Dissecting the Aggregate Market Elasticity[*]

Victor Duarte[†] Goutham Gopalakrishna[‡] Mahyar Kargar[§]

Jiacui Li[¶] Dejanir Silva[‖]

March 27, 2026

Abstract

We study aggregate stock market elasticity in a general equilibrium model with heterogeneous investors, passive demand, and financial constraints. Without frictions, aggregate elasticity for the endowment claim is infinite, as interest rate and risk premium responses offset each other. With frictions, price impact for the endowment claim remains modest (about 0.7 in our calibration). In contrast, the equity (dividend) claim exhibits large price impact (above 8), consistent with empirical evidence, as frictions dampen interest rate responses while leverage amplifies risk premium responses to portfolio flows. We introduce a state-global perturbation method that yields closed-form, state-dependent elasticities. Solving the model with deep neural networks and calibrating to Flow of Funds data, we simultaneously match the equity premium and return volatility as well as the level and countercyclical dynamics of price impact.

Keywords: Aggregate market elasticity, risk misallocation, demand shocks, demand shifters, excess volatility, asset pricing.

JEL Classification: G12, G11, G21.

---

[*] We thank Andy Atkeson, Tim Johnson, Ralph Koijen, Tyler Muir, Stavros Panageas, Yinan Su, Fabrice Tourre, and Pierre-Olivier Weill; and conference and seminar participants at AFA 2026 Annual Meeting, Liquidity in Macroeconomics Workshop at St. Louis Fed, UIUC, UCLA macro-finance lunch, and University of Zurich for valuable comments and suggestions. All errors are our own.

[†] Gies College of Business, University of Illinois Urbana-Champaign, vduarte@illinois.edu

[‡] Rotman School of Management, University of Toronto, goutham.gopalakrishna@rotman.utoronto.ca

[§] Gies College of Business, University of Illinois Urbana-Champaign, kargar@illinois.edu

[¶] David Eccles School of Business, University of Utah; jiacui.li@eccles.utah.edu

[‖] Daniels School of Business, Purdue University, dejanir@purdue.edu

1 Introduction

Why is the stock market so volatile? This question has long been central to asset pricing research. Classic tests of excess volatility attribute the puzzle to unobserved "dark matter" factors.[1] More recent evidence points instead to a demand-based channel: asset prices respond strongly to portfolio flows. Yet these flows are small relative to market capitalization, implying that markets must be highly *inelastic*, with even modest quantity shifts generating large price changes (Gabaix and Koijen, 2023). Moreover, the ratio of return volatility to flows—the "volatility multiplier"—is itself time-varying and countercyclical, rising sharply during periods of financial stress. This pattern points to a state-dependent relationship between prices and portfolio flows.

These facts raise a central question. How can small portfolio flows generate large price movements in equilibrium? More specifically, what determines the *aggregate* elasticity of the stock market, why does price impact vary over time, and which frictions are essential for matching its magnitude in the data?

To address these questions, we develop a general equilibrium model with heterogeneous investors, passive demand, and financial constraints. We study the *aggregate (macro) elasticity*, defined as the change in the total stock market value in response to a $1 flow between (short-term) bonds and stocks. Unlike micro elasticities, which capture substitution across individual stocks, aggregate elasticity reflects market-wide reallocations between risky and risk-free assets. As a result, understanding aggregate elasticity requires understanding the simultaneous response of the risk-free rate and the risk premium to changes in portfolio flows.

This joint response is particularly important when studying the elasticity of the *endowment claim*, the claim on aggregate output. Our first main contribution is to show that, in the absence of frictions, the aggregate market elasticity of this claim is *infinite*, regardless of how price-inelastic individual investors are.[2] In particular, an inflow of funds from bonds to stocks has no price impact

---

[1] See LeRoy and Porter (1981), Shiller (1981, 1992), and Cochrane (1992); and Chen, Dou, and Kogan (2024) for a recent discussion.

[2] A portfolio-flow shock is an asymmetric shock to risky-asset demand, affecting only a subset of investors and triggering portfolio reallocation in equilibrium.

in equilibrium. The key mechanism is a general equilibrium adjustment in savings behavior. When portfolio flows shift demand toward risky assets, the risk premium falls while the interest rate rises. When risk is efficiently allocated, these two forces offset exactly, leaving asset prices unchanged. As a result, aggregate elasticity is infinite even when individual demand is highly inelastic in a frictionless economy.

This result highlights a central distinction between micro and aggregate elasticities. If cross-price effects are absent, so that demand for risky assets does not respond to the interest rate, aggregate elasticity reduces to a weighted average of individual elasticities, as in Gabaix and Koijen (2023). By contrast, once interest rates adjust endogenously, aggregate elasticity is governed by general equilibrium forces rather than by the slope of individual demand. In particular, the response of aggregate savings determines whether movements in the risk-free rate can offset changes in risk premia. When portfolio flows do not alter saving incentives, movements in the risk-free rate fully offset movements in the risk premium, leaving asset prices unchanged.

Aggregate elasticity becomes finite only when risk is misallocated. In this case, portfolio reallocation changes saving incentives, preventing interest rates from fully offsetting movements in risk premia. This generates a nonzero price impact even in general equilibrium. Even with frictions—such as passive demand, preference heterogeneity, and leverage constraints—this near-cancellation remains quantitatively important, so that the endowment claim exhibits only a modest price impact of 0.7 in our calibration. We establish these results first in a transparent two-period model and then extend them to a fully dynamic setting.

Studying aggregate elasticity in models with frictions is challenging because closed-form solutions are rarely available. We introduce a tractable analytical approach based on state-global perturbations that delivers closed-form, state-dependent characterizations of aggregate elasticity. The method shows that elasticity is time-varying and depends on both the wealth distribution across active and passive investors and the allocation of risk within the active sector. Passive investors amplify price impact, preference heterogeneity raises elasticity through risk misallocation, and binding leverage constraints further amplify these effects. In this environment, excess volatility arises from the interaction between portfolio-flow shocks and a finite (state-dependent) elasticity. Inelasticity alone does not generate excess volatility, and neither do flows in infinitely elastic markets; both are required.

Our second main contribution is to show that, in contrast to the endowment claim, the *equity* (dividend) claim exhibits a substantially larger price impact, with a multiplier above 8 in our calibration, consistent with the empirical estimates of Gabaix and Koijen (2023) and Koijen, Richmond, and Yogo (2024). This difference arises because the general equilibrium offset between the risk-free rate and the risk premium that renders the endowment claim highly elastic does not carry over to the dividend claim. Dividends represent a levered share of aggregate output, so the risk premium on the dividend claim responds more strongly to portfolio flows than the risk premium on the endowment claim. As a result, changes in discount rates are not offset by the same general equilibrium forces, leading to a large price response. This result highlights that modeling the dividend claim, rather than a claim on aggregate consumption as is standard in the macro-finance literature, is essential for generating realistic levels of price impact.

A key methodological contribution is a tractable way to introduce the dividend claim in a dynamic heterogeneous-agent economy with frictions. We provide conditions under which a version of the mutual fund theorem holds: investors optimally hold risky claims at market-capitalization weights. This implies that the stochastic discount factor from the aggregate (endowment-claim) economy can be used to price the dividend claim in a second step. This separation reduces the effective dimensionality of the problem and greatly simplifies the analysis.

For the quantitative analysis, we solve the full model numerically using the deep neural network method of Duarte, Duarte, and Silva (2024) and calibrate it to Flow of Funds data. The calibrated model matches both the magnitude and countercyclical dynamics of price impact in the data, while also reproducing key asset pricing moments such as the equity premium and return volatility. These results highlight the joint role of frictions, which dampen interest rate responses, and leverage, which amplifies risk premia responses to portfolio flows.

Related literature

Our work relates to the literature on macro demand elasticity. The papers closest to ours are Johnson (2006), who finds finite elasticity even in a frictionless Lucas economy by perturbing risky asset supply, and Gabaix and Koijen (2023), who introduce a behavioral element that fixes interest rates and generates large price impacts. We build on this literature by developing a GE model with heterogeneous investors and financial frictions, showing that both GE effects and constraints are central for obtaining inelastic markets.[3] This contrasts with the larger body of work on micro elasticity, which examines relative price changes across individual stocks (e.g., Shleifer, 1986; Harris and Gurel, 1986; Chang, Hong, and Liskovich, 2015; Pavlova and Sikorskaya, 2023; Schmickler, 2020) and finds much larger elasticities, consistent with stocks being closer substitutes for each other than for bonds.

