\tau_b := \inf\{t \ge b : (W_t, X_t) \in \mathcal{S}\}. \tag{70}$$

Proof. In the proof, we suppress the cohort b in all expressions when the meaning is clear. Let $T < \infty$ and (W_t, X_t) be arbitrary. Consider first the unconstrained investor (participant) problem. Let $a = (c, \theta) \in \mathcal{A}_t^P$

be an admissible control. Let W^a be the wealth process under a. Apply Itô's formula to $J^P(W_t^a, X_t)$ to get

$$J^P(W_T^a, X_T) = J^P(W_t, X_t) + \int_t^T \mathcal{D}^{P,a} J^P(W_s^a, X_s)ds + \int_t^T \partial_x J^P(W_s^a, X_s)dA_s^{x^*} + \int_t^T \left[W_s^a \theta_s \sigma_{R,s} \partial_w J^P(W_s^a, X_s) + \sigma_{X,s}\partial_x J^P(W_s^a, X_s)\right] dZ_s,$$

where $\mathcal{D}^{P,a}$ is defined by (69) under control a. Note that, in using Itô's formula, we can ignore the set $\{s : X_s = x^*\}$ in the first integral as this set has Lebesgue measure zero. Next, because $A_s^{x^*}$ is flat off $\{s : X_s = x^*\}$, condition (iii) implies that the second integral is zero. Because of the multiplicative separable representation (67), $W_s^a \theta_s \sigma_{R,s} \partial_w J^P(W_s^a, X_s) = \theta_s \sigma_{R,s} J^P(W_s^a, X_s)$ and $\sigma_{X,s}\partial_x J^P(W_s^a, X_s) = \frac{\psi(1-\gamma)}{1-\psi}\sigma_{X,s}\frac{g_P'(X_s)}{g_P(X_s)} J^P(W_s^a, X_s)$. By condition (iv), the stochastic integral (as function of T) is then a true martingale and has conditional expectation zero (with respect to $\mathcal{F}_t$). Finally, condition (ii) implies that $\mathcal{D}^{P,a} J^P(w, x) + f(c, J^P(w, x)) \le 0$ for all (w, x) and $a \in \mathcal{A}_t^P$. Taking expectations and using these results,

$$J^P(W_t, X_t) \ge \mathbb{E}\Big[\int_t^T f(c_s, J^P(W_s^a, X_s))ds + J^P(W_T^a, X_T) \mid \mathcal{F}_t\Big].$$

Now, pick a sequence $\{T_n\}_{n=1}^\infty$ with $T_n \to \infty$ as $n \to \infty$ such that $\int_t^{T_n} \frac{c_s^{1-\psi}}{1-\psi}[(1-\gamma)J^P(W_s^a, X_s)]^{\frac{\psi-\gamma}{1-\gamma}} ds$ and $\int_t^{T_n} \frac{\rho+\pi}{1-\psi}[(1-\gamma)J^P(W_s^a, X_s)]ds$ are monotonic sequences. Such a choice is possible since J^P has an unambiguous sign, by (67). Using condition (v) and the monotone convergence theorem to take the limit as $n \to \infty$, then maximizing over feasible controls, we obtain

$$J^P(W_t, X_t) \ge \sup_{a \in \mathcal{A}_t^P} \mathbb{E}\Big[\int_t^\infty f(c_s, J^P(W_s^a, X_s))ds \mid \mathcal{F}_t\Big]. \tag{71}$$

On the other hand, consider the admissible control $a = (c, \theta)$ given by (63) and (64) and suppose this control also satisfies integrability condition (iv). Representation (67) implies that $\partial_w J^P > 0$ and $\partial_{ww} J^P < 0$, sufficient to imply that a attains the maximum in condition (ii). Then, letting τ be any stopping time, applying Itô's formula to $\int_\tau^t f(c_s, J^P(W_s^a, X_s))ds + J^P(W_t^a, X_t)$, using condition (ii), and taking expectations under the hypotheses of condition (iv), we obtain

$$J^P(W_\tau, X_\tau) = \mathbb{E}\Big[\int_\tau^T f(c_s, J^P(W_s^a, X_s))ds + J^P(W_T^a, X_T) \mid \mathcal{F}_\tau\Big].$$

Using condition (v) and the monotone convergence theorem, we can take the limit to obtain for any t

$$J^P(W_t, X_t) = \mathbb{E}\Big[\int_t^\infty f(c_s, J^P(W_s^a, X_s))ds \mid \mathcal{F}_t\Big].$$

Combining with inequality (71), this shows that control a attains the maximum in $\mathcal{A}_t^P$. In addition, we obtain an equation identical to the recursive formulation of the value function in (61). Since (61) has the unique solution $V_{t,b}^P$, we have

$$J^P(W_{t,b}, X_t) = V_{t,b}^P. \tag{72}$$

Similarly, for the non-participant, fixing admissible control $a = (c, \tau) \in \mathcal{A}_t^N$ and applying Itô's formula to $J^N(W_t^a, X_t)$, we obtain

$$J^N(W^a_{\tau\wedge T}, X_{\tau\wedge T}) = J^N(W_t, X_t) + \int_t^{\tau\wedge T} \mathcal{D}^{N,a} J^N(W_s^a, X_s)ds + \int_t^{\tau\wedge T} \partial_x J^N(W_s^a, X_s)dA_s^{x^*}$$
$$+ \int_t^{\tau\wedge T} \big[W_s^a \theta_s \sigma_{R,s} \partial_w J^N(W_s^a, X_s) + \sigma_{X,s}\partial_x J^N(W_s^a, X_s)\big] dZ_s,$$

where $\mathcal{D}^{N,a}$ is defined by (68) under control a. Now, repeat the arguments above, but also apply $J^N(w, x) \geq J^P((1-\phi)w, x)$ from condition (i) to get

$$J^N(W_t, X_t) \geq \sup_{a\in\mathcal{A}_t^N} \mathbb{E}\Big[\int_t^\tau f(c_s, J^N(W_s^a, X_s))ds + J^P((1-\phi)W_\tau^a, X_\tau) \mid \mathcal{F}_t\Big]. \tag{73}$$

On the other hand, consider the admissible control $a = (c, \tau)$ given by (63) and (70), supposing (iv) is satisfied. Again, $\partial_w J^N > 0$ and $\partial_{ww} J^N < 0$ so that c attains the maximum in (ii). Repeating the arguments above, we obtain

$$J^N(W_t, X_t) = \mathbb{E}\Big[\int_t^{\tau\wedge T} f(c_s, J^N(W_s^a, X_s))ds + J^N(W^a_{\tau\wedge T}, X_{\tau\wedge T}) \mid \mathcal{F}_t\Big],$$

since equality holds in condition (ii) for all $t < \tau$. Taking the limit $T \to \infty$ as before, and using $J^N(W_\tau^a, X_\tau) = J^P((1-\phi)W_\tau^a, X_\tau)$ from condition (i), we have

$$J^N(W_t, X_t) = \mathbb{E}\Big[\int_t^\tau f(c_s, J^N(W_s^a, X_s))ds + J^P((1-\phi)W_\tau^a, X_\tau) \mid \mathcal{F}_t\Big].$$

Finally, we can use the form of J^P in (67), our previous result (72), and the equation for the entry cost (37) to get $J^P((1-\phi)W_\tau^a, X_\tau) = (1-\phi)^{1-\gamma} J^P(W_\tau^a, X_\tau) = (1-\phi)^{1-\gamma} V_\tau^P = V_\tau^P - \Phi_\tau$. Thus,

$$J^N(W_t, X_t) = \mathbb{E}\Big[\int_t^\tau f(c_s, J^N(W_s^a, X_s))ds + V_\tau^P - \Phi_\tau \mid \mathcal{F}_t\Big].$$

Combining with inequality (73), this shows that control a attains the maximum in $\mathcal{A}_t^N$. Exactly as before, the uniqueness of $V_{t,b}^N$ as a solution to (62) implies $J^N(W_{t,b}, X_t) = V_{t,b}^N$. □

Remark A.1 (Verify conditions of Proposition 16)**.** *To verify optimality of $c_{t,b}$, $\theta_{t,b}$, and τ_b in (63), (64), and (70), it suffices to verify that the conditions of Proposition 16 are satisfied for any $x^* \in (0,1)$. In equilibrium, as proved in Proposition 14 above, X is the unique solution to the stochastic differential equation $dX_t = \mu_X(X_t)dt + \sigma_X(X_t)dZ_t + dA_t^{x^*}$. The homogeneity of problems (61)-(62) imply $V^P(w,x) = \frac{w^{1-\gamma}}{1-\gamma} g_P(x)^{\frac{\psi(1-\gamma)}{1-\psi}}$ and $V^N(w,x) = \frac{w^{1-\gamma}}{1-\gamma} g_N(x)^{\frac{\psi(1-\gamma)}{1-\psi}}$, so that (67) is satisfied by V^P and V^N. Next, if there is a solution to the ODEs (54) and (55), subject to boundary conditions (56)-(57) at $x = x^*$ and (65)-(66) at $x = 1$, then conditions (i), (ii), and (iii) of Proposition 16 are automatically satisfied, given the form of functions V^P and V^N. Thus, I assume throughout that there exists a solution to this ODE system. Condition (iv) is also satisfied: one can check that θ, σ_R, σ_X, and g_i are bounded on $[x^*, 1)$, so it suffices that the functions g_i' not explode too quickly at the boundaries, which is guaranteed by (57), (65), and (66). Finally, transversality condition (v) is verified below in Lemma A.1.*

The following lemma verifies the transversality condition (v) in Proposition 16, which completes the verification of optimality.

Lemma A.1 (Transversality condition)**.** *For $i \in \{P, N\}$,*

$$\lim_{T\to\infty} \mathbb{E}[V^i(W_{T,b}, X_T) \mid \mathcal{F}_t] = 0.$$

Proof. Note that g_P is positive and bounded, so there exists $K > 0$ such that $g_P(X_s)^{-1} \geq K$. Under the optimal participant controls (c, θ), we have

$$\begin{aligned}
+\infty &> \mathbb{E}\Big[\int_t^\infty \big|f(c_s, V_s^P)\big| ds \mid \mathcal{F}_t\Big] \\
&= \Big|\frac{1}{1-\psi}\Big| \mathbb{E}\Big[\int_t^\infty \Big|c_s^{1-\psi}[(1-\gamma)V_s^P]^{\frac{\psi-\gamma}{1-\gamma}} - (\rho+\pi)(1-\gamma)V_s^P\Big| ds \mid \mathcal{F}_t\Big] \\
&= \Big|\frac{1}{1-\psi}\Big| \mathbb{E}\Big[\int_t^\infty \Big|g_P(X_s)^{-1} g_P(X_s)^{\frac{\psi(1-\gamma)}{1-\psi}} - (\rho+\pi) g^{\frac{\psi(1-\gamma)}{1-\psi}}\Big| W_s^{1-\gamma} ds \mid \mathcal{F}_t\Big] \\
&= \Big|\frac{1-\gamma}{1-\psi}\Big| \mathbb{E}\Big[\int_t^\infty \big|g_P(X_s)^{-1} - (\rho+\pi)\big| \big|V_s^P\big| ds \mid \mathcal{F}_t\Big] \\
&\geq \max(|K-\rho-\pi|, \rho+\pi) \Big|\frac{1-\gamma}{1-\psi}\Big| \mathbb{E}\Big[\int_t^\infty \big|V_s^P\big| ds \mid \mathcal{F}_t\Big] \\
&\geq (K+\rho+\pi) \Big|\frac{1-\gamma}{1-\psi}\Big| \int_t^\infty \mathbb{E}[\big|V_s^P\big| \mid \mathcal{F}_t] ds,
\end{aligned}$$

which implies $\mathbb{E}[V_T^P \mid \mathcal{F}_t] \to 0$ as $T \to \infty$. An identical argument applies to V^N. $\square$

A.3 Free-entry equilibrium

In this section, I examine properties of the free-entry equilibrium. These are the key results for small costs ϕ. I prove these results in the context of the more general recursive utility model of Appendix A.1. One can obtain the analogous log utility model results by taking limits $\gamma, \psi \to 1$.

Proposition 17. *With $\phi = 0$, the following is an equilibrium:*

$$\begin{aligned}
r^* &:= \rho + \pi(1-\alpha) + \psi(\mu_Y + \pi\alpha) - \frac{1}{2}\gamma(\psi+1)\sigma_Y^2 \\
\eta^* &:= \gamma\sigma_Y \\
\sigma_R^* &:= \sigma_Y \\
\mu_R^* &:= r^* + \sigma_R^* \eta^* \\
p^* &:= (\mu_R^* - \mu_Y)^{-1}.
\end{aligned}$$

In particular, Proposition 17 shows that all formulas in Proposition 14 apply, putting $x = 1$ and replacing $p(X_t)/g_N(X_t)$ with $p(X_t)/g_P(X_t) = 1$, which is because newborns immediately become participants.

Proof or Proposition 17. With $\phi = 0$, agents begin participating in risky asset markets at birth. Thus, $X_t \equiv 1$ for all t. Consequently, all equilibrium objects are time-invariant and will be denoted by the symbols in the time-varying case, with the addition of stars.

Applying Itô's formula to stock returns, we obtain $\sigma_R^* = \sigma_Y$. Next, since all agents must choose $\theta_{t,b} \equiv 1$ in equilibrium, the first order condition for portfolio choice (which has no hedging demand term) can be inverted to deliver $\eta^* = \gamma\sigma_R^* = \gamma\sigma_Y$. The expected returns on the stock are simply determined by the no-arbitrage condition $\eta^* = (\sigma_R^*)^{-1}(\mu_R^* - r^*)$. Stock returns have no singular component, even with continuous entry, since p^* is constant (the singular component of returns was $\frac{p'(X_t)}{p(X_t)}dA_t^{x^*}$). Use the Gordon growth formula to determine p^*. Finally, to determine the risk-free rate, use the participants' HJB equation, which is

$$0 = \psi + \Big[-\rho - \pi + (1-\psi)\Big(r^* + \alpha\pi + \frac{1}{2\gamma}(\eta^*)^2\Big)\Big] g^*.$$

Stock market clearing, the Gordon growth formula for p^*, and the no-arbitrage condition for μ_R^* implies that

$$g^* = p^* = (\mu_R^* - \mu_Y)^{-1} = (r^* + \gamma\sigma_Y^2 - \mu_Y)^{-1}.$$

Combining this equation with the HJB equation and solving for r^* gives the result. □

Proof of Proposition 6. Denote the wealth-consumption ratios of the participants and deviating non-participant by g^* and $\tilde{g}^*$, and their expected lifetime utility by $V^*(w)$ and $\tilde{V}^*(w)$, respectively. Using the HJB equations (54)-(55) and the fact that $\mu_X = \sigma_X = 0$ we have

$$g^* = \psi\Big[\rho + \pi + (\psi - 1)\Big(r^* + \alpha\pi + \frac{1}{2\gamma}(\eta^*)^2\Big)\Big]^{-1} \quad \text{and} \quad V^*(w) = \frac{w^{1-\gamma}}{1-\gamma}(g^*)^{\frac{\psi(1-\gamma)}{1-\psi}}$$
$$\tilde{g}^* = \psi\Big[\rho + \pi + (\psi - 1)(r^* + \alpha\pi)\Big]^{-1} \quad \text{and} \quad \tilde{V}^*(w) = \frac{w^{1-\gamma}}{1-\gamma}(\tilde{g}^*)^{\frac{\psi(1-\gamma)}{1-\psi}}.$$

Using $V^*(1 - \tilde{\phi}^*) = \tilde{V}^*(1)$, we find

$$\phi^* \geq \tilde{\phi}^* = 1 - \Big(\frac{\tilde{g}^*}{g^*}\Big)^{\frac{\psi}{1-\psi}},$$

which equals 0 if and only if $\eta^* = \gamma\sigma_Y \equiv 0$, and is otherwise positive. To get the approximation result in the statement of the proposition, expand $\tilde{\phi}^*$ around $\sigma_Y^2 = 0$ and substitute the equilibrium quantities from Proposition 17. □

A.4 Proofs for the log utility model

It is straightforward to show that agents behave optimally if HJB equations (22)-(23) hold, along with $\Delta g(x) \leq \Phi$ for all x and $\Delta g(x) = \Phi$ for $x \leq x^*$. Such a proof would proceed along the lines of the verification argument of Proposition 16 for the general recursive utility model.

