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

Paymon Khorrami
Imperial College London†

October 28, 2021

Abstract

Risk concentration, due to slow-moving capital, is a prominent explanation for crisis dynamics of asset prices and macroeconomic quantities. Is this realistic? By considering costly entry in canonical limited participation and intermediary-based models, 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. Extrapolative beliefs resolves this tension: even with small participation costs, endogenous pessimism delays entry, enabling large and volatile equilibrium risk premia.

JEL Codes: D14, G11, G12.

Keywords: Limited Participation, Intermediary Asset Pricing, Entry, Extrapolative Beliefs.

*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, Alex Michaelides, 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.

†Imperial College Business School Department of Finance. Contact: p.khorrami@imperial.ac.uk

As a result of various financial frictions, aggregate risk may be concentrated on the balance sheets of a few specialized market participants. A large literature, growing rapidly after the 2008 financial crisis, has argued that such *risk concentration* can help explain crisis dynamics in macro and asset pricing. The risk concentration channel is well-captured by a standard limited participation economy: exogenously designated "participants" hold risky assets, while "non-participants" do not. In equilibrium, risky assets are 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 potentially powerful—it amplifies both the level and variability of risk premia—but is it realistic?

This paper offers a critique of the risk concentration mechanism, central to intermediary asset pricing and other literatures, by arguing the mechanism relies on implausibly costly financial frictions. To understand how I arrive at this conclusion, consider the fact that the decision to participate and hold risky assets is endogenous. If non-participants become increasingly willing to participate as risk becomes more concentrated and risk premia become more attractive, entry endogenously limits risk concentration and subdues risk price dynamics. The size of entry barriers thus determines the size of the risk concentration mechanism. Inverting this logic, the behavior of equilibrium asset prices reveals the size of entry barriers.

I study a limited participation model to uncover this mapping between asset prices and entry barriers. The key assumption: non-participants may pay a one-time cost to begin trading in risky asset markets forever after. In allowing this, I relax the notion of rigid investor "types" (e.g., "participants" or "experts" or "banks") that pervades the literature. Agents in my model are ex-ante identical, and ex-post heterogeneity arises through entry.

The main results of the paper are as follows. In the baseline results, I argue that expected return levels and dynamics are inconsistent with small entry costs and are likely associated with large barriers. Next, this class of models embeds a fundamental trade-off between risk premia levels and dynamics, such that augmenting the baseline model often improves asset pricing results in one dimension but worsens them in another. Finally, I argue extrapolative beliefs sidestep this critique, because potential entrants are endogenously pessimistic when risk premia are high.

The claim that asset prices are inconsistent with small entry costs comes from the following two results. 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, through various exercises, I find that entry costs need to be on the order of 90% of wealth to induce enough segmentation and hence empirically realistic asset prices.

The basic intuition for these 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 excess 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.

Instead, if limited participation is the mechanism generating large and volatile risk premia, 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 $1 - \frac{1/(2\%+5\%-2\%)}{1/(2\%+0.5\%-2\%)} = 90\%$ drop in aggregate wealth, consistent with the enormous implied entry costs from my model. One can think of my baseline results as formalizing these back-of-the-envelope calculations in a fully nonlinear general equilibrium environment.

Can certain auxiliary features can raise both the level and variability of risk premia, such that large entry barriers are no longer needed? For many standard extensions, no. Through a series of robustness exercises, I illustrate a trade-off in endogenous participation models between the level and variability of risk premia, a trade-off which uniquely arises when entry is endogenous. Intuitively, auxiliary features that raise unconditional risk premia tend to raise participation incentives, which mitigates risk concentration. For example, a calibration with higher risk aversion increases unconditional risk premia and thereby attenuates risk premia dynamics. I find a similar level-variability tension when introducing partial equity-issuance (by participants) and idiosyncratic risk (in returns). The level-variability tension is problematic for limited participation models, which are conventionally thought to raise *both* levels and variability of risk premia.

I conclude by showing that a limited participation economy augmented with *extrapolative beliefs* is immune to this level-variability tension. Investors who extrapolate recent past returns believe risk premia are high when they are actually low, and vice versa. As a result, in bad times, non-participants will have become endogenously pessimistic and will not want to enter risky asset markets, even with moderate entry costs. The absence of entry in bad times is precisely what allows for more extreme risk price dynamics. I show a type of "additive property" in this setting: entry frictions mainly affect risk price levels, whereas extrapolative beliefs mainly affect risk price dynamics, with no offsetting impact on each other. The other extensions explored in this paper are "sub-additive": boosting risk price levels dampens risk price dynamics, due to entry.

Quantitatively, with moderate entry costs and reasonable extrapolation parameters, the model delivers realistic magnitudes in levels and time-series variation of risk prices (as well as low and stable interest rates, and amplified return volatility). This section contributes to the literature methodologically by fully solving a model with recursive preferences, extrapolative beliefs, and limited participation with endogenous entry, which is a challenging problem.

The framework analyzed in this paper is most similar to the restricted participation model of Basak and Cuoco (1998) but with an endogenous entry margin. One key assumption I make, to sharpen the theoretical analysis, 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. Even in the case of general recursive utility, one can obtain many sharp theoretical results. That said, the quantitative results are shown to be robust to alternative entry cost formulations (in particular, I also study a fixed cost formulation, in which the entire distribution of wealth becomes an aggregate state variable). The results are also robust to including non-tradable labor income, which similarly introduces a non-homogeneity in equilibrium.

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).[1] The common thread is that arbitrageur or intermediary wealth matters for equilibrium dynamics, in a way that is very similar to the canonical limited participation model (Basak and Cuoco, 1998). The same mechanisms are central to modern macroeconomics.[2] My paper provides a transparent way to understand how large financial frictions must be in these markets to generate significant risk concentration.

In these literatures, a risk concentration channel emerges through assumed heterogeneity in investment opportunity sets or ex-ante differences in investor's types. On the other end of the spectrum, Haddad (2014) features ex-ante identical investors and focuses on an

[1] For models, see Basak and Croitoru (2000), Basak and Shapiro (2001), Kyle and Xiong (2001), Gromb and Vayanos (2002), Kondor (2009), Gârleanu and Pedersen (2011), He and Krishnamurthy (2012, 2013). For empirical evidence, see Adrian and Shin (2010), Adrian and Shin (2013), Adrian, Etula, and Muir (2014), He, Kelly, and Manela (2017), Muir (2017), Siriwardane (2019).

[2] Beginning from the seminal "financial accelerator" papers of Kiyotaki and Moore (1997) and Bernanke, Gertler, and Gilchrist (1999), there has been explosive growth in the literature on macro dynamics with financial frictions. See Gertler and Karadi (2011), Gertler and Kiyotaki (2010), Brunnermeier and Sannikov (2014), Mendoza (2010), Bianchi (2011), Gertler and Kiyotaki (2015), Christiano, Motto, and Rostagno (2014), Adrian and Boyarchenko (2012), Di Tella (2017), He and Krishnamurthy (2019).

equilibrium with free entry. 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 the literature on stock market non-participation.[3] However, applying the limited participation model to households and stock markets is somewhat problematic, as risk concentration appears modest in that context.[4] For this reason, recent literature has more often applied the mechanisms discussed in this paper to intermediated markets, a setting in which higher barriers to entry are not unreasonable.[5]

In the critical extension with extrapolative beliefs, 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.[6] 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, but at odds with reality.[7]

More recently, Krishnamurthy and Li (2020) and Maxted (2020) have embedded diagnostic beliefs (one microfoundation for extrapolation) into a similar intermediary-based model. Extrapolative intermediaries help align the models with empirical evidence that credit booms predict financial crises. My paper uses extrapolation to tackle the different but related question of how significant risk price dynamics can arise without extreme financial frictions.

The paper is organized as follows. Section 1 describes a limited participation model with entry. Section 2 argues the baseline risk concentration channel implies very large entry costs. Section 3 illustrates, through several model extensions, a tension between the level and variability of risk premia with endogenous entry. Section 4 proposes extrapolative beliefs as a remedy to this critique. Section 5 concludes. Proofs and extensions are in the appendices.

---

[3] Beginning with Mankiw and Zeldes (1991), the literature noticed only a subset of households hold stocks, and these investors' consumption is more volatile, and covaries more with stock returns, than non-stockholders' (Parker and Vissing-Jørgensen, 2009). Guvenen (2009) builds a quantitative model incorporating limited stock market participation (without entry), and Malloy, Moskowitz, and Vissing-Jørgensen (2009) provide direct tests of stockholders' Euler equations under recursive preferences.

[4] 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). Consequently, Gomes and Michaelides (2008) find that small fixed entry costs help match stock market non-participation, but have an insignificant effect on asset prices.

[5] For example, intermediaries and arbitrageurs are often highly leveraged and risk is thus highly concentrated. Separately, it is well-understood that both moral hazard and adverse selection issues contribute to high costs of intermediary equity issuances; these costly issuances play the same role as costly entry, in the sense that they affect risk-sharing. These ideas suggest higher entry costs should be entertained in the intermediary context than the household context.

[6] De Long et al. (1990) and Cutler, Poterba, and Summers (1991) are early models of price extrapolation. Section 4 for many more references to models with extrapolation.

[7] See Greenwood and Shleifer (2014); Vissing-Jorgensen (2003); Bacchetta, Mertens, and Van Wincoop (2009); Amromin and Sharpe (2014); Bordalo et al. (2020) for survey evidence.

1 Model

In this economy, time is continuous and spans the infinite past and future, $t \in \mathbb{R}$. 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 instantaneous return $r_t dt$. The model will always have a positive measure of agents trading in dynamically complete markets, so we may define a 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_t is the risk price. To ensure absence of arbitrage, it must be that[8]

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

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:

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

Later, I extend (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 $\mathcal{N}_t =$

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

$\mathcal{P}_t^c$ 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.

In the context of intermediary asset pricing, the participants are appropriately regarded as the shareholders of the intermediaries. The risk preferences of these agents are what determines intermediaries' financial positions (He and Krishnamurthy, 2012, 2013). This is why I have modeled participants just like any other agent, i.e., with a utility function over consumption.

For notational simplicity, suppose all members of a cohort are either participants or non-participants, so that there is no heterogeneity in decisions among members in the same cohort. If $b \in \mathcal{P}_t$, this means that members of cohort b are participating in risky asset markets, and conversely for $b \in \mathcal{N}_t$.[9]

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

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

where $\theta_{t,b}$ is the fraction of wealth invested in the risky asset. 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

[9] 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.

totals $\alpha \int_{-\infty}^{t} \pi e^{-\pi(t-b)} W_{t,b} db$, since dying agents are a 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. Specifically, dying agents' risky asset shares are liquidated, any loans repaid, and the balance is converted into riskless bonds, which newborn non-participants inherit as their endowment. This model only features financial wealth, but we show below that these unintended bequests are equivalent to human wealth embodied in newborn generations, as long as labor income is pledgeable and perfectly correlated with capital income. We revisit this issue in Section 2.5.

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 entry. These assumptions are unlike papers on stock market participation, e.g., Gomes and Michaelides (2008), which typically have non-homogeneous entry costs to address the wealth-participation gradient. The homogeneous specification adds tractability, but we revisit this issue in Section 2.5.

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 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

$$V_{t,b}^P = (\rho + \pi)^{-1}[\log(W_{t,b}) + G_t^P].$$

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

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

$$V_{t,b}^N = (\rho + \pi)^{-1}[\log(W_{t,b}) + G_t^N],$$

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

Non-participants compare the current payoff, $V_{t,b}^N$, against the best possible future entry payoff, $V_{t,b}^P - \Phi$, to decide when to enter. Non-participants enter when the latter dominates, or

$$\tau_b = \inf\left\{t \geq b : \log(1-\phi) + G_t^P \geq G_t^N\right\}, \tag{9}$$

which is independent of wealth and cohort b. Thus, all agents have identical entry incentives.

2 Equilibrium

An equilibrium is a set of price and allocation processes such that agents maximize utility and all markets clear. The entry decisions merit some discussion. 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. Thus, the identity of entrants is not uniquely determined in equilibrium, i.e., $\mathcal{P}_t$ is not uniquely determined. Instead, at any time point of entry, $t \in \mathcal{T}^*$, all non-participants must be indifferent between inaction and entry. As a result, entry incentives can be written:

$$\text{at an entry time } t: \ \log(1-\phi) + G_t^P = G_t^N; \tag{10}$$

$$\text{at times } t \text{ without entry}: \ \log(1-\phi) + G_t^P < G_t^N. \tag{11}$$

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.

$$\text{[Goods market]} \qquad Y_t = \int_{-\infty}^{t} \pi e^{-\pi(t-b)} c_{t,b} db \tag{12}$$

$$\text{[Stock market]} \qquad P_t = \int_{-\infty}^{t} \pi e^{-\pi(t-b)} \theta_{t,b} W_{t,b} db \tag{13}$$

$$\text{[Bond market]} \qquad 0 = \int_{-\infty}^{t} \pi e^{-\pi(t-b)} (1-\theta_{t,b}) W_{t,b} db \tag{14}$$

$$\text{[Newborn transfers]} \qquad \pi W_{t,t} = \pi(1-\alpha) \int_{-\infty}^{t} \pi e^{-\pi(t-b)} W_{t,b} db. \tag{15}$$

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

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

which represents the consumption share of the participants. This single endogenous variable is sufficient to characterize the entire equilibrium. In Markov equilibrium, we may assume entry occurs when $X_t \leq x^*$ for some point $x^* \in [0,1]$.[10] In this case, the set of entry times is

$$\mathcal{T}^* = \{t : X_t \leq x^*\}. \tag{17}$$

Thus, conjecture the following dynamics for X,

$$dX_t = \mu_X(X_t)dt + \sigma_X(X_t)dZ_t + dA_t^{x^*}, \tag{18}$$

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

Definition 1. *A* stationary Markov equilibrium *in X_t, defined in (16)-(18), 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*

[10] This is without loss of generality in a stationary Markov equilibrium. The diffusive part of X—(μ_X, σ_X), shown in Proposition 1 below—is "regular" in the following sense: without any entry, X would visit all states in $(0,1)$ in finite time, a.s. 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_t, it eventually exceeds $\overline{x}^*$, so we may take $x^* = \overline{x}^*$ as our entry point.

such that markets clear as in (12)-(15). Value processes are characterized by the functions

$$V^P(W_{t,b}, X_t) := \frac{\log(W_{t,b})}{\rho + \pi} + g_P(X_t) \quad \textit{and} \quad V^N(W_{t,b}, X_t) := \frac{\log(W_{t,b})}{\rho + \pi} + g_N(X_t). \tag{19}$$

Asset prices are characterized by $\eta_t = \eta(X_t)$, $r_t = r(X_t)$, $\sigma_{R,t} = \sigma_R(X_t)$, and $\mu_{R,t} = \mu_R(X_t)$.

2.1 Analysis of equilibrium

Assume that the equilibrium entry point x^* is given. With log utility, consumption and portfolio decisions are proportional to wealth and independent of the value functions g_P and g_N. Consequently, all asset prices and state dynamics from Definition 1 can be determined in closed form, given x^*. 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.1.

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*

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

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

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

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

$$h(x) \propto \Big(\frac{x}{1-x}\Big)^2 (1-x)^{-\frac{2\pi(1-\alpha)}{\sigma_Y^2}} \exp\Big(-\frac{2\pi(1-\alpha)}{\sigma_Y^2(1-x)}\Big), \quad x \in [x^*, 1). \tag{20}$$

Proposition 1 illustrates several key features of the model. First, we see the generic properties of limited participation: $X_t < 1$ raises the market Sharpe ratio η_t and lowers the market interest rate r_t. With limited participation, risk-bearers are levered, so they require lower borrowing costs and higher expected returns, compared to a similar economy without leverage. This is the so-called risk concentration channel referenced in the introduction.

Second, as long as there is entry such that $x^* > 0$, the economy is well-behaved in the sense that an equivalent martingale measure exists. Without entry, the equilibrium tends to explode in bad times, i.e., $r(x^*) \to -\infty$ and $\eta(x^*) \to +\infty$ as $x^* \to 0$. This can lead limited participation equilibria to have bubbles and arbitrage opportunities, as shown by

Hugonnier (2012). The same critique applies to intermediary asset pricing models such as He and Krishnamurthy (2012, 2013). Conversely, with *any amount of entry*, this explosion is prevented. Appendix A.2 shows that equilibrium with $\phi = 1$ always has bubbles, while $\phi < 1$ never does, suggesting that bubbles are not a robust feature of this class of models.

Finally, unlike models such as Basak and Cuoco (1998), the OLG environment ensures long-run stationarity. The ergodic density is given by (20).

Next, we discuss entry (in particular x^*) nd how it relates to asset prices. I apply dynamic programming to the participants' and non-participants' problems, leading to two ODEs (the HJB equations) for g_P and g_N that hold on $(x^*, 1)$:

$$0 = \log(\rho+\pi) - 1 + (\rho+\pi)^{-1}(\alpha\pi + r + \frac{1}{2}\eta^2) - (\rho+\pi)g_P + \mu_X g_P' + \frac{1}{2}\sigma_X^2 g_P'' \tag{21}$$

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

Boundary conditions for these ODEs are the following. First, equation (10) implies the *value-matching* and *smooth-pasting* conditions:[11]

$$g_P(x^*) - g_N(x^*) = \Phi \tag{23}$$

$$g_P'(x^*) = g_N'(x^*) = 0. \tag{24}$$

Finally, given $\sigma_X \to 0$ as $x \to 1$, the limits of (21)-(22) as $x \to 1$ give two more boundary conditions. These five boundary conditions suffice to solve ODEs (21)-(22) and the entry point x^*.[12]

Some simplifications can be made by noticing that the coefficients on g_P and g_N are identical in (21)-(24). Putting $\Delta g := g_P - g_N$ and taking differences between the HJB equations yields one linear ODE

$$0 = \frac{1}{2}(\rho+\pi)^{-1}\eta^2 - (\rho+\pi)\Delta g + \mu_X \Delta g' + \frac{1}{2}\sigma_X^2 \Delta g'', \quad x \in (x^*, 1), \tag{25}$$

[11] These smooth-pasting conditions (formally derived in the proof of Proposition 2 below) are actually implications of both optimality and equilibrium. Optimal entry: the inequality (11), combined with the fact that the solutions g_P, g_N will be at least continuously differentiable, implies that $g_P'(x^*) - g_N'(x^*) = 0$. Equilibrium: the fact that all non-participants are willing to enter when $x \leq x^*$ implies that $g_P(x) = g_P(x^*)$ and $g_N(x) = g_N(x^*)$ for all $x < x^*$.

[12] Proposition B.2 in Appendix B.2 provides a verification theorem that (21)-(24) are sufficient for optimality for the more general recursive utility model introduced later.

with boundary conditions

$$0 = \Delta g(x^*) - \Phi \tag{26}$$

$$0 = \Delta g'(x^*). \tag{27}$$

Proposition 2. *There is exactly one pair $(\Delta g, x^*)$ that satisfies the ODE (25) and boundary conditions in (26)-(27). Consequently, the equilibrium of Definition 1 exists uniquely.*

There are likely many ways to prove Proposition 2. Appendix A.1 demonstrates the equivalence of the model to a relatively standard variational inequality, which is portable to higher dimensions and potentially useful in other models. To economically understand entry decisions, it is informative to examine the integral representation of the value function.

Proposition 3. *The function Δg can be represented as*

$$\Delta g(x) = \mathbb{E}^x\Big[\frac{1}{2}(\rho+\pi)^{-1}\int_0^\infty e^{-(\rho+\pi)t}\eta^2(X_t)dt\Big]. \tag{28}$$

Given (28), we can now interpret Δg 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 value-matching (26), we can re-write (28) to "back out" the implied entry cost ϕ:

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

We can read (29) 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 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.

2.2 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. This is formalized in the next proposition, which implies 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,*

$$\phi^* = 1 - \exp\Big(-\frac{1}{2}(\rho+\pi)^{-1}(\eta^*)^2\Big) = \frac{1}{2}(\rho+\pi)^{-1}(\eta^*)^2 + O(\sigma_Y^4), \tag{30}$$

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

The reason for the result of Proposition 4 lies in the fact that participation strictly dominates non-participation as an investment technology. 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 (29). 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. The risk concentration mechanism, if it is behind large risk premia, must imply large entry costs.

2.3 Larger entry costs

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.

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

Table 1: Baseline calibration of model parameters.

Figure 1 plots Sharpe ratios η and the stationary density h for four different entry costs ϕ. 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.

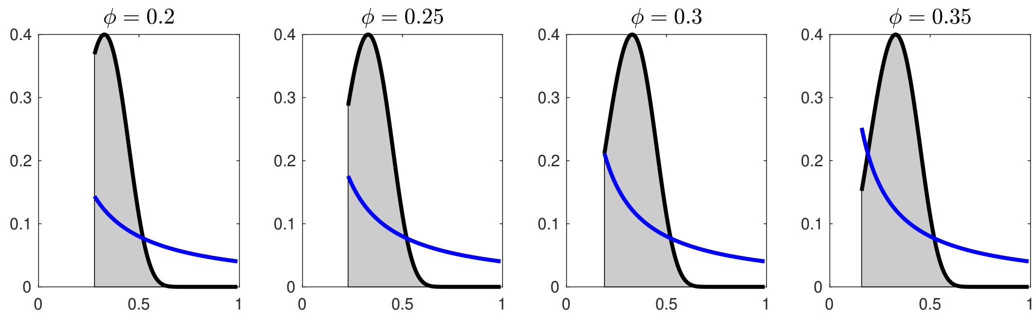

Figure 1: Market price of risk η (blue line) and stationary density for X (gray area), for four different entry costs ϕ. The horizontal axis is the participants' consumption share x. Parameters are in Table 1.

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.

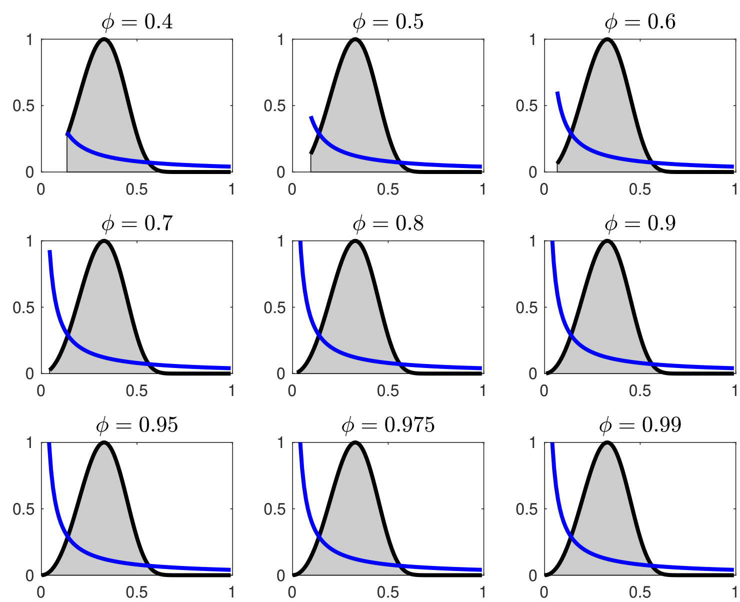

Figure 2: Market price of risk η (blue line) and stationary density for X (gray area), for nine different entry costs ϕ. The horizontal axis is the participants' consumption share x. Parameters are in Table 1.

This story is confirmed by Table 2. 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 $\mathrm{std}[\eta(X_t)]$ stabilize, even as $\max[\eta(X_t)]$ rises even more dramatically.

ϕ	0.20	0.25	0.30	0.35	0.40	0.50	0.60	0.70	0.80	0.90	0.95
$\mathbb{E}[\eta(X_t)]$	0.11	0.12	0.12	0.13	0.13	0.14	0.14	0.15	0.15	0.15	0.15
$\mathrm{std}[\eta(X_t)]$	0.02	0.03	0.03	0.04	0.05	0.06	0.07	0.08	0.08	0.09	0.09
$\max[\eta(X_t)]$	0.14	0.18	0.21	0.25	0.30	0.42	0.61	0.93	1.66	5.00	57.97

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.

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, in that it helps the economy avoid crisis states.[13]

2.4 Recursive preferences

In this section, I generalize utility to the recursive preferences of Duffie and Epstein (1992). In principle, preferences that allow for hedging demands may potentially dissuade entry, even with moderate participation costs. In addition, non-log utility allows time-varying price-dividend ratios, which can amplify risk premium variation. More detailed discussion of this recursive utility model and its equilibrium is in Appendix B.1.

Mathematically, the continuation value and associated felicity function now satisfy

$$V_{t,b} := \mathbb{E}\Big[\int_t^\infty f(c_{s,b}, V_{s,b})ds \mid \mathcal{F}_t\Big], \tag{31}$$

$$\text{where} \quad f(c,V) := \frac{1}{1-\psi}\Big(c^{1-\psi}[V(1-\gamma)]^{\frac{\psi-\gamma}{1-\gamma}} - (\rho+\pi)V(1-\gamma)\Big). \tag{32}$$

In (32), 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.

[13]For comparison, Appendix A.3 contrasts these results to an economy with "exogenous entry." There, a fraction ν of the newborns exogenously become designated participants, while $1-\nu$ fraction become non-participants. No endogenous entry is possible. The main force at play is this "buoying" effect discussed above.

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

$$\Phi_{t,b} := [1 - (1 - \phi)^{1-\gamma}] V^P_{t,b}, \tag{33}$$

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 B.1 in Appendix B.1 derives the equilibrium under these assumptions.[14]

I pick both $\gamma, \psi^{-1} > 1$, to help the model quantitatively. 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.

Figure 3 shows that the economy behaves qualitatively similar to the log utility economy for modest entry costs, but more interesting risk price dynamics are attained under extreme entry costs. 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. Large risk price variability is ruled out, similar to the log utility model. But in panels 6-9, risk prices can attain crisis dynamics.

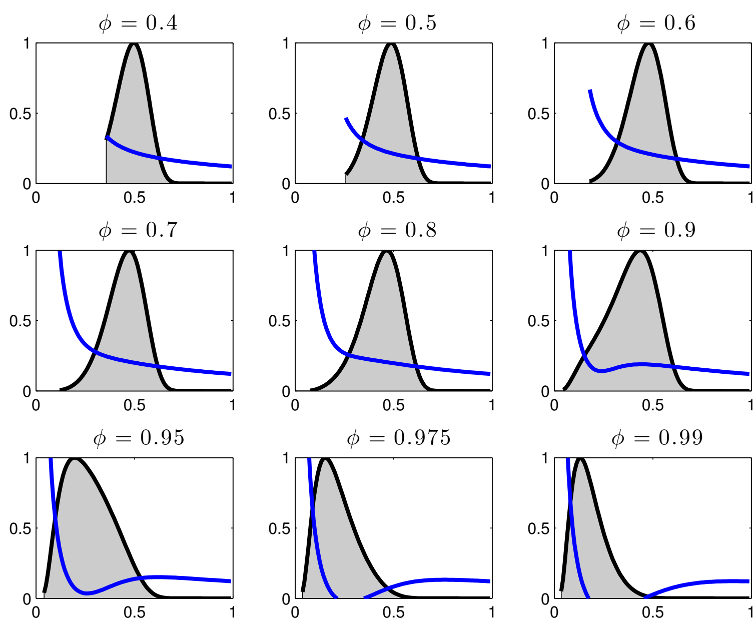

Figure 3: Market price of risk η (blue line) and stationary density for X (gray area), for nine different entry costs ϕ. The horizontal axis is the participants' consumption share x. I set $\gamma = 3$ and $\psi = 3/4$. All other parameters are as in Table 1.

[14]The equilibrium is significantly more complicated than the log utility model. We must solve for two value functions, one each for participants and non-participants, as part of a free-boundary problem for x^*. Furthermore, large nonlinearities emerge because asset prices are no longer independent of agents' value functions.

These results arise because of the hedging motives brought about by $\gamma \neq 1$. As the third row of charts show ($\phi \geq 95\%$), risk prices can become non-monotonic and even negative! Mechanically, participant profits are reduced by low risk prices, a force that reduces the drift of participants' consumption share and allows crisis states to materialize. To understand these motives, suppose x_0 is a crisis-type state with very high and volatile risk prices, i.e., $\eta(x_0)$ and $|\eta'(x_0)|$ very large. At x_1 slightly above x_0, a negative shock improves the investment opportunity set, generating a natural hedge for participants and encouraging intense risk-taking ex-ante. The intuition is "heads, I win; tails, I will win soon." Participant risk-taking pushes down $\eta(x_1)$ in equilibrium, possibly below its values in nearby states and even below zero.

On the one hand, this negative risk price result can partially provide a simple rational explanation for the empirical finding that bank equity often has negative expected returns on the eve of financial crisis (Baron and Xiong, 2017), allegedly proof of market irrationality during credit cycles. In this model, it is precisely these low and negative risk prices that allow participant net worth to attain the crisis portion of the state space (low x).

On the other hand, these features only emerge when entry costs are 90% or larger. Still, entry costs do exist such that the recursive utility model, unlike log, can match the high level and variability of risk prices. This begs the question of whether other utility parameters can match empirical asset prices for moderate entry costs, a subject we turn to in Section 3.

2.5 Robustness to fixed costs, labor income, and preference heterogeneity

For tractability and theoretical sharpness, this paper employs several simplifying assumptions: (1) entry costs are proportional to wealth; (2) all income is capital income; and (3) all agents share identical preferences. This section explores robustness to these assumptions. Details of these extensions are contained in Appendix D.

Fixed entry costs. We briefly consider how more general cost functions would affect the results. In particular, we assume a form of fixed entry costs that are completely independent of individual wealth. While the purpose of this section is to show that our baseline results are robust to deviating from proportional entry costs, such an extension with non-homogeneous entry costs is analytically and computationally non-trivial.

Because non-participants' decisions will depend on wealth in a nonlinear way, the entire distribution of wealth becomes a state variable. Appendix D.1 develops a detailed solution method, which is a "bounded rationality" procedure similar to Krusell and Smith (1998), in which the single aggregate state variable perceived by non-participants is the participant consumption share X_t (i.e., they ignore the dependence of equilibrium on the full distribution of wealth). In this context, the main advantage of continuous time is that one can still obtain quasi-analytical expressions for all equilibrium objects in terms of X_t and the cross-sectional distribution of non-participant wealth f_t; see equations (D.19)-(D.22) for solutions to $(r, \eta, \mu_X, \sigma_X)$ and see Proposition E.1 in Appendix E for solutions to the drift $\mu_{f,t}$ and diffusion $\sigma_{f,t}$ of the cross-sectional distribution f_t (note: with aggregate shocks, the cross-sectional distribution has a diffusion term, meaning it satisfies a stochastic PDE that generalizes the usual Kolmogorov Forward equation).