We also connect to demand-system asset pricing (Koijen and Yogo, 2019) and its extensions incorporating cross-asset elasticities (Fuchs, Fukuda, and Neuhann, 2023; Haddad, He, Huebner, Kondor, and Loualiche, 2025). Our framework shows that even small cross-elasticities between equities and bonds generate large effects on macro elasticities, and extends the analysis to dynamic settings where the persistence of shocks and the evolution of risk premia shape both short-run and long-run adjustments (Binsbergen, David, and Opp, 2025; He, Kondor, and Li, 2025; Davis, Kargar, Li, and Silva, 2026). Our model of passive investors also connects to the literature on intermittent rebalancing (Chien, Cole, and Lustig, 2012), where sluggish portfolio adjustments are themselves an important source of aggregate inelasticity. Finally, our paper relates to preferred-habitat models (Vayanos and Vila, 2021; Kekre, Lenel, and Mainardi, 2025; Ray, Droste, and Gorodnichenko, 2024), which emphasize intermediary frictions and market segmentation. By contrast, our framework highlights wealth effects as a distinct channel through which shifts in aggregate wealth alter portfolio demand and amplify price movements.

---

[3] For more on macro elasticity, see Johnson (2008, 2009), Deuskar and Johnson (2011), Li, Pearson, and Zhang (2020), Hartzmark and Solomon (2025), among others.

2 A Simple Model of the Aggregate Market Elasticity

In this section, we analyze the determination of the aggregate market elasticity in a simple general equilibrium, demand-based asset pricing model. To keep the discussion transparent, we specify investor demands directly, rather than deriving them from utility functions. A micro-founded version of these demands, embedded in a dynamic heterogeneous-agents model, will be presented in Section 3.

Environment. We consider a two-period economy with two assets: a risky asset and a riskless bond. There are two types of investors: passive (p) and active (a). A fraction ω_j of the population is of type $j \in \{a, p\}$. The risky asset pays a random dividend Y' in the second period. Its price is denoted by P, while the price of the riskless asset is R_f^{-1}.

The budget constraints of investor j are

$$PQ_j + R_f^{-1} B_j + C_j = W_j, \qquad C'_j = Y' Q_j + B_j,$$

where Q_j is the investor's holdings of the risky asset, B_j denotes holdings of the riskless bond, C_j is first-period consumption, and C'_j is second-period consumption. Initial wealth is given by $W_j = (P + Y) Q_{j,-1} > 0$.

We define investor j's portfolio share in the risky asset as

$$\alpha_j \equiv \frac{PQ_j}{W_j - C_j}.$$

Passive investors maintain a fixed risky asset share $\alpha_p = \overline{\alpha}_p \geq 0$. Active investors' portfolio share, in contrast, depends on the risk premium

$$\alpha_a = \overline{\alpha}_a(\pi),$$

for some increasing function $\overline{\alpha}_a(\cdot)$, where $\pi \equiv \log \frac{1}{R_f} \mathbb{E}\left[\frac{Y'}{P}\right]$ denotes the risk premium.

Consumption is specified as

$$C_j = \bar{c}_j(r, \pi) W_j,$$

where $r \equiv \log R_f$ is the log risk-free rate. The consumption-wealth ratio $\bar{c}_j(\cdot)$ depends on both r and π. If $\bar{c}_j$ decreases with r, investors save more when the interest rate is high (a substitution effect); if $\bar{c}_j$ increases with r, investors save less (an income effect). Similar logic applies to the risk premium. We assume that $\bar{c}_j(\cdot)$ does not depend on the passive portfolio share $\bar{\alpha}_p$.

The market clearing conditions are

$$\sum_{j \in \{p,a\}} \omega_j C_j = Y, \qquad \sum_{j \in \{p,a\}} \omega_j Q_j = 1, \qquad \sum_{j \in \{p,a\}} \omega_j B_j = 0,$$

with initial risky asset endowment $\sum_{j \in \{p,a\}} \omega_j Q_{j,-1} = 1$.

Market for risky assets. Let $p = \log \frac{P}{Y}$ denote the log price-dividend ratio, and let $\mu \equiv \log \frac{\mathbb{E}[Y']}{Y}$ denote the log dividend growth. The risk premium is

$$\pi = \mu - p - r. \tag{1}$$

Investor j's demand for the risky asset is

$$Q_j = \alpha_j \left[1 - \bar{c}_j(r, \pi)\right] \frac{W_j}{P}.$$

Using Equation (1) to eliminate π, we can express demand as a function of price of the risky asset, the risk-free rate, and, for passive investors, the exogenous portfolio share: $Q_j = \bar{Q}_j(p, r, \bar{\alpha}_p)$.

The market clearing condition for the risky asset is therefore

$$\underbrace{\omega_a \bar{Q}_a(p, r, \bar{\alpha}_p)}_{\text{active demand}} = \underbrace{1 - \omega_p \bar{Q}_p(p, r, \bar{\alpha}_p)}_{\text{net supply}}. \tag{2}$$

Equilibrium in the risky asset market requires that active investors absorb the net supply, defined as total supply minus the amount held by passive investors. Since active demand does not depend on $\overline{\alpha}_p$, changes in the passive portfolio share affect only the net supply curve.

Market for goods. The response of the consumption-wealth ratio to the interest rate and the risk premium is central to our analysis. In the dynamic model of Section 3, the average consumption-wealth ratio depends only on the sum $(r + \pi)$, so its sensitivity to interest rates equals its sensitivity to risk premia. Assumption 1 imposes the same structure in this two-period setting, ensuring consistency with standard micro-founded models.

Assumption 1. *The average consumption-wealth ratio is a function of the expected return on the risky asset,* $(r + \pi)$*,*

$$\sum_{j \in \{p,a\}} x_j \overline{c}_j(r, \pi) = \overline{c}(r + \pi),$$

for some function $\overline{c}(\cdot)$*, where* $x_j \equiv \frac{\omega_j W_j}{\omega_a W_a + \omega_p W_p}$ *denotes investor* j*'s wealth share.*

Intuitively, since bonds are in zero net supply, the return on investors' overall portfolios equals the return on the risky asset. From goods market clearing and Equation (1), we obtain

$$\overline{c}(\mu - p) = \frac{1}{1 + e^p}, \tag{3}$$

using the fact that $r + \pi = \mu - p$ and $\sum_j x_j \overline{c}_j = \frac{Y}{Y+P}$.

Assumption 1 implies that the system of equations determining p and r has a useful *recursive property*: condition (3) depends only on p, while condition (2) depends on both p and r.

2.1 General equilibrium implications of portfolio flows

We now examine how the price-dividend ratio p and the riskless return r respond to a portfolio flow from the riskless bond into the risky asset. Let $\overline{\alpha}_p^\star$ denote the passive portfolio share in the initial equilibrium, with $(p^\star, r^\star)$ the corresponding equilibrium prices.

For a small perturbation, $\overline{\alpha}_p = \overline{\alpha}_p^\star e^{\hat{\alpha}_p}$, we linearize demand around the initial equilibrium. Defining $\hat{p} \equiv p - p^\star$, $\hat{r} \equiv r - r^\star$, and $q_j \equiv \log(Q_j/Q_j^\star)$, we obtain

$$q_j = -\zeta_{j,p}\,\hat{p} - \zeta_{j,r}\,\hat{r} + f_j, \tag{4}$$

where the own-price and cross-price elasticities are given by $\zeta_{j,p} \equiv -\frac{\partial \log \overline{Q}_j}{\partial p}$ and $\zeta_{j,r} \equiv -\frac{\partial \log \overline{Q}_j}{\partial r}$. The *flow shock* is given by

$$f_j \equiv \frac{\partial \log \overline{Q}_j}{\partial \log \overline{\alpha}_p} \times \hat{\alpha}_p,$$

capturing an exogenous reallocation from bonds to stocks. Since active demand does not depend on $\overline{\alpha}_p$, we have $f_a = 0$. Formally, a flow shock is an asymmetric demand shock affecting some investors but not others, inducing equilibrium portfolio rebalancing.

An important implication of (4) is that demand for the risky asset depends on both the price-dividend ratio p and the bond price through the risk-free rate r. Since active investors respond to the risk premium $\pi = \mu - p - r$, their demand depends on both prices. Therefore, the prices of the risky and risk-free assets must be jointly determined in equilibrium.

Equilibrium. Let $s_j \equiv \omega_j Q_j^\star$ denote the equilibrium quantity share of the risky asset held by type j. Linearizing the market-clearing condition $\sum_j \omega_j Q_j = 1$ gives

$$s_a\left(-\zeta_{a,p}\,\hat{p} - \zeta_{a,r}\,\hat{r}\right) = s_p\left(\zeta_{p,p}\,\hat{p} + \zeta_{p,r}\,\hat{r} - f_p\right), \tag{5}$$

as $s_a q_a + s_p q_p = 0$.