Proof of Proposition 1. The entirety of the proof proceeds exactly as Proposition 14. The equilibrium expressions may alternatively be obtained by simply taking limits $\gamma, \psi \to 1$ in the expressions from Proposition 14. Although all endogenous objects are determined independently of value functions, determination of the entry point x^* requires solving the ODE for Δg. □

Proof of Proposition 3. All three equations are essentially derived from martingale arguments. Since Δg is C^2 a.e., we apply Itô's formula to $M_t := e^{-(\rho+\pi)t}\Delta g(X_t) + \frac{1}{2}(\rho+\pi)^{-1}\int_0^t e^{-(\rho+\pi)s}\eta^2(X_s)ds$. The result, for any stopping time τ, is

$$M_{T\wedge\tau} - M_0 = \frac{1}{2}(\rho+\pi)^{-1}\int_0^{T\wedge\tau} e^{-(\rho+\pi)t}\eta^2(X_t)dt + \int_0^{T\wedge\tau} \mathscr{L}[\Delta g](X_t)dt$$
$$+ \int_0^{T\wedge\tau} e^{-(\rho+\pi)t}\sigma_X(X_t)\Delta g'(X_t)dZ_t + \int_0^{T} e^{-(\rho+\pi)t}\Delta g'(X_t)dA_t^{x^*},$$

where the differential operator $\mathscr{L}$ is defined by

$$\mathscr{L} f(x) := -(\rho+\pi) f(x) + \mu_X(x) f'(x) + \frac{1}{2}\sigma_X^2(x) f''(x). \tag{74}$$

Since (26) holds on $(x^*, 1)$, and since $\{X_t = x^*\}$ has zero Lebesgue measure, the sum of the first two integrals is identically zero. As Δg is C^2 and σ_X is bounded, the stochastic integral is a martingale. As $\Delta g'(x^*) = 0$, the last integral is identically zero. Hence, M_t is a martingale, so by Doob's optional sampling, we have

$$\Delta g(x) = \mathbb{E}^x\Big[\frac{1}{2}(\rho+\pi)^{-1}\int_0^{T\wedge\tau} e^{-(\rho+\pi)t}\eta^2(X_t)dt + e^{-(\rho+\pi)(T\wedge\tau)}\Delta g(X_{T\wedge\tau})\Big]. \tag{75}$$

Using (75), we can prove (29)-(31). Equation (29) follows by picking $\tau = +\infty$, performing recursive substitution of $\Delta g(X_T)$ on the right-hand-side of (75), applying the Strong Markov property, and finally taking $T \to +\infty$ with the monotone convergence theorem. Equation (30) follows by picking $\tau = \tau_{x^*}$, noting that $\Delta g(X_{\tau_{x^*}}) = \Delta g(x^*) = \Phi$, and again taking $T \to +\infty$ with the monotone convergence theorem (also using the fact that $\tau_{x^*} < +\infty$ a.s.). Equation (31) follows by noting that $\Delta g(X_{T\wedge\tau}) \leq \Phi$, so that the objective function of (31) exceeds $\Delta g(x)$ for any choice of τ. But choosing $\tau = \tau_{x^*}$ is feasible, and so equation (30) implies equation (31). □

Proof of Proposition 4. Substitute $X_t \equiv 1$ in expression (32). To obtain the approximation in σ_Y^2, expand ϕ^* around $\sigma_Y^2 = 0$ and substitute $\eta^* := \eta(1) = \sigma_Y$. □

To prove the existence/uniqueness result of Proposition 2, we need the following two lemmas. For the remainder of the results, assume that ϕ is large enough so that (26) is not degenerate, i.e., there is not a solution with $x^* = 1$. In particular, a sufficient condition to guarantee $x^* < 1$ is that $\Phi > \frac{1}{2}(\rho+\pi)^{-2}\sigma_Y^2$, which is the result of Proposition 4.

Lemma A.2. *Any function Δg which solves (26) subject to (27)-(28) is decreasing on $(x^*, 1)$ and strictly decreasing in a neighborhood of x^*.*

Proof of Lemma A.2. For $X_t = y$ and $X_0 = x$, denote the transition density by $h_t(y, x)$. Define the transition cdf by $H_t(y, x) := \int_{x^*}^{y} h_t(x', x)dx'$. Given the continuity of the process X (besides potential positive jumps if there are entry points to the right of x^*), we necessarily have $\frac{\partial}{\partial x} H_t(y, x) \leq 0$, which says that the transition cdfs are ordered in the sense of first-order stochastic dominance as the initial point increases. First-order stochastic dominance implies $\frac{d}{dx}\mathbb{E}^x[\eta^2(X_t)] \leq 0$, as η is a decreasing function. In view of (29) above, this result implies that $\Delta g'(x) \leq 0$. Furthermore, by taking the limit of the ODE (26) as $x \to x^*$, we find

$$0 = \frac{1}{2}(\rho+\pi)^{-1}\eta^2(x^*) - (\rho+\pi)\Phi + \frac{1}{2}\sigma_X^2(x^*)\Delta g''(x^*+).$$

Taking $x \to x^*$ in representation (29) shows that $\frac{1}{2}(\rho+\pi)^{-1}\eta^2(x^*) - (\rho+\pi)\Phi > 0$, implying $\Delta g''(x^*+) < 0$. Assuming Δg is a so-called classical solution, $\Delta g''$ is continuous, so for all ε small enough, $\Delta g'(x^*+\varepsilon) < 0$. □

Now, we consider a problem which is simpler than solving (26) subject to (27)-(28). Given a fixed number

$y \in (0,1)$, and recalling the linear differential operator $\mathscr{L}$ in (74), consider instead solving

$$\begin{aligned} 0 &= \frac{1}{2}(\rho+\pi)^{-1}\eta^2 + \mathscr{L}\varphi, \quad x \in (y,1), \\ 0 &= \varphi - \Phi, \quad x \in [0,y] \\ 0 &= \frac{1}{2}(\rho+\pi)^{-1}\eta^2(1) - (\rho+\pi)\varphi(1) + \mu_X(1)\varphi'(1), \end{aligned} \tag{76}$$

which amounts to solving (26) on $(y,1)$ subject to only the value-matching condition (27), but not the smooth-pasting condition (28).

Lemma A.3. *Optimal entry and smooth-pasting are equivalent in the following sense. Let $\Delta g(x;y)$ be a solution to (76) that additionally satisfies $\frac{d}{dx}\Delta g(x;y)\big|_{x\downarrow y} = 0$. Then, $\Delta g \in C^1(0,1)$, and the optimal entry conditions $\Delta g(x;y) = \Phi$ for $x \in [0,y]$ and $\Delta g(x;y) < \Phi$ for $x \in (y,1)$ are satisfied. Conversely, let $\Delta g(x;y)$ be a C^1 solution to (76) that additionally satisfies $\Delta g(x;y) < \Phi$ for $x \in (y,1)$. Then, $\frac{d}{dx}\Delta g(x;y)\big|_{x\downarrow y} = 0$.*

Proof of Lemma A.3. Let $\Delta g(x;y)$ be a solution to (76) that additionally satisfies $\frac{d}{dx}\Delta g(x;y)\big|_{x=y} = 0$. By Lemma A.2, $\Delta g(x;y)$ is decreasing on $(y,1)$ and strictly decreasing near y. Hence $\Delta g(x;y) < \Phi$ for all $x > y$, while $\Delta g(x;y) = \Phi$ for all $x \le y$ by construction. Clearly, $\Delta g \in C^1(0,y) \cap C^1(y,1)$. Since $\lim_{x\uparrow y} \frac{d}{dx}\Delta g(x;y) = 0$, by construction, and $\lim_{x\downarrow y} \frac{d}{dx}\Delta g(x;y) = 0$, by assumption, Δg is continuously differentiable at y as well. Conversely, let $\Delta g(x;y)$ be a solution to (76) that additionally satisfies $\Delta g(x;y) < \Phi$ for $x \in (y,1)$. Since $\Delta g \in C^1(0,1)$, we have $0 = \lim_{x\uparrow y} \frac{d}{dx}\Delta g(x;y) = \lim_{x\downarrow y} \frac{d}{dx}\Delta g(x;y)$, which proves the claim. ☐

The result of Lemma A.3 suggests that, rather than solving (26)-(28), we may instead consider the following variational inequality problem. Find a function φ such that on $(0,1)$ the following hold:

$$\begin{aligned} 0 &\le \frac{1}{2}(\rho+\pi)^{-1}\eta^2 + \mathscr{L}\varphi, \\ 0 &\ge \varphi - \Phi, \\ 0 &= \Big(\frac{1}{2}(\rho+\pi)^{-1}\eta^2 + \mathscr{L}\varphi\Big)(\varphi - \Phi). \end{aligned} \tag{77}$$

In (77), the function η is taken to be the extension to $(0,1)$ of the function from Proposition 1. Similarly for μ_X and σ_X in the definition of $\mathscr{L}$. Also note that, since $\{0,1\}$ are inaccessible for X_t, the boundary conditions at those points need not be specified, which we will soon see through the stochastic representation of the solution φ.

Proof of Proposition 2. The outline of the proof is to (Step 1) show that (77) has a unique solution, assuming some boundary conditions; (Step 2) show that solutions to (26)-(27) and solutions to (77) are equivalent, as long as the appropriate boundary conditions are selected.

Step 1: Solution to (77).

Suppose $\varphi(0) = \Phi$ and $F(\varphi(1), \varphi'(1)) = 0$ for some F. Augmenting problem (77) with these boundary conditions, we have a standard variational inequality. One can consult any reference text on optimal stopping, free boundary problems, and variational inequalities to find (77) has a unique $C^1(0,1)$ solution, which is also C^2-a.e., with these boundary conditions (c.f. Bensoussan and Lions (1982), chapter 3.1, or Friedman (2010), chapter 1.2). To sidestep the issue that $\lim_{x\to 0}\eta(x) = \lim_{x\to 0}\mu_X(x) = +\infty$, we can simply construct such solutions on $(\delta, 1)$ for small enough δ. This suffices, since we will have $\varphi(x) = \Phi$ for all $x \in (0, \delta)$ as shown in Step 2.

Step 2: Equivalence of solutions to (77) and to (26)-(27).

Define the solution sets:

$$\begin{aligned}
\mathcal{S}_1 &:= \{\varphi : \varphi \text{ solves (26)-(27)}\} \\
\mathcal{S}_2 &:= \{\varphi : \varphi \text{ solves (76)}, \varphi \leq \Phi,\ \varphi \in C^1(0,1)\} \\
\mathcal{S}_3 &:= \{\varphi : \varphi \text{ solves (76)}, \varphi \leq \Phi\} \\
\mathcal{S}_4 &:= \{\varphi : \varphi \text{ solves (77)}\}.
\end{aligned}$$

Lemma A.3 implies that $\mathcal{S}_1 = \mathcal{S}_2$, and obviously $\mathcal{S}_2 \subseteq \mathcal{S}_3$. By inspection, we also have $\mathcal{S}_3 \subset \mathcal{S}_4$ as long as the function $F(\varphi(1), \varphi'(1)) = \frac{1}{2}(\rho+\pi)^{-1}\eta^2(1) - (\rho+\pi)\varphi(1) + \mu_X(1)\varphi'(1)$. Hence,

$$\mathcal{S}_1 \subseteq \mathcal{S}_4.$$

On the other hand, take $\varphi \in \mathcal{S}_4$ to be the unique solution, with this choice of F. Furthermore, let $\mathcal{X}^* \subset (0,1)$ be the set of points where $\varphi = \Phi$ (stopping set), and define $x^* := \sup \mathcal{X}^*$. Put $\Delta g(x) = \varphi(x)$ for all $x \in (x^*, 1)$ and $\Delta g(x) = \Phi$ for all $x \in (0, x^*]$. One can verify that the function Δg also solves (77) and remains $C^1(0,1)$ and C^2-a.e. (the continuity of $\Delta g'(x^*)$ follows by construction). By uniqueness, it must be that $\Delta g = \varphi$. With these properties, clearly $\Delta g \in \mathcal{S}_1$ as well. Hence,

$$\mathcal{S}_4 \subseteq \mathcal{S}_1.$$

As a result, $\mathcal{S}_1 = \mathcal{S}_4$, and both are equal to the singleton Δg. ☐

Proposition 18. *We can write*

$$\Delta g(x) = \beta^* \sum_{n=0}^{\infty} \beta_n (1-x)^n, \quad x \in (x^*, 1),$$

where $\{\beta_n\}_{n\geq 0}$ are defined recursively by

$$\begin{aligned}
\pi(1-\alpha)(n+1)\beta_{n+1} &:= \frac{1}{2\beta^*}(\rho+\pi)^{-1}\sigma_Y^2 + \Big(-\rho-\pi+2\pi(1-\alpha)n + \frac{1}{2}\sigma_Y^2 n(n-1)\Big)\beta_n \\
&\quad - \Big(-\rho-\pi+\pi(1-\alpha)(n-1) + \frac{1}{2}\sigma_Y^2(n-1)(n-4)\Big)\beta_{n-1}, \quad n \geq 2 \\
\beta_2 &:= 1 \\
\beta_1 &:= \frac{2}{\rho+\alpha\pi}\Big(\frac{1}{2\beta^*}(\rho+\pi)^{-1}\sigma_Y^2 - \pi(1-\alpha)\Big) \\
\beta_0 &:= (\rho+\pi)^{-1}\Big[\frac{1}{2\beta^*}(\rho+\pi)^{-1}\sigma_Y^2 - \frac{2\pi(1-\alpha)}{\rho+\alpha\pi}\Big(\frac{1}{2\beta^*}(\rho+\pi)^{-1}\sigma_Y^2 - \pi(1-\alpha)\Big)\Big],
\end{aligned}$$

and β^* solves

$$\beta^* = \frac{1}{2}(\rho+\pi)^{-1}\sigma_Y^2\Big[\pi(1-\alpha) - (\rho+\alpha\pi)\frac{1}{2}\sum_{n\geq 2} n\beta_n(1-x^*)^{n-1}\Big]^{-1}.$$

The optimal entry barrier x^* solves

$$-\frac{(\rho+\alpha\pi)\log(1-\phi)}{\sigma_Y^2}=\frac{1}{2}\frac{\rho+\alpha\pi}{\rho+\pi}+\Big(1-x^*-\frac{\pi(1-\alpha)}{\rho+\pi}\Big)-\frac{\frac{\pi(1-\alpha)}{\rho+\alpha\pi}\big(1-x^*-\frac{\pi(1-\alpha)}{\rho+\pi}\big)-\frac{1}{2}\sum_{n\geq 2}\beta_n(1-x^*)^n}{\frac{\pi(1-\alpha)}{\rho+\alpha\pi}-\frac{1}{2}\sum_{n\geq 2}n\beta_n(1-x^*)^{n-1}}.$$

Proof of Proposition 18. Write the ODE (26) as

$$0=Cx^{-2}+A_0\Delta g+\Big(A_1x+B_1\frac{(1-x)^2}{x}\Big)\Delta g'+B_2(1-x)^2\Delta g'',$$

where $C:=\frac{1}{2}(\rho+\pi)^{-1}\sigma_Y^2$, $A_0:=-(\rho+\pi)$, $A_1:=-\pi(1-\alpha)$, $B_1:=\sigma_Y^2$, and $B_2:=\frac{1}{2}\sigma_Y^2$. Write $\Delta g(x)=\sum_{n\geq 0}a_n(1-x)^n$. Making the change-of-variable $y=1-x$, using the Taylor series $y^{-1}=\sum_{n\geq 0}y^n$ (holds for $y<1$) and rearranging,

$$\begin{aligned}
0&=C(1-y)^{-1}+A_0(1-y)\Delta g+(A_1(1-y)^2+B_1y^2)\Delta g'+B_2(1-y)y^2\Delta g''\\
&=C\sum_{n\geq 0}y^n+A_0(1-y)\sum_{n\geq 0}a_ny^n+(A_1(1-y)^2+B_1y^2)\sum_{n\geq 0}na_ny^{n-1}+B_2(1-y)y^2\sum_{n\geq 0}n(n-1)a_ny^{n-2}\\
&=C\sum_{n\geq 0}y^n+A_0\sum_{n\geq 0}a_ny^n-A_0\sum_{n\geq 1}a_{n-1}y^n+A_1\sum_{n\geq 0}(n+1)a_{n+1}y^n+A_1\sum_{n\geq 1}(n-1)a_{n-1}y^n\\
&\quad-2A_1\sum_{n\geq 1}na_ny^n+B_1\sum_{n\geq 1}(n-1)a_{n-1}y^n+B_2\sum_{n\geq 1}n(n-1)a_ny^n-B_2\sum_{n\geq 1}(n-1)(n-2)a_{n-1}y^n\\
&=C+a_0A_0+a_1A_1+\sum_{n\geq 1}\Big\{C+A_1(n+1)a_{n+1}+[A_0-2A_1n+B_2n(n-1)]a_n\\
&\qquad+[-A_0+A_1(n-1)+B_1(n-1)-B_2(n-1)(n-2)]a_{n-1}\Big\}y^n
\end{aligned}$$

This holds for all y, so we have

$$\begin{aligned}
0&=C+a_0A_0+a_1A_1\\
0&=C+A_1(n+1)a_{n+1}+[A_0-2A_1n+B_2n(n-1)]a_n\\
&\quad+[-A_0+A_1(n-1)+B_1(n-1)-B_2(n-1)(n-2)]a_{n-1},\quad\forall n\geq 1.
\end{aligned}$$

By construction, Δg automatically satisfies the right boundary condition $0=C+A_0\Delta g(1)+A_1\Delta g'(1)$.

Given the value of a_2, solve for a_0, a_1, and $\{a_n\}_{n\geq 3}$ to get

$$\begin{aligned}
a_0&=-\frac{2}{A_1-A_0}[C+A_1a_2]\frac{A_1}{A_0}-\frac{C}{A_0}\\
a_1&=\frac{2}{A_1-A_0}[C+A_1a_2]\\
a_{n+1}&=-\frac{C}{A_1(n+1)}-\frac{1}{A_1(n+1)}\Big[(A_0-2A_1n+B_2n(n-1))a_n\\
&\quad-(-A_0+A_1(n-1)+B_1(n-1)-B_2(n-1)(n-2))a_{n-1}\Big],\quad n\geq 2.
\end{aligned}$$

Define $\beta_n:=\frac{a_n}{a_2}$, $\beta^*:=a_2$, and substitute previously defined constants to get the expressions for $\{\beta_n\}_{n\geq 0}$ in the statement of the proposition.