I consider an entry cost of the form

$$\tilde{\Phi}_t(w) := -(\rho+\pi)^{-1} \log\left[\left(1 - \phi\frac{P_t}{w}\right)^+\right], \quad \phi \in (0,1), \tag{34}$$

where w is the individual's wealth and P_t is the aggregate wealth (stock market value). For an individual of average wealth ($w = P_t$), note $\tilde{\Phi}_t(P_t) = -(\rho+\pi)^{-1}\log(1-\phi)$. This is identical to the proportional cost specification Φ used in the baseline model—see equation (6). For comparison purposes, I have intentionally specified this new fixed cost function so that the individual of average wealth perceives the same cost as in the proportional cost baseline.

ϕ	0.20	0.30	0.40	0.50
$\mathbb{E}[\eta_t]$	0.096	0.105	0.115	0.113
$\mathrm{std}[\eta_t]$	0.012	0.018	0.023	0.020
$\max[\eta_t]$	0.168	0.184	0.215	0.273

Table 3: Entry parameter ϕ (from the fixed cost economy) and different measures of market Sharpe ratios (stationary average SR, standard deviation of SR, and maximal SR). The measures are computed with Monte Carlo simulations of length 5000 years—this matters because, technically, given equilibrium entry dynamics, we have $\sup_t[\eta_t] = +\infty$. Parameters are in Table 1.

Table 3 shows some results for risk prices in this fixed cost model (solving for equilibrium becomes unstable for very large entry costs, so the table stops at $\phi = 0.50$). By and large, the results are similar to the proportional cost Table 2, but with slightly lower and less volatile risk prices. Thus, if anything, fixed costs seem to worsen the asset pricing results. The reasoning for this is connected to the selection effect that emerges with fixed costs: wealthy individuals, who have a larger impact on equilibrium dynamics, are more likely to enter than poor individuals. For a given ϕ, wealthy non-participants will be willing to enter even before risk prices reach crisis magnitudes, which tends to buoy X_t and prevent risk concentration (this can be seen in the drift $\mu_{X,t}$, which tends to be above the corresponding proportional-cost version—see Figure D.3). As ϕ increases, the mass of individuals willing to enter decreases, but the marginal entrant becomes richer and richer, meaning there is only a modest decline in the total quantity of entering wealth. This logic suggests that the fixed cost economy will tend to have less risk concentration than the proportional cost economy, and the gap should increase with ϕ.[15]

Labor income. In this paper, all income is capital income. And therefore, to ensure that newborns have positive wealth, we were forced to assume some unintended bequests from incomplete death insurance ($\alpha < 1$). In reality, approximately two-thirds of income is labor income. Here, we briefly discuss how introducing labor income affects the results.

Following Gârleanu and Panageas (2015), each newborn born at time b is endowed with no financial income, but they receive a labor tree that pays the stream $\{(1-\tilde{\alpha})Y_t\}_{t\in[b,T]}$, where T is the agent's random time of death. The total labor income in the economy is thus a fraction $1-\tilde{\alpha}$ of aggregate output. The stock market is a claim to the residual $\{\tilde{\alpha}Y_t\}$. In addition, we allow agents take full insurance against their death shocks (Blanchard, 1985; Gârleanu and Panageas, 2015).

We consider two polar cases: (1) full pledgeability of labor income and (2) non-pledgeability. In the case of full pledgeability, the non-participants would be able to effectively sell their claims to their labor endowment. To model this situation, despite the fact that non-participants typically cannot trade in the stock market, suppose all newborns have a one-time opportunity (at birth) to sell their human capital tree at market prices. The non-pledgeable case is substantially more complicated, so rather than solving the full-blown model, we provide some simple partial equilibrium calculations in this environment to get a sense of how sensitive the baseline results are to excluding labor income. Obviously, reality is somewhere in between these cases, as there exist markets to borrow against labor income, but these markets are imperfect due to frictions and ethical issues.

It turns out that the fully-pledgeable case is approximately isomorphic to the baseline model.

Lemma 1. *The model with pledgeable labor income and frictionless annuity markets is approximately isomorphic to the baseline model, in the following sense. Given capital share $\tilde{\alpha}$,*

[15] More precisely, we expect this pattern to emerge if ϕ is not too small. Indeed, for small ϕ, the proportional cost Φ induces full participation (Proposition 4), while the fixed cost $\tilde{\Phi}_t$ prevents very poor agents from participating. Thus, risk concentration is higher for the fixed cost when ϕ is small. On the other hand, as $\phi \approx 1$, the proportional cost prevents any entry at all, while the fixed cost still admits entry to agents who are significantly weather than the average. Thus, risk concentration is lower for the fixed cost when ϕ is large. These two extremes suggest there is some threshold cost parameter $\phi^\dagger$ above which the fixed cost economy generates lower and less volatile risk prices than the baseline economy. Comparing tables 2 and 3 suggests this threshold is around $\phi^\dagger \approx 20\%$. Since larger ϕ bring the model closer to data, this simple exercise seems to suggest that fixed costs are not a promising avenue forward.

suppose α is set by

$$\alpha = \frac{1}{2\pi}\Big[\sqrt{\rho^2 + 4\tilde{\alpha}\pi(\rho+\pi)} - \rho\Big]. \tag{35}$$

Then, time-paths of $\{\eta_t\}$ and $\{X_t\}$ are identical between the two models. Additionally, the time-path of $\{r_t\}$ is lower in the pledgeable labor income model by a constant level $(1-\alpha)\pi$.

Remark 1. *In view of Lemma 1, the only discussion becomes about the calibration of α. A capital share $\tilde{\alpha} = 1/3$ is typical. Using $\rho = 0.01$ and $\pi = 0.02$ as in Table 1, condition (35) suggests $\alpha = 25 \times [\sqrt{0.01^2 + 0.03 \times 0.02 \times 4/3} - 0.01] = 0.5$. This is exactly the value used in Table 1.*

The non-pledgeable case is analyzed in Appendix D.2. My calculations suggest that implied entry costs are similar under non-pledgeability; if anything, they are higher. To conclude this, I solve for asset prices from the full-participation equilibrium of this environment, and then I compute the willingness-to-pay to participate for a hypothetical non-participant. Figure D.5 shows that these willingnesses-to-pay are a bit higher with non-pledgeable income than with pledgeable income, suggesting entry costs need to be larger to keep them out of the market.

In these extensions, I have assumed that labor income only carries aggregate risk. This is done for tractability, but simple logic suggests that allowing idiosyncratic labor income risk should not dramatically alter these results. Indeed, because neither participants nor non-participants would be able to hedge the idiosyncratic risk, the increase in indirect utility from participation should depend only weakly on the amount of such risk.

Heterogeneous risk aversion. One may naturally think that participation is imperfect because of heterogeneity in risk tolerance. The more risk tolerant prefer to hold risky assets, whereas the highly risk averse stay out of the market, even for small entry costs. This is formalized in Appendix D.3, which shows that risk tolerant agents have higher willingness-to-pay to enter (Appendix D.3 also does the same exercise for heterogeneous EIS).

However, this calculation is incomplete. For concreteness, suppose two levels of risk aversion, γ_L and $\gamma_H > \gamma_L$. In equilibrium, all participants will be γ_L-agents. Furthermore, with a constant arrival of both types, there will always be a non-zero mass of γ_L-agents who are not participating. This logic implies that after a series of negative shocks, risk prices increase but only the γ_L-agents will enter. Putting these ideas together, both equilibrium risk prices *and* the entry barrier will be solely determined by the γ_L-agents. We would expect this economy to thus look very much like the baseline economy with only γ_L-agents. In that case, the only question is what choice of common risk aversion best fits the data.

3 Tension between conditional and unconditional risk premia

Although in principle limited participation models bring the promise of matching both high and variable risk premia, the previous section showed that this requires enormous entry costs. A natural hope is that implied entry costs can be reduced by adding auxiliary features which raise levels and/or variability of risk premia. The problem: when entry is endogenous, features that raise average risk premia incentivize more participation, which mitigates risk premia dynamics. This trade-off between the unconditional level and conditional dynamics of risk premia suggests a challenge in matching both with moderate entry costs. In this section, I demonstrate this trade-off for the following model extensions: increasing agents' risk aversions, allowing equity-issuance, and introducing idiosyncratic risk.

3.1 Higher risk aversion

Consider again the recursive-utility model introduced at the end of the last section. By comparing the equilibrium for different 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 incentivizes entry, thus mitigating the time-variability of risk premia. In partial equilibrium, we would expect more frequent entry when agents are more risk-tolerant. In general equilibrium, this effect is surprisingly reversed.

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 B.3.

Proposition 5. *Define $\phi^* := \sup\{\phi : x^*(\phi) = 1\}$. Then, $\phi^* > 0$ and in particular,*

$$\phi^* \geq 1 - \Big(\frac{\rho + \pi + (\psi - 1)(r^* + \alpha\pi + \frac{1}{2\gamma}(\eta^*)^2)}{\rho + \pi + (\psi - 1)(r^* + \alpha\pi)}\Big)^{\frac{\psi}{1-\psi}} = \frac{\frac{1}{2}\sigma_Y \eta^*}{\rho + \pi(1-\alpha) + \psi(\mu_Y + \pi\alpha)} + O(\sigma_Y^4), \tag{36}$$

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

The positive relationship between risk aversion and entry incentives is depicted in Figure 4. 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 4). This occurs because higher γ raises equilibrium risk prices $\eta^* = \gamma\sigma_Y$, which appear directly in (36). 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 4).

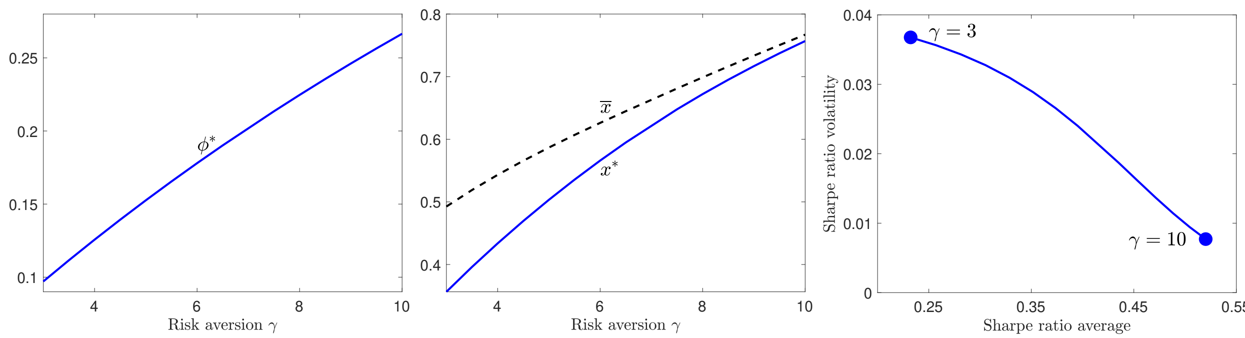

Figure 4: Left panel: full-integration cost $\phi^* := \sup\{\phi : x^*(\phi) = 1\}$, as a function of risk aversion γ. Middle panel: entry boundary, x^*, and stationary mean, $\overline{x} := \mathbb{E}X_t$, as a function of risk aversion γ. Right panel: trade-off between $\mathbb{E}[\eta(X_t)]$ and 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. The proof of the following result is in Appendix B.1.

Proposition 6. *In the recursive utility model,* $\eta(x^*)/\eta(1) = 1/x^*$.

Recall that higher γ tends to lead to increase x^* (middle panel of Figure 4). If we consider $\eta(x^*)$ a proxy for the maximal risk price and $\eta(1)$ a proxy for the minimal risk price,[16] then $\eta(x^*)/\eta(1)$ proxies for risk price variability. Thus, Proposition 6 implies 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 4): 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.[17]

[16] 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 3 show.

[17] 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 (36) 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 4 looks very similar if we were to vary σ_Y, rather than γ.

3.2 Equity issuance

In the baseline model, non-participants can only share risky asset returns through the bond market. In reality, for example asset markets in which financial intermediaries play an outsized role, participants can partially share risk by issuing equity to non-participants. In this extension, I allow partial equity issuance by participants, to facilitate a financial intermediary interpretation.[18] If non-participants can access asset returns through intermediary equity, they have less incentive to pay a cost to participate directly. Delayed entry raises the chances of extreme risk prices, bringing the model closer to data. However, as I show below, this finding comes at the expense of typical risk premia levels, which are pushed down by the improved risk-sharing achieved by equity issuance. Again, we see a trade-off between risk premia levels and dynamics.

Participants must keep a fraction $\chi_{t,b} \geq \chi^*$ of their equity risk on their books, and offload the remaining $1 - \chi_{t,b}$ of risk to financial markets. The parameter $\chi^* \in (0, 1)$ measures the tightness of the equity-issuance constraint (the baseline model coincides with $\chi^* = 1$). Complete details and derivations for this extension are in Appendix A.4.

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), which is at least as high as the one for outside equity ($\tilde{\eta}_t$).[19] In equilibrium, these risk prices are given by $\eta_t = \frac{\max(X_t,\chi^*)}{X_t}\sigma_Y$ and $\tilde{\eta}_t = \frac{1-\max(X_t,\chi^*)}{1-X_t}\sigma_Y$.

Equilibrium with equity issuance features a "safe" risk-sharing region ($X_t \geq \chi^*$) and a "vulnerable" constrained region ($X_t < \chi^*$). In the safe region, participants and non-participants perfectly share aggregate risk, because they obtain the same risk compensation and have the same risk preferences. Participants are wealthy enough that equity-issuance constraints are not binding.

However, the safe region is transient. With equalized risk prices, the only dynamic force affecting X_t is the OLG process, whereby newborn non-participants slowly replace existing participants. Thus, the economy deterministically progresses toward the constrained region ($X_t < \chi^*$), at which point perfect risk-sharing is not possible. The economy never leaves the constrained region.

Within the constrained region, dynamics are qualitatively similar for any χ^*. Risk is concentrated on participants' balance sheets, so negative fundamental shocks translate lead

[18]A similar model with equity issuance is that considered by He and Krishnamurthy (2012, 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.

[19]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., $\chi_{t,b}\sigma_{R,t}\eta_t := \mu_{R,t} - r_t - (1 - \chi_{t,b})\sigma_{R,t}\tilde{\eta}_t$.

to even more concentrated risk ($x \downarrow$), which leads to higher risk prices ($\eta \uparrow$). Despite this qualitative similarity, χ^* has opposing effects on the level of risk prices and their time-series variation.

Proposition 7. *Consider a set of alternative economies $\mathcal{E}$ parameterized by equity-issuance constraints χ^*. 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^*, almost-surely.

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

Proposition 7 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. In fact, the risk price variability ratio $\sup_t(\eta_t^{\chi^*})/\inf_t(\eta_t^{\chi^*}) = \chi^*/x^*$ would be increasing in χ^* if we held entry behavior fixed (i.e., held x^* fixed), so the tension between risk price levels and variability is fundamentally due to entry. Intuitively, higher risk price levels in good times incentivize non-participants to enter earlier, and risk prices at entry will be more moderate.

3.3 Idiosyncratic risk

One possible reason for slow-moving capital into complex risky asset markets is the presence of idiosyncratic risk embedded in the assets.[20] 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 and derivations in Appendix A.4. 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

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

[20] 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.

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

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 the sum of the aggregate plus idiosyncratic risk premia equals the total risk premium:

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

Idiosyncratic risk prices are given by the simple formula $\hat{\eta}_t = \hat{\sigma}_Y / X_t$ (Appendix A.4).

Since participants earn $\hat{\eta}$, which is increasing in $\hat{\sigma}_Y$, the presence of idiosyncratic risk makes participants wealthier in the long-run, thus leading to lower aggregate risk prices. Formally, we have the following analog of claim (i) of Proposition 7.

Proposition 8. *Consider a set of alternative economies $\mathcal{E}$ parameterized by idiosyncratic volatility $\hat{\sigma}_Y$. 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 (29), we can write the implied entry costs of this economy as

$$\phi = 1 - \exp\Big(- \frac{1}{2}\mathbb{E}^{x^*}\Big[\int_0^\infty e^{-(\rho+\pi)t}[\eta^2(X_t) + \hat{\eta}^2(X_t)]dt \Big]\Big). \tag{37}$$

With larger $\hat{\sigma}_Y$, idiosyncratic risk prices $\hat{\eta}_t = \hat{\sigma}_Y / X_t$ tend to be larger, but aggregate risk prices η_t tend to be smaller (Proposition 8). Thus, there could in principle be an ambiguous effect on entry incentives. This ambiguity disappears if we study the full-integration cost ϕ^*, analogously to Proposition 4.

Proposition 9. *Define $\phi^* := \sup\{\phi : x^*(\phi) = 1\}$. Then, $\phi^* = 1 - \exp(-\frac{1}{2}(\rho+\pi)^{-1}[\sigma_Y^2 + \hat{\sigma}_Y^2])$.*

From Proposition 9, 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. Combining Propositions 8-9, we conclude that the presence of idiosyncratic risk, while introducing idiosyncratic risk premia, can reduce both the level and variability of aggregate risk premia.

3.4 Combining all the extensions

Consider the model with recursive preferences (risk aversion γ and EIS ψ^{-1}), equity-issuance (retention share χ^*), and idiosyncratic risk (idiosyncratic volatility $\hat{\sigma}_Y$) all together. This corresponds loosely to a model like Di Tella (2017).[21] Using the same method as Propositions 5 and 9, it is straightforward to compute the maximum entry cost consistent with complete integration, $\phi^* := \sup\{\phi : x^*(\phi) = 1\}$, which is given by

$$\phi^* = \frac{\frac{1}{2}\chi^* \hat{\sigma}_Y \hat{\eta}^*}{\rho + \pi(1-\alpha) + \psi(\mu_Y + \pi\alpha)} + O(\sigma_Y^4) + O(\hat{\sigma}_Y^4), \tag{38}$$

where $\hat{\eta}^* = \gamma\hat{\sigma}_Y$ is the full-integration idiosyncratic risk price. Under parameters of Table 1, and also $\gamma = 5$, $\psi = 0.5$, $\chi^* = 0.2$, and $\hat{\sigma}_Y = 0.25$ (all exactly as in the calibration of Di Tella (2017)), this approximation delivers $\phi^* \approx 89\%$. By extending the model in these directions, the risk concentration channel seems even more reliant on enormous entry barriers.

4 Extrapolative beliefs

The results so far, that asset markets imply unreasonably high entry costs, can be rephrased in the form of a question: why is capital slow-moving, especially in crises when high risk premia prevail (Duffie, 2010b)? In this section, I introduce extrapolative beliefs a la Barberis et al. (2015) into the model to help explain why high risk premia may not induce entry.

Define agents' *sentiment* about financial markets by S_t, which is an exponentially-weighted average of previous returns, exactly as in "constant-gain learning" (Evans and Honkapohja, 2012):

$$S_t := \beta \int_{-\infty}^{t} e^{-\beta(t-s)} dR_s. \tag{39}$$

Or in changes, sentiment follows

$$dS_t = \beta(dR_t - S_t dt). \tag{40}$$

[21] Di Tella (2017) additionally considers time-variation in the idiosyncratic volatility and a full ability to contract on aggregate shocks beyond the risk-sharing parameter χ^*, although these modifications are not crucial for the point I make in this subsection (indeed, notice that the principal term in equation (38) is unaffected by aggregate risk σ_Y). The driving force behind the results here is the large value of idiosyncratic risk, which when combined with large risk aversion, presents a huge idiosyncratic risk premium that only participants can access.

Equations (39) and (40) capture the idea that a string of positive (negative) returns increases (decreases) sentiment, while sentiment mean-reverts in absence of trends.[22]

Sentiment is the basis of extrapolative expectations: all agents have biased expectations about returns, leaning in the direction of their sentiment. Mathematically, I assume, like Barberis et al. (2015), that perceived expected returns are a weighted average of actual expected returns and the level of sentiment:

$$\tilde{\mu}_{R,t} := \mu_{R,t} + \lambda(S_t - \mu_{R,t}). \tag{41}$$

Equation (41) defines a distorted probability under which agents view the economy and is the key assumption of this section. Note that $\lambda \in [0, 1]$ controls the degree of bias in agents' expectation-formation: $\lambda = 1$ is fully extrapolative, while $\lambda = 0$ is fully rational. Note also that $1/\beta$ captures the average amount of past return data (number of years) used by extrapolators.

The calibration in Barberis et al. (2015), based on survey data, delivers $\beta = 0.5$ and $\lambda = 1$, which I will use as a benchmark. Like Jin and Sui (2021), I will find that these high values of (β, λ) generate far too much risk premium variation, so I will perform some sensitivity analysis on (β, λ). In a calibration based on output growth, Maxted (2020) obtains $\beta \approx 0.12$, which produces a more reasonable amount of variation. In addition, it is natural to think that some agents in the economy are more rational, which can be captured in a reduced-form way through lower λ.

This formulation of extrapolation is potentially complex, as sentiment dynamics depend on returns, while returns depend endogenously on sentiment (by contrast, Maxted (2020) and Krishnamurthy and Li (2020) analyze simpler settings with extrapolation on an exogenous variable). With log utility, this two-way feedback will vanish; with recursive utility below, the feedback amplifies volatility and risk premia. All derivations and proofs for this section are in Appendix C.

[22]Barberis et al. (2015) model sentiment directly on past prices, rather than returns, because they have a stationary model without growth. Jin and Sui (2021) studies return extrapolation, but with a two-state regime shifting process for beliefs.

Why not build sentiment off of 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. See Barberis, Shleifer, and Vishny (1998) for an early example with dividend extrapolation, and Nagel and Xu (2019) more recently. A related literature studies learning from fundamentals based on lifetime experiences, which aggregates to a type of extrapolation, given the simultaneous existence of both young and old generations; see Collin-Dufresne, Johannes, and Lochstoer (2017), Ehling, Graniero, and Heyerdahl-Larsen (2018), Schraeder (2016), and Malmendier, Pouzo, and Vanasco (2020).

4.1 Log utility

To obtain a clean comparison with our baseline model, I start with log utility.

Proposition 10. *There exists a Markov equilibrium with sentiments, which is governed by the state variables* (X_t, S_t)*. Entry occurs whenever* $X_t \leq x^*$*, where* x^* *is identical to the entry point without sentiments. Actual and perceived market Sharpe ratios are given by, respectively,*

$$\eta_t = \frac{\sigma_Y}{X_t} - \frac{\lambda}{\sigma_Y}(S_t - \overline{s}) \quad \textit{and} \quad \tilde{\eta}_t = \frac{\sigma_Y}{X_t},$$

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

Based on the results in Proposition 10, I will now argue that sentiments can generate a type of Sharpe ratio volatility that is not curtailed by endogenous entry.

The actual Sharpe ratio $\eta(x, s) = \frac{\sigma_Y}{x} - \frac{\lambda}{\sigma_Y}(s - \overline{s})$ is equal to the perceived Sharpe ratio $\tilde{\eta}$, plus a term capturing sentiments. Computing

$$d\eta_t - \mathbb{E}_t[d\eta_t] = -\Big(\lambda\beta + \sigma_Y^2 \frac{1 - X_t}{X_t^2}\Big) dZ_t \tag{42}$$

we see the loading on the shock dZ_t is amplified by the degree of extrapolation λ and the sentiment volatility parameter β. Thus, Sharpe ratios are more countercylical and volatile with sentiments. Importantly, this result does not rely on increasing agents' consumption growth volatility ($\tilde{\eta}_t$ is participants' local consumption growth volatility, which is independent of λ).

Sentiment-driven Sharpe ratio volatility is not curbed by entry. Indeed, entry occurs when the perceived, rather than actual, participation benefits are high. Perceived benefits are given by

$$\frac{1}{2}(\rho + \pi)^{-1} \tilde{\mathbb{E}}\Big[\int_0^\infty e^{-(\rho+\pi)t} \tilde{\eta}_t^2 dt \mid \mathcal{F}_0\Big], \tag{43}$$

where $\tilde{\mathbb{E}}$ denotes expectations under the extrapolative beliefs (this formula is the irrational beliefs generalization of (28)). The perceived Sharpe ratio $\tilde{\eta}$ is independent of s. Similarly, agents perceive the dynamics of (X_t, S_t) to be independent of S_t, so the subjective forecast of $\tilde{\eta}_{t+h}$ is independent of S_t. Entry incentives are thus independent of sentiments, meaning sentiment-driven volatility of actual Sharpe ratio η_t sidesteps the conditional-unconditional tension of Section 3.

This story is confirmed in Table 4, which displays Sharpe ratio statistics for different β, λ, and ϕ. Sharpe ratio volatilities are strongly increasing in the extrapolation parameters β and λ, as suggested by the local equation (42) above, but relatively insensitive to ϕ. At the same time, average Sharpe ratios are insensitive to (β, λ) and increasing in ϕ. One can thus think of the degree of frictions (ϕ) as controlling average Sharpe ratios, while extrapolation (β, λ) as controlling Sharpe ratio volatility. In this sense, there is no tension between the mean and volatility of η_t.

Quantitatively, with $\beta = 0.5$, $\lambda = 1$, and $\phi = 0.2$, this economy generates a Sharpe ratio distribution with mean 0.11 and volatility 0.53. By contrast, in the rational model with $\phi = 0.2$, the *maximal* Sharpe ratio is only 0.14. Note that, as in the baseline model, log utility does not deliver high enough average Sharpe ratios; Section 4.3 generalizes the sentiments model to recursive utility to help address this issue.

$\beta = 0.30$ (persistent, low-volatility extrapolators)

λ	0 (rational)			0.30			0.70			1 (pure extrap.)		
ϕ	0.10	0.20	0.30	0.10	0.20	0.30	0.10	0.20	0.30	0.10	0.20	0.30
$\mathbb{E}[\eta_t]$	0.08	0.11	0.12	0.08	0.11	0.13	0.08	0.11	0.13	0.08	0.12	0.13
$\text{std}[\eta_t]$	0.01	0.02	0.03	0.13	0.14	0.14	0.29	0.30	0.30	0.41	0.42	0.42
$\max[\eta_t]$	0.08	0.14	0.20	0.53	0.58	0.64	1.11	1.17	1.23	1.56	1.61	1.67

$\beta = 0.50$ (moderate, medium-volatility extrapolators)

λ	0 (rational)			0.30			0.70			1 (pure extrap.)		
ϕ	0.10	0.20	0.30	0.10	0.20	0.30	0.10	0.20	0.30	0.10	0.20	0.30
$\mathbb{E}[\eta_t]$	0.08	0.11	0.12	0.08	0.11	0.13	0.08	0.12	0.13	0.08	0.12	0.13
$\text{std}[\eta_t]$	0.01	0.02	0.03	0.16	0.17	0.18	0.37	0.38	0.38	0.53	0.53	0.54
$\max[\eta_t]$	0.08	0.14	0.20	0.70	0.75	0.81	1.51	1.57	1.63	2.12	2.18	2.24

$\beta = 0.70$ (transitory, high-volatility extrapolators)

λ	0 (rational)			0.30			0.70			1 (pure extrap.)		
ϕ	0.10	0.20	0.30	0.10	0.20	0.30	0.10	0.20	0.30	0.10	0.20	0.30
$\mathbb{E}[\eta_t]$	0.08	0.11	0.12	0.08	0.11	0.13	0.08	0.12	0.13	0.08	0.12	0.13
$\text{std}[\eta_t]$	0.01	0.02	0.03	0.19	0.20	0.20	0.43	0.44	0.45	0.62	0.63	0.63
$\max[\eta_t]$	0.08	0.14	0.20	0.86	0.92	0.98	1.90	1.96	2.02	2.68	2.74	2.80

Table 4: Different measures of market Sharpe ratios (stationary average SR, standard deviation of SR, and maximal SR) for extrapolation parameters (β, λ) and entry cost ϕ. The measures are computed with Monte Carlo simulations of length 5000 years—this matters because, technically, given the equilibrium Ornstein-Uhlenbeck followed by S_t, we have $\sup_t[\eta_t] = +\infty$. Other parameters are in Table 1.

To more directly verify the notion that sentiments are controlling Sharpe ratio volatility, I perform the following variance decomposition of the Sharpe ratio:

$$1 = \underbrace{\frac{(\lambda\beta)^2}{\mathrm{Var}_t[d\eta_t]}}_{\text{pure sentiment}} + \underbrace{\frac{(1-X_t)^2\left(\frac{\sigma_Y}{X_t}\right)^4}{\mathrm{Var}_t[d\eta_t]}}_{\text{pure risk concentration}} + \underbrace{\frac{2\lambda\beta(1-X_t)\left(\frac{\sigma_Y}{X_t}\right)^2}{\mathrm{Var}_t[d\eta_t]}}_{\substack{\text{interaction between} \\ \text{sentiment and risk concentration}}} . \tag{44}$$

The first term is the portion of variance that arising purely due to extrapolation (i.e., the numerator represents Sharpe ratio variance in the representative agent version of this economy). As Table 5 shows, this pure sentiment term explains the lion's share of Sharpe ratio volatility.

$\beta = 0.30$ (persistent, low-volatility extrapolators)

λ	0 (rational)			0.30			0.70			1 (pure extrap.)		
ϕ	0.10	0.20	0.30	0.10	0.20	0.30	0.10	0.20	0.30	0.10	0.20	0.30
pure sentiment	0.00	0.00	0.00	0.93	0.84	0.79	0.97	0.92	0.89	0.98	0.94	0.92
risk concentration	1.00	1.00	1.00	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.00
interaction	0.00	0.00	0.00	0.07	0.15	0.20	0.03	0.08	0.10	0.02	0.06	0.08

$\beta = 0.50$ (moderate, medium-volatility extrapolators)

λ	0 (rational)			0.30			0.70			1 (pure extrap.)		
ϕ	0.10	0.20	0.30	0.10	0.20	0.30	0.10	0.20	0.30	0.10	0.20	0.30
pure sentiment	0.00	0.00	0.00	0.96	0.90	0.86	0.98	0.95	0.93	0.99	0.96	0.95
risk concentration	1.00	1.00	1.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
interaction	0.00	0.00	0.00	0.04	0.10	0.13	0.02	0.05	0.07	0.01	0.04	0.05

$\beta = 0.70$ (transitory, high-volatility extrapolators)

λ	0 (rational)			0.30			0.70			1 (pure extrap.)		
ϕ	0.10	0.20	0.30	0.10	0.20	0.30	0.10	0.20	0.30	0.10	0.20	0.30
pure sentiment	0.00	0.00	0.00	0.97	0.93	0.90	0.99	0.96	0.95	0.99	0.97	0.97
risk concentration	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
interaction	0.00	0.00	0.00	0.03	0.07	0.10	0.01	0.04	0.05	0.01	0.03	0.03

Table 5: The variance decomposition (44) for various extrapolation parameters (β, λ) and entry costs ϕ. Each term in the decomposition is computed as the average over a 5000 year Monte Carlo simulation. Other parameters are in Table 1.