The left panel of Figure 1 shows the equilibrium in the risky-asset market. The downward-sloping curve represents active demand, while the upward-sloping curve is net supply, i.e., total supply minus passive holdings.[4]

---

[4] Passive demand is $Q_p = \overline{\alpha}_p\,[1 - c_p(r,\pi)]\,\frac{W_p}{P}$, given $W_p = (P+Y)Q_{p,-1}$ and $\pi = \mu - p - r$, so its linearized contribution $-s_p q_p$ makes net supply upward sloping in p provided $\frac{\partial c_p}{\partial p}$ is not too large.

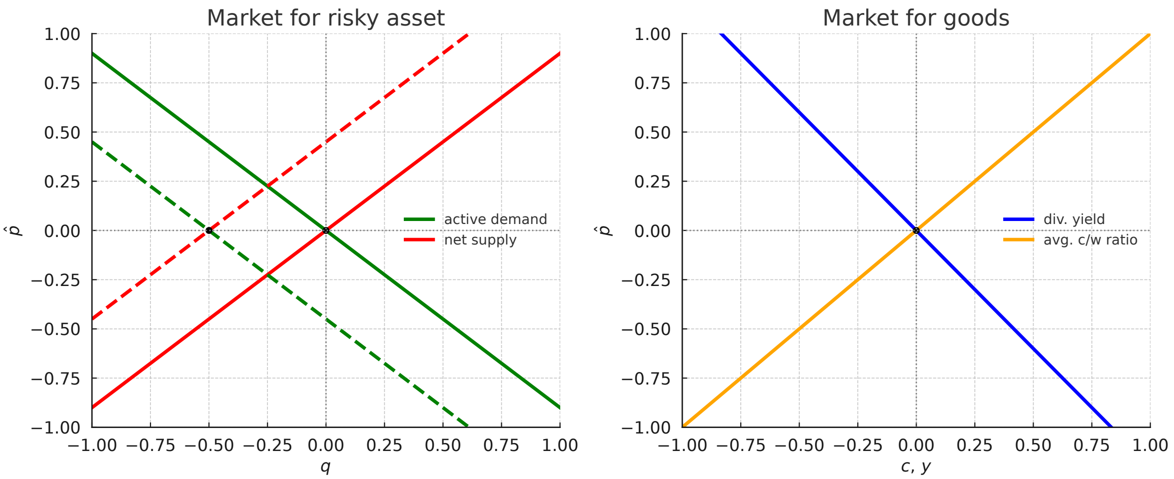


Figure 1. Equilibrium in the risky asset (left) and goods (right) markets.

Linearizing the goods-market condition $\overline{c}(\mu - p) = 1/(1+e^p)$ around $p^\star$ yields

$$\chi_p \, \hat{p} \;=\; -\frac{e^{p^\star}}{(1+e^{p^\star})^2}\,\hat{p}, \qquad \text{where} \qquad \chi_p \equiv -\overline{c}'(\mu - p^\star). \tag{6}$$

The right panel of Figure 1 shows the goods market. We focus on the case $\chi_p > 0$, so the average consumption-wealth ratio is increasing in p (investors save less when expected returns are low). The right-hand side in (6) is the (local) slope of the dividend-yield curve, which is downward sloping in p. By the recursive property, only p enters the goods market clearing condition.

The inelastic markets hypothesis. Consider a *partial equilibrium* version of the model in which we hold the interest rate fixed at $r = r^\star$ and ignore goods market clearing. Using the linearized risky asset market condition (5) with $\hat{r} = 0$, the price of the risky asset satisfies

$$\hat{p} \;=\; \underbrace{\frac{1}{\zeta_p}}_{\substack{\text{inverse market}\\ \text{elasticity}}} \times \underbrace{f}_{\text{flow shock}}, \tag{7}$$

where

$$\zeta_p \;\equiv\; s_a\,\zeta_{a,p} + s_p\,\zeta_{p,p}, \qquad f \;\equiv\; s_p\,f_p, \qquad s_j \equiv \omega_j Q_j^\star.$$

Equation (7) shows that the price-dividend ratio responds to flow shocks, with price impact given by the inverse of the (aggregate) market elasticity—an average of investors' own-price elasticities. This is the core implication of the *inelastic markets hypothesis* of Gabaix and Koijen (2023), who estimate large price responses to stock-market flows, suggesting relatively inelastic demand.

In the risky asset panel of Figure 1, an increase in the portfolio share of passive investors shifts the net supply curve to the left. The new equilibrium lies further up the active demand curve, implying lower expected returns and a higher price-dividend ratio. When demand is inelastic, even small inflows to stocks generate a sizable increase in $\hat{p}$.

The crucial role of cross-elasticity. The partial equilibrium analysis abstracts from cross-elasticity. In general equilibrium, however, both r and p adjust, and the aggregate market elasticity is fundamentally different, as the next proposition shows.

Proposition 1 (Infinite market elasticity). *Define the aggregate cross-elasticity* $\zeta_r \equiv s_a \zeta_{a,r} + s_p \zeta_{p,r}$. *If* $\zeta_r \neq 0$, *then a portfolio-flow shock has* no *price impact:* $\hat{p} = 0$. *The shock is absorbed by offsetting movements in the interest rate and the risk premium:*

$$\hat{r} = \frac{1}{\zeta_r} f, \qquad \hat{\pi} = -\frac{1}{\zeta_r} f,$$

where $\hat{\pi} \equiv \pi - \pi^{\star}$.

Proof. From Equation (6), we obtain $\hat{p} = 0$. If $\zeta_r \neq 0$, then $\hat{r} = \frac{1}{\zeta_r} f$ from Equation (5). Since $\hat{\pi} = -\hat{p} - \hat{r}$ with $\hat{p} = 0$, it follows that $\hat{\pi} = -f/\zeta_r$. $\square$

With any nonzero cross-elasticity $\zeta_r \neq 0$, the aggregate market becomes infinitely elastic, regardless of how inelastic individual demands may be. Even if $\zeta^q_{j,p} \approx 0$ for all j, a small but positive ζ_r is sufficient to make aggregate elasticity unbounded.

The disconnect between individual and aggregate elasticities arises from a general equilibrium effect. In the goods market (right panel of Figure 1), the price-dividend ratio is determined by consumption behavior. This mechanism is well known in the macro-finance literature, most clearly in models with unit EIS, where p is tied directly to the discount rate. Our framework generalizes this logic. To restore equilibrium in the risky asset market (left panel), the interest rate adjusts so that active demand equals net supply at the *original price*. Graphically, this appears as the dashed active-demand curve shifting to intersect the new net-supply curve at the initial price.

Consider a positive flow shock to the risky asset (e.g., funds shift toward stocks, so $f > 0$). Lower bond demand increases r, while increased demand for the risky asset reduces the risk premium. In equilibrium these effects offset exactly, leaving the price-dividend ratio unchanged. If the rise in r were smaller than the fall in the risk premium, there would be excess demand for goods (equivalently, excess bond supply). Simultaneous clearing of risky and risk-free asset markets therefore requires infinite aggregate elasticity in this model.

The need for a micro-founded demand system. The analysis underscores the central role of cross-price elasticity. The aggregate market can be infinitely elastic—even if individual investors are inelastic—so long as the average consumption-wealth ratio is unaffected by the flow shock f. When the flow shock shifts both the net-supply curve and the average consumption-wealth ratio, prices adjust and aggregate elasticity becomes finite.[5]

Understanding aggregate elasticity therefore requires a framework that links portfolio reallocation shocks to both risky asset demand and bond demand (i.e., savings behavior). To this end, we develop a dynamic heterogeneous-agent asset pricing model and introduce a new methodology for deriving investor demand within this setting.

3 A Dynamic Asset Pricing Model with Passive Investors

This section develops a dynamic heterogeneous-agent asset pricing model that provides a micro-founded demand system. A representative firm finances its operations by issuing two types of claims: a *dividend claim* (equity) and a *risky corporate bond*, whose payoffs sum to the aggregate

---

[5] An analogous argument applies to uncertainty. With unit EIS, the consumption–wealth ratio is fixed, so higher uncertainty leaves prices unchanged. For uncertainty to reduce prices, EIS must exceed one so that higher uncertainty lowers the consumption–wealth ratio (see Bansal and Yaron, 2004).

endowment. To capture the well-documented fact that dividends are more volatile than consumption (see, e.g., Abel 1999; Campbell 1999), we follow Menzly, Santos, and Veronesi (2004) and model the dividend as a mean-reverting stochastic share of output. In addition, investors can trade a derivative in zero net supply. The derivative plays a key role: by spanning the non-traded Brownian factor, it allows investors to achieve their desired risk exposures without taking asymmetric positions in the individual claims, and is therefore essential for the mutual fund theorem described below.