Determine β^* by the smooth-pasting condition $0 = \Delta g'(x^*) = \beta^* \sum_{n\geq 1} n\beta_n (1-x^*)^{n-1}$, i.e.,

$$C = -\beta^* \Big[A_1 + (A_1 - A_0)\frac{1}{2}\sum_{n\geq 2} n\beta_n (1-x^*)^{n-1} \Big].$$

Finally, solve for x^* via the value-matching condition $\Phi = \Delta g(x^*) = \beta^* \sum_{n\geq 0} \beta_n (1-x^*)^n$, i.e.,

$$\Phi = \Big[-1 + \frac{2A_0}{A_1 - A_0}\Big(1 - x^* - \frac{A_1}{A_0}\Big)\Big]\frac{C}{A_0} + \beta^* \Big[\frac{2A_1}{A_1 - A_0}\Big(1 - x^* - \frac{A_1}{A_0}\Big) + \sum_{n\geq 2} \beta_n (1-x^*)^n \Big].$$

Combining the last two equations to eliminate β^*, we have

$$\frac{\Phi}{C} = \Big[-1 + \frac{2A_0}{A_1 - A_0}\Big(1 - x^* - \frac{A_1}{A_0}\Big)\Big]\frac{1}{A_0} - \frac{\frac{2A_1}{A_1 - A_0}\big(1 - x^* - \frac{A_1}{A_0}\big) + \sum_{n\geq 2} \beta_n (1-x^*)^n}{A_1 + (A_1 - A_0)\frac{1}{2}\sum_{n\geq 2} n\beta_n (1-x^*)^{n-1}}.$$

This completes the proof. □

A.5 Bubbles

Hugonnier (2012) shows that limited participation economies, like the one studied here, must feature "bubbles" in both the risky and riskless asset, as a requirement for equilibrium. "Bubbles" refers to the fact that these assets have equilibrium prices that exceed the cost of any replicating portfolio, a portfolio of assets which exactly replicates the cash flows of the bubble asset. This result is surprising because participants face dynamically complete markets and can make arbitrage profits by purchasing the replicating portfolio and shorting the bubble asset. Although such trades are limited by solvency constraints, some arbitrage trading does take place in equilibrium.

The existence of these bubbles is tightly related to the explosive behavior of local risk prices that occurs when participant wealth falls to zero. With entry, participant wealth never approaches zero, so risk prices are bounded, and there exists a state-price density. In this case, all assets are priced by discounting their cash flows with the state-price density, i.e., by the replicating portfolio, which eliminates bubbles by construction. In this section, I illustrate these ideas in the model with log utility by showing (a) bubbles exist without entry, i.e., when $\phi = 1$; (b) for any $\phi < 1$, there are no bubbles.

To do this, we need to first define some concepts. Let $\mathbb{Q}$ denote the candidate equivalent local martingale measure, and let $\xi_t^* := (\frac{d\mathbb{Q}}{d\mathbb{P}})_{\mathcal{F}_t}$ be the corresponding candidate density. This is given by the exponential local martingale

$$\xi_t^* := \xi_0^* \exp\Big(-\frac{1}{2}\int_0^t \eta_s^2 ds - \int_0^t \eta_s dZ_s \Big).$$

Note that the state-price density process, if it exists, is given by $\xi_t = \exp(-\int_0^t r_s ds)\xi_t^*$. As is well known, the *replicating portfolio* for a sequence of cash flows $\{G_t\}$ has the price

$$F_t^* := \mathbb{E}_t\Big[\int_t^\infty \frac{\xi_s}{\xi_t} G_s ds \Big].$$

A *bubble* is defined by $F_t > F_t^*$, where F_t is the equilibrium price of $\{G_t\}$. We have the following proposition.

Proposition 19. *Consider the equilibrium of Proposition 1. For $\phi = 1$, the price of aggregate endowment $G = Y$ contains a bubble. For any $\phi < 1$, the economy has no bubbles. In both economies, $\mathbb{P}\{X_t > 0\, \forall t\} = 1$.*

Proposition 19 shows that bubbles are a technical issue encountered by complete absence of entry. By examining the proof below, we can see that the technicality emerges because risk prices η_t explode as participant wealth diminishes, $X_t \to 0$. Surprisingly, this is not because this event has any probability of occurring, as we also show that the boundary $\{0\}$ is unattainable for X_t.

Proof of Proposition 19. We first prove the final statement that $\mathbb{P}\{X_t > 0\ \forall t\} = 1$. It suffices to consider $\phi = 1$, in which case X_t is the pure diffusion

$$dX_t = \Big[-\pi(1-\alpha)X_t + \sigma_Y^2 \frac{(1-X_t)^2}{X_t} \Big] dt + (1-X_t)\sigma_Y dZ_t.$$

Indeed, when $\phi < 1$, the diffusive part of X_t is augmented by the weakly increasing process $A_t^{x^*}$. The result for $\phi = 1$ is proved in Lemma B.6, by substituting $\gamma = \psi = 1$.

Next, suppose ξ_t^* is a true martingale. If so, Girsanov's theorem implies that the process $dZ_t^* := dZ_t + \eta_t dt$ is a Brownian motion under $\mathbb{Q}$, which is an equivalent measure to $\mathbb{P}$. Substituting $\eta_t = \sigma_Y / X_t$, the evolution of X_t under $\mathbb{Q}$ is

$$dX_t = \Big[-\pi(1-\alpha)X_t - \sigma_Y^2(1-X_t) \Big] dt + (1-X_t)\sigma_Y dZ_t^* + dA_t^{x^*}.$$

Suppose $\phi = 1$ so that $A_t^{x^*} \equiv 0$. Given $-\pi(1-\alpha)x - \sigma_Y^2(1-x) < 0$ for $x = 0$ and $(1-x)\sigma_Y > 0$ for all $x \in (0,1)$, we see that X_t hits $\{0\}$ with positive $\mathbb{Q}$-probability in finite time, i.e., $\mathbb{Q}\{X_t > 0 \forall t\} < 1$. Hence, $\mathbb{P}$ and $\mathbb{Q}$ are mutually singular, contradicting the assumption that ξ_t^* is a true martingale. Thus, ξ_t^* is a strict local martingale, implying it is a strict super-martingale, by positivity (this means ξ_t^* is a super-martingale that is not a martingale). Hence, $\xi_t^* P_t > \mathbb{E}_t[\int_t^\infty \exp(-\int_t^s r_u du)\xi_s^* Y_s ds]$, implying $P_t > \mathbb{E}_t[\int_t^\infty \frac{\xi_s}{\xi_t} Y_s ds]$. See also Loewenstein and Willard (2000) for the strict local martingale approach to bubbles.

On the other hand, the statement for $\phi < 1$ follows from Step 3 in the proof of Proposition 14), which shows that ξ_t defines a true state-price density process. Consequently, $P_t = \mathbb{E}_t[\int_t^\infty \frac{\xi_s}{\xi_t} Y_s ds]$. □

A.6 Comparison to exogenous segmentation benchmark

Consider the following economy without endogenous entry. The setup is identical to the benchmark model except for the fact that agents are born as participants or non-participants. In particular, a fraction ν of newborns are designated participants, while $1-\nu$ are non-participants, and non-participants may not ever participate. Given the law of large numbers assumption on death shocks, each cohort b will always have ν fraction of participants. I would like to interpret this as an economy where investors have "types" (e.g., experts and non-experts; investors and households), as much of the limited participation literature.

With this modification, the goods market clearing becomes

$$Y_t = \int_{-\infty}^t \pi e^{-\pi(t-b)} \Big(\nu c_{t,b}^P + (1-\nu) c_{t,b}^N \Big) db,$$

and the consumption share state variable is defined by

$$X_t = Y_t^{-1} \nu \int_{-\infty}^t \pi e^{-\pi(t-b)} c_{t,b}^P db,$$

where $c_{t,b}^P$ is the time-t consumption of participants in cohort b, and similarly for $c_{t,b}^N$. To keep the calculations relatively simple, I specialize here to logarithmic utility, keeping other parameters as in Table 1.

This model admits a stationary equilibrium, described in Proposition 20 below. This equilibrium is very similar to that in Proposition 1, with the main difference that the expression for μ_X now adjusts for the continuously entering participants. This operates primarily to shift the stationary mean of X_t.

Proposition 20. *In the model with log utility and exogenous entry (with entry parameter ν), the following is the unique equilibrium. Asset prices are given by*

$$\begin{aligned}\eta(x) &= \frac{\sigma_Y}{x}\\ r(x) &= \rho + \pi + \mu_Y - \frac{\sigma_Y^2}{x}\\ p(x) &= (\rho+\pi)^{-1}\\ \sigma_R(x) &= \sigma_Y\\ \mu_R(x) &= r(x) + \sigma_R(x)\eta(x)\end{aligned}$$

and state dynamics by

$$\begin{aligned}\sigma_X(x) &= (1-x)\sigma_Y\\ \mu_X(x) &= -\pi(1-\alpha)(x-\nu) + \sigma_Y^2\frac{(1-x)^2}{x},\end{aligned}$$

for $x \in [0,1]$. The non-degenerate stationary density of X_t is given by

$$h_\nu(x) = \frac{K_0}{\sigma_X^2(x)}\int_0^x \Big(\frac{x}{z}\Big)^2\Big(\frac{1-x}{1-z}\Big)^{-\frac{2\pi(1-\alpha)}{\sigma_Y^2}} \exp\Big(-\frac{2\pi(1-\alpha)(1-\nu)(x-z)}{\sigma_Y^2(1-x)(1-z)}\Big)dz, \tag{78}$$

where K_0 is a constant ensuring h_ν integrates to 1.

Proof of Proposition 20. Given the similarity to Proposition 1, much of the proof is omitted. One difference is the derivation of the stationary density h_ν, so I document this below. Recall that h_ν satisfies the Kolmogorov forward equation

$$0 = -\frac{d}{dx}(\mu_X h_\nu) + \frac{1}{2}\frac{d^2}{dx^2}(\sigma_X^2 h_\nu).$$

Integrating this equation, we obtain

$$\frac{1}{2}K_0 = -\mu_X h_\nu + \frac{1}{2}\frac{d}{dx}(\sigma_X^2 h_\nu).$$

Since $\mu_X(0) = +\infty$, it follows that $h_\nu(0) = 0$. Thus, making the change of variables $\hat{h}(x) = \sigma_X^2(x)h_\nu(x)$, we must solve $K_0 = -\frac{2\mu_X}{\sigma_X^2}\hat{h} + \hat{h}'$ subject to the boundary condition $\hat{h}(0) = 0$. The solution is $\hat{h}(x) = K_0\int_0^x \exp(\int_z^x \frac{2\mu_X(y)}{\sigma_X^2(y)}dy)dz$. Lastly, the integrand on the right-hand-side can be computed explicitly by substituting μ_X and σ_X from Proposition 20, and the result is

$$\exp\Big(\int_z^x \frac{2\mu_X(y)}{\sigma_X^2(y)}dy\Big) = \Big(\frac{x}{z}\Big)^2\Big(\frac{1-x}{1-z}\Big)^{-\frac{2\pi(1-\alpha)}{\sigma_Y^2}} \exp\Big(-\frac{2\pi(1-\alpha)(1-\nu)(x-z)}{\sigma_Y^2(1-x)(1-z)}\Big).$$

To determine h_ν, we thus need only perform a single integration over z with this integrand. $\square$

A.7 Models with equity issuance and idiosyncratic risk

In this section, I extend the model in two ways – by allowing partial equity issuance by participants and introducing idiosyncratic risks into participants' risky asset returns. To illustrate the main points, it suffices to consider the model with logarithmic utility, which is simpler to analyze.

Now, participants' risky asset position is a claim to $\{\hat{Y}_t\}$, which follows

$$d\hat{Y}_t = \hat{Y}_t[\mu_Y dt + \sigma_Y dZ_t + \hat{\sigma}_Y d\hat{Z}_t],$$

where $\hat{Z}$ is an idiosyncratic Brownian motion, independent of Z. Each participant draws an independent copy of $\hat{Z}$, so that the total risky asset claims in the participant sector will be equal to Y_t, due to the Law of Large Numbers. With these cash flows, participants' risky asset return is given by

$$dR_t = \mu_{R,t} dt + \sigma_{R,t} dZ_t + \hat{\sigma}_Y d\hat{Z}_t.$$

Participants lever up this asset by the choice variable $\theta_{t,b}$.

On their liabilities side, participants keep a fraction $\chi_{t,b} \geq \chi^*$ of their equity risk on their books, with $\chi^* \in (0, 1)$, capturing partial equity issuance. They offload the remaining $1 - \chi_{t,b}$ of risk to financial markets. Here, $\chi_{t,b}$ is a choice variable. When buying participants' outside equity, non-participants optimally diversify away the embedded idiosyncratic risk, so their equity position is summarized by a single portfolio choice variable $\tilde{\theta}_{t,b}$. Finally, I allow participants to purchase long-only diversified positions in other participants' outside equity ($\tilde{\theta}_{t,b} \geq 0$ for $b \in \mathcal{P}_t$), which they might want to do if their aggregate risk exposure is too low after their equity issuance. With the introduction of equity issuance, we require the additional equilibrium equation

$$\int_{\mathcal{P}_t} \pi e^{-\pi(t-b)}(1 - \chi_{t,b})\theta_{t,b} W_{t,b} db = \int_{\mathcal{P}_t \cup \mathcal{N}_t} \pi e^{-\pi(t-b)} \tilde{\theta}_{t,b} W_{t,b} db, \tag{79}$$

which says that the equity offloaded by participants equals the equity investment of non-participants and participants.

Participants are compensated for their equity issuance constraints by additional returns (e.g., management fees), which are captured mathematically by three different risk prices: one for the aggregate risk of inside equity (η_t), one for the idiosyncratic risk of inside equity ($\hat{\eta}_t$), and one for outside equity ($\tilde{\eta}_t$). The idiosyncratic risk price $\hat{\eta}_t$ is a fictitious construct to capture the residual returns available to participants after they are fairly compensated for aggregate risk. Mathematically, we have $\mu_{R,t} - r_t - (1 - \chi_{t,b})\sigma_{R,t}\tilde{\eta}_t$ of returns available to participants after equity issuance, and we define η_t and $\hat{\eta}_t$ such that

$$\chi_{t,b}\sigma_{R,t}\eta_t + \chi_{t,b}\hat{\sigma}_Y\hat{\eta}_t := \mu_{R,t} - r_t - (1 - \chi_{t,b})\sigma_{R,t}\tilde{\eta}_t. \tag{80}$$

With these considerations, agents' budget constraints (4) and (5) are replaced by

$$\frac{dW_{t,b}}{W_{t,b}} = \Big(r_t + \tilde{\theta}_{t,b}\sigma_{R,t}\tilde{\eta}_t + \alpha\pi - \frac{c_{t,b}}{W_{t,b}}\Big)dt + \tilde{\theta}_{t,b}\sigma_{R,t} dZ_t, \quad t < \tau_b \tag{81}$$

$$\begin{aligned}\frac{dW_{t,b}}{W_{t,b}} &= \Big(r_t + \chi_{t,b}\theta_{t,b}(\sigma_{R,t}\eta_t + \hat{\sigma}_Y\hat{\eta}_t) + \tilde{\theta}_{t,b}\sigma_{R,t}\tilde{\eta}_t + \alpha\pi - \frac{c_{t,b}}{W_{t,b}}\Big)dt \\ &\quad + (\chi_{t,b}\theta_{t,b} + \tilde{\theta}_{t,b})\sigma_{R,t} dZ_t + \chi_{t,b}\theta_{t,b}\hat{\sigma}_Y d\hat{Z}_t, \quad t \geq \tau_b.\end{aligned} \tag{82}$$

Equilibrium is given by the following proposition.

Proposition 21. *Assume $(1-\chi^*)\hat{\sigma}_Y^2 < \pi(1-\alpha)$. There exists a unique equilibrium with equity issuance, which is governed by the state variable X_t. When $X_t \geq \chi^*$, aggregate risk is shared perfectly with $\sigma_{X,t} = 0$. When $X_t \in (x^*, \chi^*)$, participants are constrained in the sense that $\chi_t = \chi^*$ and $\tilde{\theta}_t = 0$. When $X_t \leq x^*$, entry occurs until $X_t \geq x^*$, where x^* is determined by solving the ODE (83). Equilibrium objects are given by the following set of functions of x which hold for $x \in [x^*, 1]$:*

$$
\begin{aligned}
\eta(x) &= \frac{\max(x,\chi^*)}{x}\sigma_Y \quad \text{and} \quad \tilde{\eta}(x) = \frac{1-\max(x,\chi^*)}{1-x}\sigma_Y \quad \text{and} \quad \hat{\eta}(x) = \frac{\chi^*}{x}\hat{\sigma}_Y \\
r(x) &= \rho + \pi + \mu_Y - \Big[x\eta^2(x) + x\hat{\eta}^2(x) + (1-x)\tilde{\eta}^2(x)\Big] \\
\sigma_X(x) &= x(1-x)\Big[\eta(x) - \tilde{\eta}(x)\Big] \\
\mu_X(x) &= -\pi(1-\alpha)x + x(1-x)\Big[\eta^2(x) - \tilde{\eta}^2(x) + \hat{\eta}^2(x)\Big] - x(1-x)[x\eta(x) + (1-x)\tilde{\eta}(x)][\eta(x) - \tilde{\eta}(x)],
\end{aligned}
$$

and the stationary density of X_t is given by

$$
h(x) = \begin{cases} \frac{K_0}{\sigma_X^2(x)} \exp\big(\int_{x^*}^{x} \frac{2\mu_X(y)}{\sigma_X^2(y)} dy\big), & x \in [x^*, \chi^*) \\ 0, & x \notin [x^*, \chi^*), \end{cases}
$$

where K_0 is a constant ensuring h integrates to 1, i.e., $\int_{x^*}^{\chi^*} h(x)dx = 1$.

Proof of Proposition 21. The proof proceeds similarly to Proposition 1. As before, conjecture $dX_t = \mu_{X,t}dt + \sigma_{X,t}dZ_t + dA_t^{x^*}$, where A^{x^*} is a continuous, increasing process corresponding to entry, i.e., A^{x^*} only increases when $X_t \leq x^*$.

From agents' HJB equations, consumption is proportional to wealth, $c_t = (\rho+\pi)W_t$. Therefore, the price-dividend ratio is given by $(\rho+\pi)^{-1}$, and hence $\sigma_R = \sigma_Y$.