An interesting way to visualize these findings is to view Sharpe ratios during a boom-bust cycle. Figure 5 plots Sharpe ratio dynamics through a 1-quarter boom, followed by a 4.5 year quiet period, and finally followed by a 1-quarter bust, in each of three models (extrapolative with frictions, rational with frictions, and extrapolative but frictionless). By comparing the blue line to the broken red line, we can see how adding extrapolation to frictions greatly amplifies the amount of Sharpe ratio variation, with far below-average Sharpe ratios in the boom (briefly even negative) and Sharpe ratio spikes in the bust. The dramatic difference

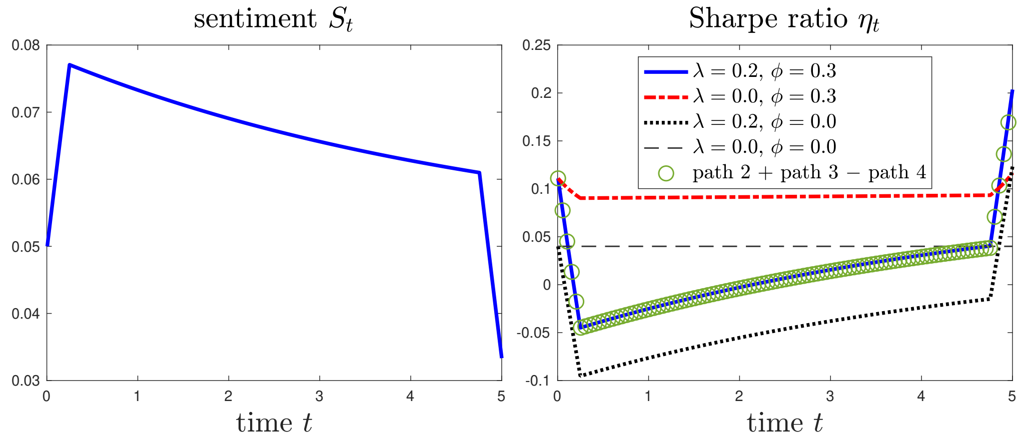

Figure 5: Dynamics through a boom-bust cycle in four models: (i) extrapolation and frictions ($\lambda = 0.2$, $\phi = 0.3$); (ii) frictions but rational ($\lambda = 0$, $\phi = 0.3$); (iii) extrapolation but frictionless ($\lambda = 0.2$, $\phi = 0$; (iv) rational and frictionless ($\lambda = \phi = 0$). The series of shocks are $Z_{t+0.01} - Z_t = \sqrt{0.01}$ for $t < 0.25$; Z_t constant for $t \in [0.25, 4.75]$; and $Z_{t+0.01} - Z_t = -\sqrt{0.01}$ for $t > 4.75$. Consequently, sentiment dynamics are identical across the models (left panel). The implied output growth rate in the 1-quarter boom is 2.5% (or 10% annualized). Sentiment mean-reversion is $\beta = 0.2$, and other parameters are in Table 1.

arises even though the extrapolation parameters are quite mild: $\lambda = 0.2$ and $\beta = 0.2$. However, these Sharpe ratio dynamics look almost identical to those of an extrapolative representative agent (dotted black line), but shifted upwards. In fact, letting $\eta^{(\lambda,\phi)}$ be the Sharpe ratio from an economy with extrapolation λ and friction ϕ, the curve with green bubbles illustrates the following approximate identity in this model:

$$\underbrace{\eta^{(\lambda,\phi)}}_{\substack{\text{extrapolative}\\\text{plus}\\\text{frictions}}} \approx \underbrace{\eta^{(0,\phi)}}_{\substack{\text{rational}\\\text{plus}\\\text{frictions}}} + \underbrace{\eta^{(\lambda,0)}}_{\substack{\text{extrapolative}\\\text{and}\\\text{frictionless}}} - \underbrace{\eta^{(0,0)}}_{\substack{\text{rational}\\\text{and}\\\text{frictionless}}} .$$

Taking expectations of this relation shows that $\mathbb{E}[\eta^{(\lambda,\phi)}] \approx \mathbb{E}[\eta^{(0,\phi)}]$.[23] From this perspective, extrapolation and frictions are additive: frictions and frictions alone control Sharpe ratio levels, while beliefs control their dynamics. This additive property stands in contrast to sub-additive nature of Section 3, in which extensions that add risk price variability reduce risk price levels.

4.2 The central importance of extrapolative participants

By and large, models with extrapolative beliefs study economies in which all agents share the same belief. Some conventional wisdom, however, posits that active participants in mar-

[23]Indeed, $\eta^{(\lambda,0)} = \sigma_Y - \frac{\lambda}{\sigma_Y}(S_t - \bar{s})$, where $\mathbb{E}[S_t] = \bar{s}$. Thus, $\mathbb{E}[\eta^{(\lambda,0)}] = \sigma_Y = \eta^{(0,0)}$.

kets will tend to learn over time and hold more rational beliefs than occasional participants or non-participants. Furthermore, the vast majority of survey evidence motivating the extrapolative belief literature is not taken from banks or other financial professionals, so it is unclear whether the assumption of extrapolative participants is well-justified.[24] One can get a sense of what participant learning would imply in my model by assuming participants are fully rational and non-participants are extrapolative (in this extreme example, agents immediately become rational when they enter, although they do not recognize this ex-ante). Through this simple extension, I will illustrate how extrapolative participants are critical to interesting Sharpe ratio dynamics. This is an important challenge for the literature to address empirically, since most of the survey evidence on beliefs does not usually come from the banks and other specialist investors that my class of models emphasizes.

Table 6 shows that this setting generates tiny Sharpe ratios, comparable to the rational representative agent economy (i.e., $\sigma_Y = 0.04$), with minuscule volatility. Two reasons drive this finding, a direct effect and an indirect effect through entry, which I explain next.

$\beta = 0.30$ (persistent, low-volatility extrapolators)

λ	0 (rational)			0.30			0.70			1 (pure extrap.)		
ϕ	0.10	0.20	0.30	0.10	0.20	0.30	0.10	0.20	0.30	0.10	0.20	0.30
$\mathbb{E}[\eta_t]$	0.08	0.11	0.12	0.04	0.05	0.08	0.04	0.04	0.04	0.04	0.04	0.04
$\mathrm{std}[\eta_t]$	0.01	0.02	0.03	0.01	0.02	0.04	0.00	0.00	0.01	0.00	0.00	0.00
$\max[\eta_t]$	0.09	0.14	0.21	0.09	0.25	0.61	0.07	0.08	0.09	0.07	0.07	0.08

$\beta = 0.50$ (moderate, medium-volatility extrapolators)

λ	0 (rational)			0.30			0.70			1 (pure extrap.)		
ϕ	0.10	0.20	0.30	0.10	0.20	0.30	0.10	0.20	0.30	0.10	0.20	0.30
$\mathbb{E}[\eta_t]$	0.08	0.11	0.12	0.04	0.04	0.05	0.04	0.04	0.04	0.04	0.04	0.04
$\mathrm{std}[\eta_t]$	0.01	0.02	0.03	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00
$\max[\eta_t]$	0.09	0.14	0.21	0.07	0.09	0.22	0.06	0.07	0.07	0.05	0.06	0.07

$\beta = 0.70$ (transitory, high-volatility extrapolators)

λ	0 (rational)			0.30			0.70			1 (pure extrap.)		
ϕ	0.10	0.20	0.30	0.10	0.20	0.30	0.10	0.20	0.30	0.10	0.20	0.30
$\mathbb{E}[\eta_t]$	0.08	0.11	0.12	0.04	0.04	0.05	0.04	0.04	0.04	0.04	0.04	0.04
$\mathrm{std}[\eta_t]$	0.01	0.02	0.03	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
$\max[\eta_t]$	0.09	0.14	0.21	0.06	0.07	0.09	0.05	0.05	0.06	0.05	0.05	0.05

Table 6: Rational participants and extrapolative non-participants. Different measures of market Sharpe ratios (stationary average SR, standard deviation of SR, and maximal SR) for extrapolation parameters (β, λ) and entry cost ϕ. The measures are computed with Monte Carlo simulations of length 5000 years—this matters because, technically, given the equilibrium Ornstein-Uhlenbeck followed by S_t, we have $\sup_t[\eta_t] = +\infty$. Other parameters are in Table 1.

[24] Two important exceptions to this are Cheng, Raina, and Xiong (2014) and Gennaioli, Ma, and Shleifer (2016), which study securitized finance managers and CFOs, respectively.

First, sentiments have no direct effect on Sharpe ratios if participants are rational. Indeed, actual and perceived risk prices are now[25]

$$\text{(only non-participants extrapolate)} \quad \eta_t = \frac{\sigma_Y}{X_t} \quad \text{and} \quad \tilde{\eta}_t = \frac{\sigma_Y}{X_t} + \frac{\lambda}{\sigma_Y}(S_t - \overline{s}). \tag{45}$$

There is a symmetry between (45) and Proposition 10, in that $\eta_t - \tilde{\eta}_t = -\frac{\lambda}{\sigma_Y}(S_t - \overline{s})$ in both. However, η_t is independent of S_t here.

Second, extrapolative non-participants are induced to enter much more often in a world with rational participants. This entry force buoys X_t and indirectly reduces Sharpe ratio levels and their dynamics. To see this, Figure 6 computes non-participants' entry incentives in this environment. From the figure, it is tempting to say that entry is delayed, since low values of sentiment require more extreme drops in participants' wealth share to induce entry. However, the correct reading is that relatively high sentiment induces lots of entry, especially for larger degrees of extrapolation, which then renders the economy "far from crisis." For example, in the right panel with $\lambda = 0.5$, the entry barrier is approximately an asymptote near $s \approx \overline{s} + \epsilon$. If sentiment rises above average, so much entry occurs that $X_t \to 1$ almost immediately. Afterward, even if sentiment declines, participants are so well-capitalized that Sharpe ratios barely move.

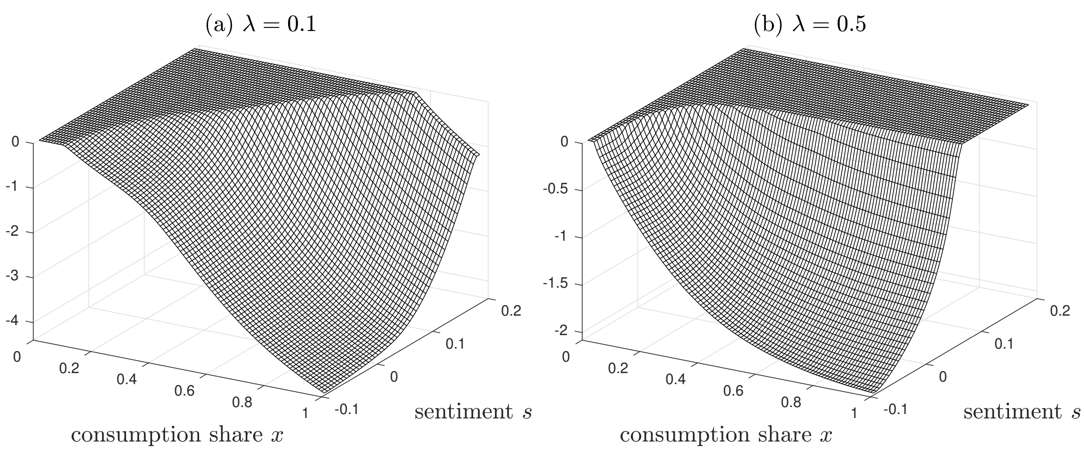

Figure 6: Net entry benefit for extrapolative non-participants, in an equilibrium with rational participants. Entry occurs whenever this net benefit is zero. In this computation, all agents additionally face a no-short-sales constraint on risky assets, which modifies the entry benefit for non-participants to (43), with $\tilde{\eta}^2$ replaced by $\max[0, \tilde{\eta}]^2$. Extrapolation parameters are $\beta = 0.5$ and $\lambda = 0.1$ (panel a) or $\lambda = 0.5$ (panel b). Entry cost is $\phi = 0.20$. Other parameters are in Table 1.

In summary, extrapolative participants are critical to generate large volatility in Sharpe ratios. This is due to both a direct effect—whereby variation in the marginal trader's beliefs

[25] With rational participants and log utility, $(r, \eta, \mu_X, \sigma_X)$ is identical to the values in Proposition 1.

translates into risk price variation—and a less-obvious indirect effect, through accelerated entry in good times.

4.3 Recursive utility

Now, we generalize preferences to recursive utility in order to obtain more reasonable quantitative magnitudes. This becomes a challenging methodological problem that has not been tackled in the literature; Proposition C.1 in Appendix C.3 characterizes equilibrium as the solution to a coupled set of nonlinear PDEs with a free boundary (entry barrier).[26]

Figure 7 displays entry incentives in this model. Although the equilibrium expressions are substantially more complex than the log case, the left panel shows that entry decisions are still well-approximated by a threshold level of X_t, as in Proposition 10 (i.e., entry incentives are roughly independent of S_t). The core intuition is that the perceived Sharpe ratio $\tilde{\eta}$ remains the crux of the entry decision, and $\tilde{\eta}$ happens to be roughly independent of S_t (right panel).

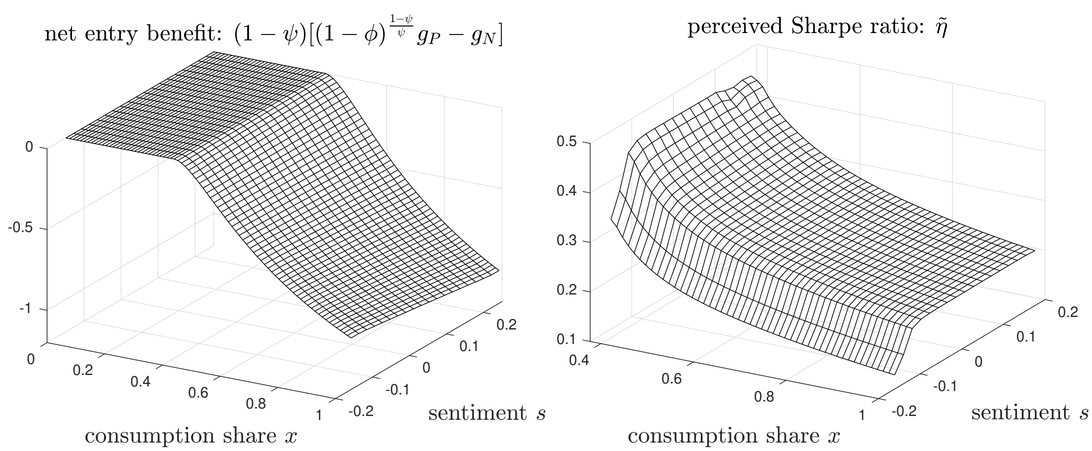

Figure 7: Entry incentives with recursive preferences and sentiment. Risk aversion is $\gamma = 3$ and EIS is $\psi^{-1} = 4/3$. Extrapolation parameters are $\beta = 0.5$ and $\lambda = 0.5$. Entry cost is $\phi = 0.45$. Other parameters are in Table 1.

Table 7 displays some quantitative results for this model. From panel A, we see the model generates a high and time-varying Sharpe ratio, a low and smooth interest rate, and return volatility above fundamental volatility. Although the increase in return volatility

[26] Jin and Sui (2021) and Nagel and Xu (2019) are the first to study extrapolation with Epstein-Zin preferences in general equilibrium, albeit in a representative agent framework—they differ by assuming return extrapolation versus fundamentals extrapolation, respectively. By contrast, Barberis et al. (2015) has CARA preferences. Our model has—in addition to return extrapolation and Epstein-Zin preferences—a role for the wealth distribution and an endogenous participation decision. As mentioned earlier, this puts our paper closer to the financial friction papers of Maxted (2020) and Krishnamurthy and Li (2020)—but note that their analyses is also limited to CRRA utility, extrapolation of an exogenous variable, and neither have entry by non-participants.

seems modest, the standard practice in the asset pricing literature is to scale this volatility by a "leverage ratio" in order to compare with equity market data; for example, Bansal and Yaron (2004) model dividends as three times more volatile than consumption. Using this admittedly aggressive procedure, $\mathbb{E}[\sigma_{R,t}] = 0.0519$ translates into equity return volatility of 15.6%, in line with the data.

The addition of extrapolative beliefs permits substantially lower entry frictions. In particular, panel A of Table 7 is generated with an entry cost of $\phi = 0.45$—half of the baseline entry cost needed in Section 2.4 to generate interesting risk price dynamics. The extrapolation parameters are $\beta = 0.5$ and $\lambda = 0.5$, which we view as reasonable (recall: Barberis et al. (2015) measure $\beta = 0.5$ and $\lambda = 1$ in surveys; $\lambda = 0.5$ is meant to capture in reduced-form the presence of some rational traders).

A. model with frictions: $\phi = 0.45$

Moment:	$\mathbb{E}[\eta_t]$	$\text{std}[\eta_t]$	$\max[\eta_t]$	$\mathbb{E}[r_t]$	$\text{std}[r_t]$	$\mathbb{E}[\sigma_{R,t}]$
Value:	0.2946	0.3075	1.5783	0.0298	0.0094	0.0519

B. frictionless model: $\phi = 0$

Moment:	$\mathbb{E}[\eta_t]$	$\text{std}[\eta_t]$	$\max[\eta_t]$	$\mathbb{E}[r_t]$	$\text{std}[r_t]$	$\mathbb{E}[\sigma_{R,t}]$
Value:	0.2071	0.2916	1.3974	0.0337	0.0087	0.0493

Table 7: Asset pricing moments with recursive preferences and sentiment. The measures are computed with Monte Carlo simulations of length 5000 years. Risk aversion is $\gamma = 3$ and EIS is $\psi^{-1} = 4/3$. Extrapolation parameters are $\beta = 0.5$ and $\lambda = 0.5$. Entry cost is $\phi = 0.45$ (panel A) or $\phi = 0$ (panel B). Other parameters are in Table 1.

For comparison, panel B shows the same results for a representative agent version of the model, similar to Jin and Sui (2021). Notice in particular that the representative agent model generates a substantially lower average Sharpe ratio, but nearly the same Sharpe ratio volatility as the model with frictions. Thus, the additive separability finding from the log utility model—that frictions modulate the average Sharpe ratio, while sentiments modulate Sharpe ratio volatility—continues to hold with recursive preferences.

5 Conclusion

Asset market data suggest enormous costs—on the order of 90% of wealth—associated to the financial frictions in limited participation and intermediary asset pricing models. Simply put, because the compensation to participating in financial markets is large, especially in crisis times, frictions that prevent participation and risk-sharing must be severe.

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 risk premia exceed actual risk premia in booms (and vice versa in busts). Among the mechanisms I consider, this is the only one which can generate both high Sharpe ratio levels and high Sharpe ratio volatility without extreme entry costs.

At a deeper level, my analysis raises the issue of how financial frictions matter for crises, cycles, and the like. For instance, I find that belief dynamics must be the primary driver of Sharpe ratio dynamics. By contrast, the policy conclusions of macro models of financial crises (e.g., time-varying capital requirements, asset purchases, etc.) rely on a strong dependence of asset price dynamics on bank leverage and other balance sheet variables. Whether these policy proposals survive in a frictional world where dynamics are belief-driven is an open question for future research.

References

Achdou, Yves, Jiequn Han, Jean Michel Lasry, Pierre Louis Lions, and Ben Moll. 2020. "Income and wealth distribution in macroeconomics: a continuous-time approach." *Review of Economic Studies (forthcoming)* .

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 and Nina Boyarchenko. 2012. "Intermediary leverage cycles and financial stability."

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.

Adrian, Tobias and Hyun Song Shin. 2010. "Liquidity and Leverage." *Journal of Financial Intermediation* 19 (3):418–437.

———. 2013. "Procyclical leverage and value-at-risk." *The Review of Financial Studies* 27 (2):373–403.

Amromin, Gene and Steven A Sharpe. 2014. "From the horse's mouth: Economic conditions and investor expectations of risk and return." *Management Science* 60 (4):845–866.

Bacchetta, Philippe, Elmar Mertens, and Eric Van Wincoop. 2009. "Predictability in financial markets: What do survey expectations tell us?" *Journal of International Money and Finance* 28 (3):406–426.

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.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny. 1998. "A model of investor sentiment." *Journal of Financial Economics* 49 (3):307–343.

Barles, Guy and Panagiotis E Souganidis. 1991. "Convergence of approximation schemes for fully nonlinear second order equations." *Asymptotic analysis* 4 (3):271–283.

Baron, Matthew and Wei Xiong. 2017. "Credit expansion and neglected crash risk." *The Quarterly Journal of Economics* 132 (2):713–764.

Basak, Suleyman and Benjamin Croitoru. 2000. "Equilibrium mispricing in a capital market with portfolio constraints." *The Review of Financial Studies* 13 (3):715–748.

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

Basak, Suleyman and Alexander Shapiro. 2001. "Value-at-risk-based risk management: optimal policies and asset prices." *The Review of Financial Studies* 14 (2):371–405.

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

Bernanke, Ben S, Mark Gertler, and Simon Gilchrist. 1999. "The financial accelerator in a quantitative business cycle framework." *Handbook of Macroeconomics* 1:1341–1393.

Bianchi, Javier. 2011. "Overborrowing and systemic externalities in the business cycle." *The American Economic Review* 101 (7):3400–3426.

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

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer. 2020. "Overreaction in Macroeconomic Expectations." *The American Economic Review (forthcoming)* .

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.

Cheng, Ing-Haw, Sahil Raina, and Wei Xiong. 2014. "Wall Street and the housing bubble." *American Economic Review* 104 (9):2797–2829.

Christiano, Lawrence J, Roberto Motto, and Massimo Rostagno. 2014. "Risk shocks." *The American Economic Review* 104 (1):27–65.

Collin-Dufresne, Pierre, Michael Johannes, and Lars A Lochstoer. 2017. "Asset pricing when 'this time is different'." *Review of Financial Studies* 30 (2):505–535.

Cutler, David M, James M Poterba, and Lawrence H Summers. 1991. "Speculative dynamics." *The Review of Economic Studies* 58 (3):529–546.

De Long, J Bradford, Andrei Shleifer, Lawrence H Summers, and Robert J Waldmann. 1990. "Positive feedback investment strategies and destabilizing rational speculation." *The Journal of Finance* 45 (2):379–395.

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

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.

Ehling, Paul, Alessandro Graniero, and Christian Heyerdahl-Larsen. 2018. "Asset prices and portfolio choice with learning from experience." *The Review of Economic Studies* 85 (3):1752–1780.

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

Evans, George W and Seppo Honkapohja. 2012. *Learning and expectations in macroeconomics*. Princeton University Press.

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

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.

Gârleanu, Nicolae and Lasse Heje Pedersen. 2011. "Margin-based asset pricing and deviations from the law of one price." *The Review of Financial Studies* 24 (6):1980–2022.

Gennaioli, Nicola, Yueran Ma, and Andrei Shleifer. 2016. "Expectations and investment." *NBER Macroeconomics Annual* 30 (1):379–431.

Gertler, Mark and Peter Karadi. 2011. "A model of unconventional monetary policy." *Journal of monetary Economics* 58 (1):17–34.

Gertler, Mark and Nobuhiro Kiyotaki. 2010. "Financial intermediation and credit policy in business cycle analysis." In *Handbook of monetary economics*, vol. 3. Elsevier, 547–599.

———. 2015. "Banking, liquidity, and bank runs in an infinite horizon economy." *The American Economic Review* 105 (7):2011–43.

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.

———. 2019. "A macroeconomic framework for quantifying systemic risk." *American Economic Journal: Macroeconomics* 11 (4):1–37.

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

Jarrow, Robert A, Philip Protter, and Kazuhiro Shimbo. 2010. "Asset price bubbles in incomplete markets." *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics* 20 (2):145–185.

Jin, Lawrence J and Pengfei Sui. 2021. "Asset pricing with return extrapolation." *Journal of Financial Economics (forthcoming)* .

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.

Kiyotaki, Nobuhiro and John Moore. 1997. "Credit cycles." *Journal of Political Economy* 105 (2):211–248.

Kondor, Peter. 2009. "Risk in dynamic arbitrage: the price effects of convergence trading." *The Journal of Finance* 64 (2):631–655.

Krishnamurthy, Arvind and Wenhao Li. 2020. "Dissecting mechanisms of financial crises: Intermediation and sentiment." Unpublished working paper.

Krusell, Per and Anthony A Smith, Jr. 1998. "Income and wealth heterogeneity in the macroeconomy." *Journal of Political Economy* 106 (5):867–896.

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.

Malmendier, Ulrike, Demian Pouzo, and Victoria Vanasco. 2020. "Investor experiences and financial market dynamics." *Journal of Financial Economics* 136 (3):597–622.

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

Maxted, Peter. 2020. "A macro-finance model with sentiment." Unpublished working paper.

Mendoza, Enrique G. 2010. "Sudden stops, financial crises, and leverage." *The American Economic Review* 100 (5):1941–66.

Muir, Tyler. 2017. "Financial crises and risk premia." *The Quarterly Journal of Economics* 132 (2):765–809.

Nagel, Stefan and Zhengyang Xu. 2019. "Asset pricing with fading memory." *Review of Financial Studies (forthcoming)* .

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

Schraeder, Stefanie. 2016. "Information processing and non-bayesian learning in financial markets." *Review of Finance* 20 (2):823–853.

Siriwardane, Emil N. 2019. "Limited investment capital and credit spreads." *The Journal of Finance* 74 (5):2303–2347.

Vissing-Jorgensen, Annette. 2003. "Perspectives on behavioral finance: Does" irrationality" disappear with wealth? Evidence from expectations and actions." *NBER macroeconomics annual* 18:139–194.

Wang, Chong, Neng Wang, and Jinqiang Yang. 2016. "Optimal consumption and savings with stochastic income and recursive utility." *Journal of Economic Theory* 165:292–331.

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.

Online Appendix:

Entry and slow-moving capital

Paymon Khorrami

October 28, 2021

The appendix proceeds as follows. Section A provides all proofs and auxiliary results for the benchmark model with log utility and its extensions. Section B provides results for the model with recursive utility. Section C provides results on the critical extension to extrapolative beliefs. Section D discusses the robustness of the model to several different extensions: a fixed entry cost, the presence of labor income, and alternative preference constellations. Finally, Section E shows how to compute the dynamic evolution of the (infinite-dimensional) wealth distribution in extensions lacking homogeneity properties (e.g., fixed entry costs).

A Log utility model

Section A.1 provides proofs of Propositions 1-4 for the benchmark model with log utility. Section A.2 discusses the fact that endogenous entry rules out bubbles and arbitrages in the economy, which are present in many limited participation economies. Section A.3 contrasts the endogenous-entry economy to one with exogenous entry. Section A.4 adds both equity issuance and idiosyncratic risks, and then derives the equilibrium.

A.1 Proofs for the benchmark log utility model

Proof of Proposition 1. The entirety of the proof proceeds exactly as Proposition B.1 (the recursive-utility generalization). Alternatively, one may simply taking limits $\gamma, \psi \to 1$ in the expressions from Proposition B.1. 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 2. In this proof, assume that ϕ is large enough so that (25) is not degenerate, i.e., there is not a solution with $x^* = 1$. (In particular, a sufficient condition to guarantee $x^* < 1$ is that $\Phi > \frac{1}{2}(\rho+\pi)^{-2}\sigma_Y^2$, from the result of Proposition 4.) In addition, the functions η, μ_X, and σ_X are taken to be the extensions to $(0,1)$ of the functions from Proposition 1 (i.e., as if $x^* \to 0$ in the expressions).

Existence. Define the linear differential operator $\mathscr{L}$ that applies to C^2-a.e. functions f on $(0,1)$:

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

Let δ be small enough, and consider the following problem. Find a function φ such that on $(\delta, 1)$ the following hold:

$$\begin{cases} 0 & \leq & \frac{1}{2}(\rho+\pi)^{-1}\eta^2 + \mathscr{L}\varphi, \\ 0 & \geq & \varphi - \Phi, \\ 0 & = & \Big(\frac{1}{2}(\rho+\pi)^{-1}\eta^2 + \mathscr{L}\varphi\Big)(\varphi - \Phi), \end{cases} \tag{A.2}$$

subject to boundary conditions $\varphi(\delta) = \Phi$ and $\frac{1}{2}(\rho+\pi)^{-1}\eta^2(1) - (\rho+\pi)\varphi(1) + \mu_X(1)\varphi'(1) = 0$. Note that none of the objects in problem (A.2) depend on the equilibrium entry point (i.e., one can think of this problem as a "partial equilibrium" problem). One can consult any reference on free boundary problems and variational inequalities to find (A.2), augmented with the boundary conditions above, has a unique $C^1(\delta, 1)$ solution, which is also C^2-a.e. (c.f. Bensoussan and Lions (1982), chapter 3.1, or Friedman (2010), chapter 1.2).

Let $\mathcal{X}^* \subset (\delta, 1)$ be the set of points where $\varphi = \Phi$ (stopping set), and define $x^* := \sup \mathcal{X}^*$. If δ was chosen small enough, then $\mathcal{X}^*$ is non-empty and so $1 > x^* > \delta > 0$. Put $\Delta g(x) = \varphi(x)$ for all $x \in (x^*, 1)$ and $\Delta g(x) = \Phi$ for all $x \in (0, x^*]$. By construction, value-matching $\Delta g(x^*) = \Phi$ holds. In addition, since Δg is $C^1(0, 1)$, and since $\Delta g'(x) = 0$ for all $x < x^*$, we have $\Delta g'(x^*) = 0$. Finally, the first line of (A.2) holds with equality for $x > x^*$, so the ODE (25) holds. This proves that $(\Delta g, x^*)$ constitute a solution to (25)-(27).