A central result of this section is a *mutual fund theorem*: despite access to separate equity and debt claims, all investors optimally hold these claims in proportion to their market-capitalization weights. Because the derivative completes markets with respect to the priced sources of risk, investors can adjust their exposure to the non-dividend-share factor through the derivative alone, and optimally hold the risky claims at market-cap weights. Consequently, the aggregate equilibrium, including the risk-free rate, risk premia, wealth dynamics, and the stochastic discount factor, is identical to that of an economy with a single claim on aggregate output. This separation implies that the perturbation analysis of Section 4 applies directly to the aggregate equilibrium derived here. Individual claim prices are then obtained by solving an auxiliary valuation equation, conditional on the aggregate equilibrium.

3.1 Environment

Financial markets. The aggregate endowment Y_t follows a geometric Brownian motion:

$$\frac{dY_t}{Y_t} = \mu\, dt + \sigma\, dZ_t, \tag{8}$$

where $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}^{1\times d}$ are constants. The process $Z = \{Z_t\}_{t\geq 0}$ is a standard d-dimensional Brownian motion on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ satisfying the usual conditions. In the baseline model, $d = 3$, capturing fundamental shocks to cash flows, portfolio demand shocks to passive investors, and shocks to the dividend share.

The firm issues two claims to finance its operations. The first is a *dividend claim* (equity), with payoff flow $D_t = s_t Y_t$. Following Menzly et al. (2004), the dividend share $s_t \in (0,1)$ follows the mean-reverting process:

$$ds_t = \kappa_s(\bar{s} - s_t)\,dt + s_t(1 - s_t)\,\sigma_s\,dZ_t,$$

given $s_0 \in (0,1)$, with $\bar{s} \in (0,1)$, $\kappa_s > 0$, and $\sigma_s \in \mathbb{R}^{1\times d}$. This specification ensures $s_t \in [0,1]$ at all times and yields a stationary share process, so consumption and dividends are cointegrated. By Itô's lemma, dividends follow:

$$\frac{dD_t}{D_t} = \mu_D(s_t)\,dt + \sigma_D(s_t)\,dZ_t,$$

where

$$\sigma_D(s) = \sigma + (1-s)\,\sigma_s, \qquad \mu_D(s) = \mu + \kappa_s\left(\frac{\bar{s}}{s} - 1\right) + (1-s)\,\sigma\sigma_s'.$$

Since $\sigma_D(s) \neq \sigma$ when $\sigma_s \neq 0$, dividends are more volatile than aggregate consumption. Moreover, dividend volatility increases after negative shocks to s_t, capturing a form of leverage effect.

The second claim is a *corporate bond* (risky debt), with payoff flow $B_t = (1 - s_t)Y_t$. Since $D_t + B_t = Y_t$, the bond payoff is smoother than aggregate consumption, reflecting the safer nature of debt.

Let $P_{D,t}$ and $P_{B,t}$ denote the ex-dividend prices of the dividend and bond claims, with returns:

$$\begin{aligned} dR_{D,t} &= \frac{dP_{D,t} + D_t\,dt}{P_{D,t}} = \mu_{R_D,t}\,dt + \sigma_{R_D,t}\,dZ_t, \\ dR_{B,t} &= \frac{dP_{B,t} + B_t\,dt}{P_{B,t}} = \mu_{R_B,t}\,dt + \sigma_{R_B,t}\,dZ_t. \end{aligned}$$

Since $D_t + B_t = Y_t$, the combined claim replicates the aggregate endowment, with price $P_{Y,t} = P_{D,t} + P_{B,t}$. The return on the aggregate claim follows:

$$dR_{Y,t} = \mu_{R_Y,t}\,dt + \sigma_{R_Y,t}\,dZ_t, \tag{9}$$

where $\mu_{R_Y,t} \equiv \frac{Y_t}{P_{Y,t}} + \mu_{P_Y,t}$ and $\sigma_{R_Y,t} = \sigma_{P_Y,t}$.

Investors also have access to a risk-free asset paying rate r_t and a derivative in zero net supply. The derivative provides risk exposure $\sigma_{R_F,t} \in \mathbb{R}^{1\times d}$, orthogonal to the aggregate claim, $\sigma_{R_F,t}\,\sigma'_{R_Y,t} = 0$, with normalization $\|\sigma_{R_F,t}\| = \|\sigma_{R_Y,t}\|$. This ensures dynamic market completeness with respect to the first two Brownian factors.

Absence of arbitrage implies the existence of a stochastic discount factor η_t with

$$\frac{d\eta_t}{\eta_t} = -r_t\,dt - \lambda_t\,dZ_t,$$

where $\lambda_t \in \mathbb{R}^{1\times d}$ is the market price of risk vector. Risk premia satisfy $\pi_{k,t} = \sigma_{R_k,t}\,\lambda'_t$ for each claim $k \in \{D, B, F\}$.

Let Σ_t denote the $3 \times d$ matrix of risk exposures for the dividend claim, bond claim, and derivative:

$$\Sigma_t = \begin{pmatrix} \sigma_{R_D,t} \\ \sigma_{R_B,t} \\ \sigma_{R_F,t} \end{pmatrix},$$

and $\pi_t \equiv [\pi_{D,t}, \pi_{B,t}, \pi_{F,t}]'$, so that $\pi_t = \Sigma_t\,\lambda'_t$.

In equilibrium, prices depend on an aggregate state vector $X_t \in \mathbb{R}^N$, so we write $r_t = r(X_t)$, $\lambda_t = \lambda(X_t)$, and return coefficients as functions of X_t. The state vector evolves as

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

with drift $\mu_{X,t} \in \mathbb{R}^N$ and exposure matrix $\sigma_{X,t} \in \mathbb{R}^{N\times d}$, both determined in equilibrium.

Demography. The economy consists of three investor types ($J = 2$), indexed by $j = 0, 1, 2$, with population mass ω_j: one passive type ($j = 0$) and two active types ($j = 1, 2$). Investors die with Poisson intensity κ, and each instant a mass $\kappa\,\omega_j$ of type-j agents are born, so the total population remains constant and normalized to one. Newborns inherit parental wealth. Mortality risk ensures a nondegenerate stationary wealth distribution.

Preferences are recursive as in Duffie and Epstein (1992), with type-j risk aversion γ_j and a common elasticity of intertemporal substitution (EIS) ψ. Each investor chooses consumption $C_{j,t}$ and portfolio allocations across the dividend claim, the bond claim, and the derivative, subject to constraints specified below.

Passive investors. Agents of type $j = 0$ are *passive investors*, who hold the dividend and bond claims in proportion to their market-capitalization weights:

$$\alpha_{0,D,t} = \overline{\alpha}_{p,t}\,\omega_{D,t}, \qquad \alpha_{0,B,t} = \overline{\alpha}_{p,t}\,\omega_{B,t}, \qquad \theta_{0,t} = 0,$$

where $\omega_{D,t} \equiv \frac{P_{D,t}}{P_{Y,t}}$ and $\omega_{B,t} \equiv \frac{P_{B,t}}{P_{Y,t}}$ are the market weights of the dividend and bond claims. The total risky share $\overline{\alpha}_{p,t}$ evolves according to an Ornstein-Uhlenbeck process:

$$d\overline{\alpha}_{p,t} \;=\; \theta_p\big(\overline{\alpha} - \overline{\alpha}_{p,t}\big)dt \;+\; \sigma_{\alpha_p}\,dZ_t, \tag{10}$$

with long-run mean $\overline{\alpha} > 0$, mean-reversion $\theta_p > 0$, and loading $\sigma_{\alpha_p} \in \mathbb{R}^{1\times d}$ on the d-dimensional Brownian motion Z_t. Innovations to $\overline{\alpha}_{p,t}$ represent *portfolio flow shocks*, i.e., fluctuations in the fraction of wealth passive investors allocate to risky claims. As in Section 2, a flow shock is an asymmetric demand shock that affects only the passive investors and triggers portfolio reallocation.

Portfolio flows may comove with cash-flow shocks or arise independently. For example, with $d = 3$, take $\sigma = (\sigma_1, 0, 0)$ for endowment risk and $\sigma_{\alpha_p} = (\sigma_{\alpha_p,1}, \sigma_{\alpha_p,2}, 0)$ for the passive-share process: the first factor generates correlation with fundamentals, while the second delivers portfolio-flow shocks orthogonal to cash-flow risk. The third factor captures dividend-share shocks.