Portfolio choices are as follows. Because non-participants are unconstrained in their choice of $\tilde{\theta}$, they optimally choose $\tilde{\theta} = \frac{\tilde{\eta}}{\sigma_Y}$. Taking participants' first-order conditions with respect to θ, $\tilde{\theta}$, and χ, we have

$$
\begin{aligned}
[\,\theta\,] &: 0 = \mu_R - r - (1-\chi)\sigma_Y\tilde{\eta} - \chi(\chi\theta + \tilde{\theta})\sigma_Y^2 - \chi(\chi\theta)\hat{\sigma}_Y^2 \\
[\,\tilde{\theta}\,] &: 0 \geq \sigma_Y\tilde{\eta} - (\chi\theta + \tilde{\theta})\sigma_Y^2 \\
[\,\chi\,] &: 0 \geq \sigma_Y\tilde{\eta} - (\chi\theta + \tilde{\theta})\sigma_Y^2 - \chi\theta\hat{\sigma}_Y^2.
\end{aligned}
$$

It is clear that the FOCs for $\tilde{\theta}$ and χ cannot simultaneously hold with equality (unless $\hat{\sigma}_Y = 0$). Furthermore, assuming $\theta > 0$ as will be verified in equilibrium, the FOC for χ must always be slack, so we set $\chi = \chi^*$ (even when $\hat{\sigma}_Y = 0$, we may without loss of generality assume $\chi = \chi^*$ because of the available choice of $\tilde{\theta}$). Let $\mathcal{U}$ denote the region where participants choose $\tilde{\theta} > 0$. We will characterize the equilibrium separately on $\mathcal{U}$ and $[0,1]\backslash\mathcal{U}$.

On $\mathcal{U}$, participants' $\tilde{\theta}$ FOC holds with equality. Using asset market clearing, we have $\theta = x^{-1}$, so $\tilde{\theta} = \frac{\tilde{\eta}}{\sigma_Y} - \frac{\chi^*}{x}$. Substitute this result and non-participants' $\tilde{\theta}$ choice into market clearing for participants' outside equity, equation (79). This yields $1-\chi^* = x(\frac{\tilde{\eta}}{\sigma_Y} - \frac{\chi^*}{x}) + (1-x)\frac{\tilde{\eta}}{\sigma_Y} = \frac{\tilde{\eta}}{\sigma_Y} - \chi^*$. Therefore, $\tilde{\eta} = \sigma_Y$ on $\mathcal{U}$. Substituting this back into participants' choice, we find $\tilde{\theta} = 1 - \frac{\chi^*}{x}$, implying $\mathcal{U} = \{x : x \geq \chi^*\}$. Next, substituting $\theta = x^{-1}$ into participants' FOC for θ, and using equation (80), we obtain $\frac{\chi^*}{x}\hat{\sigma}_Y^2 = \sigma_Y(\eta - \tilde{\eta}) + \hat{\sigma}_Y\hat{\eta}$. One solution to this equation is to set $\eta = \tilde{\eta}$ and $\hat{\eta} = \frac{\chi^*}{x}\hat{\sigma}_Y$. This choice is unique in the sense that it exactly corresponds to participants' shadow risk prices and continues to hold when $\hat{\sigma}_Y = 0$.

On $[0,1]\backslash\mathcal{U}$, participants choose $\tilde{\theta}=0$. Outside equity market clearing yields $\tilde{\eta}=\frac{1-\chi^*}{1-x}\sigma_Y$. Applying asset market clearing to participants' θ choice, we have $\frac{\chi^*}{x}(\sigma_Y^2+\hat{\sigma}_Y^2)=\sigma_Y\eta+\hat{\sigma}_Y\hat{\eta}$. One solution to this equation, which is consistent with the result for $\hat{\eta}$ on $\mathcal{U}$, is to set $\eta=\frac{\chi^*}{x}\sigma_Y$ and $\hat{\eta}=\frac{\chi^*}{x}\hat{\sigma}_Y$. As with the choices of $(\eta,\hat{\eta})$ on $\mathcal{U}$, this choice is unique in the sense that it exactly corresponds to participants' shadow risk prices and continues to hold when $\hat{\sigma}_Y=0$.

The results above directly determine the drifts and diffusions on the wealths of participants and non-participants:

$$\begin{aligned}\frac{dW_t^P}{W_t^P}&=\Big(r_t+\alpha\pi-\rho-\pi+\eta_t^2+\hat{\eta}_t^2\Big)dt+\eta_t dZ_t+\hat{\eta}_t d\hat{Z}_t\\ \frac{dW_t^N}{W_t^N}&=\Big(r_t+\alpha\pi-\rho-\pi+\tilde{\eta}_t^2\Big)dt+\tilde{\eta}_t dZ_t\end{aligned}$$

The dynamics of (μ_X,σ_X) of $X_t:=Y_t^{-1}\int_{\mathcal{P}_t}\pi e^{-\pi(t-b)}c_{t,b}db=P_t^{-1}\int_{\mathcal{P}_t}\pi e^{-\pi(t-b)}W_{t,b}db$ are determined by applying Itô's formula to this definition. Applying Itô's formula to the goods market clearing equation, and substituting previous results, we obtain an equation for r.

We solve for the entry point x^* as before, by solving the following ODE on $(x^*,1)$:

$$0=\frac{1}{2}(\rho+\pi)^{-1}\Big[\eta^2+\hat{\eta}^2-\tilde{\eta}^2\Big]-(\rho+\pi)\Delta g+\mu_X\Delta g'+\frac{1}{2}\sigma_X^2\Delta g'',\quad \Delta g(x^*)=\Phi,\quad \Delta g'(x^*)=0. \tag{83}$$

This is derived by taking the difference between participants' and non-participants HJB equations, as in the discussion leading up to equation (26).

Finally, the stationary distribution is computed using the Kolmogorov Forward Equation, with the reflecting boundary condition at $x=x^*$. Under the assumed parameter restriction $(1-\chi^*)\hat{\sigma}_Y^2<\pi(1-\alpha)$, (x^*,χ^*) constitutes the unique ergodic region. This is because $\mu_X(x)<0$ and $\sigma_X(x)=0$ for all $x\geq\chi^*$. From $x=1$, reaching $x\leq\chi^*$ takes a finite amount of time T that satisfies $1-\chi^*=-\int_0^T\mu_X(X_t)dt$. As μ_X and σ_X are continuous, X_t can never reach χ^* from below. This completes the proof. ☐

Proof of Proposition 8. See Proposition 21 with $\hat{\sigma}_Y=0$. ☐

Proof of Proposition 10. See Proposition 21 with $\chi^*=1$. ☐

Lemma A.4. *Entry occurs at a time τ when the following holds:*

$$\Phi=\mathbb{E}\Big[\frac{1}{2}(\rho+\pi)^{-1}\int_\tau^\infty e^{-(\rho+\pi)(t-\tau)}\Big(\eta_t^2+\hat{\eta}_t^2-\tilde{\eta}_t^2\Big)dt\mid\mathcal{F}_\tau\Big]. \tag{84}$$

Proof of Lemma A.4. The proof proceeds by using equation (83) and proceeding as in Proposition 3. ☐

Proof of Proposition 9. Let $\omega_t^{\chi^*}:=X_t/\chi^*$ be a revised state variable. The revised state dynamics are

$$\begin{aligned}\mu_\omega(\omega)&=-\pi(1-\alpha)\omega+(1-\omega)\Big[\frac{1-\omega}{\omega}+\frac{1-\chi^*}{1-\chi^*\omega}\Big]\sigma_Y^2\\ \sigma_\omega(\omega)&=(1-\omega)\sigma_Y.\end{aligned}$$

Since $\omega_t^{\chi^*}\leq 1$, standard diffusion comparison theorems (see, e.g., Karatzas and Shreve (1991)) imply the path $\{\omega_t^{\chi^*}:t\leq T\}$, for any $T\leq\inf_{\mathcal{E}}(\tau_{x^*}^{\chi^*})$, is uniformly decreasing in χ^*, almost-surely. Hence, $\eta_t=\sigma_Y/\omega_t^{\chi^*}$ is uniformly increasing in χ^* until time T.

Next, define $\tilde{\omega}_t^{\chi^*} := 1 - (1 - \chi^* \omega_t^{\chi^*})/(1 - \chi^*)$. Notice that $(1 - \chi^*)(1 - \tilde{\omega}_t^{\chi^*})/\chi^*$ has the same dynamics as $-\omega_t^{\chi^*}$. Consequently, the process $\{(1 - \chi^*)(1 - \tilde{\omega}_t^{\chi^*})/\chi^* : t \leq T\}$ is uniformly increasing in χ^*. Since $\chi^*/(1 - \chi^*)$ is increasing in χ^*, we have shown that $1 - \tilde{\omega}_t^{\chi^*}$ is also uniformly increasing in χ^*. Hence, $\tilde{\eta}_t = \sigma_Y/(1 - \tilde{\omega}_t^{\chi^*})$ is uniformly decreasing in χ^* until time T.

To prove (ii) and (iii), use the result of Lemma A.4 with $\hat{\sigma}_Y = 0$, so that $\hat{\eta}_t \equiv 0$. Now, we argue by contradiction. Assume, leading to contradiction, $\eta_{\max}^{\chi^*} := \sup_t(\eta_t^{\chi^*})$ is increasing in χ^*.

Because the function $\eta(x; \chi^*) = \sigma_Y \chi^*/x$ is strictly decreasing in x, we have $\eta_{\max}^{\chi^*} = \eta(x^*(\chi^*); \chi^*)$, where $x^*(\chi^*)$ is the equilibrium entry point in the χ^*-economy. This, plus our assumption that $\eta_{\max}^{\chi^*}$ is increasing in χ^*, implies that $\omega^*(\chi^*) := x^*(\chi^*)/\chi^*$ is decreasing in χ^*. Since $\omega_t^{\chi^*} \geq \omega^*(\chi^*)$, which is a reflecting boundary, it can easily be verified that $\{\omega_t^{\chi^*} : t \in \mathbb{R}\}$ is uniformly decreasing in χ^* (i.e., if the reflecting boundary is decreasing in χ^*, we can replace $T = +\infty$ from the previous comparison result). Therefore, $\eta_t = \sigma_Y/\omega_t^{\chi^*}$ is uniformly increasing in χ^* for all t. Similarly, it can be shown that $\tilde{\eta}_t = \sigma_Y/(1 - \tilde{\omega}_t^{\chi^*})$ is uniformly decreasing in χ^* for all t. As a result, the right-hand-side of equation (84) is strictly increasing in χ^*. This is a contradiction, as Φ is unaffected by χ^*. Hence, $\eta_{\max}^{\chi^*}$ is decreasing in χ^*. The proof of (ii) follows from this fact and the fact that $\inf_t(\eta_t^{\chi^*}) = \eta(\chi^*; \chi^*) = \sigma_Y$ is independent of σ_Y.

Finally, we prove (iii). Since $\eta_{\max}^{\chi^*} = \sigma_Y/\omega^*(\chi^*)$ is decreasing in χ^*, we have $\omega^*(\chi^*)$ is increasing in χ^*. Combined with the fact that the path $\{\omega_t^{\chi^*} : t \leq T\}$ is decreasing in χ^*, we have $\inf\{t \geq 0 : \omega_t^{\chi^*} \leq \omega^*(\chi^*)\}$ decreasing in χ^*. But we know that this is the same as $\tau_{x^*}^{\chi^*} = \inf\{t \geq 0 : X_t \leq x^*(\chi^*)\}$. □

Proof of Proposition 11. Consider an increase in $\hat{\sigma}_Y$. This increases $\mu_X(x)$, through $\hat{\eta}(x)$, and leaves $\sigma_X(x)$ unchanged. Standard diffusion comparison theorems (see, e.g., Karatzas and Shreve (1991)) imply that the equilibrium process $\{X_t^{\hat{\sigma}_Y} : t \leq T\}$ is uniformly increasing in $\hat{\sigma}_Y$, almost-surely, where $T \leq \inf_{\mathcal{E}}(\tau_{x^*}^{\hat{\sigma}_Y})$. Therefore, $\eta_t = \sigma_Y/X_t^{\hat{\sigma}_Y}$ is uniformly decreasing in $\hat{\sigma}_Y$. □

Proof of Proposition 12. Substitute $X_t \equiv 1$ in expression (41). □

A.8 Adding exogenous state variables

What happens to entry decisions when additional state variables govern aggregate dynamics? To answer this question, consider an exogenous diffusion $S_t \in \mathbb{R}^n$ following

$$dS_t = \mu_S(S_t)dt + \sigma_S(S_t)d\hat{Z}_t, \tag{85}$$

where $\hat{Z}$ is a d-dimensional Brownian motion independent from Z. Generalize aggregate endowment dynamics (1) to depend on S_t,

$$\frac{dY_t}{Y_t} = \mu_Y(S_t)dt + \sigma_Y(S_t)dZ_t. \tag{86}$$

Assume that conditions are satisfied such that (85) and (86) have unique strong solutions. For example, the long-run risk setting analyzed in Bansal and Yaron (2004) fits into this setup.[19]

To ensure asset markets are dynamically complete for participants, allow them to trade a full set of zero-net-supply futures contracts tracking the Brownian shocks. Non-participants cannot trade these contracts. The contract tracking Brownian motion $\hat{Z}^{(j)}$ for $j \in \{1, \dots, d\}$ delivers returns $dR_t^{(j)} = \hat{\eta}_t^{(j)}dt + d\hat{Z}_t^{(j)}$. The unique state-price density is given by $\overline{\eta}_t := (\eta_t, \hat{\eta}_t^{(1)}, \dots, \hat{\eta}_t^{(d)})$, where η is the risk price for the Z shock. No

[19]This would be accomplished by setting $n = d = 2$, where

$$\frac{dY_t}{Y_t} = \left[\overline{\mu}_Y + S_t^{(1)}\right]dt + \sqrt{S_t^{(2)}}\overline{\sigma}_Y dZ_t$$

arbitrage requires expected excess returns satisfy $\mu_R - r = \sigma_R \cdot \overline{\eta}$, where σ_R is now a $d \times 1$ diffusion vector for the return on the aggregate endowment claim.

In a Markov equilibrium, the consumption share of participants X_t remains the unique endogenous state variable, which follows

$$dX_t = \mu_X(X_t, S_t)dt + \sigma_X(X_t, S_t)dZ_t.$$

All equilibrium objects can be written as functions of (X_t, S_t), e.g., $\eta_t = \eta(X_t, S_t)$. We have the following proposition, a generalization of Proposition 3.

Proposition 22. *With the general exogenous state dynamics in (85)-(86), the optimal entry time solves*

$$\inf_\tau \mathbb{E}\Big[\frac{1}{2}(\rho+\pi)^{-1}\int_0^\tau e^{-(\rho+\pi)t}\eta_t^2 dt + e^{-(\rho+\pi)\tau}\Phi \mid \mathcal{F}_0\Big].$$

Consequently, entry patterns can be fully characterized by the equilibrium process for $\{\eta_t\}_{t\in\mathbb{R}}$.

Proof of Proposition 22. Omitted, as it resembles very closely the proofs of Propositions 1 and 3. □

The result of Proposition 22 states that the logic behind entry patterns in the simpler one-dimensional log utility model are robust to adding more dimensions. With log utility, the auxiliary risk prices are all zero, $\hat{\eta}^{(j)} = 0$ for all j. In this case, entry can be summarized by the one-dimensional object η_t^2 and its dynamics, which, of course, are affected by S_t, since $\eta_t = \sigma_Y(S_t)/X_t$ in equilibrium. Another consequence of Proposition 22 is that we may "back out" ϕ as we did before in equation (32). We have

$$\phi = 1 - \exp\Big(-\frac{1}{2}\mathbb{E}^{\eta_\tau}\Big[\int_\tau^\infty e^{-(\rho+\pi)t}\eta_t^2 dt\Big]\Big),$$

where η_τ is the risk price prevailing at some entry time τ. Again, entry costs are given by the expected present discounted value of the squared market Sharpe ratios, starting from an extreme state of the world.

A.9 Model with extrapolative expectations

Here, I derive the equilibrium under extrapolative expectations, described in section 5 of the text.

Proof of Proposition 13. Posit an equilibrium in the two state variables (X_t, S_t). From equation (44), $dS_t = \beta(\mu_{R,t} - S_t)dt + \beta\sigma_{R,t}dZ_t$. As before, conjecture $dX_t = \mu_{X,t}dt + \sigma_{X,t}dZ_t + dA_t^{x^*}$, where A^{x^*} is a continuous, increasing process corresponding to entry. Suppose $A_t^{x^*}$ only increases when $X_t \leq x^*(S_t)$ for some function $x^*(\cdot)$. Below, we show that $x^*(\cdot)$ is a constant function and denote its level by x^*.