Uniqueness. Let (φ_1, x_1^*) and (φ_2, x_2^*) be two distinct solutions to (25)-(27), and suppose $x_1^* < x_2^*$ without loss of generality. We have $\varphi_1(x_1^*) = \varphi_2(x_2^*) = \Phi$ by (26). Using (A.4), we then have

$$\mathbb{E}^{x_1^*}\Big[\int_0^\infty e^{-(\rho+\pi)t}\eta^2(X_t^{(1)})dt\Big] = \mathbb{E}^{x_2^*}\Big[\int_0^\infty e^{-(\rho+\pi)t}\eta^2(X_t^{(2)})dt\Big], \tag{A.3}$$

where $X_t^{(1)}$ and $X_t^{(2)}$ are processes for X_t with reflecting barriers at x_1^* and x_2^* respectively. In the above, the initial values are $X_0^{(1)} = x_1^*$ and $X_0^{(2)} = x_2^* > X_0^{(1)}$. Path-by-path comparison implies $X_t^{(1)} \leq X_t^{(2)}$ almost-surely, with the inequality strict on a positive-Lebesgue-measure subset of times. Since $\eta(x)$ is a decreasing function, we therefore have $\eta^2(X_t^{(1)}) \geq \eta^2(X_t^{(2)})$, with strict inequality on a positive-measure subset. This fact contradicts (A.3) above and implies there cannot be two solutions. □

Proof of Proposition 3. The result is proved in equation (A.4) in Proposition A.1 below. □

Proof of Proposition 4. Substitute $X_t \equiv 1$ in expression (29). To obtain the approximation in σ_Y^2, expand ϕ^* around $\sigma_Y^2 = 0$ and substitute $\eta^* := \eta(1) = \sigma_Y$. □

Proposition A.1. *The function Δg can be equivalently represented in the following three ways:*

$$\Delta g(x) = \mathbb{E}^x\Big[\frac{1}{2}(\rho+\pi)^{-1}\int_0^\infty e^{-(\rho+\pi)t}\eta^2(X_t)dt\Big] \tag{A.4}$$

$$= \mathbb{E}^x\Big[\frac{1}{2}(\rho+\pi)^{-1}\int_0^{\tau_{x^*}} e^{-(\rho+\pi)t}\eta^2(X_t)dt + e^{-(\rho+\pi)\tau_{x^*}}\Phi\Big] \tag{A.5}$$

$$= \inf_\tau \mathbb{E}^x\Big[\frac{1}{2}(\rho+\pi)^{-1}\int_0^{\tau} e^{-(\rho+\pi)t}\eta^2(X_t)dt + e^{-(\rho+\pi)\tau}\Phi\Big], \tag{A.6}$$

where $\tau_{x^*} := \inf\{t \geq 0 : X_t = x^*\}$, the minimization in (A.6) 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 $X_0 = x$.

Proof of Proposition A.1. 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

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

where the differential operator $\mathscr{L}$ is defined by (A.1). Since ODE (25) holds on $(x^*, 1)$, and since $\{t : X_t = x^*\}$ has zero Lebesgue measure a.s., the sum of the first two integrals is zero a.s. As Δg is C^2 and σ_X is bounded, the stochastic integral is a martingale. As $\Delta g'(x^*) = 0$ by (27), the last integral is zero. Hence, M_t is a martingale, so by Doob's optional sampling, we have

$$\Delta g(x) = \mathbb{E}^x\Big[\frac{1}{2}(\rho+\pi)^{-1}\int_0^{T\wedge\tau} e^{-(\rho+\pi)t}\eta^2(X_t)dt + e^{-(\rho+\pi)(T\wedge\tau)}\Delta g(X_{T\wedge\tau})\Big]. \tag{A.7}$$

Using (A.7), we can prove (A.4)-(A.6). Result (A.4) follows by picking $\tau = +\infty$, performing recursive substitution of $\Delta g(X_T)$ on the right-hand-side of (A.7), applying the Strong Markov property, and finally taking $T \to +\infty$ with the monotone convergence theorem. Result (A.5) 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.). Result (A.6) follows by noting that $\Delta g(X_{T\wedge\tau}) \leq \Phi$ (see equation (11)), so that the objective function of (A.6) exceeds $\Delta g(x)$ for any choice of τ. But choosing $\tau = \tau_{x^*}$ is feasible, and so equation (A.5) implies equation (A.6). □

A.2 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 the cheapest replicating portfolio. 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 *fundamental value*, or cheapest replicating cost, for a sequence of cash flows $\{G_t\}$ is

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

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

Proposition A.2. *Consider the equilibrium of Proposition 1. For $\phi = 1$, the economy contains bubbles. For any $\phi < 1$, the economy has no bubbles. In both economies, $\mathbb{P}\{X_t > 0\, \forall t\} = 1$.*

Proposition A.2 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 A.2. 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.3, 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 the fact that it is positive and non-constant (this means ξ_t^* is a super-martingale that is not a martingale). This immediately implies that the risk-free asset has a bubble: investing 1 in the risk-free bond at time t delivers $\exp(\int_t^T r_s ds)$ at time T, whereas the fundamental value of time-T cash flow $\exp(\int_t^T r_s ds)$ is $F_t^* = \mathbb{E}_t[\frac{\xi_T}{\xi_t} \exp(\int_t^T r_s ds)] = \mathbb{E}_t[\frac{\xi_T^*}{\xi_t^*}] < 1$. In other words, the risk-free asset is more expensive than its fundamental value. A related argument can be applied to the price of the aggregate endowment, $G = Y$. See Loewenstein and Willard (2000) and Jarrow, Protter, and Shimbo (2010) 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 B.1, which shows that ξ_t defines a true state-price density process, with bounded market price of risk, i.e., ξ_t^* is a true martingale. Consequently, there cannot be any arbitrages or bubbles. □

A.3 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 $c_{t,b}^N$.

This model admits a stationary equilibrium, described in Proposition A.3 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 A.3. *In the model with log utility and exogenous entry (with entry parameter ν), the following is the unique equilibrium. Asset prices are given by*

$$\eta(x) = \frac{\sigma_Y}{x} \quad \text{and} \quad r(x) = \rho + \pi + \mu_Y - \frac{\sigma_Y^2}{x}, \quad x \in (0,1)$$

and state dynamics by

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

The non-degenerate stationary density of X_t is given by

$$h_\nu(x) = \frac{K_0}{\sigma_X^2(x)} \int_0^x \Big(\frac{x}{z}\Big)^2 \Big(\frac{1-x}{1-z}\Big)^{-\frac{2\pi(1-\alpha)}{\sigma_Y^2}} \exp\Big(-\frac{2\pi(1-\alpha)(1-\nu)(x-z)}{\sigma_Y^2(1-x)(1-z)}\Big) dz, \tag{A.8}$$

where K_0 is a constant ensuring h_ν integrates to 1.

Proof of Proposition A.3. 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 A.3, 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. $\square$

Figure A.1 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.

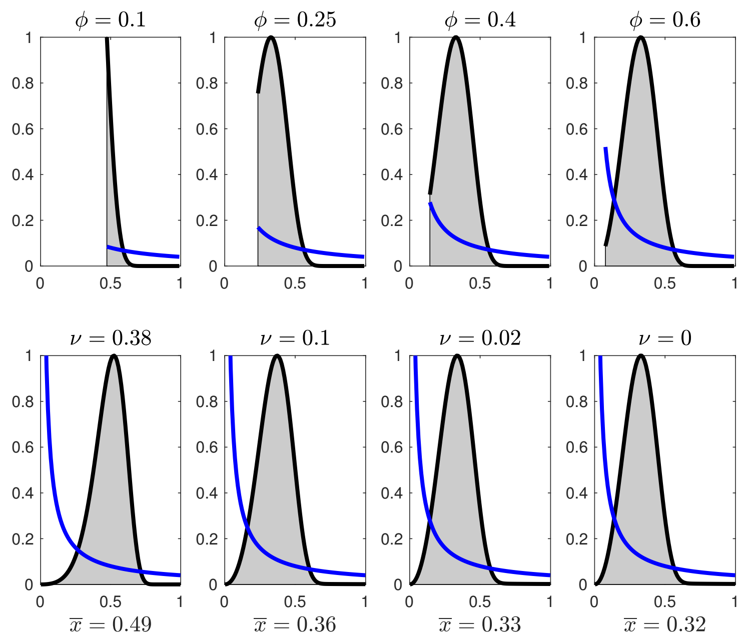

Figure A.1: Each plot features the market price of risk η (blue line) and stationary density for X (gray area). The horizontal axis is the participants' consumption share x. Top four panels: Each plot corresponds to a different entry cost ϕ. Bottom four panels: Each plot corresponds to a different participant 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.

For relatively small costs (e.g., $\phi = 0.10$, 0.25), Figure A.1 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 much less time in low-x states.

For larger costs (e.g., $\phi = 0.40$, 0.60), there is less of a distinction between endogenous and exogenous entry. Asset prices become increasingly similar between the models as entry is eliminated.

A.4 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.

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

$$\int_{\mathcal{P}_t} \pi e^{-\pi(t-b)}(1-\chi_{t,b})\theta_{t,b}W_{t,b}db = \int_{\mathcal{P}_t \cup \mathcal{N}_t} \pi e^{-\pi(t-b)}\tilde{\theta}_{t,b}W_{t,b}db, \tag{A.9}$$

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

$$\chi_{t,b}\sigma_{R,t}\eta_t + \chi_{t,b}\hat{\sigma}_Y\hat{\eta}_t := \mu_{R,t} - r_t - (1-\chi_{t,b})\sigma_{R,t}\tilde{\eta}_t. \tag{A.10}$$

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

$$\frac{dW_{t,b}}{W_{t,b}} = \Big(r_t + \tilde{\theta}_{t,b}\sigma_{R,t}\tilde{\eta}_t + \alpha\pi - \frac{c_{t,b}}{W_{t,b}}\Big)dt + \tilde{\theta}_{t,b}\sigma_{R,t}dZ_t, \quad t < \tau_b \tag{A.11}$$

$$\begin{aligned}\frac{dW_{t,b}}{W_{t,b}} &= \Big(r_t + \chi_{t,b}\theta_{t,b}(\sigma_{R,t}\eta_t + \hat{\sigma}_Y\hat{\eta}_t) + \tilde{\theta}_{t,b}\sigma_{R,t}\tilde{\eta}_t + \alpha\pi - \frac{c_{t,b}}{W_{t,b}}\Big)dt \\ &\quad + (\chi_{t,b}\theta_{t,b} + \tilde{\theta}_{t,b})\sigma_{R,t}dZ_t + \chi_{t,b}\theta_{t,b}\hat{\sigma}_Y d\hat{Z}_t, \quad t \geq \tau_b. \end{aligned} \tag{A.12}$$

Equilibrium is given by the following proposition.

Proposition A.4. *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 (A.13). Equilibrium objects are given by the following set of functions of x which hold for $x \in [x^*, 1]$:*

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

and the 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 $x \in [x^*, \chi^*)$ and $h(x) = 0$ for $x \notin [x^*, \chi^*)$, where K_0 is a constant ensuring h integrates to 1, i.e., $\int_{x^*}^{\chi^*} h(x)dx = 1$.

Proof of Proposition A.4. 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

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

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 (A.9). 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 (A.10), 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:

$$dW_t^P = W_t^P\Big[\Big(r_t + \alpha\pi - \rho - \pi + \eta_t^2 + \hat{\eta}_t^2\Big)dt + \eta_t dZ_t + \hat{\eta}_t d\hat{Z}_t\Big]$$
$$dW_t^N = W_t^N\Big[\Big(r_t + \alpha\pi - \rho - \pi + \tilde{\eta}_t^2\Big)dt + \tilde{\eta}_t dZ_t\Big].$$

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 $(x^*, 1)$:

$$0 = \frac{1}{2}(\rho+\pi)^{-1}\Big[\eta^2+\hat{\eta}^2-\tilde{\eta}^2\Big]-(\rho+\pi)\Delta g+\mu_X\Delta g'+\frac{1}{2}\sigma_X^2\Delta g'', \quad \Delta g(x^*) = \Phi, \quad \Delta g'(x^*) = 0. \tag{A.13}$$

This is derived by taking the difference between participants' and non-participants HJB equations, as in the discussion leading up to equation (25).

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. □

Lemma A.1. *Entry occurs at a time τ when the following holds:*

$$\Phi = \mathbb{E}\Big[\frac{1}{2}(\rho+\pi)^{-1}\int_\tau^\infty e^{-(\rho+\pi)(t-\tau)}\Big(\eta_t^2 + \hat{\eta}_t^2 - \tilde{\eta}_t^2\Big)dt \mid \mathcal{F}_\tau\Big]. \tag{A.14}$$

Proof of Lemma A.1. Start with equation (A.13) and proceed as in Proposition 3. □

Proof of Proposition 7. Let $\omega_t^{\chi^*} := X_t/\chi^*$ be a revised state variable. The revised state dynamics are

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

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), start with the result of Lemma A.1 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 $\omega_t^{\chi^*} \geq \omega^*(\chi^*)$, the latter of which is a reflecting boundary, it is obvious 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, the path of $\eta_t = \sigma_Y/\omega_t^{\chi^*}$ is uniformly increasing in χ^* for all t. Similar arguments show $\tilde{\eta}_t = \sigma_Y/(1-\tilde{\omega}_t^{\chi^*})$ is uniformly decreasing in χ^* for all t. As a result, the right-hand-side of equation (A.14) 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 χ^*.

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

Proof of Proposition 8. 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 9. Substitute $X_t \equiv 1$ in expression (37). □

B Recursive utility model

In Section B.1, I generalize the model by introducing recursive preferences and derive the equilibrium. Section B.2 provides a verification theorem, proving that the HJB equations and associated boundary conditions are sufficient for optimality in individual agents' control problems. Section B.3 provides details on the full-integration equilibrium (i.e., for ϕ small enough such that $x^* = 1$). Section B.4 provides details on the equilibrium with asymptotically large entry costs (i.e., $\phi \to 1$). Section B.5 reviews the quantitative appeal of adding recursive Epstein-Zin preferences to the model.

B.1 Details and equilibrium for recursive utility generalization

This section presents more details on the model environment under the recursive utility generalization. The utility function is defined by (31)-(32) in the text, which we restate here for convenience:

$$V_{t,b} := \mathbb{E}\Big[\int_t^\infty f(c_{s,b}, V_{s,b})ds \mid \mathcal{F}_t\Big], \tag{B.1}$$

where the felicity function f is defined by

$$f(c, V) := \frac{1}{1-\psi}\Big(c^{1-\psi}[V(1-\gamma)]^{\frac{\psi-\gamma}{1-\gamma}} - (\rho+\pi)V(1-\gamma)\Big). \tag{B.2}$$

In (B.2), parameter γ is the coefficient of relative risk aversion (RRA), and ψ^{-1} is the elasticity of intertemporal substitution (EIS). Assume $\gamma, \psi \neq 1$. The death rate π simply augments the subjective discount rate ρ, as shown by Gârleanu and Panageas (2015). For reference, we also restate here the modified participation cost $\Phi_{t,b}$, which now has a time and cohort dimension and is given by

$$\Phi_{t,b} := [1-(1-\phi)^{1-\gamma}]V^P_{t,b}. \tag{B.3}$$

In the equilibrium of this model, the returns on the risky asset will need to be modified to

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

where A^R is a non-decreasing, singularly continuous process.[27] The bond pays an instantaneous return of $r_t dt + dA^R_t$.[28] The presence of the singular term dA^R_t, though unusual, is due to the equilibrium entry in the model in conjunction with these more general preferences. The state-price

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

[28]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.

density process is modified to

$$\xi_t := \exp\Big\{ - \int_{-\infty}^{t} \Big(r_s + dA_s^R + \frac{1}{2}\eta_s^2 \Big) ds - \int_{-\infty}^{t} \eta_s dZ_s \Big\}. \tag{B.4}$$

Participants' wealth dynamics are now given by

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

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

$$V_{t,b}^P = \sup_{c,\theta} \mathbb{E}\Big[\int_t^\infty f(c_{s,b}, V_{s,b}^P) ds \mid \mathcal{F}_t \Big], \tag{B.6}$$

subject to (B.5). Non-participants solve

$$V_{t,b}^N = \sup_{c,\tau} \mathbb{E}\Big[\int_t^\tau f(c_{s,b}, V_{s,b}^N) ds + V_{\tau,b}^P - \Phi_{\tau,b} \mid \mathcal{F}_t \Big], \tag{B.7}$$

where wealth dynamics are given by

$$dW_{t,b} = (r_t W_{t,b} + \alpha\pi W_{t,b} - c_{t,b})dt + W_{t,b} dA_t^R, \quad t < \tau_b, \quad W_{b,b} > 0 \quad \text{given.} \tag{B.8}$$

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

$$V_{t,b}^P = \frac{W_{t,b}^{1-\gamma}}{1-\gamma} G_t^P \quad \text{and} \quad V_{t,b}^N = \frac{W_{t,b}^{1-\gamma}}{1-\gamma} G_t^N,$$

where G^P and G^N are processes independent of agents' wealth. The key to achieving this is the homogeneity of the entry cost (B.3), so that the payoff to an entrant at time $t \geq b$ is $V_{t,b}^P - \Phi_{t,b} = (1-\phi)^{1-\gamma} V_{t,b}^P = \frac{((1-\phi)W_{t,b})^{1-\gamma}}{1-\gamma} G_t^P$. 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

$$t \in \mathcal{T}^* : \ (1-\phi)^{1-\gamma} G_t^P = G_t^N; \tag{B.9}$$

$$t \notin \mathcal{T}^* : \ (1-\phi)^{1-\gamma} G_t^P (1-\gamma)^{-1} < G_t^N (1-\gamma)^{-1}, \tag{B.10}$$

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

Solving for Markov equilibrium. In a Markov equilibrium with state variable X_t (participants' consumption share), the dynamics of X_t are given by

$$dX_t = \mu_X(X_t)dt + \sigma_X(X_t)dZ_t + dA_t^{x^*}. \tag{B.11}$$

In addition, there are functions g_P and g_N such that $G_t^P = g_P(X_t)^{\frac{\psi(1-\gamma)}{1-\psi}}$ and $G_t^N = g_N(X_t)^{\frac{\psi(1-\gamma)}{1-\psi}}$. 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:

$$\begin{aligned}
0 &= \psi + \Big[-\rho - \pi + (1-\psi)\Big(r + \alpha\pi + \frac{1}{2\gamma}\eta^2\Big)\Big] g_P + \Big[\psi\mu_X + \frac{\psi}{\gamma}(1-\gamma)\eta\sigma_X\Big] g_P' + \frac{1}{2}\psi\sigma_X^2 g_P'' \\
&\quad + \frac{1}{2}\frac{\psi(\psi-\gamma)}{\gamma(1-\psi)}\sigma_X^2 \frac{(g_P')^2}{g_P} \qquad \text{(B.12)} \\
0 &= \psi + \Big[-\rho - \pi + (1-\psi)(r + \alpha\pi)\Big] g_N + \psi\mu_X g_N' + \frac{1}{2}\psi\sigma_X^2 g_N'' \\
&\quad + \frac{1}{2}\Big(\frac{\psi(\psi-\gamma)}{\gamma(1-\psi)} - \frac{\psi^2(1-\gamma)^2}{\gamma(1-\psi)}\Big)\sigma_X^2 \frac{(g_N')^2}{g_N}. \qquad \text{(B.13)}
\end{aligned}$$

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 (B.9) at entry times, i.e., times where $X_t = x^*$, implying the *value-matching* condition

$$(1-\phi)^{\frac{1-\psi}{\psi}} g_P(x^*) = g_N(x^*). \qquad \text{(B.14)}$$

Next, the *smooth-pasting* conditions

$$g_P'(x^*) = g_N'(x^*) = 0 \qquad \text{(B.15)}$$

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 (B.12)-(B.13) as $x \to 1$. In Appendix B.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 B.2).

In Proposition B.1, I solve for all equilibrium objects, up to the solutions g_P and g_N to (B.12)-(B.13). 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 B.1. *There exists a stationary Markov equilibrium defined by asset prices*

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

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*

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

and functions g_P *and* g_N, *with endogenous entry point* x^*, *satisfy the ordinary differential equations (B.12) and (B.13) subject to boundary conditions given by (B.14), (B.15), (B.23), and (B.24), assuming these ODEs have a solution. In that case,* (X, A^{x^*}) *is the unique strong solution to (B.11). 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. With recursive utility, equilibrium dynamics are more complicated, but strong effects of limited participation remain.

Indeed, notice that equilibrium objects can be decomposed into terms from a frictionless economy, terms arising due to limited participation ("LP effects"), and terms due to recursive preferences. For example, the market price of risk can be understood this way:

$$\eta(x) \quad = \underbrace{\gamma\sigma_Y}_{\text{frictionless}} + \underbrace{\gamma\sigma_Y\Big(\frac{1-x}{x}\Big)\omega(x)}_{\text{LP effects}} + \underbrace{\Big(\frac{\gamma-\psi}{1-\psi}\Big)\sigma_Y(1-x)\frac{g_P'(x)}{g_P(x)}}_{\text{recursive preferences}}. \tag{B.16}$$

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:

$$r(x) = \underbrace{\rho + \psi\mu_Y - \frac{1}{2}\gamma(\psi+1)\sigma_Y^2}_{\text{frictionless}} + \underbrace{\pi(1-\alpha) + \psi\pi - \psi\pi(1-\alpha)\frac{p(x)}{g_N(x)}}_{\text{OLG effects}} - \underbrace{\frac{1}{2}\gamma(\psi+1)\sigma_Y^2\Big(\frac{1-x}{x}\Big)\omega^2(x)}_{\text{LP effects}}$$
$$+ \underbrace{\frac{1}{2}\Big(\frac{\gamma-\psi}{1-\psi}\Big)\sigma_Y^2(1-x)(1-\omega(x))^2 - \Big(\frac{\gamma-\psi}{1-\psi}\Big)\sigma_Y^2\Big[x\zeta(x) + \frac{x}{2}\zeta^2(x) + (1-x)\omega(x)\zeta(x)\Big]}_{\text{recursive preferences}}. \tag{B.17}$$

The contributions of the "LP effects" terms are generally to lower r and increase η, qualitatively the same as the log utility model.

Note on entry and singular asset prices. Examining the proof of Proposition B.1 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 $dR_t = \mu_{R,t}dt + \sigma_{R,t}dZ_t + dA_t^R$.[29] 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.

Below, I prove Proposition B.1. The proof takes as given agents' optimal controls to construct

[29] The fact that this singular component is non-zero can be attributed to the non-zero derivative of the price-dividend ratio at x^*. Combining $p(x) = xg_P(x) + (1-x)g_N(x)$ with the value-matching and smooth-pasting conditions (B.14)-(B.15),

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

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.

an equilibrium. These optimal controls are contained in (B.21), (B.22), and (B.28) in Appendix B.2. For verification of the optimality of these controls, see Proposition B.2 in Appendix B.2.

Proof of Proposition B.1. 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.

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

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

for participants, and

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

for non-participants. Substituting the second derivatives from the HJB equations (B.12)-(B.13), we have

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

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

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

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

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

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 $p_t := P_t/Y_t$, 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 (B.11), 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 (B.11), 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.3 in Appendix B.4, we could prove that any solution to (B.11) must satisfy $\mathbb{P}\{X_t < 1, \forall t \geq 0\} = 1$ (although Lemma B.3 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 (B.15), 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 (B.23)-(B.24) 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 (B.4) exists and is consistent with absence of arbitrage.

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

$$0 = -\mu_X h + \frac{1}{2}\frac{d}{dx}(\sigma_X^2 h). \tag{B.18}$$

Equation (B.18) 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 (B.18) 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 $[x^*, 1)$. □

Proof of Proposition 6. Take $x \to x^*$ and $x \to 1$ in Proposition B.1, using the smooth-pasting condition (B.15) to find $\eta(x^*) = \gamma\sigma_Y/x^*$ and $\eta(1) = \gamma\sigma_Y$. □

B.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 (B.5) 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 (B.2). 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

$$b \in \mathcal{P}_t : \ V_{t,b}^P = \sup_{c,\theta\in\mathcal{A}_t^P} \mathbb{E}\Big[\int_t^\infty f(c_{s,b}, V_{s,b}^P)ds \mid \mathcal{F}_t\Big] \tag{B.19}$$

$$b \in \mathcal{N}_t : \ V_{t,b}^N = \sup_{c,\tau\in\mathcal{A}_t^N} \mathbb{E}\Big[\int_t^\tau f(c_{s,b}, V_{s,b}^N)ds + V_{\tau,b}^P - \Phi_{\tau,b} \mid \mathcal{F}_t\Big], \tag{B.20}$$

where $\Phi_{t,b}$ is given by (B.3).

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 $\theta \equiv 0$):

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

where the second line uses the homogeneity property $V(w,x) = \frac{w^{1-\gamma}}{1-\gamma}g(x)^{\frac{\psi(1-\gamma)}{1-\psi}}$. The FOCs for c

and θ imply

$$c_{t,b} = \begin{cases} W_{t,b} g_P(X_t)^{-1}, & \text{if } b \in \mathcal{P}_t; \\ W_{t,b} g_N(X_t)^{-1}, & \text{if } b \in \mathcal{N}_t. \end{cases} \tag{B.21}$$

and

$$\theta_{t,b} = \frac{\mu_{R,t} - r_t}{\gamma \sigma_{R,t}^2} + \frac{\sigma_{X,t}}{\sigma_{R,t}} \frac{\psi(1-\gamma)}{\gamma(1-\psi)} \frac{g_P'(X_t)}{g_P(X_t)} \quad \text{if } b \in \mathcal{P}_t. \tag{B.22}$$

Substituting these back into the HJB equations, we get two ODEs for the wealth-consumption ratios, given by (B.12) and (B.13).

Boundary conditions.

The boundary conditions for (B.12)-(B.13) at $x = x^*$ are given by the value-matching and smooth-pasting conditions (B.14)-(B.15) 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 $x \to 1$,

$$0 = \psi + \Big[-\rho - \pi + (1-\psi)\Big(r(1) + \alpha\pi + \frac{1}{2\gamma}\eta^2(1) \Big) \Big] g_P(1) + \psi \mu_X(1) g_P'(1) \tag{B.23}$$

$$0 = \psi + \Big[-\rho - \pi + (1-\psi)\Big(r(1) + \alpha\pi \Big) \Big] g_N(1) + \psi \mu_X(1) g_N'(1). \tag{B.24}$$

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

Proposition B.2 (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*

$$J^i(w, x) = \frac{w^{1-\gamma}}{1-\gamma} g_i(x)^{\frac{\psi(1-\gamma)}{1-\psi}}, \tag{B.25}$$

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*

$$\mathcal{D}^N J^N := w\big(r + \alpha\pi - \tfrac{c}{w}\big)\partial_w J^N + \mu_X \partial_x J^N + \tfrac{1}{2}\sigma_X^2 \partial_{xx} J^N. \tag{B.26}$$

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

$$\begin{aligned}\mathcal{D}^P J^P &:= w\big(r + \alpha\pi + \theta(\mu_R - r) - \tfrac{c}{w}\big)\partial_w J^P + \mu_X \partial_x J^P \\ &\quad + \tfrac{1}{2}w^2\theta^2\sigma_R^2 \partial_{ww} J^P + \tfrac{1}{2}\sigma_X^2 \partial_{xx} J^P + w\theta\sigma_R\sigma_X \partial_{wx} J^P.\end{aligned} \tag{B.27}$$

(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* $\mathcal{H}^2 := \{h : \mathbb{E}\int_{-T}^T |h_t|^2 dt < +\infty, \text{ for all } T\}$.

(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 $\{V_{t,b}^P\}_{t\geq b}$ *and* $\{V_{t,b}^N\}_{t\geq b}$ *are unique solutions to (B.19)-(B.20), 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 (B.21),* $\theta_{t,b}$ *in (B.22), and*

$$\tau_b := \inf\{t \geq b : (W_t, X_t) \in \mathcal{S}\}. \tag{B.28}$$

Proof of Proposition B.2. 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

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

where $\mathcal{D}^{P,a}$ is defined by (B.27) 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 (B.25), $W_s^a\theta_s\sigma_{R,s}\partial_w J^P(W_s^a, X_s) = (1-\gamma)\theta_s\sigma_{R,s}J^P(W_s^a, X_s)$ 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 $\mathcal{D}^{P,a}J^P(w,x) + f(c, J^P(w,x)) \leq 0$ for all (w,x) and $a \in \mathcal{A}_t^P$. Taking expectations and using these results,

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

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 (B.25). Using condition (v) and the monotone convergence theorem to take the limit as $n \to \infty$, then maximizing over feasible controls, we obtain

$$J^P(W_t, X_t) \geq \sup_{a \in \mathcal{A}_t^P} \mathbb{E}\Big[\int_t^\infty f(c_s, J^P(W_s^a, X_s))ds \mid \mathcal{F}_t\Big]. \tag{B.29}$$

On the other hand, consider the admissible control $a = (c, \theta)$ given by (B.21) and (B.22) and suppose this control also satisfies integrability condition (iv). Representation (B.25) 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 (B.29), 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 (B.19). Since (B.19) has the unique solution $V_{t,b}^P$, we have

$$J^P(W_{t,b}, X_t) = V_{t,b}^P. \tag{B.30}$$

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

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

where $\mathcal{D}^{N,a}$ is defined by (B.26) 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

$$J^N(W_t, X_t) \geq \sup_{a \in \mathcal{A}_t^N} \mathbb{E}\Big[\int_t^\tau f(c_s, J^N(W_s^a, X_s))ds + J^P((1-\phi)W_\tau^a, X_\tau) \mid \mathcal{F}_t\Big]. \tag{B.31}$$

On the other hand, consider the admissible control $a = (c, \tau)$ given by (B.21) and (B.28), 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

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

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

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

Finally, we can use the form of J^P in (B.25), our previous result (B.30), and the equation for the entry cost (33) to get $J^P((1-\phi)W^a_\tau, X_\tau) = (1-\phi)^{1-\gamma}J^P(W^a_\tau, X_\tau) = (1-\phi)^{1-\gamma}V^P_\tau = V^P_\tau - \Phi_\tau$. Thus,

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

Combining with inequality (B.31), this shows that control a attains the maximum in $\mathcal{A}^N_t$. Exactly as before, the uniqueness of $V^N_{t,b}$ as a solution to (B.20) implies $J^N(W_{t,b}, X_t) = V^N_{t,b}$. □

Remark B.1 (Verify conditions of Proposition B.2)**.** *To verify optimality of $c_{t,b}$, $\theta_{t,b}$, and τ_b in (B.21), (B.22), and (B.28), it suffices to verify that the conditions of Proposition B.2 are satisfied for any $x^* \in (0,1)$. First, as proved in Proposition B.1, the SDE $dX_t = \mu_X(X_t)dt + \sigma_X(X_t)dZ_t + dA^{x^*}_t$ indeed has a unique strong solution in equilibrium. Second, the homogeneity of problems (B.19)-(B.20) 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 (B.25) is satisfied by V^P and V^N. Third, if there is a solution to the ODEs (B.12)-(B.13), subject to boundary conditions (B.14)-(B.15) at $x = x^*$ and (B.23)-(B.24) at $x = 1$, then conditions (i), (ii), and (iii) of Proposition B.2 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. Fourth, 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 (B.15), (B.23), and (B.24). Finally, transversality condition (v) is verified below in Lemma B.1.*

Lemma B.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 of Lemma B.1. 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. $\square$

B.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. One can obtain the analogous log utility model results by taking limits $\gamma, \psi \to 1$.