The exposure of the passive share to aggregate shocks can reflect *imperfect rebalancing* or *performance-induced flows*. Without rebalancing, a positive return on the risky asset raises the passive share when $\overline{\alpha}_{p,t} < 1$; similarly, investors may increase risky exposure after good performance. This specification aligns with evidence on passive behavior: households exhibit substantial portfolio inertia (Ameriks and Zeldes, 2004; Brunnermeier and Nagel, 2008); flows chase performance (Chevalier and Ellison, 1997; Dannhauser and Pontiff, 2024); and regulatory changes can sharply increase equity allocations via TDF adoption (Parker, Schoar, Cole, and Simester, 2022), illustrating portfolio flow shocks not tied to cash flow fundamentals.

Active investors. Investors of type $j = 1, \dots, J$ are *active investors*, who rebalance portfolios subject to a leverage constraint on total risk exposure:

$$\|\Sigma_{j,t}\| \leq \overline{\sigma},$$

where $\Sigma_{j,t} = \alpha_{j,D,t}\, \sigma_{R_D,t} + \alpha_{j,B,t}\, \sigma_{R_B,t} + \theta_{j,t}\, \sigma_{R_F,t}$ is the investor's total risk exposure and $\|\cdot\|$ denotes the Euclidean norm. This constraint resembles a Value-at-Risk (VaR) limit, often applied to banks and leveraged institutions (see, e.g., Adrian and Shin 2014 and Brunnermeier and Pedersen 2009).

Investors' problem. An investor of type $j = 0, \dots, J$ solves

$$V_{j,t} = \max_{[C_j, \alpha_{j,D}, \alpha_{j,B}, \theta_j]} \mathbb{E}_t \left[\int_t^\infty f_j(C_{j,s}, V_{j,s})\, ds \right], \qquad s.t. \tag{11}$$

$$dW_{j,t} = \left[(r_t + \pi_{D,t}\, \alpha_{j,D,t} + \pi_{B,t}\, \alpha_{j,B,t} + \pi_{F,t}\, \theta_{j,t}) W_{j,t} - C_{j,t} \right] dt + \Sigma_{j,t}\, W_{j,t}\, dZ_t,$$

with $W_{j,t} \geq 0$ and $(\alpha_{j,D,t}, \alpha_{j,B,t}, \theta_{j,t}) \in \Omega_{j,t}$. For passive investors ($j = 0$), $\Omega_{0,t} = \{(\alpha_{0,D}, \alpha_{0,B}, \theta_0) : \alpha_{0,D} = \overline{\alpha}_{p,t}\, \omega_{D,t},\ \alpha_{0,B} = \overline{\alpha}_{p,t}\, \omega_{B,t},\ \theta_0 = 0\}$. For active investors ($j = 1, \dots, J$), $\Omega_{j,t} = \{(\alpha_{j,D}, \alpha_{j,B}, \theta_j) : \|\Sigma_{j,t}\| \leq \overline{\sigma}\}$.

Preferences are recursive with aggregator

$$f_j(C, V) = \rho \frac{(1-\gamma_j)V}{1-\psi^{-1}} \left[\left(\frac{C}{((1-\gamma_j)V)^{\frac{1}{1-\gamma_j}}} \right)^{1-\psi^{-1}} - 1 \right],$$

where the discount factor $\rho \equiv \hat{\rho} + \kappa$ reflects both time impatience $\hat{\rho}$ and the death probability κ.

Market clearing and equilibrium. We define equilibrium as follows:

Definition 1. *A competitive equilibrium consists of stochastic processes adapted to the filtration generated by* Z_t*: the aggregate endowment* Y*, the prices of the dividend and bond claims* P_D *and* P_B*, the risk-free rate* r*; and, for each investor* $j \in \{0, \ldots, J\}$*, wealth* W_j*, consumption* C_j*, and portfolio allocations* $(\alpha_{j,D}, \alpha_{j,B}, \theta_j)$*, such that*

(i) The aggregate endowment evolves according to (8)*, with* $Y_0 > 0$*.*

(ii) Given $\{P_{D,t}, P_{B,t}, r_t\}$*, the choices* $(C_j, \alpha_{j,D}, \alpha_{j,B}, \theta_j)$ *solve investor* j*'s problem in* (11)*.*

(iii) Markets clear for goods, the dividend claim, the bond claim, riskless bonds, and the derivative:

$$\sum_{j=0}^{J} \omega_j C_{j,t} = Y_t, \quad \sum_{j=0}^{J} \omega_j Q_{j,D,t} = 1, \quad \sum_{j=0}^{J} \omega_j Q_{j,B,t} = 1, \quad \sum_{j=0}^{J} \omega_j B_{j,t} = 0, \quad \sum_{j=0}^{J} \omega_j W_{j,t} \theta_{j,t} = 0,$$

where $Q_{j,D,t} = \alpha_{j,D,t} W_{j,t} / P_{D,t}$ *and* $Q_{j,B,t} = \alpha_{j,B,t} W_{j,t} / P_{B,t}$ *denote units held of the dividend and bond claims, and* $B_{j,t} = (1 - \alpha_{j,D,t} - \alpha_{j,B,t}) W_{j,t}$ *denotes riskless holdings.*

3.2 Equilibrium characterization

We first present the investors' optimality conditions in the multi-asset economy, then establish a mutual fund theorem that reduces the aggregate equilibrium to a single-claim problem.

Value function and consumption-wealth ratio. With homothetic preferences, investor j's value function can be written as

$$V_{j,t} = \left(\frac{c_{j,t}}{\rho^{\psi}} \right)^{\frac{1-\gamma_j}{1-\psi}} \frac{W_{j,t}^{1-\gamma_j}}{1-\gamma_j}, \qquad \text{where} \qquad \frac{dc_{j,t}}{c_{j,t}} = \mu_{c_j,t}\, dt + \sigma_{c_j,t}\, dZ_t,$$

and $c_{j,t} = c_j(X_t)$ corresponds to the consumption-wealth ratio: $C_{j,t}/W_{j,t} = c_{j,t}$.

From the Hamilton-Jacobi-Bellman equation, the consumption-wealth ratio satisfies

$$c_{j,t} = \psi\rho + (1-\psi) \left[r_t + \pi_{D,t}\, \alpha_{j,D,t} + \pi_{B,t}\, \alpha_{j,B,t} + \pi_{F,t}\, \theta_{j,t} - \frac{\gamma_j}{2} \|\Sigma_{j,t}\|^2 \right] + \xi_{j,t}, \tag{12}$$

where

$$\xi_{j,t} \equiv \mu_{c_j,t} + (1-\gamma_j)\sigma_{c_j,t}\Sigma'_{j,t} + \frac{\psi - \gamma_j}{1-\psi}\frac{\|\sigma_{c_j,t}\|^2}{2}.$$

Optimal portfolio and derivative exposure. An active investor chooses $(\alpha_{j,D,t}, \alpha_{j,B,t}, \theta_{j,t})$ to maximize the right-hand side of (12) subject to $\|\Sigma_{j,t}\| \leq \overline{\sigma}$. Let $\nu_{j,t} \geq 0$ denote the Lagrange multiplier on the leverage constraint. The first-order conditions imply

$$\lambda_{j,t} = \lambda_t,$$

where $\lambda_{j,t} \equiv \gamma_j \left[\Sigma_{j,t} + \frac{1-\gamma_j^{-1}}{1-\psi}\sigma_{c_j,t}\right] + \frac{\nu_{j,t}}{1-\psi}\Sigma_{j,t}$ is investor j's marginal price of risk. That is, each investor's marginal valuation of risk is equalized to the market price of risk λ_t.

Mutual fund theorem. The availability of three tradable assets (the dividend claim, the bond claim, and the derivative) raises the question of whether the decomposition of the aggregate endowment into separate claims affects equilibrium allocations. The following proposition shows that it does not.

Proposition 2 (Mutual fund theorem)**.** *Let $(c^*_{j,t}, \alpha^*_{j,t}, \theta^*_{j,t})$ denote the equilibrium allocation in the economy with a single claim on the aggregate endowment and a derivative. Then the equilibrium of the economy with separate dividend and bond claims is given by:*

(i) $c_{j,t} = c^*_{j,t}$ for all j;

(ii) $\alpha_{j,D,t} = \omega_{D,t}\,\alpha^*_{j,t}$ and $\alpha_{j,B,t} = \omega_{B,t}\,\alpha^*_{j,t}$ for all active j;

(iii) $\theta_{j,t} = \theta^*_{j,t}$ for all j.

All investors hold the dividend and bond claims in market-capitalization weights.