Due to log utility, participants and non-participants with wealth level W_t achieve values $V^i(W_t, X_t, S_t) = (\rho+\pi)^{-1}\log W_t + g_i(X_t, S_t)$ for $i \in \{P, N\}$. HJB equations for g_P and g_N are derived similarly to Proposition 1, except under the probability measure induced by the extrapolative expectations, $\tilde{\mathbb{P}}$. By Girsanov's theorem, this change-of-measure only serves to adjust the drifts of each stochastic process. In particular,

and the state variables follow

$$dS_t^{(1)} = -\lambda_1(S_t^{(1)} - \kappa_1)dt + \sqrt{S_t^{(2)}}\omega_1 d\hat{Z}_t^{(1)}$$
$$dS_t^{(2)} = -\lambda_2(S_t^{(2)} - \kappa_2)dt + \sqrt{S_t^{(2)}}\omega_2 d\hat{Z}_t^{(2)}.$$

Here, $S^{(1)}$ is a growth-rate process, and $S^{(2)}$ is a stochastic volatility process.

given the beliefs in equation (45), agents perceive $\tilde{Z}_t$ as the driving Brownian motion, which in fact follows

$$d\tilde{Z}_t = dZ_t + \Big(\frac{\mu_{R,t} - \tilde{\mu}_{R,t}}{\sigma_{R,t}}\Big)dt,$$

so all drifts are modified to reflect the term $\frac{\mu_{R,t}-\tilde{\mu}_{R,t}}{\sigma_{R,t}}dt$. Denoting these modified drifts by $\tilde{\mu}_X$ and $\tilde{\mu}_S$, and using $\tilde{\eta} = \frac{\tilde{\mu}_R - r}{\sigma_R}$ for the perceived Sharpe ratio under $\tilde{\mathbb{P}}$, the HJB equation for participants is

$$\begin{aligned}0 &= \log(\rho+\pi) - 1 + (\rho+\pi)^{-1}[\alpha\pi + r + \frac{1}{2}\tilde{\eta}^2] - (\rho+\pi)g_P \\ &\quad + \tilde{\mu}_X \frac{\partial}{\partial x} g_P + \tilde{\mu}_S \frac{\partial}{\partial s} g_P + \frac{1}{2}\sigma_X^2 \frac{\partial^2}{\partial x^2} g_P + \frac{1}{2}\sigma_S^2 \frac{\partial^2}{\partial s^2} g_P + \sigma_X \sigma_S \frac{\partial^2}{\partial x \partial s} g_P\end{aligned}$$

and similarly for g_N (with $\tilde{\eta}$ replaced by 0). As in Proposition 1, it suffices to consider $\Delta g := g_P - g_N$, which solves

$$0 = \frac{1}{2}(\rho+\pi)^{-1}\tilde{\eta}^2 + \Big[\tilde{\mu}_X \frac{\partial}{\partial x} + \tilde{\mu}_S \frac{\partial}{\partial s} + \frac{1}{2}\sigma_X^2 \frac{\partial^2}{\partial x^2} + \frac{1}{2}\sigma_S^2 \frac{\partial^2}{\partial s^2} + \sigma_X\sigma_S \frac{\partial^2}{\partial x \partial s} - (\rho+\pi)\Big]\Delta g$$

subject to the value-matching and smooth-pasting conditions $\Delta g(x^*(s), s) = \Phi$ and $\frac{\partial}{\partial x}\Delta g(x^*(s), s) = 0$.

To solve for equilibrium objects, we use a similar procedure as before. First, since all agents' consumption-wealth ratios equal $\rho + \pi$, the asset market clearing condition implies the price-dividend ratio and stock volatility are constant: $p_t = (\rho+\pi)^{-1}$ and $\sigma_{R,t} = \sigma_Y$. This gives us $\sigma_S = \beta\sigma_Y$. Next, apply Itô's formula to the definition of X_t and the goods market clearing condition, noting that participants' consumption growth volatility equals $\tilde{\eta}$, while it is 0 for non-participants. This yields a system of equations in $\tilde{\eta}$, r, μ_X, and σ_X, and the solution is given in the expressions of Proposition 13. Then, $\frac{\sigma_Y}{x} = \tilde{\eta} := \frac{\tilde{\mu}_R - r}{\sigma_R}$ can be inverted for $\tilde{\mu}_R$ and then combined with equation (45) to solve for $\mu_R = \rho + \pi + \mu_Y$. Hence, $\mu_S = \beta(\rho + \pi + \mu_Y - s)$. Lastly, no-arbitrage with extrapolative agents implies discounted returns must be local martingales under $\tilde{\mathbb{P}}$, which gives us the equation $\tilde{\mu}_R - r = \eta(\sigma_R + \tilde{\eta} - \frac{\mu_R - r}{\sigma_R})$. This equation can be solved for η.

Having solved for the equilibrium objects, it remains to determine equilibrium entry, $x^*(s)$. Substitute all equilibrium objects back into the PDE for Δg and observe that the only term depending on s is $\tilde{\mu}_S = (1-\lambda)\beta(\rho+\pi+\mu_Y - s)$. Guess $\frac{\partial}{\partial s}\Delta g \equiv 0$ identically, so that Δg only depends on x. Thus, the PDE simplifies to an ODE, which shows that $x^*(s) = x^*$ constant. The ODE we need to solve to determine x^* is

$$0 = \frac{1}{2}(\rho+\pi)^{-1}\tilde{\eta}^2 - (\rho+\pi)\Delta g + \tilde{\mu}_X \Delta g' + \frac{1}{2}\sigma_X^2 \Delta g'', \quad \Delta g(x^*) = \Phi, \quad \Delta g'(x^*) = 0, \tag{87}$$

where $\tilde{\eta} = \frac{\sigma_Y}{x}$, $\tilde{\mu}_X = -\pi(1-\alpha)x + \sigma_Y^2 \frac{(1-x)^2}{x}$, and $\sigma_X = (1-x)\sigma_Y$. By inspection, this exactly the same problem as (26), which implies that the entry point x^* is independent of λ and β.

Finally, we could apply an identical existence/uniqueness theory as in Appendix A.4, Proposition 2, so the equilibrium exists and is unique. $\square$

References

Ackerman, Robert Argento, Gerhard Fries, and Richard A Windle. 2012. "Changes in US family finances from 2007 to 2010: Evidence from the Survey of Consumer Finances." *Federal Reserve Bulletin* 100:1–80.

Adrian, Tobias, Erkko Etula, and Tyler Muir. 2014. "Financial Intermediaries and the Cross-Section of Asset Returns." *The Journal of Finance* 69 (6):2557–2596.

Bansal, Ravi and Amir Yaron. 2004. "Risks for the long run: A potential resolution of asset pricing puzzles." *The Journal of Finance* 59 (4):1481–1509.

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer. 2015. "X-CAPM: An extrapolative capital asset pricing model." *Journal of Financial Economics* 115 (1):1–24.

Basak, Suleyman and Domenico Cuoco. 1998. "An Equilibrium Model with Restricted Stock Market Participation." *Review of Financial Studies* 11 (2):309–341.

Bensoussan, Alain and J-L Lions. 1982. *Applications of variational inequalities in stochastic control*, vol. 12. Elsevier.

Bhandari, Anmol, Jaroslav Borovička, and Paul Ho. 2016. "Identifying ambiguity shocks in business cycle models using survey data."

Blanchard, Olivier J. 1985. "Debt, Deficits, and Finite Horizons." *Journal of Political Economy* 93 (2):223–47.

Borovička, Jaroslav and Lars Peter Hansen. 2016. "Term structure of uncertainty in the macroeconomy." *Handbook of Macroeconomics* 2:1641–1696.

Borovička, Jaroslav, Lars Peter Hansen, Mark Hendricks, and José A Scheinkman. 2011. "Risk-price dynamics." *Journal of Financial Econometrics* 9 (1):3–65.

Borovička, Jaroslav, Lars Peter Hansen, and José A Scheinkman. 2014. "Shock elasticities and impulse responses." *Mathematics and Financial Economics* 8 (4):333–354.

Brunnermeier, Markus K and Yuliy Sannikov. 2014. "A Macroeconomic Model with a Financial Sector." *The American Economic Review* 104 (2):379–421.

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

Cao, H Henry, Tan Wang, and Harold H Zhang. 2005. "Model uncertainty, limited market participation, and asset prices." *Review of Financial Studies* 18 (4):1219–1251.

Di Tella, Sebastian. 2017. "Uncertainty shocks and balance sheet recessions." *Journal of Political Economy* 125 (6):2038–2081.

Dow, James and Sergio Ribeiro da Costa Werlang. 1992. "Uncertainty aversion, risk aversion, and the optimal choice of portfolio." *Econometrica: Journal of the Econometric Society* :197–204.

Duffie, Darrell. 2010a. *Dynamic Asset Pricing Theory*. Princeton University Press.

———. 2010b. "Presidential Address: Asset Price Dynamics with Slow-Moving Capital." *The Journal of finance* 65 (4):1237–1267.

Duffie, Darrell and Larry G Epstein. 1992. "Stochastic differential utility." *Econometrica* 60 (2):353–394.

Eisfeldt, Andrea L, Hanno N Lustig, and Lei Zhang. 2017. "Complex Asset Markets."

Epstein, Larry G and Martin Schneider. 2007. "Learning under ambiguity." *The Review of Economic Studies* 74 (4):1275–1303.

Friedman, Avner. 2010. *Variational principles and free-boundary problems.* Courier Corporation.

Fuster, Andreas, Benjamin Hebert, and David Laibson. 2011. "Natural expectations, macroeconomic dynamics, and asset pricing."

Gârleanu, Nicolae and Stavros Panageas. 2015. "Young, Old, Conservative, and Bold: The Implications of Heterogeneity and Finite Lives for Asset Pricing." *Journal of Political Economy* 123 (3):670–685.

Gomes, Francisco and Alexander Michaelides. 2008. "Asset pricing with limited risk sharing and heterogeneous agents." *Review of Financial Studies* 21 (1):415–448.

Greenwood, Robin and Andrei Shleifer. 2014. "Expectations of returns and expected returns." *Review of Financial Studies* 27 (3):714–746.

Gromb, Denis and Dimitri Vayanos. 2002. "Equilibrium and Welfare in Markets with Financially Constrained Arbitrageurs." *Journal of Financial Economics* 66 (2):361–407.

Guvenen, Fatih. 2009. "A Parsimonious Macroeconomic Model for Asset Pricing." *Econometrica* 77 (6):1711–1750.

Haddad, Valentin. 2014. "Concentrated Ownership and Equilibrium Asset Prices."

He, Zhiguo, Bryan Kelly, and Asaf Manela. 2017. "Intermediary asset pricing: New evidence from many asset classes." *Journal of Financial Economics* 126 (1):1–35.

He, Zhiguo and Arvind Krishnamurthy. 2012. "A Model of Capital and Crises." *The Review of Economic Studies* 79 (2):735–777.

———. 2013. "Intermediary Asset Pricing." *The American Economic Review* 103 (2):732–70.

———. 2014. "A Macroeconomic Framework for Quantifying Systemic Risk."

Hugonnier, Julien. 2012. "Rational Asset Pricing Bubbles and Portfolio Constraints." *Journal of Economic Theory* 147 (6):2260–2302.

Karatzas, Ioannis, John P Lehoczky, and Steven E Shreve. 1991. "Equilibrium Models with Singular Asset Prices." *Mathematical Finance* 1 (3):11–29.

Karatzas, Ioannis and Steven E Shreve. 1991. "Brownian Motion and Stochastic Calculus." *Graduate Texts in Mathematics* 113.

———. 1998. *Methods of Mathematical Finance*, vol. 39. Springer Science & Business Media.

Kaustia, Markku and Samuli Knüpfer. 2012. "Peer performance and stock market entry." *Journal of Financial Economics* 104 (2):321–338.

Kyle, Albert S and Wei Xiong. 2001. "Contagion as a Wealth Effect." *The Journal of Finance* 56 (4):1401–1440.

Loewenstein, Mark and Gregory A Willard. 2000. "Rational equilibrium asset-pricing bubbles in continuous trading models." *Journal of Economic Theory* 91 (1):17–58.

Luttmer, Erzo GJ and Thomas Mariotti. 2003. "Subjective discounting in an exchange economy." *Journal of Political Economy* 111 (5):959–989.

Malloy, Christopher J, Tobias J Moskowitz, and Annette Vissing-Jørgensen. 2009. "Long-run Stockholder Consumption Risk and Asset Returns." *The Journal of Finance* 64 (6):2427–2479.

Mankiw, N Gregory and Stephen P Zeldes. 1991. "The Consumption of Stockholders and Nonstockholders." *Journal of Financial Economics* 29 (1):97–112.

Parker, Jonathan A and Annette Vissing-Jørgensen. 2009. "Who Bears Aggregate Fluctuations and How?" *The American Economic Review* 99 (2):399.

Polkovnichenko, Valery. 2004. "Limited Stock Market Participation and the Equity Premium." *Finance Research Letters* 1 (1):24–34.

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

Zhang, Tu-Sheng. 1994. "On the strong solutions of one-dimensional stochastic differential equations with reflecting boundary." *Stochastic processes and their applications* 50 (1):135–147.

B Internet Appendix for "Entry and slow-moving capital"

In this online appendix, we explore several technical and auxiliary results. First, in section B.1, I discuss the quantitative appeal of adding recursive Epstein-Zin preferences in the model. Section B.2 presents some further quantitative asset pricing explorations, including long-run asset price moments, the term structure of interest rates, and the term structure of risk prices. These explorations are meant to further understand the shortcomings of the limited participation model. Section B.3 provides details on the equilibrium with asymptotically large entry costs (i.e., $\phi \to 1$). In section B.4, I derive the hitting time densities for X and apply them to gauge the time it takes for this economy to reach crisis states. Finally, section B.5 discusses the robustness of the model to several different extensions.

B.1 Benefits of recursive preferences

In this section, we discuss the benefits of recursive Epstein-Zin preferences in the model. To understand the quantitative appeal of recursive utility, I compare asset prices from Proposition 14 (with parameters given in table 1) to an identical economy with CRRA preferences, that is setting $\psi = \gamma$. The results of this comparison are plotted in figure 10.

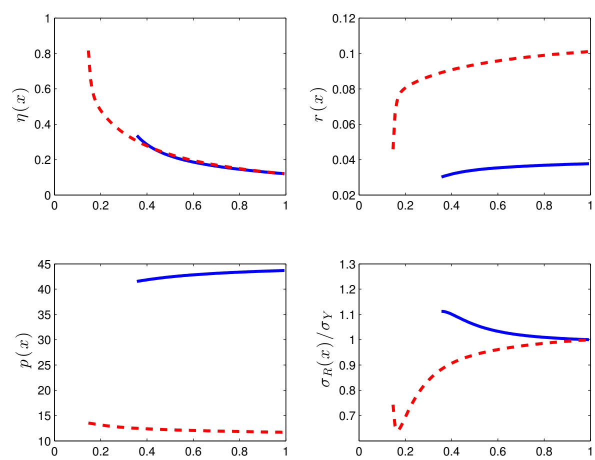

Figure 10: Asset prices in the benchmark limited participation economy with recursive preferences (blue) versus CRRA preferences (dashed red). The horizontal axis is the participants' consumption share x.

First, notice that the risk-free rate is substantially lower under recursive utility, helping resolve the risk-free rate puzzle. This resolution is primarily due to reducing the contribution of the term $\psi(\mu_Y + \pi + \pi(1-\alpha)\frac{p(x)}{g_N(x)})$ in (59). Indeed, if the terms in parentheses are approximately 3%, then lowering ψ from $\psi = \gamma = 3$ to $\psi = 1$ reduces r by 6%, which explains two-thirds of the fall in r seen in figure 10.

Second, with $\psi < 1$, we find that the price-dividend ratio is procyclical, $p'(x) > 0$, meaning that positive cash flow shocks translate to higher prices as we see in the data. For the same reason, asset volatility can be made higher than fundamental volatility, $\sigma_R > \sigma_Y$, and countercyclical, $\sigma_R'(x) < 0$. The reason for both effects relates to the standard intuition that cash flow effects (or income effects) dominate discount rate effects (or substitution effects) when the EIS is larger than 1. In particular, when there is a negative shock to dividends, cash flows are permanently lower while discount rates also fall, as dividends are consumed in general equilibrium. The net effect in a representative agent model is to lower the price-dividend ratio if EIS is larger than 1 and raise the price-dividend ratio if the EIS is smaller than 1. When the EIS exceeds 1, the price-dividend ratio responds in the same direction as the cash flow shock, which amplifies returns, thereby increasing volatility. Figure 10 shows graphically, and Lemma B.1 demonstrates analytically, that this intuition carries through, in some sense, to limited participation models.[20]

Lemma B.1. *Suppose $\psi < 1$. Then, for all x close enough to x^*, $\sigma_R(x) > \sigma_R(1) = \sigma_Y$.*

Finally, the following lemma emphasizes another quantitative virtue of recursive preferences in a limited participation model: expected returns on risky assets can be made substantially more volatile than riskless returns. In particular, one can pick parameters such that nearly all of the volatility of risk premia comes from expected return variation rather than risk-free rate variation, matching the data. In symbols, we have by Itô's formula,

$$\frac{\mathrm{std}[\mu_R(x) - r(x)]}{\mathrm{std}[r(x)]} = \frac{|\eta(x)\sigma_R'(x) + \eta'(x)\sigma_R(x)|}{|r'(x)|}. \tag{88}$$

So the high risk premium volatility is due to high and time-varying risk prices (volatility and conditional heteroskedasticity of the SDF), along with high and time-varying risk quantities (volatility and conditional heteroskedasticity of returns).

With log utility, it's the opposite: all variation in risk premia is due to variation in the riskless rate, as in the model of Basak and Cuoco (1998), and (88) is equal to 1.[21] With more general CRRA utility, a similar result holds numerically, as one can see in figure 9. Indeed, observe that $r'(x)$ is higher with CRRA than recursive preferences, making the denominator of (88) higher, while $\sigma_R'(x) > 0$ with CRRA, making the numerator lower (since $\eta'(x) < 0$). Recursive preferences are needed: the proof of Lemma B.2 suggests that $\psi < 1$ and $\gamma > 1$ are requisite parameter choices to generate the result.

Lemma B.2. *There exist choices for the parameters such that*

$$\frac{std[\mu_R(x) - r(x)]}{std[r(x)]}$$

can be made arbitrarily large. In particular, given ϕ small enough, there exist choices for the other parameters making (88) arbitrarily large.