Proposition B.3. *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 B.3 shows that all formulas in Proposition B.1 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 of Proposition B.3. 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 5. 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 (B.12)-(B.13) 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 B.3. □

B.4 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.2 and B.3.

As Lemma B.2 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.2 shows that the market price of risk process η increases without bound as x approaches 0.

Lemma B.2. *Let $\{\phi_n\}_{n=1}^{\infty}$ be a sequence of entry costs such that $x_n^* > 0$ for all n, and such that $\phi_n \to 1$ and consequently $x_n^* \to 0$ as $n \to \infty$. 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.

Proof of Lemma B.2. 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 B.1 at $x = \kappa x^*$, for small x^*. To do this, first note that

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

Then, we obtain

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

and

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

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. □

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,

$$\begin{aligned} &(\text{near } x=1) \qquad \lim_{x\uparrow 1}\sigma_X(x) = 0 \quad \text{and} \quad \lim_{x\uparrow 1}\mu_X(x) < 0 \\ &(\text{near } x=0) \qquad \lim_{x\downarrow 0}\sigma_X(x) < +\infty \quad \text{and} \quad \lim_{x\downarrow 0}\mu_X(x) = +\infty. \end{aligned}$$

The top two conditions are enough to ensure that X_t does not reach 1; the bottom 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.2. The question of stationarity is addressed formally in Lemma B.3, which gives an affirmative answer. In particular, Lemma B.3 implies that the complete segmentation model stays away from $x = 0$ with probability 1, i.e., entry never occurs.

Lemma B.3. *Let the assumptions of Lemma B.2 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.*

Proof of Lemma B.3. 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.2 shows that X_t evolves approximately with

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

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

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

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

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

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

B.5 Benefits of recursive preferences for asset pricing

In this section, we discuss the (quantitative) benefits of recursive Epstein-Zin preferences in the model. The effects of this utility specification on asset prices are relatively well-understood in the representative agent asset pricing literature, and several of those well-known features carry over to this limited participation setting. 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. Below, I verify some of these properties analytically (see Lemmas B.4 and B.5 below) as well as in numerical calculations.

To begin the discussion, I compare asset prices from Proposition B.1 to an identical economy with CRRA preferences, that is setting $\gamma = \psi$. The results of this comparison are plotted in Figure B.1.

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 "growth term" $\psi(\mu_Y + \pi + \pi(1-\alpha)\frac{p(x)}{g_N(x)})$ in (B.17). Indeed, if the terms in parentheses are approximately 3%, then lowering ψ from $\psi = \gamma = 3$ to $\psi = 3/4$ reduces r by 7%, which explains the entire fall in r from Figure B.1.

Second, with $\psi < 1$ (EIS > 1), 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 is higher than fundamental volatility, $\sigma_R > \sigma_Y$, and countercyclical, $\sigma_R'(x) < 0$. The reason for all these effects relates to the standard intuition that cash flow effects (income effects) dominate discount rate effects (substitution effects) when the EIS is larger than 1. In particular, when there is a negative shock to dividends, cash flows are permanently lower but discount rates also fall. The discount rate effect comes from the fact that participant consumption falls now but is

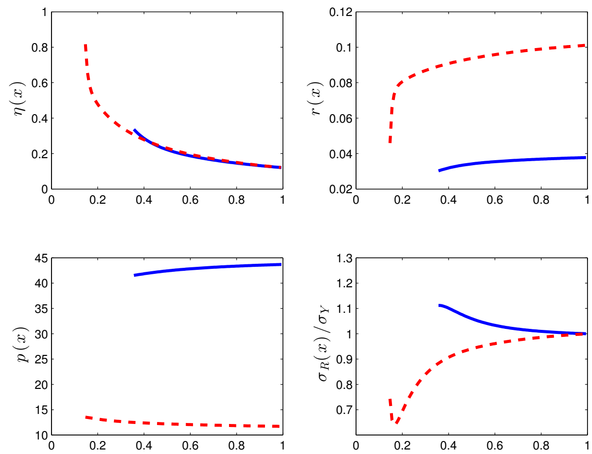

Figure B.1: Asset prices in the benchmark limited participation economy with recursive preferences (blue: $\gamma = 3$ and $\psi = 3/4$) versus CRRA preferences (dashed red: $\gamma = \psi = 3$). Other parameters given in Table 1. The horizontal axis is the participants' consumption share x.

expected to rebound in the future; high participant consumption growth (thus low marginal utility growth) implies low discount rates, as participants are the marginal risky asset pricers. The net effect 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). Figure B.1 shows this result graphically, and Lemma B.4 demonstrates it analytically, that this intuition carries through to limited participation models.

Lemma B.4. *Suppose $\psi < 1$ (EIS > 1). Then, for all x close enough to x^*, $\sigma_R(x) > \sigma_R(1) = \sigma_Y$.*

Proof of Lemma B.4. First, note that

$$p'(x^*) = x^* g'_P(x^*) + (1 - x^*) g'_N(x^*) + g_P(x^*) - g_N(x^*) = g_P(x^*)\left[1 - (1 - \phi)^{\frac{1-\psi}{\psi}}\right],$$

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

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

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

Finally, expected returns on risky assets can be made substantially more volatile than riskless returns In particular, one can pick utility 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,

$$\frac{\text{std}[\mu_R(x) - r(x)]}{\text{std}[r(x)]} = \frac{|\eta(x)\sigma'_R(x) + \eta'(x)\sigma_R(x)|}{|r'(x)|}. \tag{B.32}$$

We will show that the numerator of (B.32) dominates the denominator, implying high risk premium volatility owes much more to expected risky asset returns than to risk-free rate variation.

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 (B.32) is equal to 1. With more general CRRA utility, the result is quantitatively similar to 1 as well. Recursive preferences are needed: the proof of Lemma B.5 below suggests that $\psi < 1$ and $\gamma > 1$ are requisite parameter choices to generate the result.

Lemma B.5. *Given ϕ small enough, there exist choices for the other parameters such that (B.32) can be made arbitrarily large.*

Proof of Lemma B.5. Approximating the equilibrium objects described in Proposition B.1 for x near x^*, we obtain

$$\begin{aligned}
\eta(x) &= \gamma\sigma_Y + O(1-x^*) \\
\sigma_R(x) &= \sigma_Y + O(1-x^*) \\
\eta'(x) &= -\gamma\sigma_Y + O(1-x^*) \\
\sigma_R'(x) &= \sigma_Y\big[1-(1-\phi)^{\frac{1-\psi}{\psi}}\big] + \sigma_Y O(1) + O(1-x^*) \\
r'(x) &= \psi\pi(1-\alpha)\big[1-(1-\phi)^{\frac{\psi-1}{\psi}}\big] + \frac{1}{2}\gamma(\psi+1)\sigma_Y^2 + O(1-x^*).
\end{aligned}$$

To derive these expressions, we have used the facts that

$$\begin{aligned}
p(x^*) &= \big[x^*(1-\phi)^{\frac{\psi-1}{\psi}} + (1-x^*)\big]g_N(x^*) \\
p'(x^*) &= \big[(1-\phi)^{\frac{\psi-1}{\psi}} - 1\big]g_N(x^*) \\
(1-x^*)p''(x^*) &= (1-x^*)x^* g_P''(x^*) + (1-x^*)^2 g_N''(x^*),
\end{aligned}$$

as well as the result in Lemma B.6 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. $\square$

Lemma B.6. *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 of Lemma B.6. 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 (B.15) 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 (B.23) and (B.24) 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. $\square$

C Sentiments model

This appendix adds extrapolative beliefs to the economy. Section C.1 first establishes some general preliminaries for how sentiments govern belief distortions and impacts other aspects of the economy. Section C.2 then adds sentiments to the log utility economy. Section C.3 repeats the exercise for the recursive utility economy. Section C.4 discusses the numerical solution method.

C.1 Preliminaries with sentiments

Recall the actual return process $dR_t = \mu_{R,t}dt + \sigma_{R,t}dZ_t + dA_t^R$, where Z is a one-dimensional Brownian motion and A^R is a continuous, weakly increasing process (which only increases at times of entry). Thus, $\mu_{R,t}$ is the true expected return. By contrast, equation (41) defines agents' perceived expected return, restated here for convenience:

$$\tilde{\mu}_{R,t} = (1-\lambda)\mu_{R,t} + \lambda S_t, \tag{C.1}$$

where recall that sentiment S_t follows

$$dS_t = \beta(dR_t - S_t dt). \tag{C.2}$$

Equation (C.1) effectively defines agents misspecified beliefs. Based on the true and perceived expected returns, define the actual and perceived Sharpe ratios

$$\underbrace{\eta_t := \frac{\mu_{R,t} - r_t}{\sigma_{R,t}}}_{\text{actual Sharpe ratio}} \quad \text{and} \quad \underbrace{\tilde{\eta}_t := \frac{\tilde{\mu}_{R,t} - r_t}{\sigma_{R,t}}}_{\text{perceived Sharpe ratio}}. \tag{C.3}$$

By Girsanov's theorem, in order to justify $\tilde{\mu}_{R,t}$ as the "perceived expected return", agents must think of $\tilde{Z}_t$ as the driving Brownian motion, which in fact follows

$$d\tilde{Z}_t := dZ_t + \Big(\frac{\mu_{R,t} - \tilde{\mu}_{R,t}}{\sigma_{R,t}}\Big)dt = dZ_t + (\eta_t - \tilde{\eta}_t)dt. \tag{C.4}$$

Indeed, substituting (C.4) into the return process gives $dR_t = \tilde{\mu}_{R,t}dt + \sigma_{R,t}d\tilde{Z}_t$ as desired.

The object $\eta_t - \tilde{\eta}_t$ serves as a change-of-measure from the true probability $\mathbb{P}$ to agents' perceived probability $\tilde{\mathbb{P}}$. By Girsanov's theorem, this change-of-measure only adjusts the drifts of each stochastic process. In what follows, we will put tildes on all drifts to indicate drifts under the perceived probability $\tilde{\mathbb{P}}$. Mathematically, if M is any stochastic process adapted to Z—i.e., we can write it as $dM_t = \mu_{M,t}dt + \sigma_{M,t}dZ_t$—then we will denote $\tilde{\mu}_M$ as its perceived drift. Due to (C.4), we have

$$\tilde{\mu}_{M,t} = \mu_{M,t} + \sigma_{M,t}(\tilde{\eta}_t - \eta_t). \tag{C.5}$$

The same formula holds if the dynamics of M are specified in geometric form, e.g., $dM_t = M_t[\mu_{M,t}dt + \sigma_{M,t}dZ_t]$. We can use equation (C.5) to recover ex-post true drifts from their perceived counterparts, or vice-versa. For example, aggregate output Y, which has true drift $Y_t\mu_Y$, possesses the following geometric drift under $\tilde{\mathbb{P}}$:

$$\tilde{\mu}_{Y,t} := \mu_Y + \sigma_Y(\tilde{\eta}_t - \eta_t). \tag{C.6}$$

In the remainder of this appendix, we will posit an equilibrium in the two state variables (X_t, S_t). The state processes follow

$$dX_t = \tilde{\mu}_X(X_t, S_t)dt + \sigma_X(X_t, S_t)d\tilde{Z}_t + dA_t^X \tag{C.7}$$

$$dS_t = \tilde{\mu}_S(X_t, S_t)dt + \sigma_S(X_t, S_t)d\tilde{Z}_t + dA_t^S, \tag{C.8}$$

where A^X and A^S are continuous, weakly increasing processes. These processes will increase only when $(X_t, S_t) \notin \Omega$, where the set $\Omega \subset (0,1) \times \mathbb{R}$ corresponds to when entry is not occurring (Ω is the relevant region for equilibrium, as the economy will spend zero measure of time outside of Ω).

We now derive a few formulas for returns that will apply to the rest of this appendix. First, given that we need the true Sharpe ratio η in order to perform the change-of-drift in (C.5), we solve for it here. By combining (C.3) with (C.1), we obtain

$$\eta_t = \tilde{\eta}_t + \frac{\lambda}{1-\lambda}\Big(\tilde{\eta}_t + \frac{r_t - S_t}{\sigma_{R,t}}\Big), \quad \lambda \neq 1, \tag{C.9}$$

where if $\lambda = 1$, we instead have the relation $r_t = S_t - \sigma_{R,t}\tilde{\eta}_t$.

Next, if $p_t = p(X_t, S_t)$ denotes the equilibrium price-dividend ratio, then the definition of returns as dividend yield plus capital gains delivers $dR_t = \frac{1}{p_t}dt + \frac{d(p_tY_t)}{p_tY_t}$. Using Itô's formula to write the capital gain term and matching the result to $dR_t = \tilde{\mu}_{R,t}dt + \sigma_{R,t}d\tilde{Z}_t + dA_t^R$ delivers the following relations

$$\sigma_{R,t} = \sigma_Y + \Big(\frac{\partial}{\partial x}\log p_t\Big)\sigma_{X,t} + \Big(\frac{\partial}{\partial s}\log p_t\Big)\sigma_{S,t} \tag{C.10}$$

$$dA_t^R = \Big(\frac{\partial}{\partial x}\log p_t\Big)dA_t^X + \Big(\frac{\partial}{\partial s}\log p_t\Big)dA_t^S. \tag{C.11}$$

On the other hand, substituting dR_t into equation (C.2), and then matching terms with (C.8), we obtain

$$\sigma_{S,t} = \beta\sigma_{R,t} \tag{C.12}$$

$$dA_t^S = \beta dA_t^R. \tag{C.13}$$

Equations (C.10)-(C.11) and (C.12)-(C.13) specify a two-way feedback between (σ_R, dA^R) and

(σ_S, dA^S). Solving this two-way feedback, we obtain

$$\sigma_{R,t} = \frac{\sigma_Y + \sigma_{X,t}\frac{\partial}{\partial x}\log p_t}{1 - \beta\frac{\partial}{\partial s}\log p_t} \tag{C.14}$$

$$dA_t^R = \frac{\frac{\partial}{\partial x}\log p_t}{1 - \beta\frac{\partial}{\partial s}\log p_t} dA_t^X \tag{C.15}$$

We can substitute (C.14)-(C.15) back into (C.12)-(C.13) to obtain (σ_S, dA^S).

Finally, we may use (C.3) to write true and perceived expected returns in terms of the interest rate and Sharpe ratio: $\mu_R = r + \sigma_R\eta$ and $\tilde{\mu}_R = r + \sigma_R\tilde{\eta}$. Given these formulas, we can obtain $\mu_S = \beta(\mu_R - s)$ and $\tilde{\mu}_S = \beta(\tilde{\mu}_R - s)$ from (C.2). Given these results, along with (C.9) and (C.14), it only remains to solve for $(r, \tilde{\eta}, p)$.[30]

C.2 Extrapolative beliefs and log utility

Proof of Proposition 10. Posit an equilibrium in the two state variables (X_t, S_t), and assume the entry boundary is given by $x^*(S_t)$ for some function $x^*(\cdot)$ (i.e., entry occurs whenever $X_t \leq x^*(S_t)$). 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 similar to the rational model, except under the probability measure induced by the extrapolative beliefs, $\tilde{\mathbb{P}}$. The HJB equation for participants is

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

and similarly for g_N (with $\tilde{\eta}$ replaced by 0). As in the rational model, 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 $\frac{\partial}{\partial x}\Delta g(x^*(s), s) = \frac{\partial}{\partial s}\Delta g(x^*(s), s) = 0$.

To solve for equilibrium objects, we use an identical procedure as in Proposition C.1. First, since all agents' consumption-wealth ratios equal $\rho + \pi$, the asset market clearing condition implies the price-dividend ratio is constant, $p_t = (\rho+\pi)^{-1}$. From (C.10) and (C.12), we thus have $\sigma_R = \sigma_Y$ and $\sigma_S = \beta\sigma_Y$. Next, apply Itô's formula to the definition of X_t and goods market clearing, noting that participants' consumption growth volatility is $\tilde{\eta}$, while it is 0 for non-participants. This yields

[30] If $\lambda = 1$, then we instead have already a formula for r and instead need to solve for $(\eta, \tilde{\eta}, p)$.

a system of equations in $\tilde{\eta}$, r, $\tilde{\mu}_X$, and σ_X, which is

$$\begin{aligned}
\tilde{\mu}_Y &= \pi(1-\alpha) - \pi + x[r + \alpha\pi - \rho - \pi + \tilde{\eta}^2] + (1-x)[r + \alpha\pi - \rho - \pi] \\
\sigma_Y &= x\tilde{\eta} \\
x\tilde{\mu}_Y + \tilde{\mu}_X + \sigma_Y\sigma_X &= x[r + \alpha\pi - \rho - \pi - \pi + \tilde{\eta}^2] \\
x\sigma_Y + \sigma_X &= x\tilde{\eta}.
\end{aligned}$$

Substitute $\tilde{\mu}_Y = \mu_Y + \frac{\lambda}{1-\lambda}(s - r - \sigma_Y\tilde{\eta})$, which is (C.39) with $\sigma_R = \sigma_Y$ substituted, and then solve the four equations to get

$$\begin{aligned}
\tilde{\eta} &= \sigma_Y/x \\
r &= (1-\lambda)[\rho + \pi + \mu_Y] + \lambda s - \sigma_Y^2/x \\
\sigma_X &= (1-x)\sigma_Y \\
\tilde{\mu}_X &= -\pi(1-\alpha)x + \sigma_Y^2(1-x)^2/x.
\end{aligned}$$

Next, we solve $\tilde{\mu}_R = r + \sigma_R\tilde{\eta} = (1-\lambda)[\rho + \pi + \mu_Y] + \lambda s$. Combined this with equation (C.1) to solve for $\mu_R = \rho + \pi + \mu_Y$. Hence, $\mu_S = \beta(\rho + \pi + \mu_Y - s)$. At this point, it is clear that $\overline{s} = \rho + \pi + \mu_Y$, since dS_t is an Ornstein-Uhlenbeck process. Then substitute μ_R, r, and σ_R to obtain $\eta := \frac{\mu_R - r}{\sigma_R}$. Lastly, no-arbitrage with extrapolative agents implies discounted returns—discounted by the process $d\xi_t = -\xi_t[\mu_{\xi,t}dt + \sigma_{\xi,t}d\tilde{Z}_t]$—must be local martingales under $\tilde{\mathbb{P}}$, which identifies $\mu_\xi = r$ and $\sigma_\xi = \tilde{\eta}$.

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 $\tilde{\mu}_S = (1-\lambda)\beta(\overline{s} - s)$ (this is obtained using (C.5) in conjunction with μ_S, $\tilde{\eta}$, and η). 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 (if a solution exists). The ODE we need to solve to determine x^* is

$$0 = \frac{1}{2}(\rho + \pi)^{-1}\tilde{\eta}^2 - (\rho + \pi)\Delta g + \tilde{\mu}_X\Delta g' + \frac{1}{2}\sigma_X^2\Delta g'', \quad \Delta g(x^*) = \Phi, \quad \Delta g'(x^*) = 0, \tag{C.16}$$

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 (25), which implies that the entry point x^* is independent of λ and β. □

C.3 Extrapolative beliefs and recursive preferences

Proposition C.1. *Assuming existence of a solution (g_P, g_N) to the variational inequality (C.22)-(C.23), there exists a stationary Markov equilibrium with sentiments, governed by (X_t, S_t).*

Proof of Proposition C.1. As in the rational model with recursive preferences, due to the homogeneity properties of the utility function and budget constraints, participants and non-participants with wealth level W_t achieve values $V^i(W_t, X_t, S_t) = (1-\gamma) W_t^{1-\gamma} g_i(X_t, S_t)^{\frac{\psi(1-\gamma)}{1-\psi}}$ for $i \in \{P, N\}$, where functions g_i correspond to agent's wealth-consumption ratios. To save on notation, write the process for these wealth-consumption ratios as

$$dg_i(X_t, S_t) = g_i(X_t, S_t)\Big[\tilde{\mu}_{g_i}(X_t, S_t)dt + \sigma_{g_i}(X_t, S_t)d\tilde{Z}_t\Big], \quad i \in \{P, N\}.$$

Due to Itô's formula, we have the identities

$$\tilde{\mu}_{g_i} = \frac{1}{g_i}\Big[\tilde{\mu}_X \frac{\partial}{\partial x} g_i + \tilde{\mu}_S \frac{\partial}{\partial s} g_i + \frac{1}{2}\sigma_X^2 \frac{\partial^2}{\partial x^2} g_i + \frac{1}{2}\sigma_S^2 \frac{\partial^2}{\partial s^2} g_i + \sigma_X \sigma_S \frac{\partial^2}{\partial x \partial s} g_i\Big] \tag{C.17}$$

$$\sigma_{g_i} = \frac{1}{g_i}\Big[\sigma_X \frac{\partial}{\partial x} g_i + \sigma_S \frac{\partial}{\partial s} g_i\Big]. \tag{C.18}$$

We derive HJB equations for g_P and g_N similarly to Appendix B.2, except under the perceived probability $\tilde{\mathbb{P}}$. We merely state the result here. Define the operators $\mathcal{D}_N$ and $\mathcal{D}_P$ by

$$\begin{aligned}\mathcal{D}_i g_i := \psi/g_i - (\rho+\pi) + (1-\psi)\Big[r + \alpha\pi + \frac{1}{2\gamma}\tilde{\eta}^2 \mathbf{1}_{\{i=P\}}\Big] + \psi\frac{1-\gamma}{\gamma}\sigma_{g_i}\tilde{\eta}\mathbf{1}_{\{i=P\}} \\ + \psi\tilde{\mu}_{g_i} + \frac{\psi}{2}\Big(\frac{\psi(1-\gamma)}{1-\psi} - 1 + \frac{\psi(1-\gamma)^2}{(1-\psi)\gamma}\mathbf{1}_{\{i=P\}}\Big)\sigma_{g_i}^2.\end{aligned} \tag{C.19}$$

Due to (C.17)-(C.18), $\mathcal{D}_i$ is a differential operator. Then, optimal behavior requires (where recall Ω is the region of the state space where entry is not taking place)

$$(x,s) \in \Omega \;:\; g_N^{\frac{\psi}{1-\psi}} - (1-\phi)g_P^{\frac{\psi}{1-\psi}} > 0 \quad \text{and} \quad (1-\psi)^{-1}\mathcal{D}_N g_N = (1-\psi)^{-1}\mathcal{D}_P g_P = 0 \tag{C.20}$$

$$(x,s) \notin \Omega \;:\; g_N^{\frac{\psi}{1-\psi}} - (1-\phi)g_P^{\frac{\psi}{1-\psi}} = 0 \quad \text{and} \quad (1-\psi)^{-1}\mathcal{D}_N g_N \leq (1-\psi)^{-1}\mathcal{D}_P g_P = 0. \tag{C.21}$$

The equations (C.20)-(C.21) can be compactly summarized by

$$\min\Big\{g_N^{\frac{\psi}{1-\psi}} - (1-\phi)g_P^{\frac{\psi}{1-\psi}},\, (\psi-1)^{-1}\mathcal{D}_N g_N\Big\} = 0 \tag{C.22}$$

$$\mathcal{D}_P g_P = 0. \tag{C.23}$$

Along the way to deriving (C.22)-(C.23), we also obtain the optimal portfolios $\theta_{N,t} = 0$ (by the non-participation constraint) and

$$\theta_P = \frac{\tilde{\eta}}{\gamma\sigma_R} + \frac{\psi(1-\gamma)}{(1-\psi)\gamma}\frac{\sigma_{g_P}}{\sigma_R}. \tag{C.24}$$

One can use straightforward arguments to extend Proposition B.2 to show that these conditions characterize individual optimality; we omit this for brevity.

To solve for equilibrium objects, we use a similar procedure as in Proposition B.1 but under $\tilde{\mathbb{P}}$. First, since (g_P, g_N) are wealth-consumption ratios, and since x represents the participant consumption share, the price-dividend ratio $p(x,s)$ must satisfy (from combining asset market clearing equations (13)-(14))

$$p(x,s) = xg_P(x,s) + (1-x)g_N(x,s). \tag{C.25}$$

Since we are solving the equilibrium up to the functions (g_P, g_N) and their derivatives, (C.25) implies we may solve up to (g_P, g_N, p) and their derivatives.

Second, suppose the consumption dynamics of type $i \in \{P, N\}$ agents can be written

$$\frac{dc_{i,t}}{c_{i,t}} = \tilde{\mu}_{c_i,t}dt + \sigma_{c_i,t}d\tilde{Z}_t + dA_t^{c_i}.$$

Apply Itô's formula to the definition of X_t and goods market clearing, then match terms to obtain the following equations:

$$\tilde{\mu}_Y = \pi(1-\alpha)p/g_N - \pi + x\tilde{\mu}_{c_P} + (1-x)\tilde{\mu}_{c_N} \tag{C.26}$$

$$\sigma_Y = x\sigma_{c_P} + (1-x)\sigma_{c_N} \tag{C.27}$$

$$x\tilde{\mu}_Y + \tilde{\mu}_X + \sigma_Y\sigma_X = x\tilde{\mu}_{c_P} - x\pi \tag{C.28}$$

$$x\sigma_Y + \sigma_X = x\sigma_{c_P}. \tag{C.29}$$

To replace $(\tilde{\mu}_{c_i}, \sigma_{c_i})$, we obtain consumption dynamics as follows. Letting W_i denote the wealth of a type $i \in \{P, N\}$ agent, Itô's formula gives the dynamics of $c_i = W_i/g_i$ as

$$\begin{aligned} \frac{dc_{i,t}}{c_{i,t}} &= \frac{dW_{i,t}}{W_{i,t}} - \frac{dg_{i,t}}{g_{i,t}} + \frac{d[g_i]_t}{g_{i,t}^2} - \frac{d[W_i, g_i]_t}{W_{i,t}g_{i,t}} \\ &= \Big[\underbrace{r_t + \alpha\pi - g_{i,t}^{-1} + \theta_{i,t}\sigma_{R,t}(\tilde{\eta}_t - \sigma_{g_i,t}) - \tilde{\mu}_{g_i,t} + \sigma_{g_i,t}^2}_{=\tilde{\mu}_{c_i,t}}\Big]dt + \underbrace{(\theta_{i,t}\sigma_{R,t} - \sigma_{g_i,t})}_{=\sigma_{c_i,t}}d\tilde{Z}_t + dA_t^R. \end{aligned} \tag{C.30}$$

I will solve (C.26)-(C.29) for $(\sigma_X, \tilde{\eta}, r, \tilde{\mu}_X)$, taking as given (g_P, g_N, p) and their derivatives.

First consider equations (C.27) and (C.29). Substituting (C.30), then using (C.18), we have

$$\sigma_Y = x[\theta_P\sigma_R - \sigma_{g_P}] - (1-x)\sigma_{g_N} \tag{C.31}$$

$$x\sigma_Y + \sigma_X = x[\theta_P\sigma_R - \sigma_{g_P}]. \tag{C.32}$$

Combining (C.31) with (C.32) to eliminate $\theta_P\sigma_R - \sigma_{g_P}$, then using (C.18) to substitute σ_{g_N}, then using (C.14) in (C.12) to plug σ_S, and finally rearranging, we obtain

$$\sigma_X = \frac{(1-x)(1+\beta\zeta\frac{\partial}{\partial s}\log g_N)}{1-(1-x)[\frac{\partial}{\partial x}\log g_N + \beta\zeta(\frac{\partial}{\partial s}\log g_N)(\frac{\partial}{\partial x}\log p)]}\sigma_Y, \tag{C.33}$$

where

$$\zeta := \big(1 - \beta \frac{\partial}{\partial s} \log p\big)^{-1}$$

Given the solution for σ_X, we now have solutions to σ_R and σ_S from (C.14) and (C.12). Consequently, armed with both σ_X and σ_S, we have solutions for both σ_{g_P} and σ_{g_N} in (C.18). Thus, substituting θ_P from (C.24) into (C.31), we obtain

$$\tilde{\eta} = \frac{\gamma}{x}\Big[\sigma_Y + x\sigma_{g_P} + (1-x)\sigma_{g_N}\Big] + \frac{\psi(\gamma-1)}{1-\psi}\sigma_{g_P}. \tag{C.34}$$

Now consider equations (C.26) and (C.28). Substituting $(\tilde{\mu}_{c_P}, \tilde{\mu}_{c_N})$ from (C.30), the system of equations becomes

$$\begin{aligned} \tilde{\mu}_Y = \pi(1-\alpha)(p/g_N - 1) + r + x\Big[-g_P^{-1} + \theta_P\sigma_R(\tilde{\eta} - \sigma_{g_P}) - \tilde{\mu}_{g_P} + \sigma^2_{g_P}\Big] \quad &(C.35) \\ + (1-x)\Big[-g_N^{-1} + \theta_N\sigma_R(\tilde{\eta} - \sigma_{g_N}) - \tilde{\mu}_{g_N} + \sigma^2_{g_N}\Big] & \\ x\tilde{\mu}_Y + \tilde{\mu}_X + \sigma_Y\sigma_X = x\Big[r + \alpha\pi - \pi - g_P^{-1} + \theta_P\sigma_R(\tilde{\eta} - \sigma_{g_P}) - \tilde{\mu}_{g_P} + \sigma^2_{g_P}\Big]. \quad &(C.36) \end{aligned}$$

To solve equations (C.35)-(C.36), we need formulas for $(\tilde{\mu}_{g_P}, \tilde{\mu}_{g_N})$. These formulas come from setting $\mathcal{D}_N g_N = 0$ and $\mathcal{D}_P g_P = 0$, which recall is valid when entry is not taking place (i.e., $\mathcal{D}_N g_N = \mathcal{D}_P g_P = 0$ holds on Ω; see equation (C.20)). The results are

$$\tilde{\mu}_{g_P} = -1/g_P + \frac{\rho+\pi}{\psi} + \frac{\psi-1}{\psi}\big(r + \alpha\pi + \frac{1}{2\gamma}\tilde{\eta}^2\big) + \frac{\gamma-1}{\gamma}\sigma_{g_P}\tilde{\eta} - \frac{1}{2}\Big(\frac{\psi(1-\gamma)}{\gamma(1-\psi)} - 1\Big)\sigma^2_{g_P} \tag{C.37}$$

$$\tilde{\mu}_{g_N} = -1/g_N + \frac{\rho+\pi}{\psi} + \frac{\psi-1}{\psi}(r+\alpha\pi) - \frac{1}{2}\Big(\frac{\psi(1-\gamma)}{1-\psi} - 1\Big)\sigma^2_{g_N}. \tag{C.38}$$

Let us solve equation (C.35). We treat the cases $\lambda \neq 1$ and $\lambda = 1$ separately. If $\lambda \neq 1$, then by combining (C.6) and (C.9), we obtain

$$\tilde{\mu}_Y = \mu_Y + \sigma_Y\frac{\lambda}{1-\lambda}\Big(\frac{s-r}{\sigma_R} - \tilde{\eta}\Big), \quad \lambda \neq 1. \tag{C.39}$$

Into (C.35), substitute (C.37)-(C.38)-(C.39) and portfolio choices (θ_N, θ_P), then rearrange to obtain

$$\begin{aligned} r = \Big(\frac{\lambda\frac{\sigma_Y}{\sigma_R}}{(1-\lambda)\psi^{-1} + \lambda\frac{\sigma_Y}{\sigma_R}}\Big)[s - \sigma_R\tilde{\eta}] + \Big(\frac{(1-\lambda)\psi^{-1}}{(1-\lambda)\psi^{-1} + \lambda\frac{\sigma_Y}{\sigma_R}}\Big)\Big[\rho + \pi(1-\alpha)(1 - \psi\frac{p}{g_N}) + \psi(\pi + \mu_Y) \\ - \frac{\psi x}{2}\Big(\frac{1+\psi}{\gamma\psi}\tilde{\eta}^2 + \Big(1 - \frac{\psi(1-\gamma)}{\gamma(1-\psi)}\Big)\sigma^2_{g_P} - 2\Big(1 - \frac{\psi(1-\gamma)}{\gamma(1-\psi)}\Big)\tilde{\eta}\sigma_{g_P}\Big) \\ - \frac{\psi(1-x)}{2}\Big(\frac{\psi(1-\gamma)}{1-\psi} + 1\Big)\sigma^2_{g_N}\Big], \quad \lambda \neq 1. \end{aligned} \tag{C.40}$$

With r in hand, we can also compute $\tilde{\mu}_Y$ from (C.39) and η from (C.9).