The proof proceeds by verification. Under the conjectured allocation, each investor's total risk exposure becomes $\Sigma_{j,t} = \alpha^*_{j,t}\,\sigma_{R_Y,t} + \theta^*_{j,t}\,\sigma_{R_F,t}$, which has zero loading on the dividend-share shock (the third Brownian component). Since the risk premium on the aggregate claim satisfies

$\pi_{Y,t} = \omega_{D,t}\,\pi_{D,t}+\omega_{B,t}\,\pi_{B,t}$ and the aggregate return exposure satisfies $\sigma_{R_Y,t} = \omega_{D,t}\,\sigma_{R_D,t}+\omega_{B,t}\,\sigma_{R_B,t}$, it follows that $c_{j,t} = c^*_{j,t}$ satisfies the HJB equation. The first-order conditions and all market clearing conditions are also satisfied, as the value-weighted sum of market-cap-weighted claims replicates the aggregate endowment.

Proposition 2 has two important implications. First, the stochastic discount factor, risk-free rate, and aggregate risk premium depend only on the aggregate state X_t and are independent of the dividend share s_t. The perturbation analysis of Section 4 therefore applies directly to the aggregate equilibrium. Second, individual claim prices can be computed in a second step, by solving a valuation equation conditional on the aggregate equilibrium, as described in Section 3.3.

In light of Proposition 2, we characterize the aggregate equilibrium in terms of the total risky share $\alpha_{j,t} \equiv \alpha_{j,D,t} + \alpha_{j,B,t}$, the risk premium on the aggregate claim $\pi_t \equiv \mu_{R_Y,t} - r_t$, and the return exposure $\sigma_{R,t} \equiv \sigma_{R_Y,t}$.

Optimal portfolio share. The optimal total risky share for an active investor is

$$\alpha_{j,t} = \min\left\{\frac{\pi_t}{\gamma_j\|\sigma_{R,t}\|^2} + \varsigma_{j,t},\ \frac{\overline{\sigma}}{\|\sigma_{R,t}\|}\right\}, \tag{13}$$

where $\varsigma_{j,t} \equiv \frac{1-\gamma_j^{-1}}{\psi-1}\frac{\sigma_{c_j,t}\sigma'_{R,t}}{\|\sigma_{R,t}\|^2}$ is the hedging component. The myopic demand, $\frac{\pi_t}{\gamma_j\|\sigma_{R,t}\|^2}$, reflects the investor's risk tolerance, while the hedging demand captures the covariance between $c_{j,t}$ and risky returns. The leverage constraint caps exposure at $\overline{\sigma}/\|\sigma_{R,t}\|$.

From the Hamilton-Jacobi-Bellman equation, the consumption-wealth ratio is

$$c_{j,t} = \psi\rho + (1-\psi)\left[r_t + \pi_t\alpha_{j,t} - \frac{\gamma_j}{2}\|\sigma_{R,t}\|^2\alpha_{j,t}^2\right] + \xi_{j,t}, \tag{14}$$

where

$$\xi_{j,t} \equiv \mu_{c_j,t} + (1-\gamma_j)\sigma_{c_j,t}\sigma'_{R,t}\alpha_{j,t} + \frac{\psi-\gamma_j}{1-\psi}\frac{\|\sigma_{c_j,t}\|^2}{2}$$

is a forward-looking term capturing how expected shifts in $c_{j,t}$ affect consumption choices. The consumption–wealth ratio depends both on current investment opportunities, summarized by $r_t + \pi_t \alpha_{j,t} - \frac{\gamma_j}{2}\alpha_{j,t}^2 \|\sigma_{R,t}\|^2$, and on future opportunities, summarized by $\xi_{j,t}$.

Pricing condition. Let $p_t \equiv P_{Y,t}/Y_t$ denote the price-dividend ratio of the aggregate claim. Since the expected return is $r_t + \pi_t = \frac{1}{p_t} + \mu_{P,t}$, the price-dividend ratio satisfies

$$\frac{1}{p_t} = r_t + \pi_t - (\mu + \mu_{p,t} + \sigma\sigma_{p,t}'), \tag{15}$$

where $\sigma_{R,t} = \sigma + \sigma_{p,t}$ and $(\mu_{p,t}, \sigma_{p,t})$ follow from Itô's lemma.

Aggregate state variable. Let $x_{j,t}$ denote the wealth share of type-j investors:

$$x_{j,t} \equiv \frac{\omega_j W_{j,t}}{P_{Y,t}}.$$

The aggregate state is $X_t = (x_t, \overline{\alpha}_{p,t})$, where $x_t = (x_{1,t}, x_{2,t}, \ldots, x_{J,t})$ and $\overline{\alpha}_{p,t}$ evolves as in (10). By Itô's lemma, the law of motion for $x_{j,t}$ is

$$\frac{dx_{j,t}}{x_{j,t}} = \left[r_t + \pi_t \alpha_{j,t} - c_{j,t} - \mu_{P,t} + (1 - \alpha_{j,t})\|\sigma_{R,t}\|^2 + \kappa \frac{\omega_j - x_{j,t}}{x_{j,t}}\right] dt + (\alpha_{j,t} - 1)\, \sigma_{R,t}\, dZ_t.$$

Risk premium and interest rate. Let $\mathcal{J}_t^u$ and $\mathcal{J}_t^c$ denote the sets of unconstrained and constrained active investors at time t, with wealth shares $x_{u,t} \equiv \sum_{j\in\mathcal{J}_t^u} x_{j,t}$ and $x_{c,t} \equiv \sum_{j\in\mathcal{J}_t^c} x_{j,t}$. From market clearing for the risky asset, demand from unconstrained investors satisfies

$$\underbrace{\sum_{j\in\mathcal{J}_t^u} x_{j,t}\alpha_{j,t}}_{\text{unconstrained active demand}} = \underbrace{1 - x_{0,t}\overline{\alpha}_{p,t} - x_{c,t}\overline{\alpha}_{c,t},}_{\text{net supply}} \tag{16}$$

where $\overline{\alpha}_{c,t} \equiv \overline{\sigma}/\|\sigma_{R,t}\|$ is the portfolio share of constrained agents.

Combining this condition with the optimal portfolios yields the risk premium:

$$\pi_t = \frac{\gamma_{u,t}\|\sigma_{R,t}\|^2}{x_{u,t}}\left[1 - x_{0,t}\,\overline{\alpha}_{p,t} - x_{c,t}\,\overline{\alpha}_{c,t} - x_{u,t}\,\varsigma_t\right], \tag{17}$$

where $\gamma_{u,t} \equiv [\sum_{j\in\mathcal{J}_t^u}\frac{x_{j,t}}{x_{u,t}}\frac{1}{\gamma_j}]^{-1}$ is the average risk aversion and $\varsigma_t \equiv \sum_{j\in\mathcal{J}_t^u}\frac{x_{j,t}}{x_{u,t}}\varsigma_{j,t}$ the average hedging demand of unconstrained investors. Equation (17) shows that the risk premium rises with higher risk aversion, lower passive demand, tighter leverage constraints, or weaker hedging motives.

Goods market clearing requires

$$\sum_{j=0}^{J} x_{j,t}c_{j,t} = \frac{1}{p_t}.$$

Combining this with (14) gives the equilibrium interest rate:

$$r_t = \rho + \psi^{-1}(\mu_{P,t} + \xi_t) + (1-\psi^{-1})\sum_{j=0}^{J} x_{j,t}\frac{\gamma_j\alpha_{j,t}^2}{2}\|\sigma_{R,t}\|^2 - \pi_t, \tag{18}$$

where $\xi_t \equiv \sum_{j=0}^{J} x_{j,t}\xi_{j,t}$. The first two terms reflect impatience and intertemporal substitution, while the third term captures precautionary savings.

Endogenous volatility. Return volatility has both an exogenous and an endogenous component, $\sigma_{R,t} = \sigma + \sigma_{p,t}$, where σ reflects cash-flow volatility and $\sigma_{p,t}$ arises from movements in the price-dividend ratio p_t. By Itô's lemma,

$$\sigma_{p,t} = \frac{p_x(X_t)}{p(X_t)}\sigma_{x,t} + \frac{p_{\overline{\alpha}_p}(X_t)}{p(X_t)}\sigma_{\alpha_p}. \tag{19}$$

Equation (19) shows that endogenous volatility depends on (i) the sensitivity of p_t to changes in the wealth distribution $x_{j,t}$ and the response of wealth shares to shocks, $\sigma_{x,t}$, and (ii) the sensitivity of p_t to changes in passive portfolios and the response of $\overline{\alpha}_{p,t}$ to shocks. The first term is standard in heterogeneous-agent models (see Panageas, 2020), while the second arises only with exogenous portfolio flow shocks and is quantitatively relevant when markets are sufficiently inelastic (i.e., when $\frac{p_{\overline{\alpha}_p}(X_t)}{p(X_t)}$ is large). Thus, (19) links market elasticity to return volatility $\|\sigma + \sigma_{p,t}\|$.