Proof of Lemma B.1. First, note that

$$\begin{aligned} p'(x^*) &= x^* g_P'(x^*) + (1 - x^*) g_N'(x^*) + g_P(x^*) - g_N(x^*) \\ &= g_P(x^*) \Big[1 - (1 - \phi)^{\frac{1-\psi}{\psi}} \Big], \end{aligned}$$

which is strictly positive for $\psi < 1$. As a result, there exists $\delta_1 > 0$ such that $p'(x) > 0$ for all $x \in [x^*, x^* + \delta_1)$. Similarly, since $\omega(x^*) = 1$, there exists $\delta_2 > 0$ such that $\omega(x) > 0$ for all $x \in [x^*, x^* + \delta_2)$. Letting $\delta := \delta_1 \wedge \delta_2$, we have

$$\sigma_R(x) = \sigma_Y \Big[1 + (1 - x)\omega(x) \frac{p'(x)}{p(x)} \Big] > \sigma_Y = \sigma_R(1)$$

for all $x \in [x^*, x^* + \delta)$. □

[20] In numerical solutions, whenever $\psi < 1$, we always find stock volatilities to be decreasing in the state.

[21] The model with log utility is solved in Appendix A.4, and one can easily compute that (88) equals 1 using those formulas.

Proof of Lemma B.2. Approximating the equilibrium objects described in Proposition 14 for x near x^*, we obtain

$$\eta(x) = \gamma\sigma_Y + O(1-x^*)$$
$$\sigma_R(x) = \sigma_Y + O(1-x^*),$$

and similarly for the first-derivatives,

$$\eta'(x) = -\gamma\sigma_Y + O(1-x^*)$$
$$\sigma_R'(x) = \sigma_Y\Big[1-(1-\phi)^{\frac{1-\psi}{\psi}}\Big] + \sigma_Y O(1) + O(1-x^*)$$
$$r'(x) = \psi\pi(1-\alpha)\Big[1-(1-\phi)^{\frac{\psi-1}{\psi}}\Big] + \frac{1}{2}\gamma(\psi+1)\sigma_Y^2 + O(1-x^*),$$

where we have used the facts that

$$p(x^*) = \Big[x^*(1-\phi)^{\frac{\psi-1}{\psi}} + (1-x^*)\Big]g_N(x^*),$$

$$p'(x^*) = \Big[(1-\phi)^{\frac{\psi-1}{\psi}} - 1\Big]g_N(x^*),$$

and

$$(1-x^*)p''(x^*) = (1-x^*)x^*g_P''(x^*) + (1-x^*)^2 g_N''(x^*),$$

as well as the result in Lemma B.3 below, which implies that $(1-x^*)p''(x^*)$ converges to a finite constant, possibly zero, as $x^* \to 1$ (this constant is denoted by $O(1)$ above).

Next, observe that there exist parameters such that $r'(x) = O(1-x^*)$, i.e., such that

$$\psi\pi(1-\alpha)\Big[1-(1-\phi)^{\frac{\psi-1}{\psi}}\Big] + \frac{1}{2}\gamma(\psi+1)\sigma_Y^2 = 0.$$

Indeed, choosing ψ arbitrarily close to 0 makes the left-hand-side diverge to $-\infty$, while choosing $\psi > 1$ makes the left-hand-side positive. Choosing parameters in such a way, we find $r'(x) \to 0$ as $\phi \to \phi^* := \sup\{\phi : x^*(\phi) = 1\}$. On the other hand,

$$\eta'(x)\sigma_R(x) + \eta(x)\sigma_R'(x) = -\gamma\sigma_Y^2(1-\phi)^{\frac{1-\psi}{\psi}} + \gamma\sigma_Y^2 O(1) + O(1-x^*),$$

which converges to a non-zero constant as $\phi \to \phi^*$. Hence,

$$\lim_{\phi\to\phi^*}\frac{|\eta(x)\sigma_R'(x) + \eta'(x)\sigma_R(x)|}{|r'(x)|} = +\infty$$

under these parameter choices, proving the claim. □

Lemma B.3. *Let $\{\phi_n\}_{n=1}^{\infty}$ be any sequence converging to $\phi^* := \inf\{\phi : x^*(\phi) = 1\}$ such that $\phi_n > \phi^*$ for all n, and let x_n^* denote the corresponding equilibrium entry threshold. For any $x_n \in [x_n^*, 1]$, we have $\limsup_n g_i''(x_n)(1-x_n) < +\infty$ for $i \in \{P, N\}$.*

Proof. Let ϕ_n and x_n^* be the entry cost and entry point, respectively. Approximating $g_P'(1) - g_N'(1)$ by its value at point $x_n \in (x_n^*, 1)$, then taking the limit $x_n \to x_n^*$ and using the smooth-pasting conditions (25) results in

$$g_i'(1) = g_i''(x_n^*+)(1 - x_n^*) + o(1 - x_n^*), \quad i \in \{P, N\}.$$

Because the boundary conditions (65) and (66) imply that $g_i'(1)$ is uniformly bounded (for any ϕ and x^*), the result that $g_i''(x_n)(1 - x_n)$ converges uniformly is immediate. □

B.2 Other asset price explorations

Asset price moments.

Another way to evaluate the quantitative impact of the entry cost ϕ is to study different moments of asset prices, which are directly comparable with the data. Figure 11 below depicts various unconditional asset pricing moments as a function of ϕ. Risk-free rates (panel 1) are relatively low, but their volatility (panel 4) is far too low compared to the data. Only once $\phi > 0.9$ does the riskless rate volatility begin to rise, although the level of the risk-free rate also rises. The annualized equity premium (panel 3) and annualized equity volatility (panel 6) are relatively small as the volatility of dividends is close to the volatility of output σ_Y. By accounting for leverage in firms' capital structures, we would scale these quantities by the average firm leverage ratio and do better to match the data on the equity premium. However, even after making such an adjustment, the equity premium and its volatility are far too low for moderate entry costs.

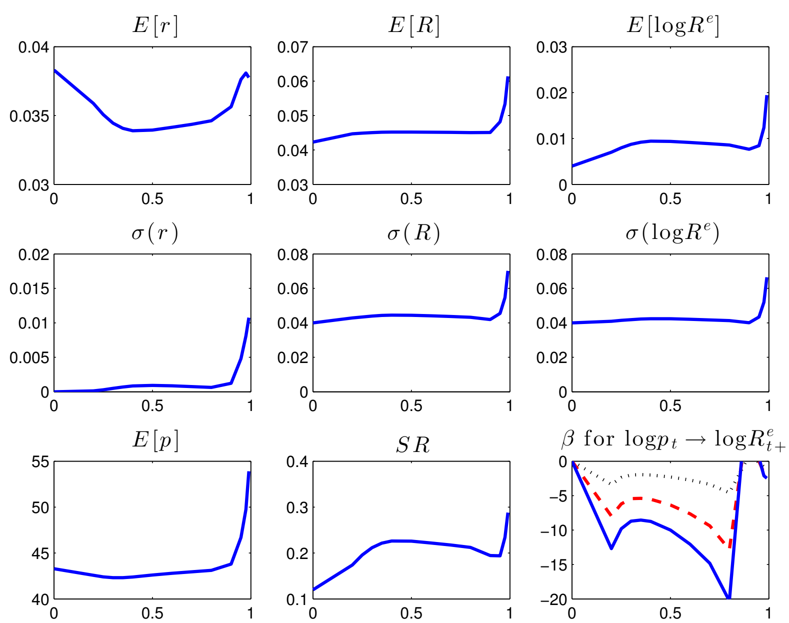

Figure 11: Stationary averages for various asset pricing objects, as a function of ϕ (horizontal axis). To construct the data, I simulate 100 economies for 50 years each. To construct annual data, I time-aggregate the returns: r refers to the instantaneous riskless rate at the beginning of each year; R refers to the annual gross return on equity; $\log R^e$ refers to log excess returns, i.e., $\log R - \log R^f$, where R^f is the annual gross return on continuously reinvesting at the short-term rate r for one year. The β's in panel 9 refer to coefficients from a forecasting regression of k-year log excess returns $\log R^e_{t+k}$ on a constant and the log price-dividend ratio at the beginning $\log p_t$.

One direct consequence of relatively low risk prices η is that Sharpe ratios (panel 8) produced by the model are too small, and these Sharpe ratios are unaffected by any capital structure adjustments discussed above. Compared to a market Sharpe ratio of approximately 0.4 in the data, the model with $\phi < 0.9$

generates Sharpe ratios on the order of 0.2. Only by cranking up entry costs to unreasonable levels can we increase the market Sharpe ratio.

Finally, the model produces excess returns that are far too predictable (panel 9). Panel 9 shows the slope coefficient from regression of future excess returns on the current price-dividend ratio; the dotted, dashed, and solid lines are for $k = 1$-, 3-, and 5-year ahead returns, respectively. Return predicability goes the right direction, but it is an order of magnitude too large relative to the data. Intuitively, for most levels of entry costs, the price-dividend ratio (p) is globally negatively correlated with risk premia. In addition, because p is much more stable than risk premia in the model, small changes in p translate into much larger changes in risk premia.[22]

Term structure of interest rates.

Let $q_{t,T}$ be the time-t price of a zero-coupon bond paying one unit of the numeraire at time T. Then, pricing implies that $\xi_t q_{t,T} = \mathbb{E}[\xi_{t+\tau} q_{t+\tau,T} \mid \mathcal{F}_t]$, so that $\xi_t q_{t,T}$ is a martingale. At the same time, there is a function q of the bond maturity and Markov state such that $q(T-t, X_t) = q_{t,T}$, for $T - t \geq 0$, the time-to-maturity. Applying Itô's formula, these observations lead to a PDE for q on $(x^*, 1)$:

$$rq = -\frac{\partial q}{\partial s} + (\mu_X - \eta\sigma_X)\frac{\partial q}{\partial x} + \frac{1}{2}\sigma_X^2 \frac{\partial^2 q}{\partial x^2}, \quad s > 0, \quad x \in (x^*, 1]. \tag{89}$$

This PDE is solved backward in time, or forward in maturity, subject to the initial condition $q(0,x) = 1$ and the boundary condition $\frac{\partial_x q(s,x^*)}{q(s,x^*)} = \frac{p'(x^*)}{p(x^*)}$, which arises by no-arbitrage.[23]

Figure 12 shows the yield curves coming out of the model. Yields are defined by $y_{t,T} := -\frac{1}{T-t}\log q_{t,T}$, as usual. First, notice that the curves are shifted down in low-x states, as interest rates are lower. Second, notice that yield curves are typically downward-sloping in high-x states and downward-sloping in low-x states. The intuition is that the risk-free rate $r(x)$ is increasing and $\mu_X(x)$ decreasing in x, simply due to an expectations-hypothesis-like argument. For example, if X_t is high, so that r_t is high and $\mu_{X,t}$ is low (negative), then X_{t+s} is expected to be lower than X_t, implying r_{t+s} is expected to be below r_t. In addition, the function $r(x)$ is concave in x, implying lower yield curve slopes than without this concavity, by a Jensen's inequality argument. This is why the high-x yield curves are more downward-sloping than the low-x yield curves are upward-sloping. Finally, the dominant characteristic of the yield curves is that their slopes are very small in magnitude. This is due to the fact that the model does not attain extreme low-x states, whether due to entry or the buoying effect of high risk premia on stockholder wealth. The economy largely operates in the high-x region, where $r(x)$ is nearly flat and dynamics (μ_X, σ_X) are tame.

Term structure of risk prices.

Some recent papers have characterized the term structure of risk prices, arguing that it is an alternative and useful way to depict asset prices in dynamic models.[24] Assume that, consistent with this literature, any cash flow process $\{G_t\}$ can be constructed from the Markov state $\{X_t\}$ (participants' consumption share) as

[22] As $\phi \to 1$, return predictability is significantly muted. The reason for this is the region of negative risk prices due to hedging demands, as explained in the main text. Because of this feature, risk prices are procyclical in part of the state space, which muddies the relationship between price-dividend ratios and future returns.

[23] In other words, this is necessary in order for bond returns to have the same singular components as stock returns. One can also see the necessity of this condition by applying Itô's formula to $\xi_t q_{t,T}$ and examining the singular terms $-\xi_t q(s, X_t) dA_t^R + \xi_t \partial_x q(s, X_t) dA_t^{x^*}$. Since $dA_t^R = \frac{p'(x^*)}{p(x^*)} dA_t^{x^*}$, the singular terms vanish if and only if the boundary condition in the text holds.

[24] See Borovička et al. (2011), Borovička, Hansen, and Scheinkman (2014), and Borovička and Hansen (2016) for details on the methodology and interpretation.

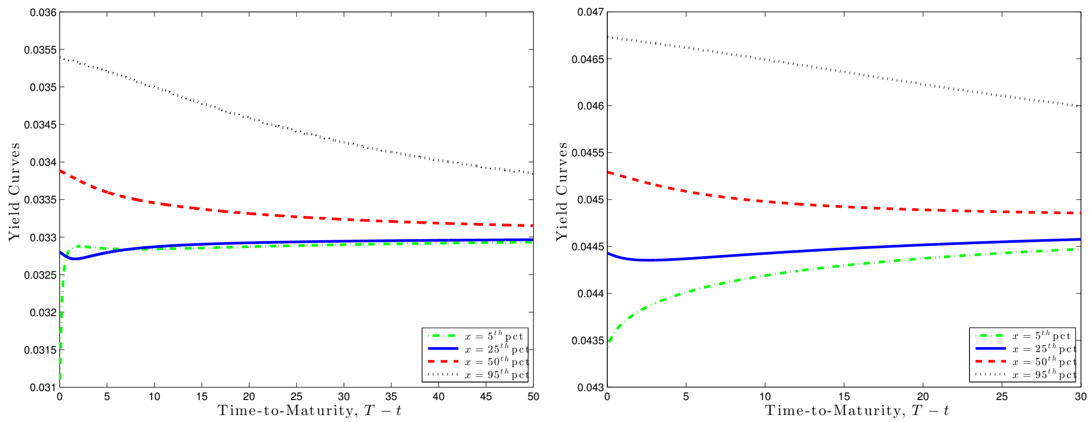

Figure 12: Yield curves for alternative initial states, $X_t = x$. Left panel uses parameters in table 1. Right panel uses log utility, $\gamma = \psi = 1$.

$d \log G_t = \mu_G(X_t)dt + \sigma_G(X_t)dZ_t$ and that this cash flow is prices by the SDF $\{\xi_t\}$ in (2). Given cash flow and stochastic discounting processes $\{G_t\}$ and $\{\xi_t\}$, the term structure of risk prices depicts an expected return for every maturity $T \geq t$, similar to the yield curve. The expected return is scaled to be in units of a Sharpe ratio (per unit of risk) or an elasticity. Because of this unitless feature, and an analogy to non-linear impulse response functions, the term structure is sometimes called a *shock-price elasticity*.

Formally, define the *shock-price elasticity* at time t for cash flow G_T to be

$$\varphi(T-t,x) := \lim_{s\to 0}\frac{1}{s}\log \mathbb{E}\Big[\Big(\frac{G_T}{G_t}\Big)H^s_{t,T} \mid X_t = x\Big] - \lim_{s\to 0}\frac{1}{s}\log\mathbb{E}\Big[\Big(\frac{\xi_T}{\xi_t}\Big)\Big(\frac{G_T}{G_t}\Big)H^s_{t,T} \mid X_t = x\Big], \qquad (90)$$

where $H^s_{t,T} := \exp\big([Z_{T\wedge(t+s)} - Z_t] - \frac{1}{2}[(T-t)\wedge s]\big)$, $s > 0$, is an exponential martingale that acts to perturb a process it multiplies near time t. To see this, let $T > t+s$ and look at

$$\log\Big(\frac{G_T}{G_t}\Big)H^s_{t,T} = \int_t^{t+s}\big[\mu_G(X_u)+1\big]du + \int_t^{t+s}\big[\sigma_G(X_u)+1\big]dZ_u + \int_{t+s}^T \mu_G(X_u)du + \int_{t+s}^T \sigma_G(X_u)dZ_u.$$

On the interval $[t, t+s]$, this process has perturbed exposure $\sigma_G(X_u)+1$ to the Brownian shock dZ_u. Thus, multiplying by $H^s_{t,T}$ for small s has the effect of increasing the exposure of a process to the underlying shock, which is akin to giving the system a larger shock at time t. Then the function $\varphi(s,x)$ traces out the effect of this instantaneous increase in shock exposure over a horizon s, much like an impulse response function. The fact that it is a unit impulse justifies the use of the term elasticity.

The shock-price elasticities are computed by applying Malliavin calculus, which is beyond the scope of this Appendix. The result is (see the papers cited above)

$$\varphi(T-t,x) = \eta(x) + \sigma_X(x)\frac{d}{dx}\Big(\log\mathbb{E}\Big[\Big(\frac{G_T}{G_t}\Big) \mid X_t = x\Big] - \log\mathbb{E}\Big[\Big(\frac{\xi_T}{\xi_t}\Big)\Big(\frac{G_T}{G_t}\Big) \mid X_t = x\Big]\Big). \qquad (91)$$

Note that the term in parentheses is the log expected return on G_T between t and T. This gives (91) a nice intuition. A unit increase in the risk exposure of G to the shock at time t increases the expected excess return to G_T through two channels. First, risk premia respond to an increase in exposure by the amount of the current risk price $\eta(X_t)$. Second, the risk that the state variable may now adjust more than before needs to be priced; the increase in state variable risk is given by $\sigma_X(X_t)$, and the adjustment to expected returns is given by the derivative term.