On the other hand, if $\lambda = 1$, equation (C.9) implies $r = s - \sigma_R\tilde{\eta}$. Given r, we view (C.35)—after combining with (C.37)-(C.38)—as an equation for $\tilde{\mu}_Y$. Given $\tilde{\mu}_Y$, we can compute η by rearranging equation (C.6) to obtain $\eta = \tilde{\eta} + \frac{\mu_Y - \tilde{\mu}_Y}{\sigma_Y}$. Note that the equilibrium solution for r from (C.40) is thus continuous as $\lambda \to 1$ (hence, the rest of the equilibrium too).

Finally, in either case ($\lambda \neq 1$ or $\lambda = 1$), rearrange (C.36) to get

$$\tilde{\mu}_X = -\sigma_Y \sigma_X - x\tilde{\mu}_Y + x\Big[r + \alpha\pi - \pi - g_P^{-1} + \theta_P \sigma_R(\tilde{\eta} - \sigma_{g_P}) - \tilde{\mu}_{g_P} + \sigma_{g_P}^2\Big]. \tag{C.41}$$

This completes the derivation of equilibrium, since all other equilibrium objects can be obtained using the results in Appendix C.1. □

C.4 Numerical method with sentiments

In the baseline version of the model, where X_t is the only state variable, I use the matlab package "bvp4c" to solve the boundary value problem characterizing equilibrium. This package projects the solution onto Chebyshev polynomials and works well in one dimension.

In the model with sentiments, as well as the extension with a fixed cost in Appendix D.1, a second state variable arises in agents' optimization problems. To solve this type of problem (e.g., the variational inequality (C.22)-(C.23)), I use a finite difference scheme augmented with a false transient (fake time derivative). I provide a brief verbal description of the solution algorithm, with details available in the numerical code.

The participant PDE (C.23) is solved using a standard iterative implicit finite difference scheme as in Achdou et al. (2020). This scheme first treats any nonlinear terms as known functions of the state (using the previous time-iteration's value functions), and uses "upwinding" of the first-order derivative terms. Because sentiment and the consumption share are correlated, there are cross-derivatives in (C.23), which technically breaks the sufficient conditions for the scheme's convergence (i.e., the "monotonicity" property of Barles and Souganidis (1991)). However, in practice, I find that the scheme converges. The boundary conditions used are reflecting at high and low values of s, reflecting at the lower boundary $x \approx 0$ (because entry occurs at that point for sure), and "natural" at the upper boundary $x \approx 1$ (because the drift $\mu_X < 0$ while $\sigma_X \to 0$ as $x \to 1$). In sum, taking a time-step in the implicit scheme involves solving a system of linear equations.

The non-participant PDE (C.22) is complicated by the entry decision. In this case, taking a time-step in the implicit scheme cannot be done by solving a system of linear equations. However, it turns out that, after treating nonlinear terms as known from the previous time-iteration, taking a time-step in (C.22) is equivalent to solving a "linear complementarity problem" (LCP). I use standard code to solve this problem at each time-step.

To test this scheme, I apply it to the log utility model, which has a quasi-analytical solution even with sentiments (see Proposition 10). I also test the same methodology on the various one-dimensional models (e.g., the baseline model with recursive utility but without sentiments). In all cases, I find that the numerical described above converges to the correct solution.

D Other extensions

This section presents some robustness exercises. Section D.1 discusses and solves another formulation of entry costs that are not homogeneous in wealth. Section D.2 discusses and solves an extension with labor income. Section D.3 discusses other types of preferences (heterogeneous risk aversion; heterogeneous EIS; and hyperbolic discounting).

D.1 Fixed non-homogeneous entry cost

In this section, we consider a different formulation of entry costs that are not a fixed proportion of wealth. We take the other extreme and assume entry costs are completely independent of individual wealth. Such non-homogeneous fixed costs have been emphasized in the household finance literature (Campbell, 2006) to achieve consistency with two important stock market non-participation facts: (i) wealthier households tend to participate and poorer households do not; (ii) participation tends to be procyclical. In the context of intermediation, it also makes some sense that setup and market research costs would have a component which is largely size-independent. While the purpose of this section is to show that our baseline results are robust to deviating from proportional entry costs, such an extension with non-homogeneous entry costs is analytically and computationally non-trivial. We develop a detailed solution method below, and then present the results.

Consider an entry cost of the form

$$\tilde{\Phi}_t(w) := -(\rho+\pi)^{-1} \log\left[\left(1-\phi\frac{P_t}{w}\right)^+\right], \quad \phi \in (0,1), \tag{D.1}$$

where w is the individual's wealth and P_t is the aggregate wealth (stock market value). For an individual of average wealth ($w = P_t$), note $\tilde{\Phi}_t(P_t) = -(\rho+\pi)^{-1}\log(1-\phi)$. This is identical to the proportional cost specification Φ (perceived as ϕ fraction of wealth) used in the baseline model (see equation (6)). For comparison purposes, I have intentionally specified this new fixed cost function so that the individual of average wealth perceives the same cost as in the proportional cost baseline.

But with a fixed cost specification, entry incentives depend on financial wealth levels. For an individual of below average wealth ($w < P_t$), the cost is perceived as something greater than ϕ fraction of wealth. For an individual with very low wealth ($w \leq \phi P_t$), entry is perceived as infinitely costly ($\tilde{\Phi}_t = +\infty$). Finally, for a very wealthy individual ($w \to +\infty$), entry is perceived as costless ($\tilde{\Phi}_t \to 0$).

Boundedly-rational beliefs. Given the discussion above, non-participant entry decisions (and thus also their consumption decisions) will be non-homogeneous in their wealth. Because of this feature, the full cross-sectional distribution becomes an aggregate state variable, leading to infinite-dimensional optimization problems for our agents. To circumvent this complexity, we assume a form of "bounded rationality" a la Krusell and Smith (1998): suppose agents attempt to summarize aggregate dynamics by a single state variable. As in our baseline model, assume that agents use participants' consumption share X_t as the relevant aggregate state variable. Dynamics of X_t are given by

$$dX_t = \mu_{X,t}dt + \sigma_{X,t}dZ_t,$$

for some μ_X and σ_X to be determined. (Note the absence of a singularly continuous term here, as it will turn out that entry will occur at the dt order in this extension.) Conditioning solely on X_t will be flawed, as many equilibrium objects will now depend on more than just X_t; we will allow this dependence in deriving equilibrium but agents will suboptimally ignore it.

Bounded rationality means agents perceive asset prices and state dynamics $(r_t, p_t, \eta_t, \sigma_{R,t}, \mu_{X,t}, \sigma_{X,t})$ as governed solely by X_t, and we write their approximating functions $(\hat{r}(x), \hat{p}(x), \hat{\eta}(x), \hat{\sigma}_R(x), \hat{\mu}_X(x), \hat{\sigma}_X(x))$. Agents will use the latter objects in decision-making. At the end of this subsection, we will describe how agents observe data and design the fitted values (objects with hats), but for now just note that it will be a type of regression procedure.[31]

Participant optimality. Because they are unconstrained and have log utility, participants make similar portfolio and consumption decisions as in the baseline model (i.e., with log utility, they display the familiar "myopic" behavior and ignore the dynamics of aggregate states). For this reason, we may actually continue to assume participants are fully rational (i.e., we do not need them to approximate equilibrium dynamics solely with X_t in order to solve their optimization problem).[32] One can use standard martingale techniques for unconstrained log utility investors to find that participants consume $\rho + \pi$ fraction of their wealth and hold exposure η_t to the aggregate shock. Under these policies, optimal participant consumption dynamics are

$$\begin{aligned} & dc_{t,b} = c_{t,b}\big[\mu_{c_P,t}dt + \sigma_{c_P,t}dZ_t\big], \quad b \in \mathcal{P}_t \\ \text{where} \quad & \mu_{c_P,t} := r_t - \rho - \pi + \alpha\pi + \eta_t^2 \\ & \sigma_{c_P,t} := \eta_t. \end{aligned} \tag{D.2}$$

These are the true consumption dynamics; perceived consumption dynamics simply replace (r_t, η_t) by $(\hat{r}_t, \hat{\eta}_t)$.

[31] In continuous time Brownian models, this type of bounded rationality procedure delivers a distorted expectation $\hat{\mathbb{E}}$ which is mutually singular to the true expectation $\mathbb{E}$ (i.e., the change-of-measure cannot be justified via Girsanov's theorem). For example, our derivation below shows that the diffusion σ_X will depend on more than X_t, whereas the perceived diffusion $\hat{\sigma}_X$ only depends on X_t (to use Girsanov's theorem, these diffusions must agree). That being said, our numerical results will show that $\hat{\sigma}_X \approx \sigma_X$, so we view this as a minor issue. Relatedly, the riskless bond return cannot be misestimated under an equivalent distorted belief, and technically these agents will misestimate it; however, we will show numerically that $\hat{r} \approx r$ in equilibrium. We leave for future research the task of formalizing the present type of bounded-rationality procedure as a Girsanov change-of-measure.

[32] An additional advantage of this strategy is that the risk price η_t will show up directly in participants' consumption growth volatility, as opposed to their perceived risk price $\hat{\eta}_t$. As a result, market clearing will pin down η_t rather than the belief $\hat{\eta}_t$, which is inconsistent with algorithms like Krusell and Smith (1998), in which beliefs are specified prior to imposing market clearing.

To later solve the non-participants' entry problem, it is important to calculate the participant utility under the approximate beliefs $\hat{\mathbb{E}}$. Under these beliefs, a participant with time-t wealth w obtains indirect utility

$$
\begin{aligned}
&V_t^P(w) := \hat{\mathbb{E}}_t\Big[\int_t^\infty e^{-(\rho+\pi)(s-t)}\log(c_{s,b})ds\Big] \\
&= \hat{\mathbb{E}}_t\Big[\int_t^\infty e^{-(\rho+\pi)(s-t)}\Big(\log(\rho+\pi)+\log(w)+\int_t^s[\alpha\pi-\rho-\pi+\hat{r}_u+\frac{1}{2}\hat{\eta}_u^2]du+\int_t^s\hat{\eta}_u dZ_u\Big)ds\Big] \\
&= (\rho+\pi)^{-1}\log(w)+(\rho+\pi)^{-1}\log(\rho+\pi)+\hat{\mathbb{E}}_t\Big[\int_t^\infty e^{-(\rho+\pi)(s-t)}\int_t^s[\alpha\pi-\rho-\pi+\hat{r}_u+\frac{1}{2}\hat{\eta}_u^2]duds\Big] \\
&:= (\rho+\pi)^{-1}\log(w)+\hat{g}_{P,t},
\end{aligned}
$$

where the last line defines $\hat{g}_{P,t}$ as a stochastic process independent of individual behavior. In principle, this process can depend arbitrarily on aggregate states, but under the approximate beliefs, it will be solely a function of X_t, i.e., $\hat{g}_{P,t}=\hat{g}_P(X_t)$. In summary, we still have for participants the additively separable characterization of the value function $V_t^P(w)=(\rho+\pi)^{-1}\log(w)+\hat{g}_P(X_t)$.

Given the distorted belief, we can solve for the function $\hat{g}_P$ via the differential equation (HJB equation)

$$
0=\hat{\mathcal{D}}_P\hat{g}_P, \tag{D.3}
$$

where the differential operator $\hat{\mathcal{D}}_P$ is defined on functions of x by (given $\hat{r}$, $\hat{\eta}$, $\hat{\mu}_X$, and $\hat{\sigma}_X$)

$$
\hat{\mathcal{D}}_P g := -(\rho+\pi)g+\log(\rho+\pi)-1+(\rho+\pi)^{-1}[\hat{r}+\alpha\pi+\frac{1}{2}\hat{\eta}^2]+\hat{\mu}_X\frac{\partial}{\partial x}g+\frac{1}{2}\hat{\sigma}_X^2\frac{\partial^2}{\partial x^2}g.
$$

This is the same HJB equation as in the baseline model, but with prices and state dynamics approximated to be functions of X_t only.

Non-participant approximate optimality. Non-participants' problem is significantly complicated by the combination of their constraints and non-homogeneous entry cost in (D.1). In particular, non-participants' problem becomes inhomogeneous in wealth, and we can no longer write their value function as $V_t^N(w)=(\rho+\pi)^{-1}\log(w)+g_{N,t}$ for a process $g_{N,t}$ independent of wealth. This failure precludes several of the convenient properties of the baseline model.

What matters for decisions is not only does the level of an individual non-participant's wealth, but also the ratio of his wealth to some aggregate. We use the following wealth-to-income ratio to capture this relative wealth:

$$
\omega_{t,b} := \frac{W_{t,b}}{Y_t}, \quad b\le t. \tag{D.4}
$$

In addition to aggregate states, the non-participant value function will depend on both wealth w and wealth-to-aggregate-income ω, and we write this as $V_t^N(w,\omega)$. It turns out that ω is a convenient choice for the second individual state variable, as it allows us to obtain the characterization $V_t^N(w,\omega) = (\rho+\pi)^{-1}\log(w) + \hat{g}_{N,t}(\omega)$ for some function $\hat{g}_{N,t}(\omega)$ independent of wealth w (once accounting for relative wealth ω). Under bounded rationality, this latter term will only depend on relative wealth and participants' consumption share; i.e., $\hat{g}_{N,t}(\omega) = \hat{g}_N(\omega, X_t)$. For the future, we derive the dynamics of ω for non-participants. Recall that $dW_{t,b} = W_{t,b}[r_t + \alpha\pi - c_{t,b}/W_{t,b}]dt$ and $dY_t = Y_t[\mu_Y dt + \sigma_Y dZ_t]$, so that

$$d\omega_{t,b} = \omega_{t,b}\Big[r_t + \alpha\pi - \mu_Y - \frac{c_{t,b}}{W_{t,b}} + \sigma_Y^2\Big]dt - \omega_{t,b}\sigma_Y dZ_t, \quad b \in \mathcal{N}_t. \tag{D.5}$$

The HJB equation for a non-participant with time-t wealth w and wealth-to-income ω is

$$\begin{aligned}\log(w) + (\rho+\pi)\hat{g}_N = \max_c\ & \log(c) + (\rho+\pi)^{-1}[\hat{r} + \alpha\pi - \frac{c}{w}] + \omega[\hat{r} + \alpha\pi - \mu_Y - \frac{c}{w} + \sigma_Y^2]\frac{\partial}{\partial\omega}\hat{g}_N \\ & + \hat{\mu}_X\frac{\partial}{\partial x}\hat{g}_N + \frac{1}{2}\hat{\sigma}_X^2\frac{\partial^2}{\partial x^2}\hat{g}_N + \sigma_\omega\hat{\sigma}_X\frac{\partial^2}{\partial x\partial\omega}\hat{g}_N + \frac{1}{2}\sigma_\omega^2\frac{\partial^2}{\partial\omega^2}\hat{g}_N,\end{aligned} \tag{D.6}$$

where $\sigma_\omega(\omega) := -\omega\sigma_Y$ is the diffusion of individual level ω. Importantly, the HJB equation (D.6) is written under the non-participant beliefs. Optimal consumption is thus given by $c = (\rho + \pi)m(\omega,x)w$, where

$$m(\omega,x) := \frac{1}{1 + (\rho+\pi)\omega\frac{\partial}{\partial\omega}\hat{g}_N(\omega,x)}. \tag{D.7}$$

Note in passing that, following the same steps as for participants above, we can substitute this consumption rule into the definition of V_t^N to verify our conjectured form, in particular that $\hat{g}_N$ is indeed independent of wealth w. After substituting m into the HJB above, we obtain

$$\begin{aligned}(\rho+\pi)\hat{g}_N = \log(\rho+\pi) + \log(m) + (\rho+\pi)^{-1}[\hat{r} + \alpha\pi] - m + \omega[\hat{r} + \alpha\pi - \mu_Y - (\rho+\pi)m + \sigma_Y^2]\frac{\partial}{\partial\omega}\hat{g}_N \\ + \hat{\mu}_X\frac{\partial}{\partial x}\hat{g}_N + \frac{1}{2}\hat{\sigma}_X^2\frac{\partial^2}{\partial x^2}\hat{g}_N + \sigma_\omega\hat{\sigma}_X\frac{\partial^2}{\partial x\partial\omega}\hat{g}_N + \frac{1}{2}\sigma_\omega^2\frac{\partial^2}{\partial\omega^2}\hat{g}_N.\end{aligned}$$

These equations hold when a non-participant delays entry and chooses to remain a non-participant.

Finally, when a non-participant decides to enter at some time T, we have the value-matching condition $V_T^N(w) = V_T^P(w) - \tilde{\Phi}_T(w)$. Substituting the entry cost from (D.1), the definition of ω from (D.4), and the various bounded rationality approximations, this condition can be written

$$\hat{g}_N(\omega,x) = \hat{g}_P(x) + (\rho+\pi)^{-1}\log\Big[\Big(1 - \phi\frac{\hat{p}(x)}{\omega}\Big)^+\Big]. \tag{D.8}$$

Before entry, $\hat{g}_N(\omega,x) > \hat{g}_P(x) + (\rho+\pi)^{-1}\log[(1-\phi\frac{\hat{p}(x)}{\omega})^+]$, implying entry occurs when $\omega \geq \omega^*(x)$ for some function ω^*. We can thus summarize the non-participant decision problem as solving the variational inequality

$$0 = \min\Big\{ -\hat{\mathcal{D}}_N \hat{g}_N,\ \hat{g}_N - \hat{g}_P - (\rho+\pi)^{-1}\log\big[\big(1-\phi\frac{\hat{p}}{\omega}\big)^+\big]\Big\}, \tag{D.9}$$

where the differential operator $\hat{\mathcal{D}}_N$ is defined on functions of ω and x by (given $\hat{r}$, $\hat{\mu}_X$, and $\hat{\sigma}_X$)

$$\begin{aligned}\hat{\mathcal{D}}_N g := &-(\rho+\pi)g + \log(\rho+\pi) - \log\Big(1+(\rho+\pi)\omega\frac{\partial}{\partial\omega}g\Big) + (\rho+\pi)^{-1}[\hat{r}+\alpha\pi] - \Big(1+(\rho+\pi)\omega\frac{\partial}{\partial\omega}g\Big)^{-1}\\ &+ \omega\Big[\hat{r}+\alpha\pi-\mu_Y - \frac{\rho+\pi}{1+(\rho+\pi)\omega\frac{\partial}{\partial\omega}g} + \sigma_Y^2\Big]\frac{\partial}{\partial\omega}g + \hat{\mu}_X\frac{\partial}{\partial x}g + \frac{1}{2}\hat{\sigma}_X^2\frac{\partial^2}{\partial x^2}g + \sigma_\omega\hat{\sigma}_X\frac{\partial^2}{\partial x\partial\omega}g + \frac{1}{2}\sigma_\omega^2\frac{\partial^2}{\partial\omega^2}g.\end{aligned}$$

Equilibrium. We derive the equilibrium similarly to the baseline model, but the expressions will be substantially more complicated, given non-participant consumption and entry decisions depend on their ω. To characterize equilibrium, we will thus need the distribution of non-participant wealth, i.e., $f_t(\omega)$ that satisfies the following expression for any function φ in an appropriate function space (see Definition E.1 in Appendix E below):

$$\int_0^\infty \varphi(\omega, X_t) f_t(\omega)d\omega = \frac{1}{(1-X_t)Y_t}\int_{\mathcal{N}_t}\pi e^{-\pi(t-b)}c_{t,b}\varphi(\omega_{t,b},X_t)db. \tag{D.10}$$

We take as given knowledge of $f_t(\omega)$ for now. In fact, as part of equilibrium, f_t will satisfy a stochastic PDE, derived in Proposition E.1 below. We also introduce the notation $\mathcal{M}_t(\varphi) := \int \varphi_t(\omega)f_t(\omega)d\omega$ to denote the time-t cross-sectional average (among non-participants) of φ_t.

First, combining optimal consumption policies with asset market clearing, we obtain

$$p_tY_t = (\rho+\pi)^{-1}\Big[X_tY_t + \int_{\mathcal{N}_t}\pi e^{-\pi(t-b)}\frac{c_{t,b}}{m(\omega_{t,b},X_t)}db\Big].$$

Using equation (D.10), we can then obtain an equation for the price-dividend ratio

$$p_t = (\rho+\pi)^{-1}\Big[X_t + (1-X_t)\mathcal{M}_t(\frac{1}{m})\Big]. \tag{D.11}$$

It is clear that p_t depends not only on X_t, but also on the entire density f_t, as foreshadowed earlier.

Second, given p_t and all the other equilibrium objects (which we solve for below), we may compute σ_R by applying the infinite-dimensional version of Itô's formula to $p_t = p(X_t, f_t)$. To do this, note that $\mathcal{L}[h] := \int_0^\infty \frac{1}{m(\omega,x)}h(\omega)d\omega$ defines a linear operator $\mathcal{L}$, which thus has Fréchet derivative $\mathcal{L}' = \mathcal{L}$. Using this facts, and the evolution of $f_t(\omega)$ derived in Proposition E.1, we have

$$\begin{aligned}\sigma_{R,t} &= \sigma_Y + \sigma_{X,t}\frac{\partial}{\partial x}\log p_t + (\rho+\pi)^{-1}p_t^{-1}(1-X_t)\mathcal{M}_t(\frac{\sigma_f}{m})\\ &= (\rho+\pi)^{-1}p_t^{-1}[X_t\sigma_Y + \sigma_{X,t}].\end{aligned} \tag{D.12}$$

The result on the second line is derived after tedious algebra. Given solutions for (σ_R, r, η), the latter two of which are derived below, we obtain $\mu_{R,t} = r_t + \sigma_{R,t}\eta_t$ from no-arbitrage.

Next, given non-participants consume $c_{t,b} = (\rho+\pi)m(\omega_{t,b}, X_t)W_{t,b}$, we have the following form for consumption dynamics for non-participants (these are the true consumption dynamics):

$$dc_{t,b} = c_{t,b}\big[\mu_{c_N,t}(\omega_{t,b})dt + \sigma_{c_N,t}(\omega_{t,b})dZ_t\big], \quad b \in \mathcal{N}_t \tag{D.13}$$

$$\mu_{c_N,t}(\omega) := r_t + \alpha\pi - (\rho+\pi)m(\omega, X_t) + \frac{1}{m(\omega,X_t)}\Big[\mu_{X,t}\frac{\partial}{\partial x} + \mu_{\omega,t}(\omega)\frac{\partial}{\partial \omega} + \frac{1}{2}\sigma_{X,t}^2\frac{\partial^2}{\partial x^2} + \frac{1}{2}\sigma_\omega(\omega)^2\frac{\partial^2}{\partial \omega^2} + \sigma_{X,t}\sigma_\omega(\omega)\frac{\partial^2}{\partial x \partial \omega}\Big]m(\omega, X_t)$$

$$\sigma_{c_N,t}(\omega) := \frac{1}{m(\omega,X_t)}\Big[\sigma_{X,t}\frac{\partial}{\partial x} + \sigma_\omega(\omega)\frac{\partial}{\partial \omega}\Big]m(\omega, X_t),$$

where

$$\mu_{\omega,t}(\omega) := \omega[r_t + \alpha\pi - \mu_Y - (\rho+\pi)m(\omega, X_t) + \sigma_Y^2]$$
$$\sigma_\omega(\omega) := -\omega\sigma_Y.$$

We already derived the participant consumption dynamics in (D.2). Exactly as in the baseline model, time-differentiate goods market clearing $Y_t = \int_{-\infty}^t \pi e^{-\pi(t-b)}c_{t,b}db$ and the definition $X_t := Y_t^{-1}\int_{\mathcal{P}_t}\pi e^{-\pi(t-b)}c_{t,b}db$ to obtain the following four equations:

$$\sigma_Y = X_t\sigma_{c_P,t} + (1-X_t)\mathcal{M}_t(\sigma_{c_N}) \tag{D.14}$$

$$X_t\sigma_Y + \sigma_{X,t} = X_t\sigma_{c_P,t} \tag{D.15}$$

$$\mu_Y = -\pi + \pi(\rho+\pi)\omega_{t,t}m_{t,t} + X_t\mu_{c_P,t} + (1-X_t)\mathcal{M}_t(\mu_{c_N}) \tag{D.16}$$

$$X_t\mu_Y + \mu_{X,t} + \sigma_Y\sigma_{X,t} = X_t\mu_{c_P,t} - X_t\pi + (1-X_t)n_t, \tag{D.17}$$

where newborns enter with relative wealth $\omega_{t,t} := (1-\alpha)p_t$, where $m_{t,t} := \mathbf{1}_{\{\omega_{t,t}<\omega^*(X_t)\}}m(\omega_{t,t}, X_t) + \mathbf{1}_{\{\omega_{t,t}\geq\omega^*(X_t)\}}$ denotes the newborns' scaled consumption-wealth ratio depending on whether they enter as non-participants or participants, and where $(1-X_t)n_tdt$ represents the change in X_t exclusively due to expansions in $\mathcal{P}_t$ (i.e., entry by non-participants). In the proof of Proposition E.1, we showed that (see equation (E.7) and use $g \equiv 1$ and the boundary condition (E.6))

$$n_t = -\frac{1}{2}\Big[\sigma_\omega(\omega^*(X_t))^2 + \big(\sigma_{X,t}\frac{\partial}{\partial x}\omega^*(X_t)\big)^2\Big]\frac{\partial}{\partial \omega}f_t(\omega)\Big|_{\omega=\omega^*(X_t)}. \tag{D.18}$$

Using the above results, and taking f_t and the functions m and ω^* as given, equations (D.14)-(D.17) become four equations in the four unknowns $(r_t, \eta_t, \mu_{X,t}, \sigma_{X,t})$, which we can solve to obtain

$$\sigma_{X,t} = \frac{(1-X_t)[\sigma_Y - \mathcal{M}_t(\sigma_\omega \frac{\partial}{\partial \omega}\log m)]}{1+(1-X_t)\mathcal{M}_t(\frac{\partial}{\partial x}\log m)} \tag{D.19}$$

$$\eta_t = \sigma_Y + \frac{\sigma_{X,t}}{X_t} \tag{D.20}$$

$$r_t = \Big(1 + (1-X_t)(\mathcal{M}_t[X_t \frac{\frac{\partial}{\partial x}m}{m}] + \mathcal{M}_t[\omega \frac{\frac{\partial}{\partial \omega}m}{m}])\Big)^{-1} \tag{D.21}$$

$$\times \Big(\mu_Y + (1-\alpha)\pi - \pi(\rho+\pi)\omega_{t,t}m_{t,t} + (\rho+\pi)(X_t + (1-X_t)\mathcal{M}_t[m]) - X_t\eta_t^2 - (1-X_t)\mathcal{M}_t[q]\Big)$$

$$\mu_{X,t} = (1-X_t)n_t - X_t\pi + X_t\Big[r_t + \alpha\pi - \rho - \pi - \mu_Y + \eta_t^2\Big] - \sigma_Y\sigma_{X,t}, \tag{D.22}$$

where

$$\begin{aligned} q_t(\omega) := {}& \omega\Big[\alpha\pi - \mu_Y + \sigma_Y^2 - (\rho+\pi)m(\omega, X_t)\Big]\frac{\frac{\partial}{\partial \omega}m(\omega,X_t)}{m(\omega,X_t)} \\ &+ \Big[(1-X_t)n_t - X_t\pi + X_t(\alpha\pi - \rho - \pi - \mu_Y + \eta_t^2) - \sigma_Y\sigma_{X,t}\Big]\frac{\frac{\partial}{\partial x}m(\omega,X_t)}{m(\omega,X_t)} \\ &+ \frac{1}{2}\frac{1}{m(\omega,X_t)}\Big[\sigma_{X,t}^2\frac{\partial^2}{\partial x^2} + \sigma_\omega(\omega)^2\frac{\partial^2}{\partial \omega^2} + 2\sigma_{X,t}\sigma_\omega(\omega)\frac{\partial^2}{\partial x \partial \omega}\Big]m(\omega, X_t). \end{aligned}$$

Thus, we have solved for $(r_t, \eta_t, p_t, \mu_{X,t}, \sigma_{X,t})$ explicitly, given beliefs, knowledge of f_t and X_t, and the policy functions m and ω^*.