Markov equilibrium. Equations (17) and (18) allow us to express π_t and r_t in terms of $c_{j,t}$, p_t, and their derivatives, once $(\mu_{c_j,t}, \mu_{p,t})$ and $(\sigma_{c_j,t}, \sigma_{p,t})$ are written as functions of $(\mu_{X,t}, \sigma_{X,t})$ via Itô's lemma. Substituting these into (14) and (15), together with the laws of motion for the state variables, yields a system of four PDEs in $\{c_j(X_t)\}_{j=0}^{2}$ and $p(X_t)$. These functions depend on three state variables: the two active wealth shares $x_{1,t}$ and $x_{2,t}$, and the passive portfolio share $\overline{\alpha}_{p,t}$.

3.3 Pricing the dividend and risky debt claims

By Proposition 2, the stochastic discount factor depends only on the aggregate state X_t and is independent of the dividend share s_t. Individual claim prices can therefore be determined in a second step, conditional on the equilibrium objects $r(X_t)$ and $\lambda(X_t)$ from the aggregate economy.

Let $\hat{X}_t = (X_t, s_t)$ denote the extended state vector. Define the valuation ratio of claim $k \in \{D, B\}$ as $p_{k,t} \equiv P_{k,t}/Y_t = p_k(X_t, s_t)$. Since $P_{k,t} = p_{k,t} Y_t$, the return on claim k satisfies

$$\sigma_{R_k,t} = \sigma + \sigma_{p_k,t}, \qquad \mu_{R_k,t} = \frac{s_{k,t}}{p_{k,t}} + \mu + \mu_{p_k,t} + \sigma\,\sigma_{p_k,t}',$$

where $s_{D,t} = s_t$, $s_{B,t} = 1 - s_t$, and the diffusion of the valuation ratio is

$$\sigma_{p_k,t} = \frac{1}{p_{k,t}}\left(p_{k,X}\,\sigma_X + p_{k,s}\,\sigma_s\right),$$

with $p_{k,X} \equiv \frac{\partial p_k}{\partial X}$ and $p_{k,s} \equiv \frac{\partial p_k}{\partial s}$. Substituting into the pricing condition $\mu_{R_k,t} - r_t = \sigma_{R_k,t}\lambda_t'$ yields the partial differential equation for the valuation ratio $p_k(X, s)$:

$$\begin{aligned} 0 = \; & p_{k,X}\,\mu_X + p_{k,s}\,\mu_s + \frac{1}{2}\Big[\mathrm{tr}\big(\sigma_X\sigma_X'\,p_{k,XX}\big) + 2\,p_{k,Xs}\,(\sigma_X\sigma_s') + (\sigma_s\sigma_s')\,p_{k,ss}\Big] \\ & - (r(X) - \mu)\,p_k(X,s) + s_k(s) \\ & + p_k(X,s)\,\sigma\,\sigma_{p_k}(X,s)' - p_k(X,s)\,\sigma\,\lambda(X)' - p_k(X,s)\,\sigma_{p_k}(X,s)\,\lambda(X)'. \end{aligned}$$

This equation can be solved numerically given the equilibrium functions $r(X)$ and $\lambda(X)$ from the aggregate economy. Since $p_D(X, s) + p_B(X, s) = p(X)$ by construction, only one of the two claim prices needs to be solved independently.

4 The Determinants of the Aggregate Market Elasticity

In this section, we analyze how the *aggregate market elasticity* governs the price impact of portfolio flows. The elasticity is determined by the coupled PDE system (14)–(15), which generally has no closed-form solution. To study its determinants, we extend the perturbation method of Kogan and Uppal (2001) and derive asymptotic closed-form approximations.

State-global perturbations Our perturbation approach studies economies in a neighborhood of a benchmark that admits a closed-form solution. A convenient benchmark arises when preferences are homogeneous, $\gamma_j = \gamma$, and passive investors are fully invested in the risky asset, $\overline{\alpha}_{p,t} = 1$. In this case, the economy collapses to a Lucas economy, as shown in Lemma 1. Throughout, a superscript $\star$ denotes benchmark values.

Lemma 1 (Benchmark economy)**.** *Suppose* $\gamma_j = \gamma$, $\overline{\alpha}_{p,t} = 1$, *and* $\rho > \left(1 - \psi^{-1}\right)\left(\mu - \frac{\gamma\|\sigma\|^2}{2}\right)$. *Then,*

$$\pi^\star = \gamma\|\sigma\|^2, \qquad r^\star = \rho + \psi^{-1}\mu - \left(1 + \psi^{-1}\right)\frac{\gamma\|\sigma\|^2}{2}, \qquad p^\star = \left[\rho - \left(1 - \psi^{-1}\right)\left(\mu - \frac{\gamma\|\sigma\|^2}{2}\right)\right]^{-1},$$

$c_j^\star = (p^\star)^{-1}$, *and* $\alpha_j^\star = 1$ *for* $j = 0, \ldots, J$.

We analyze small deviations from this benchmark. Formally, consider a family of economies indexed by $\epsilon > 0$, where ϵ scales departures from the benchmark along the following dimensions.

First, investors have heterogeneous risk aversion:

$$\gamma_j = \gamma\left(1 + \hat{\gamma}_j\epsilon\right), \qquad \sum_{j=0}^{J}\omega_j\hat{\gamma}_j = 0, \tag{20}$$

so that γ is the population-weighted average and $\hat{\gamma}_j$ measures deviations. When $\epsilon = 0$, preferences are homogeneous; when $\epsilon = 1$, we recover the heterogeneous-investor economy.

Second, passive investors need not be fully invested in the risky asset. Their portfolio share is

$$\overline{\alpha}_{p,t} = 1 + \hat{\alpha}_{p,t}\,\epsilon, \tag{21}$$

where $\hat{\alpha}_{p,t}$ measures deviations from full investment benchmark and evolves as an Ornstein-Uhlenbeck process, consistent with Equation (10).

Third, the leverage constraint coefficient is perturbed according to

$$\overline{\sigma} = \|\sigma\|\big(1 + \hat{\sigma}\,\epsilon\big),$$

so that the tightness of the constraint is of order $\mathcal{O}(\epsilon)$, matching the order of leverage demand. Finally, we assume a small mortality rate, $\kappa = \hat{\kappa}\,\epsilon$.

Equilibrium objects now depend on both the state X_t and the parameter ϵ. We expand them to second order:

$$p(X,\epsilon) = p_0(X) + p_1(X)\epsilon + p_2(X)\epsilon^2 + \mathcal{O}(\epsilon^3),$$
$$c_j(X,\epsilon) = c_{j,0}(X) + c_{j,1}(X)\epsilon + c_{j,2}(X)\epsilon^2 + \mathcal{O}(\epsilon^3),$$

where $p_k(X)$ and $c_{j,k}(X)$, $k \in \{0,1,2\}$, are functions of the state X.

Unlike standard DSGE perturbations, which linearize in both X and ϵ, our method is *state-global*: we solve directly for functions of X rather than coefficients around a steady state.[6] We refer to this approach as a *state-global perturbation*. This approach builds on the perturbation method of Kogan and Uppal (2001) and, more closely, the extension by Silva (2020). This method allows us to characterize how aggregate market elasticity varies with the state of the economy.

We proceed in two steps. First, we compute the first-order corrections $p_1(X)$ and $c_{j,1}(X)$. Second, we solve for the second-order terms $p_2(X)$ and $c_{j,2}(X)$. Zeroth-order terms are given by Lemma 1, i.e., $p_0(X) = p^\star$ and $c_{j,0}(X) = c_j^\star$.

---

[6] See Kargar, Passadore, and Silva (2023) for a comparison with standard perturbation methods.

4.1 The first-order demand system

Demand for risky asset. Investor j's demand for the risky asset is

$$Q_j(X;\epsilon) = \alpha_j(X;\epsilon)\,\frac{x_j}{\omega_j},$$

where x_j is the wealth share. Define the proportional deviation from the benchmark as

$$q_{j,t} \;\equiv\; \frac{Q_j(X_t;\epsilon) - Q_j(X_t;0)}{Q_j(X_t;0)} \;=\; \alpha_{j,1}(X_t)\,\epsilon + O(\epsilon^2),$$

where $\alpha_{j,1}(X)$ is the first-order correction to the portfolio weight.