To compute each of the conditional expectations in (91) numerically, we solve a PDE very similar to the bond price PDE (89). The derivation of this is as follows. For $M = G$ or $M = \xi G$, define $\tilde{\varphi}(t, x) := \mathbb{E}[\frac{M_{t+s}}{M_s}\tilde{\varphi}(0, X_{t+s}) \mid X_s = x]$. Due to time-homogeneity of X, this definition is independent of the choice of s. Then, using the law of iterated expectations, followed by the definition of $\tilde{\varphi}$, we have $\tilde{\varphi}(t, x) = \mathbb{E}[\frac{M_u}{M_0}\mathbb{E}[\frac{M_t}{M_u}\tilde{\varphi}(0, X_t) \mid X_u] \mid X_0 = x] = \mathbb{E}[\frac{M_u}{M_0}\tilde{\varphi}(t-u, X_u) \mid X_0 = x]$. Hence, $\{M_t\tilde{\varphi}(T-t, X_t)\}_{t\in[0,T]}$ is a martingale and must have zero drift. Applying Itô's formula gives a PDE for $\tilde{\varphi}$ in $(T-t, x)$, i.e.,

$$0 = -\frac{\partial\tilde{\varphi}}{\partial s} + \Big(\mu_M + \frac{1}{2}\sigma_M^2\Big)\tilde{\varphi} + \big(\mu_X + \sigma_M\sigma_X\big)\frac{\partial\tilde{\varphi}}{\partial x} + \frac{1}{2}\sigma_X^2\frac{\partial^2\tilde{\varphi}}{\partial x^2},$$

which is solved subject to the boundary condition $\frac{\partial_x\tilde{\varphi}(s,x^*)}{\tilde{\varphi}(s,x^*)} = 0$ or $\frac{\partial_x\tilde{\varphi}(s,x^*)}{\tilde{\varphi}(s,x^*)} = \frac{p'(x^*)}{p(x^*)}$ (according to whether $M = G$ or $M = \xi G$, respectively) and the initial condition $\tilde{\varphi}(0, x) \equiv 1$, all exactly as in the bond price equation. The solution is $\varphi(T-t, x)$ due to the initial condition.

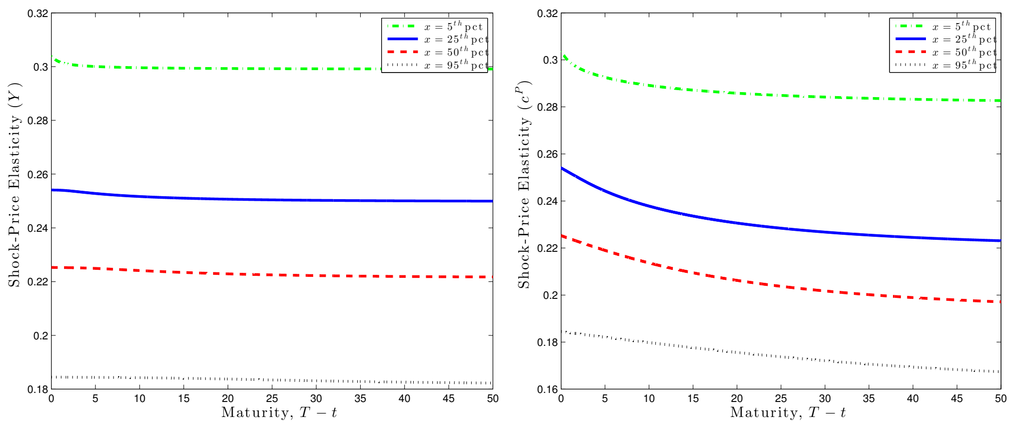

Figure 13: This figure shows risk price term structures for alternative initial states $X_t = x$. Left panel: the cash flow is $G = Y$, aggregate output. Right panel: the cash flow is $G = c^P$, participant consumption. Parameters are from table 1.

The left panel of figure 13 shows the shock-price elasticities for $G = Y$, the aggregate endowment, for the benchmark parameters given in table 1. As expected, the curves are ordered according to the initial condition $X_0 = x$. In low-x states, risk prices are high, and the resulting shock-price elasticities start out very high at low maturities. But strikingly, the shock-price elasticities do not decay much over time. This result, that shock-price elasticities are flat, is literally true in the model with log utility, as the following lemma shows.

Lemma B.4. *In the model with log utility, shock-price elasticities for $G = Y$ are constant over maturity $(T-t)$ for any initial state (x). In particular $\varphi(T-t, x; Y) = \eta(x)$ for all t and x.*

Proof of Lemma B.4. Let $G = Y$. Applying Itô's formula to $\log(\xi_t Y_t)$ and substituting $r(x)$ and $\eta(x)$ from Proposition 1, we obtain

$$\xi_t Y_t = \xi_0 Y_0 e^{-(\rho+\pi)t}\exp\Big\{-\frac{1}{2}\int_0^t \sigma_Y^2\Big(\frac{1-X_s}{X_s}\Big)^2 ds - \int_0^t \sigma_Y\Big(\frac{1-X_s}{X_s}\Big)dZ_s\Big\}.$$

The last exponential is a martingale and has expectation 1. Hence, $\mathbb{E}[\frac{\xi_t Y_t}{\xi_0 Y_0} \mid X_0 = x] = e^{-(\rho+\pi)t}$, which is independent of x. Similarly for $\mathbb{E}[\frac{Y_t}{Y_0} \mid X_0 = x]$. From (91), $\varphi(T-t, x) = \eta(x)$ for all t and x. □

This has two interpretations: (a) the term structure of risk prices is nearly flat, and approximately equal to the short-term risk price; or (b) the shock-price elasticities don't decay over time, implying shocks to risk exposure are perceived as permanent shocks by investors. Flat shock-price elasticities, consistently over multiple parameter configurations, seems to implicate the structure of the model. An open question is which features are key determinants of the term structure's shape, in models with financial frictions.

Why are the shock-price elasticities so flat in this model? There are two important reasons. First, the cash flow process $G = Y$ only experiences permanent shocks. When constructing shock-price elasticities for alternative cash flows that experience transitory shocks as well, we see more decay in the elasticities. For instance, shock-price elasticities associated with $G = c^P$, participants' consumption process, are depicted in the right panel of figure 13. This process naturally embeds some transitory dynamics: when x is low, participants' wealth and consumption will tend to be low, but high risky asset returns and/or entry imply that their consumption will be higher in the future.

Second, the stochastic discount factor lacks a significant transitory component. This could be seen in the yield curves in figure 12, which are very flat themselves. When the SDF has no transitory component at all, yield curves are literally flat, and when the transitory component is small, yield curves are nearly flat. As discussed in the text surrounding figure 12, this flatness can be attributed to the primary conclusions of this paper: the economy does not attain low-x states, in which asset price dynamics are extreme but will mean-revert.

Given these two reasons, we can infer that a model would be more likely to demonstrate quickly decaying shock-price elasticities under two conditions. First, a model needs to embed more cyclicality in cash flows, especially the type of cyclicality that arises endogenously in production, rather than endowment, economies. Second, an economy needs to spend more time in crisis states, in which extreme asset price dynamics are possible, but will ultimately be transitory.

B.3 Economy for asymptotically large entry costs

In this section, I consider what happens to the economy and the stationary distribution as entry costs become infinitely large, i.e., $\phi \to 1$. The main results are Lemmas B.5 and B.6.

As Lemma B.5 below shows, without entry, the economy has a singularity at $x = 0$. This is typical of limited participation models without entry. For example, Lemma B.5 shows that the market price of risk process η increases without bound as x approaches 0, which creates difficulties when attempting to show that a well-behaved equilibrium exists.

Lemma B.5. *Let $\{\phi_n\}_{n=1}^{\infty}$ be a sequence of entry costs such that $\phi_n \to 1$, and consequently $x_n^* \to 0$, as $n \to \infty$, and such that $x_n^* > 0$ for all n, assuming such a sequence exists. Assume that $(x_n^*)^2(1-\phi_n)^{\frac{\psi-1}{\psi}} \to 0$ as $n \to \infty$. Assume finally that $(x_n^*)^2 g_P''(x_n^*+) \to 0$ and $x_n^* g_N''(x_n^*+) \to 0$ as $x_n^* \to 0$. Then, for n large enough, the equilibrium is given approximately by*

$$\begin{aligned} x\eta(x) &= a_\eta + O(x) \\ xr(x) &= a_r + O(x) \\ x\mu_X(x) &= a_\mu + O(x) \\ x\sigma_X(x) &= O(x), \end{aligned}$$

where $a_\eta > 0$, $a_r < 0$, and $a_\mu > 0$ are constants depending only on parameters.

The important issue of long-run stationarity arises in this limiting economy. Although the OLG environment implies that the measure of participants vanishes asymptotically without entry, their wealth and consumption shares do not necessarily vanish. Intuitively, there are two forces: non-participants tend to replace participants in the birth-death process, while an individual participants' wealth expands faster than a non-participants'. The memoryless and independent death shocks ensure that these two forces offset in some sense. Heuristically,

$$\lim_{x\uparrow 1}\sigma_X(x) = 0 \quad \text{and} \quad \lim_{x\uparrow 1}\mu_X(x) < 0$$
$$\lim_{x\downarrow 0}\sigma_X(x) < +\infty \quad \text{and} \quad \lim_{x\downarrow 0}\mu_X(x) = +\infty.$$

The first two conditions are enough to ensure that X_t does not reach 1; the second two are essentially enough to ensure that X_t does not reach 0, and they hold under parameter restrictions given below. The reason for $\mu_X(0) = +\infty$ and $\sigma_X(0) < +\infty$ is that risky assets become infinitely attractive as participants' wealth share dwindles, as demonstrated in Lemma B.5. The question of stationarity is addressed formally in Lemma B.6, which gives an affirmative answer.

Lemma B.6. *Let the assumptions of Lemma B.5 hold. If, in addition, $\gamma > \frac{\psi}{\psi+1}$, then the limiting economy with $\phi \to 1$ is stationary in the sense that X_t never reaches 0 or 1, almost surely.*

Lemma B.6 implies that the complete segmentation model stays away from $x = 0$ with probability 1, i.e., entry never occurs. Without entry, this model is similar to the rest of the limited participation literature, and we may use it as a laboratory to understand how much non-participants would be willing to pay to participate in risky asset markets. This gives us a first look at the quantitative robustness of limited participation models. If non-participants are willing to pay very little to participate, small fixed costs may quantitatively justify assumptions of complete segmentation. If non-participants are willing to pay large amounts to participate, large fixed costs are needed to prevent entry. As we show in the analysis below, limited participation models are more consistent with the latter assertion that very large fixed costs are needed to deter entry.

Proof of Lemma B.5. Let ϕ and x^* be arbitrary members of the sequence $\{\phi_n, x_n^*\}_{n=1}^{\infty}$. Let $x \in [x^*, \kappa x^*]$ for some constant $\kappa > 1$. As we take $x^* \to 0$, we keep fixed the constant κ, so that $x \to 0$ as well. Under the stated assumption that such sequences exist such that $(x_n^*)^2(1-\phi_n)^{\frac{\psi-1}{\psi}} \to 0$ as $n \to \infty$, we then have

$$x^*(x - x^*)(1-\phi)^{\frac{\psi-1}{\psi}} = O(x^*).$$

Next, we approximate the equilibrium objects from Proposition 14 at $x = \kappa x^*$, for small x^*. To do this, first note that

$$\omega(x^*) = 1, \quad \omega'(x^*) = -(1-x^*)\frac{g_N''(x^*)}{g_N(x^*)}, \quad \zeta(x^*) = 0, \quad \zeta'(x^*) = (1-x^*)\frac{g_P''(x^*)}{g_P(x^*)},$$
$$x^*\frac{p(x^*)}{g_N(x^*)} = x^*\Big[x^*(1-\phi)^{\frac{\psi-1}{\psi}} + 1 - x^*\Big] + O(x^*) = O(x^*),$$
$$x^*\frac{p'(x^*)}{g_N(x^*)}(x - x^*) = x^*(x-x^*)\Big[(1-\phi)^{\frac{\psi-1}{\psi}} - 1\Big] + O(x^*) = O(x^*).$$

Then, we obtain

$$\begin{aligned}
\eta(\kappa x^*) &= \frac{\gamma\sigma_Y}{x^*} + (1-x^*)\Big[\Big(\frac{\gamma-\psi}{1-\psi}\Big)\sigma_Y\frac{g_P''(x^*)}{g_P(x^*)} - \frac{g_N''(x^*)}{g_N(x^*)}\Big](\kappa-1)x^* + \frac{O(x^*)}{x^*} + O(x^*)\\
r(\kappa x^*) &= \rho + \psi\mu_Y + \pi(1-\alpha) + \psi\pi + \frac{1}{2}(\kappa-1)\gamma(\psi+1)\sigma_Y^2 - \frac{1}{2}\frac{\gamma(\psi+1)\sigma_Y^2}{x^*} + \frac{O(x^*)}{x^*}\\
&\quad + \sigma_Y^2(1-x^*)\Big[\frac{\gamma(\psi+1)}{2}\frac{g_N''(x^*)}{g_N(x^*)} - \Big(\frac{\gamma-\psi}{1-\psi}\Big)\frac{g_P''(x^*)}{g_P(x^*)}\Big](\kappa-1)x^* + O(x^*)\\
\mu_X(\kappa x^*) &= -\sigma_Y^2 - \frac{\gamma(\psi+1)}{\psi}\sigma_Y^2\Big[\frac{\kappa}{2} + (\kappa-1)(1-x^*)^2(1-2x^*)\frac{g_N''(x^*)}{g_N(x^*)}\Big] + \frac{1}{2}\frac{\gamma(\psi+1)}{\psi}\frac{\sigma_Y^2}{x^*} + \frac{O(x^*)}{x^*}\\
&\quad + \sigma_Y^2\Big[\Big(\frac{\gamma-\psi}{\psi(1-\psi)}\Big)(1-x^*)^2\frac{g_P''(x^*)}{g_P(x^*)} - \Big(\frac{\gamma(\psi+1)}{\psi} - 1\Big)\Big(1 + (1-x^*)^2\frac{g_N''(x^*)}{g_N(x^*)}\Big)\Big](\kappa-1)x^* + O(x^*)\\
\sigma_X(\kappa x^*) &= \sigma_Y - \sigma_Y\Big[1 + (1-x^*)^2\frac{g_N''(x^*)}{g_N(x^*)}\Big](\kappa-1)x^* + O(x^*).
\end{aligned}$$

Under the assumptions in the statement of the lemma, $x^* g_N''(x^*) \to 0$ and $(x^*)^2 g_P''(x^*) \to 0$ as $x^* \to 0$, so we have

$$\begin{aligned}
x^*\eta(\kappa x^*) &= \gamma\sigma_Y + O(x^*)\\
x^* r(\kappa x^*) &= -\frac{1}{2}\gamma(\psi+1)\sigma_Y^2 + O(x^*)\\
x^*\mu_X(\kappa x^*) &= \frac{1}{2\psi}\gamma(\psi+1)\sigma_Y^2 + O(x^*)\\
x^*\sigma_X(\kappa x^*) &= O(x^*)
\end{aligned}$$

Since $\kappa > 1$ is arbitrary, and since x^* and x converge to zero together (recall that $x \in [x^*, \kappa x^*]$), we may replace both x^* and κx^* with x, which completes the proof. □

Proof of Lemma B.6. Let $(\underline{x}_n, \overline{x}_n)$ be a sequence of intervals converging to $(0,1)$ as $\phi_n \to 1$, in such a way that $\underline{x}_n > 0$ and $\overline{x}_n < 1$ for each n. Letting $T_n := \inf\{t : X_t \notin (\underline{x}_n, \overline{x}_n)\}$ and $T := \lim_{n\to\infty} T_n$, we want to show that $\mathbb{P}\{T = \infty\} = 1$. By Feller's theory of explosions, cf. Karatzas and Shreve (1991) section 5.5.C, it suffices to show that $v(0+) = v(1-) = +\infty$, where the function v is defined by

$$v(x) := \int_c^x \int_c^y \exp\Big(-2\int_z^y \frac{\mu_X(u)}{\sigma_X^2(u)}du\Big)\frac{1}{\sigma_X^2(z)}dzdy,$$

for some fixed $c \in (0,1)$.

For ϕ_n large enough, hence x_n^* small enough, and for x near enough to x_n^*, the analysis of Lemma B.5 shows that X_t evolves approximately with

$$\begin{aligned}
\mu_X(x) &= K + \frac{a_\mu}{x} + O(x)\\
\sigma_X(x) &= a_\sigma + O(x),
\end{aligned}$$

for constants K, $a_\mu := \frac{1}{2\psi}\gamma(\psi+1)\sigma_Y^2$, and $a_\sigma := \sigma_Y$. For x small enough, we can evaluate the integrals in the definition of v using these approximate formulas for μ_X and σ_X. We obtain:

$$
\begin{aligned}
-\int_z^y \frac{\mu_X(u)}{\sigma_X^2(u)}du &= -\int_z^y \frac{K + a_\mu u^{-1} + O(u)}{a_\sigma^2 + O(u)}du \\
&= -\int_z^y \frac{K}{a_\sigma^2}du - \int_z^y \frac{a_\mu}{a_\sigma^2 u}du - \int_z^y O(u)du \\
&= -\frac{K}{a_\sigma^2}(y-z) - \frac{a_\mu}{a_\sigma^2}(\log y - \log z) - O(y),
\end{aligned}
$$

$$
\begin{aligned}
\int_c^y \exp\Big(-2\int_z^y \frac{\mu_X(u)}{\sigma_X^2(u)}du\Big)\frac{1}{\sigma_X^2(z)}dz &= \int_c^y \exp\Big(-\frac{2a_\mu}{a_\sigma^2}\log(y/z) + O(y)\Big)\Big(\frac{1}{a_\sigma^2} + O(z)\Big)dz \\
&= \int_c^y (1 + O(y))\Big(\frac{y}{z}\Big)^{-2a_\mu/a_\sigma^2}\Big(\frac{1}{a_\sigma^2} + O(z)\Big)dz \\
&= y^{-2a_\mu/a_\sigma^2}(a_\sigma^2 + 2a_\mu)^{-1}\Big(y^{1+2a_\mu/a_\sigma^2} - c^{1+2a_\mu/a_\sigma^2}\Big) + O(y),
\end{aligned}
$$

and

$$
\begin{aligned}
v(x) &= \int_c^x y^{-2a_\mu/a_\sigma^2}(a_\sigma^2 + 2a_\mu)^{-1}\Big(y^{1+2a_\mu/a_\sigma^2} - c^{1+2a_\mu/a_\sigma^2}\Big)dy + O(x) \\
&= -(a_\sigma^2 + 2a_\mu)^{-1}c^{1+2a_\mu/a_\sigma^2}\int_c^x y^{-2a_\mu/a_\sigma^2}dy + O(x) \\
&= -(a_\sigma^2 + 2a_\mu)^{-1}c^{1+2a_\mu/a_\sigma^2}\Big(1 - \frac{2a_\mu}{a_\sigma^2}\Big)^{-1}\Big(x^{1-2a_\mu/a_\sigma^2} - c^{1-2a_\mu/a_\sigma^2}\Big) + O(x) \\
&= O(1) - (a_\sigma^2 + 2a_\mu)^{-1}\Big(1 - \frac{2a_\mu}{a_\sigma^2}\Big)^{-1}c^{1+2a_\mu/a_\sigma^2}x^{1-2a_\mu/a_\sigma^2}.
\end{aligned}
$$

If $\gamma > \frac{\psi}{\psi+1}$, then

$$
\operatorname{sgn}\Big(1 - \frac{2a_\mu}{a_\sigma^2}\Big) = \operatorname{sgn}(\psi - \gamma(\psi+1)) < 0.
$$

As a result, $v(x) \to +\infty$ as $x \to 0$.