Fixed point for beliefs and equilibrium. It remains to form agents' boundedly rational beliefs, namely to project all relevant functions into the space of functions depending only on x. This is similar to Krusell and Smith (1998), except that our asset pricing objects (r, η, p) are not solely functions of x (they depend on the distribution f), so they must undergo a projection.

The iterative procedure to obtain approximate beliefs and solve for a boundedly rational equilibrium is as follows:

0. **Initial guess:** Start with guesses $(\hat{r}^{(0)}, \hat{p}^{(0)}, \hat{\mu}_X^{(0)}, \hat{\sigma}_X^{(0)})$—solely functions of participants' consumption share X_t. Although later iterations of these objects will belong to a particular class of functions (e.g., polynomials), it is not necessary that these initial guesses belong to the same class.

1. **Participant value:** Using projections $(\hat{r}^{(i)}, \hat{\eta}^{(i)}, \hat{\mu}_X^{(i)}, \hat{\sigma}_X^{(i)})$ in place of $(\hat{r}, \hat{\eta}, \hat{\mu}_X, \hat{\sigma}_X)$, obtain $\hat{g}_P^{(i)}$ by solving the approximate participant HJB equation (D.3). The result will be solely a function of x.

2. **Non-participant value:** Given projections $(\hat{g}_P^{(i)}, \hat{r}^{(i)}, \hat{\mu}_X^{(i)}, \hat{\sigma}_X^{(i)})$, obtain $\hat{g}_N^{(i)}$ by solving variational inequality (D.9). The result will be solely a function of ω and x.

3. **Policy functions:** obtain the consumption-wealth ratio $m^{(i)}$ via (D.7); obtain entry threshold $(\omega^*)^{(i)}$ via the solution to (D.9), i.e., $(\omega^*)^{(i)} := \inf\{\omega :$ equation (D.8) holds with $\hat{g}_N^{(i)}$, $\hat{g}_P^{(i)}$, and $\hat{p}^{(i)}$ $\}$.

4. **Simulate:** Using the calculated objects from iteration (i), simulate time-paths for $(X_t)_{t\in[0,T]}$, $(f_t)_{t\in[0,T]}$, and $(r_t, \eta_t, p_t, \mu_{X,t}, \sigma_{X,t})_{t\in[0,T]}$ using equations (D.11), (D.19)-(D.22), and Proposition E.1, as well as a simulation scheme (e.g., first-order Euler). Burn the first τ length of time from the simulations, chosen large enough such that the remaining paths are approximately drawn from the ergodic distribution.

5. **Project:** For each time-path $(z_t)_{t\in[\tau,T]} \in \{r_t, \eta_t, p_t, \mu_{X,t}, \sigma_{X,t} : t \in [\tau, T]\}$, obtain an approximating function $\hat{z} \in \{\hat{r}^{(i+1)}, \hat{\eta}^{(i+1)}, \hat{p}^{(i+1)}, \hat{\mu}_X^{(i+1)}, \hat{\sigma}_X^{(i+1)}\}$ by running a univariate nonlinear regression as follows

$$\min_{\hat{z}\in\mathcal{Z}} \frac{1}{T-\tau} \int_\tau^T \Big(z_t - \hat{z}(X_t)\Big)^2 dt, \tag{D.23}$$

where $\mathcal{Z}$ is a class of functions (e.g., polynomials in x; step functions on an x-grid does "nonparametric regression"). We impose certain boundary conditions when selecting the function class $\mathcal{Z}$, which is necessary to preserve global well-posedness of the optimization problems. In particular, we impose that as $x \to 1$, the economy (and thus beliefs) converges to the complete-integration economy. At the other boundary, we impose asymptotic conditions that are known to hold in the proportional cost economy. Mathematically,

$$\begin{aligned}
\hat{r}(1) &= \mu_Y + \rho + \pi - \sigma_Y^2 &\quad \text{and} \quad& \lim_{x\to 0} x\hat{r}(x) = \text{constant}\\
\hat{\eta}(1) &= \sigma_Y &\quad \text{and} \quad& \lim_{x\to 0} x\hat{\eta}(x) = \text{constant}\\
\hat{\mu}_X(1) &= -(1-\alpha)\pi &\quad \text{and} \quad& \lim_{x\to 0} x\hat{\mu}_X(x) = \text{constant}\\
\hat{\sigma}_X(1) &= 0 &\quad \text{and} \quad& \hat{\sigma}_X(0) = \sigma_Y\\
\hat{p}(1) &= (\rho+\pi)^{-1} &\quad \text{and} \quad& \hat{p}(0) = \text{constant}.
\end{aligned}$$

For $\hat{\sigma}_X$, we also make sure it always takes strictly positive values by imposing it be at least 0.001. In forming the updated projections of these objects, we dampen the updating by taking a weighted average of the old estimates and new estimates, e.g., we set $\hat{\eta}^{(i+1)}$ as a weighted average of $\hat{\eta}^{(i)}$ and the solution to (D.23).

6. **Iterate:** Stop if the new and old estimates are close, i.e., if $(\hat{r}^{(i+1)}, \hat{\eta}^{(i+1)}, \hat{p}^{(i+1)}, \hat{\mu}_X^{(i+1)}, \hat{\sigma}_X^{(i+1)})$ is close to $(\hat{r}^{(i)}, \hat{\eta}^{(i)}, \hat{p}^{(i)}, \hat{\mu}_X^{(i)}, \hat{\sigma}_X^{(i)})$ in some metric.

Along the way to solving the model, we obtain a measure of fitness in step 5 (e.g., R-squared) for each equilibrium object and each iteration. Krusell and Smith (1998) includes an additional layer, whereby more moments from the cross-sectional distribution are added (and the space $\mathcal{Z}$

is expanded to functions with a higher-dimensional domain), if the goodness-of-fit in step 5 is inadequate.

Results. First, just to get a sense of how optimal behavior works in this model, Figure D.1 plots a comparison of the proportional and fixed cost economies, both with $\phi = 0.4$. There is non-trivial dependence of the non-participant value function on their wealth-to-income ratio ω. Higher ω is beneficial to non-participants, but much more so if participants' consumption share x is low, which is when risk prices are particularly high. Relatedly, the "net entry benefit," defined as $\hat{g}_P - \hat{g}_N - \tilde{\Phi}$ depends non-trivially on both ω and x. To see it a different way, Figure D.2 plots the wealth-to-income entry threshold $\omega^*(x)$ (plotted relative to the perceived price-dividend ratio $\hat{p}(x)$ so that the threshold is interpreted as "how many times richer than average is the marginal entrant").

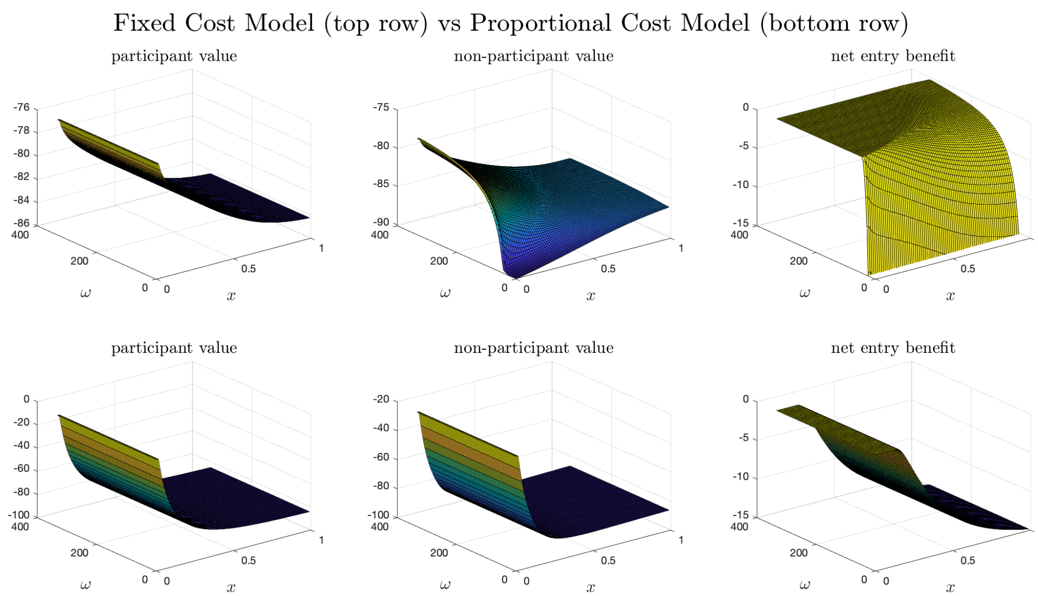

Figure D.1: A comparison of the proportional cost economy to the fixed cost economy, with entry parameter $\phi = 0.4$ in both. The value functions in the top row are $\hat{g}_P$ and $\hat{g}_N$. All other parameters are as in Table 1.

The resulting equilibrium objects are plotted in Figure D.3. The blue dots are outcomes of a simulation, whereas the red lines are their projections into the space of functions of X_t. In particular, we fit the simulated data, suitably adjusted so that boundary conditions in step 5 above are automatically satisfied, to quadratic functions of X_t. For instance, if $(X_t, \eta_t)_{t \geq 0}$ are simulated data, we first fit the quadratic function $\hat{a}(X_t) := \hat{a}_0 + \hat{a}_1 X_t + \hat{a}_2 X_t^2$ to $(\eta_t - \sigma_Y)\frac{X_t}{1-X_t}$, which ensures that $\hat{\eta}(x) := \sigma_Y + \frac{1-x}{x}\hat{a}(x)$ satisfies the appropriate boundary conditions. As in Krusell and Smith (1998), we can compute the R-squared from the projection, and we obtain the following for the fitting of $\hat{r}$, $\hat{\eta}$, $\hat{\mu}_X$, $\hat{\sigma}_X$, and $\hat{p}$, respectively: 0.9887, 0.9993, 0.9526, 0.9378, and 0.9401. Note in particular how tight the fit of $\hat{\eta}$ is, a result which can also be seen in Figure D.3. This is

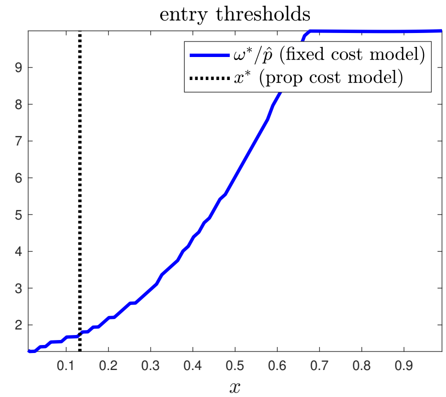

Figure D.2: Entry threshold in the fixed cost economy, in units of multiples of the average perceived wealth, i.e., $\omega^*/\hat{p}$ is the wealth of the marginal entrant relative to the average perceived wealth. The grid for the model solution is truncated at $\omega = 10(\rho+\pi)^{-1}$, which is why the entry threshold becomes flat at high values of x. For comparison, the vertical dotted line is the entry threshold from the proportional cost economy. Both economies have entry parameter $\phi = 0.4$. All other parameters are as in Table 1.

reassuring, given entry benefits depend particularly strongly on η_t, an intuition we made exact in the proportional cost model through formula (29).

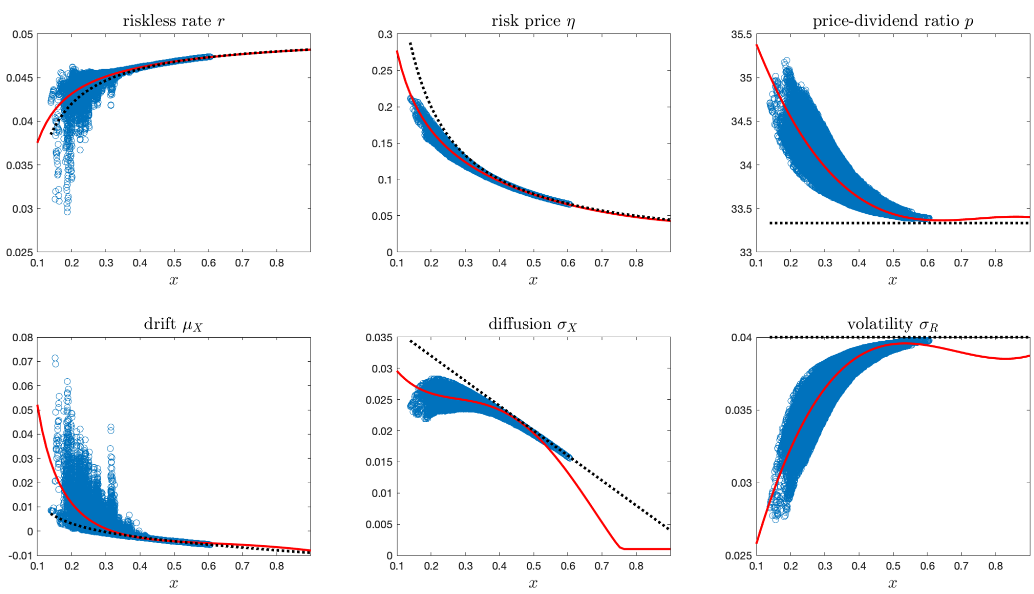

Figure D.3: Equilibrium objects (blue dots, from a 5000 year simulation) and their fitted perceived values (red lines) in the fixed cost economy. The entry cost parameter is $\phi = 0.4$. For comparison, the black dotted line plots the equilibrium objects in the proportional cost economy with the same ϕ. All other parameters are as in Table 1.

For comparison, the dotted black lines in Figure D.3 represent the equilibrium objects from the corresponding proportional cost economy (i.e., with the same $\phi = 0.4$). Notice that risk prices

η_t are slightly lower with fixed costs, and the dynamics of X_t are less volatile (lower $\sigma_{X,t}$) and less conducive to reaching bad states (i.e., $\mu_{X,t}$ is higher for low X_t than in the proportional cost economy). Although slightly tangential to the main arguments of this paper, note also that fixed entry costs induce some counterfactual behavior in the price-dividend ratio and return volatility (in particular, p_t becomes countercyclical, $\sigma_{R,t}$ becomes procyclical, and $\sigma_{R,t}$ is reduced relative to the proportional cost economy).

D.2 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 the model if agents also receive labor income endowments? For simplicity, and to focus on the relevant issues, we proceed under the assumption that labor income is perfectly correlated with capital income (i.e., it is only subject to the aggregate shock dZ_t and carries no idiosyncratic risk).

In the model with labor income, each newborn born at time b is endowed with no financial income, but they receive a labor tree that pays the stream $\{(1-\tilde{\alpha})Y_t\}_{t\in[b,T]}$, where T is the agent's random time of death. The total labor income in the economy is thus a fraction $1-\tilde{\alpha}$ of aggregate output. The stock market is a claim to the residual $\{\tilde{\alpha}Y_t\}$, i.e., capital accounts for $\tilde{\alpha}$ fraction of total income.

In addition, we assume as in Blanchard (1985) and Gârleanu and Panageas (2015) that all agents have full access to actuarily-fair annuity markets for insurance against their Poisson death shocks—our baseline model limits the degree of participation in these annuity markets to a fraction α of wealth, so that $1-\alpha$ fraction of dying wealth is redistributed as "unintended bequests" to newborns.

Next, we consider two cases: (1) full pledgeability of labor income and (2) non-pledgeability. Obviously, reality is somewhere in between these two cases, as there exist markets to borrow against labor income, but these markets are imperfect due to financial frictions and ethical issues.

Full pledgeability. First, I prove Lemma 1, which says that my baseline setup is essentially equivalent to a setting with labor income, *if the labor income is pledgeable.*

Proof of Lemma 1. Because the tree is risky, and because agents are risk-averse, all newborns will use this opportunity to sell the entirety of their tree. Indeed, they receive a fair market price, namely participants' willingness-to-pay; this value is always higher than non-participants' shadow valuation of their own labor income, because participants can hedge labor income risks and non-participants cannot. Thus, after selling their tree, a time-t newborn begins life with $(1-\tilde{\alpha})Y_t\ell_t$ units of financial wealth, where

$$\ell_t := \mathbb{E}_t\Big[\int_0^\infty e^{-\pi s}\frac{\xi_{t+s}}{\xi_t}\frac{Y_{t+s}}{Y_t}ds\Big] \tag{D.24}$$

is a scaled measure of human wealth-income, which is computed using participants' marginal utility ξ_t (i.e., the stochastic discount factor for the economy). As in the baseline model, the financial wealth received by non-participants must be held in riskless bonds until the participation cost is paid. Non-participants will still consume $\rho+\pi$ fraction of this financial wealth, but they receive flow payments of π for their annuities, meaning an individual non-participant's consumption dynamics are

$$dc_{N,t} = c_{N,t}[r_t - \rho]dt.$$

Newborn consumption equals $(\rho+\pi)(1-\tilde{\alpha})Y_t\ell_t$.

Given the discussion above, participants hold, in addition to their own financial wealth, all claims to (alive) human wealth in the economy, whose total value is $\ell_t Y_t$ at time t. Nevertheless, the participant optimization problem remains the same, as they face dynamically complete markets: because participants have unconstrained access to both the stock market and the market for this human capital, and since both values respond to the same shock dZ_t, participants can effectively pick their exposure to dZ_t in an unconstrained fashion, earning the risk price η_t, as in the baseline model. Consequently, participants will consume a fraction $\rho+\pi$ of their "total wealth" (sum of their financial capital and the human capital they have purchased), and their consumption dynamics are

$$dc_{P,t} = c_{P,t}[r_t - \rho + \eta_t^2]dt + c_{P,t}\eta_t dZ_t.$$

Given these consumption policies, the price-dividend ratio p for financial capital is determined via

$$\tilde{\alpha}p_t + (1-\tilde{\alpha})\ell_t = (\rho+\pi)^{-1}. \tag{D.25}$$

In the baseline model without labor, we had $p_t = (\rho+\pi)^{-1}$; any level or cyclicality modifications to p_t are inherited oppositely by ℓ_t.

Because the participant consumption share X_t remains the sole state variable, we follow a similar procedure as in the baseline model to determine endogenous objects r, η, μ_X, and σ_X. The solutions $\eta = \sigma_Y/x$ and $\sigma_X = (1-x)\sigma_Y$ remain the same as in Proposition 1, independent of $\tilde{\alpha}$. By contrast, r and μ_X are now given by

$$\begin{aligned} r &= \rho + \pi + \mu_Y - \pi(1-\tilde{\alpha})(\rho+\pi)\ell - \frac{\sigma_Y^2}{x} \\ \mu_X &= -x\pi(1-\tilde{\alpha})(\rho+\pi)\ell + \frac{(1-x)^2}{x}\sigma_Y^2. \end{aligned}$$

Comparing to the expressions in Proposition 1, we see that μ_X will remain unchanged if and only if

$$(1-\tilde{\alpha})(\rho+\pi)\ell_t = 1-\alpha. \tag{D.26}$$

The only remaining object needed for equilibrium is the human wealth-income ratio ℓ_t. Standard arguments can be used to obtain an ODE for $\ell_t = \ell(X_t)$, i.e., at times of non-entry we have

$$0 = 1 + [\mu_Y - \eta\sigma_Y - \pi - r]\ell + [\mu_X - (\eta - \sigma_Y)\sigma_X]\ell' + \frac{1}{2}\sigma_X^2\ell''. \tag{D.27}$$

Conjecture a constant solution $\ell_t = \bar{\ell}$. Setting $\ell' = \ell'' = 0$ and plugging in r and η, we then obtain the quadratic equation

$$0 = 1 - [\rho + 2\pi]\bar{\ell} + \pi(1 - \tilde{\alpha})(\rho + \pi)\bar{\ell}^2,$$

which has the two roots

$$\bar{\ell}_{\pm} := \frac{1}{2\pi(1-\tilde{\alpha})(\rho+\pi)}\Big[\rho + 2\pi \pm \sqrt{(\rho+2\pi)^2 - 4\pi(1-\tilde{\alpha})(\rho+\pi)}\Big]. \tag{D.28}$$

One can show that $p_t > \ell_t$ is required in equilibrium (because the cash flows of capital have an infinite horizon, unlike labor income that disappears at rate π). One can also show that $\bar{\ell}_+ > (\rho+\pi)^{-1} > \bar{\ell}_-$ (if and only if $\tilde{\alpha} < 1$). Thus, equation (D.25) implies the smaller root $\bar{\ell}_-$ must be chosen. This proves that a unique constant solution $\ell_t = \bar{\ell}_-$ satisfies equilibrium.

Substituting $\bar{\ell}_-$ back into the condition (D.26), we find that we must set α according to (35). If we make this choice, then (η, μ_X, σ_X) are identical, as functions of X_t, in the baseline model and the present model with labor income. Furthermore, following identical logic as in Proposition 3, non-participant entry decisions are determined solely by the participation cost ϕ and the dynamics of η_t (i.e., the riskless rate r_t plays no role). Thus, the entry barrier x^* remains identical. This proves that risk price time-paths $\{\eta_t\}$ are isomorphic between the two models, under the parameter choice (35). Similarly, given an initial condition X_0, the time-paths of $\{X_t\}$ are identical between the two models. Finally, substituting condition (35) into the expression for r_t and comparing to Proposition 1, we see that the riskless rate r_t is lower by the constant level $(1-\alpha)\pi$ in the present labor income model, relative to the baseline model. □

Non-pledgeability. Solving the model with non-pledgeable human capital is substantially more complicated, because non-participant consumption will not only depend on total wealth, but also the split between financial and human wealth. As this ratio is heterogeneous across non-participants, the entire distribution of non-participant financial wealth-to-income ratios becomes an aggregate state variable. The solution is even more complicated than the infinite-dimensional equilibrium of Appendix D.1, because the non-participant optimization problem is not separable in financial wealth and labor income. Instead of tackling this problem directly, we study a simplified partial equilibrium environment that nevertheless allows some comparisons to the baseline model.

In particular, we suppose asset prices are given by their full-integration benchmark (in which markets are complete). We will also allow agents to have recursive utility with risk aversion γ and EIS ψ^{-1}, as in Appendix B. In this environment, we will compute measures of willingness-to-pay (to enter) of a hypothetical non-participant. We will also compute the entry willingness-to-pay of a hypothetical non-participant who additionally has pledgeable labor income. By comparing these cases, we will get a sense of how non-pledgeability of labor income affects entry incentives.

By Proposition 1 of Gârleanu and Panageas (2015), equilibrium of our full-integration economy satisfies

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

The equation for r is a quadratic equation, and we must select the larger of the two roots so that aggregate wealth is well-defined. Indeed, we can immediately calculate (using the Gordon growth formula)

$$
\begin{aligned}
\ell &= (r + \pi - \mu_Y + \sigma_Y\eta)^{-1} \\
p &= (r - \mu_Y + \sigma_Y\eta)^{-1},
\end{aligned}
$$

where the labor valuation ℓ is defined in (D.24). Thus, the riskless rate is

$$
\begin{aligned}
r &= \frac{1}{2}\Big[B + \Big(B^2 - 4\pi(1-\tilde{\alpha})(\rho + \psi\pi + \frac{\gamma}{2}(\psi-1)\sigma_Y^2) + 4r^*(\pi + \gamma\sigma_Y^2 - \mu_Y)\Big)^{1/2}\Big] \\
\text{where} \quad B &:= r^* - \pi - \gamma\sigma_Y^2 + \mu_Y + \pi(1-\tilde{\alpha})(\psi-1) \\
\text{and} \quad r^* &:= \rho + \psi(\mu_Y + \pi) - \frac{\gamma}{2}(\psi+1)\sigma_Y^2
\end{aligned}
$$

Next, consider participants' optimization problem, which is unaffected by the non-tradability of their labor endowment, because of their access to complete financial markets. Indeed, we only need to use "total wealth" (sum of financial plus human wealth)

$$
\tilde{W}_{t,b} := W_{t,b} + (1-\tilde{\alpha})\ell Y_t,
$$

in place financial wealth as the relevant individual state variable. Omitting the details (one can also see the Appendix A of Gârleanu and Panageas (2015)), a participant with total wealth $\tilde{w}$ obtains indirect utility

$$
\begin{aligned}
V^P(\tilde{w}) &= \frac{(A(\eta)\tilde{w})^{1-\gamma}}{1-\gamma}, \\
\text{where} \quad A(\eta) &:= \psi^{-\frac{\psi}{\psi-1}}\Big[\rho + \pi + (\psi-1)(r + \pi + \frac{1}{2}\frac{\eta^2}{\gamma})\Big]^{\frac{\psi}{\psi-1}}.
\end{aligned}
$$

We explicitly write the dependence of A on η because a hypothetical non-participant will obtain indirect utility that depends on $A(0)$.

To solve a hypothetical non-participant's optimization problem, we follow the analysis of Wang, Wang, and Yang (2016). Define the financial wealth-to-labor-income ratio $\omega_t := W_t/(1-\tilde{\alpha})Y_t$, which will be the relevant state variable. Recall the financial wealth of this non-participant evolves as

$$dW_t = [W_t(r+\pi) - c_t + (1-\tilde{\alpha})Y_t]dt, \quad W_0 = 0.$$

Thus, the dynamics of ω_t are given by

$$d\omega_t = \Big[\omega_t(r+\pi-\mu_Y+\sigma_Y^2) - \frac{c_t}{(1-\tilde{\alpha})Y_t} + 1\Big]dt - \omega_t\sigma_Y dZ_t, \quad \omega_0 = 0.$$

Following Wang, Wang, and Yang (2016), we assume a borrowing constraint $W_t \geq 0$ for all t, which implies $\omega_t \geq 0$. As shown in Proposition 2 of Wang, Wang, and Yang (2016), the value function of this agent then satisfies

$$V^N(Y_t, \omega_t) = \frac{(A(0)(1-\tilde{\alpha})Y_t q(\omega_t))^{1-\gamma}}{1-\gamma},$$

where the function q is a measure of "certainty equivalent wealth" and solves the following ODE:

$$0 = \Big(-\frac{(A(0)q')^{1-\psi^{-1}} - \psi^{-1}(\rho+\pi)}{1-\psi^{-1}} + \mu_Y - \frac{\gamma}{2}\sigma_Y^2\Big)q + [1+\omega(r+\pi-\mu_Y+\gamma\sigma_Y^2)]q' + \frac{1}{2}(\omega\sigma_Y)^2\Big[q'' - \gamma\frac{(q')^2}{q}\Big]$$

subject to the boundary conditions $\lim_{\omega\to\infty} q'(\omega) = 1$ and $A(0)^{\psi-1}q(0)^{\psi} \leq q'(0)$. The latter boundary condition binds if and only if the borrowing constraint ever binds, since the consumption-to-income ratio is $A(0)^{1-\psi^{-1}}q(\omega)q'(\omega)^{-\psi^{-1}}$. For the parameters we use below, the borrowing constraint never binds, so instead we may simply use as a boundary condition the ODE itself in the limit as $\omega \to 0$ (i.e., the "natural boundary" condition). Finally, note that we can also write the participant value function V^P in the same form as V^N, which facilitates comparison:

$$V^P(Y_t, \omega_t) = \frac{(A(\eta)(1-\tilde{\alpha})Y_t q^*(\omega_t))^{1-\gamma}}{1-\gamma},$$
$$\text{where} \quad q^*(\omega) := \omega + \ell.$$

Figure D.4 plots a comparison of participants to the hypothetical non-participant in this environment. Note that non-participants have a concave consumption function (middle panel), although the concavity is very mild given labor income risk $\sigma_Y = 0.04$ is small. Notice that participants' marginal value of wealth (right panel) is greater than one, not because of any financial friction, but because they are uniquely able to earn the Sharpe ratio $\eta > 0$.

Figure D.5 plots a non-participant's willingness-to-pay (WTP) to become a participant. The WTP is defined in various ways: as a fraction of financial wealth; as a fraction of total wealth; and as a fraction of aggregate total wealth. It is straightforward to show that these are correctly

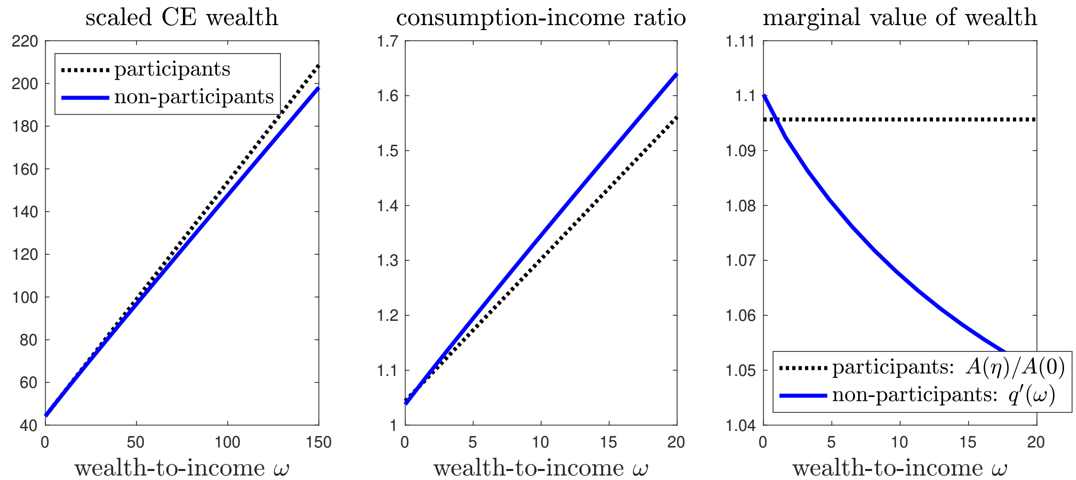

Figure D.4: Equilibrium objects in the non-pledgeable labor income economy. Non-participants (blue lines) cannot hedge labor income shocks. Participants (dashed black lines) can hedge shocks, and they additionally receive/pay risk price η when trading this shock. Risk aversion and EIS are set to $\gamma = 3$ and $\psi^{-1} = 4/3$. The labor share is set to $1 - \tilde{\alpha} = 2/3$. All other parameters are as in Table 1.

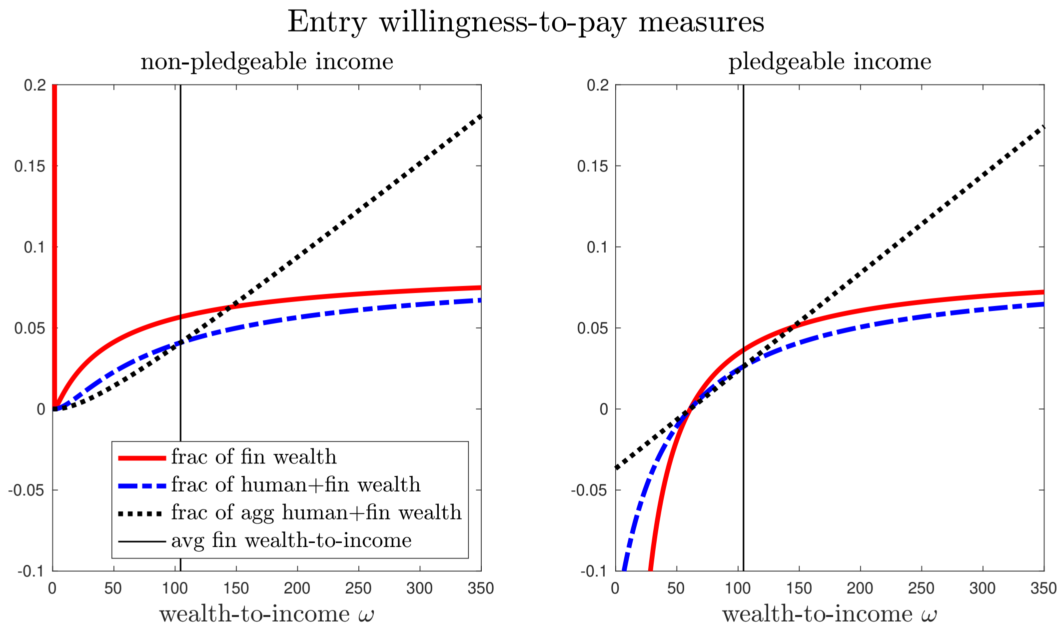

Figure D.5: The willingness-to-pay of a non-participant to become a participant in the full-integration economy outlined above. Left panel: the non-participant has non-pledgeable labor income. Right panel: the non-participant has pledgeable labor income. Risk aversion and EIS are set to $\gamma = 3$ and $\psi^{-1} = 4/3$. The labor share is set to $1 - \tilde{\alpha} = 2/3$. All other parameters are as in Table 1.

defined as functions $e(\omega)$ that satisfy

$$\begin{aligned}
\text{(fraction of financial wealth):} \quad & V^N(Y, \omega) = V^P(Y, (1 - e(\omega))\omega) \\
\text{(fraction of total wealth):} \quad & V^N(Y, \omega) = V^P((1 - e(\omega))Y, \omega) \\
\text{(fraction of aggregate total wealth):} \quad & V^N(Y, \omega) = V^P(Y, \omega - \frac{\tilde{\alpha} p + (1 - \tilde{\alpha})\ell}{1 - \tilde{\alpha}} e(\omega))
\end{aligned}$$

For comparison, we also compute the same WTP measures but for a hypothetical non-participant whose labor income is pledgeable, as in the baseline model. Figure D.5 shows that WTPs are uniformly higher with non-pledgeable labor income, as compared to pledgeable income. Surprisingly, at low levels of ω, a non-participant with pledgeable income would actually need to be paid to participate: this finding arises because poor non-participants are using financial markets to hedge their labor income, and entry implies they will have to pay a non-zero Sharpe ratio η to continue performing this hedge. At high values of ω, where hedging becomes relatively unimportant, the WTP measures converge. Overall, this comparison suggests that non-pledgeability of labor income would raise implied participation costs, if anything.

D.3 Other preference constellations

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 γ, respectively, we can understand how much stronger the entry incentives are for less risk-averse agents. The result of this analysis, displayed in the left panel of Figure D.6, shows that $e(x; \hat{\gamma}) > e(x; \gamma)$ for $\hat{\gamma} < \gamma$, and vice versa for $\hat{\gamma} > \gamma$, as conjectured. Intuitively, 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.

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

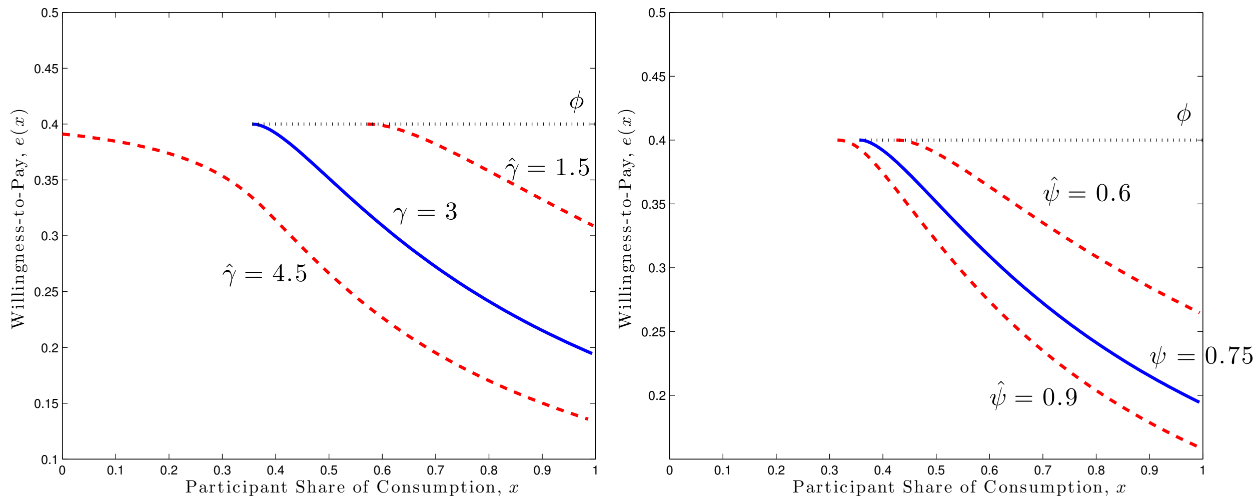

Figure D.6: Left panel: Functions $e(x; \hat{\gamma})$ and $e(x; \gamma)$ are entry willingnesses-to-pay, as a fraction of wealth, for two agents with risk aversions $\hat{\gamma}$ and γ, living in an economy populated by γ-agents. Right panel: Functions $e(x; \hat{\psi})$ and $e(x; \psi)$ are entry 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.

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. The right panel of Figure D.6 repeats the analysis from the left panel 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 (30) 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 (Luttmer and Mariotti, 2003). If, in addition, we hold the interest rate fixed, this back-of-the-envelope calculation suggests that hyperbolic discounting can dramatically lower participation incentives and thus induce more risk concentration.

E Dynamics of the cross-sectional distribution of wealth

Some of the extensions we consider (e.g., fixed entry costs) do not possess the same homogeneity properties as the baseline model. Consequently, the equilibria of these extensions cannot be summarized by the aggregate consumption share of participants (a scalar state variable). Instead, these equilibria depend on the entire cross-sectional distribution of wealth. This appendix develops analytical formulae for the dynamics of this distribution, formulae which will apply to all extensions of interest. These formulae are analogous to the Kolmogorov Forward equation, but with the addition of aggregate shocks. Thus, instead of a partial differential equation (PDE), the result will be a stochastic partial differential equation (SPDE).

Define the ratio of cohort-level wealth to aggregate income:

$$\omega_{t,b} := \frac{W_{t,b}}{Y_t}, \quad b \leq t. \tag{E.1}$$

As before, agents are identical within a cohort, so $\omega_{t,b}$ can be thought of as an individual level variable. Because participants are unconstrained, their policies are homogeneous in their wealth, and ω plays no role for them. As in the baseline model, participants can be aggregated into a representative participant. By contrast, non-participants decisions (consumption and entry) will depend non-trivially on their individual wealth-to-income ω.

The key assumption we make is that non-participant decisions can be written as functions of individual wealth-to-income ω and a *finite set of aggregate states*. In a rational expectations equilibrium, the full set of aggregate states would include the full cross-sectional distribution of $\omega_{t,b}$, because decisions depend non-trivially on ω. This distribution is an infinite dimensional object, which precludes a numerical solution. For the purposes of this paper, we follow the spirit of Krusell and Smith (1998) and take a finite set of "moments" of the distribution as summary statistics; this can be understood as a kind of "bounded rationality."

Furthermore, in our computations and the analysis below, we will specialize the finite set of moments to a single moment (although the analysis is very easily generalized to any finite number of moments). For this purpose, we continue to let X_t denote the participant consumption share, which is assumed to have Itô dynamics

$$dX_t = \mu_{X,t}dt + \sigma_{X,t}dZ_t.$$

(Note the absence of a singularly continuous term here, as it will turn out that entry will occur at the dt order in the relevant extensions where the full distribution of wealth matters.) Going forward, we will use X_t as the aggregate state variable agents use in approximating their more complex environment.[33] Nevertheless, even though agents will be performing this dimension-reduction,

[33] This is similar to the first moment used by Krusell and Smith (1998), for the following reason. Note that X_tY_t represents aggregate participant consumption, so that $(1 - X_t)Y_t$ represents aggregate non-participant consumption. As in the baseline model, all equilibrium objects will either scale with Y_t or remain independent

many equilibrium objects will depend on more than just X_t. To compute such an equilibrium, we require dynamics equations for the entire distribution of wealth-to-income.

Now, we proceed more formally. In particular, let $m(\omega, x)$ denote the (scaled) non-participant consumption-to-wealth ratio, i.e.,

$$m(\omega_{t,b}, X_t) := (\rho + \pi)^{-1} \frac{c_{t,b}}{W_{t,b}}. \tag{E.2}$$

In the baseline version of the model, we had $m \equiv 1$. Additionally, let $\omega^*(x)$ denote the non-participant entry threshold: non-participants with $\omega_{t,b} \geq \omega^*(X_t)$ pay the entry cost to become participants; those with $\omega_{t,b} < \omega^*(X_t)$ remain non-participants. Thus, the indicator function $e(\omega, X_t) := \mathbf{1}_{\{\omega \geq \omega^*(X_t)\}}$ represents entry decisions.

Given any policy function $g(\omega, x)$ of this form, and a cross-sectional density $f_t(\omega)$ of non-participant wealth-to-income ratios, we compute non-participant aggregates as

$$\int_0^\infty g(\omega, X_t) f_t(\omega) d\omega = \frac{1}{(1 - X_t) Y_t} \int_{\mathcal{N}_t} \pi e^{-\pi(t-b)} c_{t,b} g(\omega_{t,b}, X_t) db, \tag{E.3}$$

where recall $\mathcal{N}_t$ denotes the set of non-participants. Note that the density f_t is required to integrate to one, i.e., $\int_0^\infty f_t(\omega) d\omega = 1$. Similarly, on the right-hand-side, the expression $\pi e^{-\pi(t-b)} \frac{c_{t,b}}{(1-X_t)Y_t}$ integrates to one, making it also a density. Therefore, equation (E.3) is an equality between two cross-sectional expected values; the particular type of integral appearing on the right-hand-side recurs repeatedly in derivation of equilibrium, which is why we seek a density satisfying (E.3). In fact, equation (E.3) serves as the definition of f_t, since it needs to hold for every policy function g in an appropriate space of functions. For our purposes, the following space suffices.

Definition E.1. *The space $\mathcal{G}$ consists of functions $g : \mathbb{R}_+ \times (0, 1) \mapsto \mathbb{R}$ that are twice continuously differentiable and have compact support.*

To derive the appropriate dynamic equation, we will time-differentiate (E.3), giving the following result.

Proposition E.1. *Assume equation (E.3) holds for almost all t and for each function $g \in \mathcal{G}$. Assume that non-participant policies depend only on aggregate states through X_t; furthermore, assume the non-participant consumption-wealth ratio $m \in \mathcal{G}$, and assume ω^* is twice continuously differentiable. Then, the density $f_t(\omega)$ satisfies the following stochastic partial differential equation*

of Y_t, so X_t can be thought of as capturing the first moment of the consumption distribution.

on $\omega < \omega^*(X_t)$, subject to $f_t(\omega) = 0$ for all $\omega \geq \omega^*(X_t)$:

$$
\begin{aligned}
df_t(\omega) &= \mu_{f,t}(\omega)dt + \sigma_{f,t}(\omega)dZ_t, \quad \textit{where} \\
\mu_{f,t}(\omega) &:= \Big[-\pi + \frac{\mu_{X,t}}{1-X_t} - \mu_Y + \frac{\sigma_{X,t}\sigma_Y}{1-X_t} + \frac{\sigma_{X,t}^2}{(1-X_t)^2} + \sigma_Y^2 + \mu_{c_N,t}(\omega) + \Big(\frac{\sigma_{X,t}}{1-X_t} - \sigma_Y\Big)\sigma_{c_N,t}(\omega)\Big] f_t(\omega) \\
&\quad - \frac{\partial}{\partial \omega}\Big[\Big(\mu_{\omega,t}(\omega) + \Big(\frac{\sigma_{X,t}}{1-X_t} - \sigma_Y + \sigma_{c_N,t}(\omega)\Big)\sigma_{\omega,t}(\omega)\Big) f_t(\omega)\Big] + \frac{1}{2}\frac{\partial^2}{\partial \omega^2}[\sigma_{\omega,t}(\omega)^2 f_t(\omega)] \\
&\quad + \frac{\pi(\rho+\pi)\omega_{t,t} m(\omega_{t,t}, X_t)}{1-X_t}\mathrm{Dirac}_{\omega_{t,t}}(\omega) \\
\sigma_{f,t}(\omega) &:= \Big(\frac{\sigma_{X,t}}{1-X_t} - \sigma_Y + \sigma_{c_N,t}(\omega)\Big) f_t(\omega) - \frac{\partial}{\partial \omega}[\sigma_{\omega,t}(\omega) f_t(\omega)],
\end{aligned}
$$

where $(\mu_{c_N}, \sigma_{c_N})$ denote the geometric drift and diffusion of non-participant consumption (reported in (E.5)), where $(\mu_\omega, \sigma_\omega)$ denote the drift and diffusion of non-participant wealth-to-income ratios (reported in (E.4)), and where $\mathrm{Dirac}_{\omega_{t,t}}(\omega)$ *is the Dirac delta function at the time-t newborn level of wealth-to-income, $\omega_{t,t}$.*

PROOF OF PROPOSITION E.1. Before obtaining the results, we state some preliminaries. First, recall the dynamics of $W_{t,b}$ and Y_t are

$$
\begin{aligned}
dW_{t,b} &= W_{t,b}\big[r_t + \alpha\pi - (\rho+\pi)m(\omega_{t,b}, X_t)\big]dt, \quad W_{t,t} \text{ given} \\
dY_t &= Y_t\big[\mu_Y dt + \sigma_Y dZ_t\big].
\end{aligned}
$$

By Itô's formula, we obtain the dynamics of the wealth-to-income ratio:

$$
\begin{aligned}
d\omega_{t,b} &= \mu_{\omega,t}(\omega_{t,b})dt + \sigma_{\omega,t}(\omega_{t,b})dZ_t, \quad \omega_{t,t} \text{ given}, \quad \text{where} \\
\mu_{\omega,t}(\omega) &:= \omega\Big[r_t + \alpha\pi - \mu_Y - (\rho+\pi)m(\omega, X_t) + \sigma_Y^2\Big] \\
\sigma_{\omega,t}(\omega) &:= -\omega\sigma_Y.
\end{aligned} \tag{E.4}
$$

Next, given $g \in \mathcal{G}$ (see Definition E.1), we have the following useful evolution equation:

$$
\begin{aligned}
dg(\omega_{t,b}, X_t) &= \mu_{g,t}(\omega_{t,b})dt + \sigma_{g,t}(\omega_{t,b})dZ_t \\
\mu_{g,t}(\omega) &:= \Big[\mu_{X,t}\frac{\partial}{\partial x} + \mu_{\omega,t}(\omega)\frac{\partial}{\partial \omega} + \frac{1}{2}\sigma_{X,t}^2\frac{\partial^2}{\partial x^2} + \frac{1}{2}\sigma_{\omega,t}(\omega)^2\frac{\partial^2}{\partial \omega^2} + \sigma_{X,t}\sigma_{\omega,t}(\omega)\frac{\partial^2}{\partial x \partial \omega}\Big] g(\omega, X_t) \\
\sigma_{g,t}(\omega) &:= \Big[\sigma_{X,t}\frac{\partial}{\partial x} + \sigma_{\omega,t}(\omega)\frac{\partial}{\partial \omega}\Big] g(\omega, X_t).
\end{aligned}
$$

The dynamics of $c_{t,b}$ can also be obtained via Itô's formula on $c_{t,b} = (\rho+\pi)m(\omega_{t,b}, X_t)W_{t,b}$, assuming that $m \in \mathcal{G}$ as well. In particular, if $\mu_{m,t}(\omega)$ and $\sigma_{m,t}(\omega)$ are the drift and diffusion coefficients of

m, analogously to g above, then

$$dc_{t,b} = c_{t,b}\big[\mu_{c_N,t}(\omega_{t,b})dt + \sigma_{c_N,t}(\omega_{t,b})dZ_t\big], \quad \text{where} \tag{E.5}$$
$$\mu_{c_N,t}(\omega) := r_t + \alpha\pi - (\rho+\pi)m(\omega, X_t) + \frac{\mu_{m,t}(\omega)}{m(\omega, X_t)}$$
$$\sigma_{c_N,t}(\omega) := \frac{\sigma_{m,t}(\omega)}{m(\omega, X_t)}.$$

Finally, note that the optimal entry policy $e(\omega, x) = \mathbf{1}_{\{\omega \geq \omega^*(x)\}}$, substituted into (E.3), implies the boundary condition

$$f_t(\omega^*(X_t)) = 0. \tag{E.6}$$

Consequently, all integrals of the form $\int_0^\infty g(\omega, X_t)f_t(\omega)d\omega$ can be truncated as $\int_0^{\omega^*(X_t)} g(\omega, X_t)f_t(\omega)d\omega$. At the other boundary, we will also make use of the fact that $f_t(0) = \frac{\partial}{\partial\omega}f_t(0) = 0$ naturally due to the dynamics of $\omega_{t,b}$. With these preliminaries, we can determine the dynamics of f_t.

Time-differentiate (E.3) to obtain

$$\int_0^{\omega^*(X_t)} \Big[g(\omega, X_t)df_t(\omega) + f_t(\omega)dg(\omega, X_t) + \sigma_{X,t}\sigma_{f,t}(\omega)\frac{\partial}{\partial x}g(\omega, X_t)dt\Big]d\omega$$
$$+ \frac{1}{2}\Big(\sigma_{X,t}\frac{\partial}{\partial x}\omega^*(X_t)\Big)^2 \frac{\partial}{\partial\omega}[g(\omega, X_t)f_t(\omega)]\big|_{\omega=\omega^*(X_t)}dt$$

$$= \frac{\pi c_{t,t}g(\omega_{t,t}, X_t)\mathbf{1}_{\{\omega_{t,t}<\omega^*(X_t)\}}dt}{(1-X_t)Y_t} + \frac{1}{(1-X_t)Y_t}\int_{\mathcal{N}_t} \pi e^{-\pi(t-b)}c_{t,b}g(\omega_{t,b}, X_t)\Big[-\pi dt + d\Big(\frac{(1-X_t)^{-1}Y_t^{-1}}{(1-X_t)^{-1}Y_t^{-1}}\Big)$$
$$+ \frac{dg(\omega_{t,b}, X_t)}{g(\omega_{t,b}, X_t)} + \frac{dc_{t,b}}{c_{t,b}} + \sigma_{c_N,t}(\omega_{t,b})\frac{\sigma_{g,t}(\omega_{t,b})}{g(\omega_{t,b}, X_t)}dt + \Big(\frac{\sigma_{X,t}}{1-X_t} - \sigma_Y\Big)\Big(\sigma_{c_N,t}(\omega_{t,b}) + \frac{\sigma_{g,t}(\omega_{t,b})}{g(\omega_{t,b}, X_t)}\Big)dt\Big]db - n_t(g)dt,$$

where $\frac{\pi c_{t,t}g(\omega_{t,t}, X_t)\mathbf{1}_{\{\omega_{t,t}<\omega^*(X_t)\}}dt}{(1-X_t)Y_t}$ represents newborns that become non-participants, and $n_t(g)dt$ represents changes due to $\mathcal{N}_t$ (i.e., entry into participation), which are to be determined. Note that in deriving the left-hand-side, several terms involving $f_t(\omega^*(X_t))$ arise, and these terms vanish due to (E.6). The equation above must hold separately for the Brownian dZ_t terms and the drift dt terms.

Look first at the Brownian terms:

$$\int_0^{\omega^*(X_t)} \Big[g(\omega, X_t)\sigma_{f,t}(\omega) + f_t(\omega)\sigma_{X,t}\frac{\partial}{\partial x}g(\omega, X_t)\Big]d\omega$$
$$= \frac{1}{(1-X_t)Y_t}\int_{\mathcal{N}_t} \pi e^{-\pi(t-b)}c_{t,b}g(\omega_{t,b}, X_t)\Big(\frac{\sigma_{X,t}}{1-X_t} - \sigma_Y + \sigma_{c_N,t}(\omega_{t,b}) + \frac{\sigma_{g,t}(\omega_{t,b})}{g(\omega_{t,b}, X_t)}\Big)db.$$

We can use equation (E.3), which holds for any function $g \in \mathcal{G}$, and then substitute $\sigma_{g,t}$ and collect terms in order to rewrite the previous expression as

$$\int_0^{\omega^*(X_t)} g(\omega,X_t)\sigma_{f,t}(\omega)d\omega = \int_0^{\omega^*(X_t)} \Big(\frac{\sigma_{X,t}}{1-X_t} - \sigma_Y + \sigma_{c_N,t}(\omega) + \frac{\sigma_{\omega,t}(\omega)\frac{\partial}{\partial\omega}g(\omega,X_t)}{g(\omega,X_t)}\Big) g(\omega,X_t)f_t(\omega)d\omega.$$

For the term $\sigma_{\omega,t}(\omega)\frac{\partial}{\partial\omega}g(\omega,X_t)$, we can integrate-by-parts to obtain

$$\int_0^{\omega^*(X_t)} g(\omega,X_t)\sigma_{f,t}(\omega)d\omega = \int_0^{\omega^*(X_t)} \Big[\Big(\frac{\sigma_{X,t}}{1-X_t} - \sigma_Y + \sigma_{c_N,t}(\omega)\Big)f_t(\omega) - \frac{\partial}{\partial\omega}[\sigma_{\omega,t}(\omega)f_t(\omega)]\Big] g(\omega,X_t)d\omega.$$

Since g is arbitrary, this equation holds ω-by-ω without the integral, providing the desired expression for $\sigma_{f,t}(\omega)$.

Look next at the drift terms, as before using equation (E.3) to convert the db integrals to $d\omega$ integrals, and then substituting expressions for $\mu_{g,t}$ and $\sigma_{g,t}$, we have

$$\begin{aligned}
&\int_0^{\omega^*(X_t)} \Big[g(\omega,X_t)\mu_{f,t}(\omega) + \sigma_{f,t}(\omega)\sigma_{X,t}\frac{\partial}{\partial x}g(\omega,X_t)\Big]d\omega + \frac{1}{2}\Big(\sigma_{X,t}\frac{\partial}{\partial x}\omega^*(X_t)\Big)^2 \frac{\partial}{\partial\omega}[g(\omega,X_t)f_t(\omega)]\big|_{\omega=\omega^*(X_t)} \\
&= \int_0^{\omega^*(X_t)} \Big[-\pi + \frac{\mu_{X,t}}{1-X_t} - \mu_Y + \frac{\sigma_{X,t}\sigma_Y}{1-X_t} + \frac{\sigma_{X,t}^2}{(1-X_t)^2} + \sigma_Y^2 + \mu_{c_N,t}(\omega) + \Big(\frac{\sigma_{X,t}}{1-X_t} - \sigma_Y\Big)\sigma_{c_N,t}(\omega)\Big] g(\omega,X_t)f_t(\omega)d\omega \\
&\quad + \int_0^{\omega^*(X_t)} \Big[\mu_{\omega,t}(\omega)\frac{\partial}{\partial\omega}g(\omega,X_t) + \frac{1}{2}\sigma_{\omega,t}(\omega)^2\frac{\partial^2}{\partial\omega^2}g(\omega,X_t) + \sigma_{X,t}\sigma_{\omega,t}(\omega)\frac{\partial^2}{\partial x\partial\omega}g(\omega,X_t)\Big] f_t(\omega)d\omega \\
&\quad + \int_0^{\omega^*(X_t)} \Big(\frac{\sigma_{X,t}}{1-X_t} - \sigma_Y + \sigma_{c_N,t}(\omega)\Big)\Big[\sigma_{X,t}\frac{\partial}{\partial x}g(\omega,X_t) + \sigma_{\omega,t}(\omega)\frac{\partial}{\partial\omega}g(\omega,X_t)\Big] f_t(\omega)d\omega \\
&\quad + \frac{\pi c_{t,t}g(\omega_{t,t},X_t)\mathbf{1}_{\{\omega_{t,t}<\omega^*(X_t)\}}}{(1-X_t)Y_t} - n_t(g).
\end{aligned}$$

Substituting the result for $\sigma_{f,t}$ derived above, we have

$$\begin{aligned}
&\int_0^{\omega^*(X_t)} \Big[g(\omega,X_t)\mu_{f,t}(\omega) - \frac{\partial}{\partial\omega}[\sigma_{\omega,t}(\omega)f_t(\omega)]\sigma_{X,t}\frac{\partial}{\partial x}g(\omega,X_t)\Big]d\omega + \frac{1}{2}\Big(\sigma_{X,t}\frac{\partial}{\partial x}\omega^*(X_t)\Big)^2 \frac{\partial}{\partial\omega}[g(\omega,X_t)f_t(\omega)]\big|_{\omega=\omega^*(X_t)} \\
&= \int_0^{\omega^*(X_t)} \Big[-\pi + \frac{\mu_{X,t}}{1-X_t} - \mu_Y + \frac{\sigma_{X,t}\sigma_Y}{1-X_t} + \frac{\sigma_{X,t}^2}{(1-X_t)^2} + \sigma_Y^2 + \mu_{c_N,t}(\omega) + \Big(\frac{\sigma_{X,t}}{1-X_t} - \sigma_Y\Big)\sigma_{c_N,t}(\omega)\Big] g(\omega,X_t)f_t(\omega)d\omega \\
&\quad + \int_0^{\omega^*(X_t)} \Big[\mu_{\omega,t}(\omega)\frac{\partial}{\partial\omega}g(\omega,X_t) + \frac{1}{2}\sigma_{\omega,t}(\omega)^2\frac{\partial^2}{\partial\omega^2}g(\omega,X_t) + \sigma_{X,t}\sigma_{\omega,t}(\omega)\frac{\partial^2}{\partial x\partial\omega}g(\omega,X_t)\Big] f_t(\omega)d\omega \\
&\quad + \int_0^{\omega^*(X_t)} \Big(\frac{\sigma_{X,t}}{1-X_t} - \sigma_Y + \sigma_{c_N,t}(\omega)\Big)\sigma_{\omega,t}(\omega)f_t(\omega)\frac{\partial}{\partial\omega}g(\omega,X_t)d\omega \\
&\quad + \frac{\pi c_{t,t}g(\omega_{t,t},X_t)\mathbf{1}_{\{\omega_{t,t}<\omega^*(X_t)\}}}{(1-X_t)Y_t} - n_t(g).
\end{aligned}$$

We integrate-by-parts the third and fourth lines (twice for the term with $\frac{\partial^2}{\partial\omega^2}$), using the compact support assumption on g and the boundary conditions for f_t, to obtain

$$\begin{aligned}
&\int_0^{\omega^*(X_t)} g(\omega, X_t)\mu_{f,t}(\omega)d\omega + \frac{1}{2}\Big(\sigma_{X,t}\frac{\partial}{\partial x}\omega^*(X_t)\Big)^2 \frac{\partial}{\partial \omega}[g(\omega, X_t)f_t(\omega)]\big|_{\omega=\omega^*(X_t)} \\
&= \int_0^{\omega^*(X_t)} \Big[-\pi + \frac{\mu_{X,t}}{1-X_t} - \mu_Y + \frac{\sigma_{X,t}\sigma_Y}{1-X_t} + \frac{\sigma_{X,t}^2}{(1-X_t)^2} + \sigma_Y^2 + \mu_{c_N,t}(\omega) + \Big(\frac{\sigma_{X,t}}{1-X_t} - \sigma_Y\Big)\sigma_{c_N,t}(\omega)\Big] g(\omega, X_t)f_t(\omega)d\omega \\
&\quad - \int_0^{\omega^*(X_t)} g(\omega, X_t)\frac{\partial}{\partial \omega}[\mu_{\omega,t}(\omega)f_t(\omega)]d\omega + \frac{1}{2}\int_0^{\omega^*(X_t)} g(\omega, X_t)\frac{\partial^2}{\partial \omega^2}[\sigma_{\omega,t}(\omega)^2 f_t(\omega)]d\omega \\
&\quad - \int_0^{\omega^*(X_t)} g(\omega, X_t)\frac{\partial}{\partial \omega}\Big[\Big(\frac{\sigma_{X,t}}{1-X_t} - \sigma_Y + \sigma_{c_N,t}(\omega)\Big)\sigma_{\omega,t}(\omega)f_t(\omega)\Big]d\omega - \frac{1}{2}g(\omega^*(X_t), X_t)\frac{\partial}{\partial \omega}[\sigma_{\omega,t}(\omega)^2 f_t(\omega)]\big|_{\omega=\omega^*(X_t)} \\
&\quad + \frac{\pi c_{t,t} g(\omega_{t,t}, X_t)\mathbf{1}_{\{\omega_{t,t}<\omega^*(X_t)\}}}{(1-X_t)Y_t} - n_t(g).
\end{aligned}$$

From this result, we can see that for the density to be finite at $\omega^*(X_t)$, it must be the case that

$$n_t(g) = -\frac{1}{2}\Big[g(\omega, X_t)\frac{\partial}{\partial \omega}[\sigma_{\omega,t}(\omega)^2 f_t(\omega)] + \Big(\sigma_{X,t}\frac{\partial}{\partial x}\omega^*(X_t)\Big)^2 \frac{\partial}{\partial \omega}[g(\omega, X_t)f_t(\omega)]\Big]\Big|_{\omega=\omega^*(X_t)}. \tag{E.7}$$

Note also that, using (E.1) and (E.2), we may rewrite the newborn entry term as

$$\frac{\pi c_{t,t} g(\omega_{t,t}, X_t)\mathbf{1}_{\{\omega_{t,t}<\omega^*(X_t)\}}}{(1-X_t)Y_t} = \frac{\pi(\rho+\pi)\omega_{t,t} m(\omega_{t,t}, X_t)}{1-X_t}\int_0^{\omega^*(X_t)} \mathrm{Dirac}_{\omega_{t,t}}(\omega)g(\omega, X_t)d\omega. \tag{E.8}$$

Using (E.7)-(E.8), and then the fact that g is arbitrary, we obtain the desired result for $\mu_{f,t}$. $\square$