Near 1, $\sigma_X(x) = O(1-x)$ while $\mu_X(x) = -\pi(1-\alpha)\frac{p(1-)}{g_N(1-)} + O(1-x) < 0$. Using these facts, it is easily verified that $v(1-) = +\infty$. First,

$$
\exp\Big(-2\int_z^y \frac{\mu_X(u)}{\sigma_X^2(u)}du\Big) \geq 1,
$$

and so

$$
\begin{aligned}
v(x) &\geq \int_c^x \int_c^y \frac{1}{\sigma_X^2(z)}dzdy \\
&\geq \int_c^x \int_c^y \frac{1}{(1-z)^2}dzdy \\
&= \int_c^x \Big(\frac{1}{1-y} - \frac{1}{1-c}\Big)dy \\
&= -\log(1-x) + \log(1-c) - \frac{x-c}{1-c},
\end{aligned}
$$

which approaches $+\infty$ as $x \to 1$. $\square$

B.4 Hitting times

Besides the stationary density, we may wish to characterize state transitions. In the following lemma, we make partial progress by calculating expected hitting times, beginning from the entry point $x = x^*$ and from the full participation point $x = 1$.

Lemma B.7. *Define the mean hitting times*

$$m(x; x^*) := \mathbb{E}[\inf\{t : X_t = x\} \mid X_0 = x^*]$$

and

$$m(x; 1) := \mathbb{E}[\inf\{t : X_t = x\} \mid X_0 = 1].$$

Viewed as a function of x, either function m solves the ordinary differential equation

$$0 = 1 - \mu_X m' - \frac{1}{2}\sigma_X^2 m''. \tag{92}$$

The boundary conditions for $m(x; x^*)$ are $m(x^*; x^*) = 0$ and $m'(x^*; x^*) = 0$. The boundary conditions for $m(x; 1)$ are $m(1; 1) = 0$ and $m'(1; 1) = 1/\mu_X(1)$.

In particular, we may be interested in the expected amount of time it takes to reach a "crisis" from a typical point. Since bad times in this model correspond to a series of negative shocks and low values of x (this is because $\sigma_X \geq 0$), we might want to compute $m(x^*; \overline{x})$, the time to reach x^* from the stationary mean $\overline{x} := \mathbb{E}X_t$. Given $m(x; 1)$, this is straightforward, as equation (93) below implies that $m(x^*; \overline{x}) = m(x^*; 1) - m(\overline{x}; 1)$. More generally, given the functions $m(x; 1)$ and $m(x; x^*)$ satisfying (92), we are able to similarly compute $m(x'; x)$ for any x', x. Figure 14 plots $m(x; \overline{x})$ for the parameters in table 1 and shows that it takes a very long time to for entry to occur.

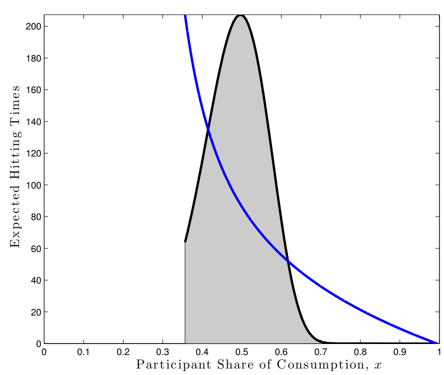

Figure 14: Expected hitting times $m(x; 1)$ (blue line) and the stationary density of X (gray area).

Proof of Lemma B.7. For every $(z, y) \in [x^*, 1]$, define the expected hitting time $m(z; y) := \mathbb{E}^y[\tau_z]$, where $\tau_z := \inf\{t \geq 0 : X_t = z\}$. Viewed as a function of y, $m(z; y)$ satisfies the following ordinary differential equation, by Feynman-Kac:

$$0 = 1 + \mu_X(y)\frac{d}{dy}m(z; y) + \frac{1}{2}\sigma_X^2(y)\frac{d^2}{dy^2}m(z; y).$$

Next, note that since X is continuous, it must pass through every point x in between y and z prior to reaching z. Thus, the following identity holds for

$$m(z;y) = m(x;y) + m(z;x). \tag{93}$$

Finally, setting $y = x^*$, equation (93) implies $m(x;x^*) = m(z;x^*) - m(z;x)$. Taking derivatives, $m'(x;x^*) = -\frac{d}{dx}m(z;x)$ and $m''(x;x^*) = -\frac{d^2}{dx^2}m(z;x)$ for any z. Plugging these derivatives into the ODE derived above, we arrive at (92). The boundary conditions are $m(x^*;x^*) = 0$ and $m'(x^*;x^*) = 0$, the first because it takes zero time to reach x^* from x^* and the latter because x^* is a reflecting barrier. Since z and y are not ordered, the exact same arguments hold for $m(x;1)$, so that (92) holds there as well. The boundary condition for $m'(1;1)$ is derived by taking the limit $x \to 1$ in (92) and noting that $\sigma_X(1) = 0$. □

B.5 Extensions

Although the model presented here is quite simple, straightforward arguments illustrate the effects that several extensions would have on risk concentration and asset price dynamics. I briefly discuss labor income; other fixed costs; heterogeneous risk aversion; heterogeneous EIS; hyperbolic discounting; and ambiguity aversion.

Labor income.

In this paper, all income is capital income. In reality, approximately two-thirds of income is labor income. What would happen to asset price dynamics in my model if agents also receive labor income endowments?

First note that, if human capital is tradable and perfectly correlated with physical capital, the setup I consider is essentially without loss of generality. Indeed, consider a setup where all agents are born with no financial wealth but with a human capital "tree" that they immediately "sell" at market prices. If the size of the tree is $1 - \alpha$, then newborns begin life with $(1 - \alpha)P_t$ units of financial wealth, exactly as in the present model. Since entry costs are proportional to wealth, the size of the human capital tree is also irrelevant.[25]

If human capital is non-tradable, but still perfectly correlated with physical capital, the model's conclusions are still qualitatively unchanged. Indeed, asset prices would be similar to a model in which participants can issue some equity to non-participants (an extension discussed in the text). Young non-participants would have little financial wealth but large claims to future inflows from their wages. The stock-like feature of wages (perfect correlation of labor and capital) means that non-participants essentially hold some risky assets, much like a model with equity issuance. Although this reduces non-participants' incentives to enter financial markets, generating more financial market segmentation with a smaller entry cost ϕ, the qualitative message of the paper is unchanged. See Appendix A.7, which demonstrates that partial equity issuance does not alter the main conclusions here.

However, wages are also not perfectly correlated with capital market returns. Indeed, the results of Polkovnichenko (2004) suggest that the addition of labor income attenuates the asset pricing results of typical limited participation models (without entry). Empirically, since labor income has low volatility and relatively low correlation with capital income, a realistically calibrated limited participation model generates much lower consumption growth volatility for market participants with the inclusion of labor income than

[25] In this case, the only discussion becomes about the calibration of α. I set $\alpha = 0.5$ in table 1 to incorporate some notion of labor income. Indeed, under the earlier life insurance interpretation, α measures the degree of insurance, and rational agents should set $\alpha = 1$. Under the labor income interpretation, α measures the share of capital income in total income, which is approximately one-third in the data. I split the difference by setting $\alpha = 0.5$.

without it. Consequently, one of two things must occur: either risk prices must be lower in such a model than one with only financial wealth; or participation incentives are too high to justify large risk concentration.

Other fixed cost specifications.

In the context of stock market non-participation, we might want a fixed entry cost independent of wealth, in order to be consistent with two of the most salient features of household finance data: (1) wealthier households tend to participate and poorer households do not; (2) participation tends to be procyclical (see, e.g., Campbell (2006)). Consider a non-pecuniary cost of the form

$$\hat{\Phi}_t = [1 - (1-\phi)^{1-\gamma}] G_t^P \frac{Y_t^{1-\gamma}}{1-\gamma}$$

For a household with average wealth $W_{t,b} = P_t$, the cost $\hat{\Phi}_t$ is perceived as paying a fraction ϕp_t^{-1} of their wealth. Households with above average wealth perceive this as paying a smaller fraction, and vice versa for households with below average wealth. One might conjecture that, in equilibrium, wealthy households would be more likely to participate in financial markets, consistent with the household finance data. In addition, if the price-dividend ratio p_t is procyclical, households will perceive the fixed cost as a smaller fraction of their wealth in good times, leading to procyclical participation.

That said, the asset pricing results in this paper are robust and could even be made worse under such an entry cost. Indeed, for a large enough ϕ, the cost $\hat{\Phi}_t$ could induce a greater amount of wealth to participate in stock markets than does the cost $\Phi_{t,b}$ in (6), because of the additional selection effect.[26] Through asset market clearing, this alleviates the risk concentration channel even further, attenuating asset price dynamics. Limited participation could be even less likely to explain asset prices under more realistic costs like $\hat{\Phi}_t$.

Heterogeneous risk aversion.

Gârleanu and Panageas (2015) – which features OLG, recursive preferences, and heterogeneity in γ – showed that heterogeneous risk aversion can potentially help resolve multiple aggregate asset price puzzles. For example, with negative shocks, less risk-averse agents (who are levered) lose wealth faster than more risk-averse agents. As they liquidate some of their risky asset position, more risk-averse agents have to buy, which generates amplification in the risk price dynamics.

With limited participation and an entry mechanism, their results are not likely carry over to this model. Indeed, the dominant effect of risk aversion heterogeneity is for risk-tolerant agents to select into financial markets, while risk-averse agents stay out. First, this selection depresses risk prices on average. Second, because the more risk-averse agents choose not to participate, they do not buy the liquidated positions of risk-tolerant agents in bad times, shutting down any amplification of risk price dynamics.

To formalize this conjecture, I consider a single agent whose risk aversion is $\hat{\gamma} \neq \gamma$. Define $e(x)$ to be non-participants' willingness-to-pay function, which says how much wealth non-participants are willing to give up in order to participate forever after in risky asset markets. The function $e(x)$ solves the equation $V^P(1-e(x), x) = V^N(1, x)$. By computing $e(x; \hat{\gamma})$ and $e(x; \gamma)$ for non-participants with risk aversions $\hat{\gamma}$ and

[26] For small ϕ, the proportional cost $\Phi_{t,b}$ in (6) induces full participation, as in Proposition 6, while the fixed cost $\hat{\Phi}_t$ prevents relatively poor agents (e.g., newborns) from participating. Thus, the fraction of wealth participating is lower for the fixed cost when ϕ is small. On the other hand, when ϕ approaches 1, the proportional cost $\Phi_{t,b}$ prevents any entry at all, while the fixed cost $\hat{\Phi}_t$ still admits entry to agents who are significantly weather than the average. For example, if an agent is four times wealthier than average, $\phi \approx 1$ is perceived as a 25% of wealth entry cost under the $\hat{\Phi}_t$ specification. Therefore, I conjecture there is some intermediate level $\hat{\phi}$ such that $\Phi_{t,b}$ and $\hat{\Phi}_t$ generate the same equilibrium amount of participant wealth, with less risk concentration occurring under $\hat{\Phi}_t$ for all $\phi \geq \hat{\phi}$.

γ, respectively, we can understand how much stronger the entry incentives are for less risk-averse agents. The result of this analysis, in figure 15, shows that $e(x;\hat{\gamma}) > e(x;\gamma)$ for $\hat{\gamma} < \gamma$, and vice versa for $\hat{\gamma} > \gamma$, as conjectured. Simply, for $\hat{\gamma} < \gamma$, the $\hat{\gamma}$-agent finds risk prices very attractive, as they are an equilibrium outcome from an economy full of γ-agents.

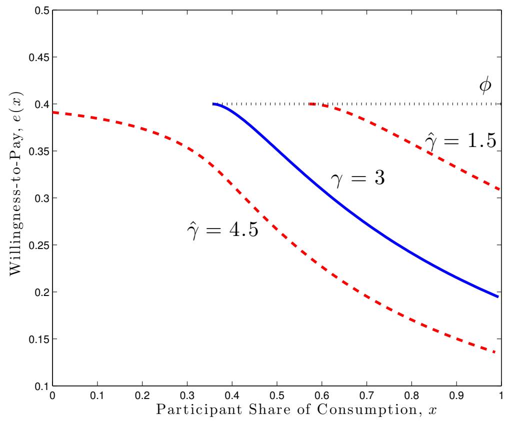

Figure 15: Functions $e(x;\hat{\gamma})$ and $e(x;\gamma)$ are entry willingnesses-to-pay, as a fraction of wealth, for two agents with risk aversions $\hat{\gamma}$ and γ, living in an economy populated by γ-agents.

Heterogeneous elasticity of intertemporal substitution.

Guvenen (2009) builds a model with exogenous limited participation in which the EIS of participants exceeds the EIS of non-participants and shows how these two features interact to generate large risk prices and risk price dynamics. Bondholders exogenously have a low EIS, meaning they want to save in good times and borrow in bad times. The fact that stockholders have a larger EIS means they are willing to tolerate this resulting larger consumption volatility over time, which amplifies risk prices and their dynamics.

With endogenous participation through entry, this channel survives only if the high EIS agents are more willing to participate. Figure 16 repeats the analysis from the figure 15 by constructing willingness-to-pay functions $e(x;\psi)$ and $e(x;\hat{\psi})$ for $\hat{\psi} \neq \psi$. Indeed, the higher EIS agents prefers to enter risky asset markets sooner than the lower EIS agents.

An agent with high EIS (low $\hat{\psi}$) living in an economy full of low EIS individuals finds the volatility of risky assets to be low, and she is willing to take a levered position in such an asset to achieve a tolerable level of consumption growth volatility. This force induces earlier entry by the $\hat{\psi}$-agent. This analysis suggests that heterogeneity in EIS is a promising ingredient in limited participation models, even with entry.

Hyperbolic discounting.

Hyperbolic discounting, under which agents display present bias and procrastinating behavior, could mitigate entry incentives and thus increase risk concentration. The idea is that agents excessively discount the large lifetime benefit from holding risky assets that pay substantial premia, so they are less likely to enter. In order to get an cursory understanding of how hyperbolic discounting might affect the results, consider using $\rho = 0.3$ as opposed to 0.01, keeping all other parameters fixed as in table 1. With this adjustment, the estimate of ϕ^* in (34) becomes 1% of wealth, rather than 10% of wealth. Since hyperbolic discounting has a modest effect on risk attitudes, we don't expect such a model to imply very different risk prices (e.g., Luttmer and Mariotti (2003)). If, in addition, we hold the interest rate fixed, this back-of-the-envelope calculation

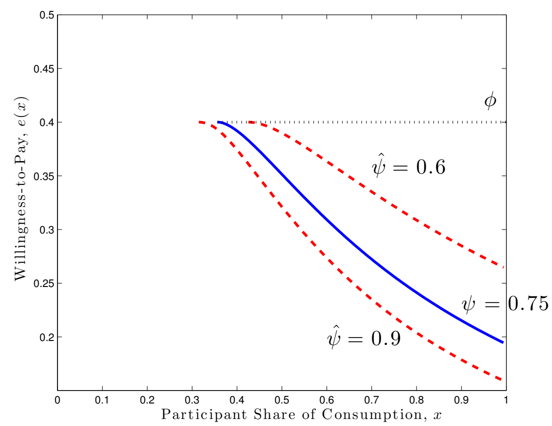

Figure 16: Functions $e(x;\hat{\psi})$ and $e(x;\psi)$ are entry willingnesses-to-pay, as a fraction of wealth, for two agents with EIS $\hat{\psi}^{-1}$ and ψ^{-1}, living in an economy populated by ψ-agents.

suggests that hyperbolic discounting can dramatically lower participation incentives and thus induce more risk concentration.

Ambiguity aversion.

Even without entry costs, ambiguity aversion can lead to limited participation. For example, in Dow and Werlang (1992), agents endogenously hesitate to trade because they are unsure about the probability distribution of returns. Thus, we expect such preferences to generate significant risk concentration and high risk premia, even with small entry costs. On the other hand, since less ambiguity-averse investors will tend to select into participation, risk premia can be lower than a full-participation economy (e.g., Cao, Wang, and Zhang (2005)). Finally, an exogenous decrease in ambiguity, which one might reasonably conjecture to be a characteristic of good times, tends to increase participation (e.g., Epstein and Schneider (2007)). Consequently, ambiguity aversion can also be consistent with procyclical entry, as in a model with extrapolative expectations in section 5. This link to extrapolation may not be surprising, given the results of Bhandari, Borovička, and Ho (2016), who show that time-varying concerns for robustness can look like extrapolative expectations in surveys. See also Fuster, Hebert, and Laibson (2011), who show that some model misspecification about a hump-shaped time series can result in extrapolative-looking expectations.