Three results determine $\alpha_{j,1}(X)$. First, the volatility of the price-dividend ratio is second order:

$$\sigma_{p,t} = \frac{p_X(X_t)}{p(X_t)}\,\sigma_X(X_t) = O(\epsilon^2),$$

since $p_{X,0} = \sigma_{X,0} = 0$ (Lemma 1). Second, hedging demand is also $O(\epsilon^2)$ because $\sigma_{c_j}(X_t)$ is second order. Third, expanding the pricing condition (15) gives

$$\hat{\pi}_t = -\,\hat{r}_t - \frac{1}{p^\star}\hat{p}_t + O(\epsilon^2),$$

where $\hat{\pi}_t \equiv \pi_t - \pi^\star$, $\hat{r}_t \equiv r_t - r^\star$, and $\hat{p}_t \equiv \frac{p_t - p^\star}{p^\star}$, using $\mu_{p,t} = O(\epsilon^2)$.

Substituting into (13), the first-order demand system is

$$q_{j,t} = -\,\zeta_{j,p}\,\hat{p}_t - \zeta_{j,r}\,\hat{r}_t + f_{j,t} + O(\epsilon^2), \tag{22}$$

where, for unconstrained investors,

$$\zeta_{j,p} = \frac{1}{p^\star}\frac{1}{\gamma\|\sigma\|^2}, \qquad \zeta_{j,r} = \frac{1}{\gamma\|\sigma\|^2}, \qquad f_{j,t} = -\,\hat{\gamma}_j\,\epsilon.$$

For passive and constrained investors, own- and cross-price elasticities vanish, $\zeta_{j,p} = \zeta_{j,r} = 0$, with demand shifters $f_{j,t} = \hat{\sigma}\,\epsilon$ for constrained investors and $f_{0,t} = \hat{\alpha}_{p,t}\,\epsilon$ for passive investors.

Equation (22) shows that the first-order approximation to the dynamic heterogeneous agent model in Section 3 provides microfoundations for the demand system in Section 2. As in the simple model, unconstrained active investors exhibit nonzero own- and cross-price elasticities, here explicitly linked to structural parameters. Shocks to $\overline{\alpha}_{p,t}$ move the passive investors' demand shifter $f_{0,t}$—the dynamic analog of the comparative statics in Section 2. In addition, the model generates further demand shifts from occasionally binding leverage constraints faced by active investors.

Consumption-wealth ratio. A first-order expansion of (14) yields

$$c_{j,t} \;=\; \psi\rho \;+\; (1-\psi)\left[r_t + \pi_t - \frac{\gamma_j\|\sigma\|^2}{2}\right] \;+\; O(\epsilon^2).$$

Consistent with Assumption 1, the risk-free rate and the risk premium enter symmetrically at first order. Substituting the expansion $\hat{\pi}_t = -\hat{r}_t - p^{\star-1}\hat{p}_t + O(\epsilon^2)$ gives

$$\hat{c}_{j,t} \;\equiv\; c_{j,t} - c_j^{\star} \;=\; \chi_{j,0} \;+\; \chi_{j,p}\,\hat{p}_t \;+\; O(\epsilon^2),$$

with coefficients

$$\chi_{j,p} = (\psi - 1)\,\frac{1}{p^{\star}}, \qquad \chi_{j,0} = (\psi - 1)\,\frac{\gamma\|\sigma\|^2}{2}\,\hat{\gamma}_j\,\epsilon.$$

Hence, conditional on p_t, the consumption-wealth ratio is independent of r_t, as in Section 2. Its slope with respect to p_t depends on the EIS: it rises with p_t when $\psi > 1$, the standard assumption in macro-finance models. Finally, to first order, the consumption–wealth ratio is unaffected by $\overline{\alpha}_{p,t}$, a property that will be central for aggregate market elasticity.

Aggregate market elasticity. Given the demand system, equilibrium prices follow from market clearing in the risky-asset and goods markets:

$$\sum_{j=0}^{J} x_{j,t}\,q_{j,t} = 0, \qquad \sum_{j=0}^{J} x_{j,t}\,\hat{c}_{j,t} = -\frac{1}{p^{\star}}\,\hat{p}_t.$$

Aggregating across investors, and dropping the j subscript on coefficients that are common across types, yields the linear system

$$\begin{bmatrix} \zeta_p & \zeta_r \\ \chi_p + \frac{1}{p^\star} & 0 \end{bmatrix} \begin{bmatrix} \hat{p}_t \\ \hat{r}_t \end{bmatrix} = \begin{bmatrix} f_t \\ -\chi_{0,t} \end{bmatrix},$$

where $f_t \equiv \sum_{j=0}^{J} x_{j,t} f_{j,t}$ and $\chi_{0,t} \equiv \sum_{j=0}^{J} x_j \chi_{j,0}$.

Definition 2 (Aggregate market elasticity)**.** *The aggregate market elasticity $\epsilon_{M,t}$ is defined as the inverse of the proportional change in the price of the risky asset in response to a flow shock f_t:*

$$\epsilon_{M,t} \equiv \left[\frac{\partial \hat{p}_t}{\partial f_t} \right]^{-1}.$$

Proposition 3 (First-order impact of portfolio flows)**.** *Suppose $\rho > \left(1 - \psi^{-1}\right)\left(\mu - \frac{\gamma \|\sigma\|^2}{2}\right)$. Then:*

(i) The price impact is zero to first order: $\epsilon_{M,t}^{-1} = \frac{\partial \hat{p}_t}{\partial f_t} = 0$.

(ii) The interest rate and risk premium responses satisfy

$$\frac{\partial \hat{r}_t}{\partial f_t} = \frac{1}{\zeta_r}, \qquad \frac{\partial \hat{\pi}_t}{\partial f_t} = -\frac{1}{\zeta_r}.$$

Proposition 3 shows that the main lesson from the simple model in Section 2 extends to the dynamic economy: portfolio inflows into the risky asset reduce the risk premium, while the corresponding shift out of bonds raises the risk-free rate. To first order, these effects offset exactly, leaving the price-dividend ratio unchanged. Hence, the aggregate market is infinitely elastic, regardless of the individual investors' own-price elasticities $\zeta_{j,p}$.

Why is the market infinitely elastic? As shown in Figure 1, price impact is absent as long as the consumption-wealth ratio does not respond directly to the portfolio-flow shock. Otherwise, price changes would generate excess demand or supply in the goods market.

To see why flows leave consumption unchanged to first order, differentiate (14) with respect to $\alpha_{j,t}$:

$$\frac{\partial c_{j,t}}{\partial \alpha_{j,t}} = (1-\psi)\,\gamma_j\,\|\sigma_{R,t}\|^2 \left[\frac{\pi_t}{\gamma_j \|\sigma_{R,t}\|^2} + \frac{1-\gamma_j^{-1}}{\psi - 1}\,\frac{\sigma_{c_{j,t}}\,\sigma_{R,t}^{\top}}{\|\sigma_{R,t}\|^2} - \alpha_{j,t} \right].$$

The term in brackets is equal to zero at the privately optimal portfolio, as implied by the investor's first-order condition. By the envelope theorem, small changes in $\alpha_{j,t}$ around the optimum have negligible welfare effects on welfare and hence do not affect $c_{j,t}$.

In the first-order approximation, this derivative is evaluated at the benchmark Lucas economy, where all investors hold optimal portfolios. The derivative is thus zero, and by Proposition 3, price impact is negligible for small (first-order) deviations from the frictionless allocation: the consumption-wealth ratio does not react to flow shocks. In the simple model of Section 2, this was imposed as an *assumption*; here it emerges as a *result* of expanding around the efficient allocation.

This reasoning also clarifies when aggregate elasticity becomes finite: when the initial allocation of risk is *inefficient*. To study such economies requires larger departures from the benchmark, captured by a second-order approximation of the demand system.

4.2 Inefficient passive demand

We next turn to a second-order approximation of the demand system to study how risk allocation affects aggregate market elasticity. To focus on the mechanism most clearly, consider the special case without preference heterogeneity or leverage constraints.

The role of risk misallocation. In this setting, demand for the risky asset is still given by (22) up to second order. Without heterogeneity, the relevant demand shifter is

$$f_t \equiv x_{0,t}\left(\overline{\alpha}_{p,t} - 1\right),$$

where $x_{0,t}$ denotes the wealth share of passive investors.

The key difference relative to the benchmark arises from the consumption-wealth ratio.

Aggregating (14) across investors, letting $c_t \equiv \sum_{j=0}^{J} x_{j,t} c_{j,t}$, and imposing market clearing $\sum_{j=0}^{J} x_{j,t} \alpha_{j,t} = 1$, yields

$$c_t \;=\; \psi\rho \;+\; (1-\psi)\left[r_t + \pi_t - \frac{\gamma\|\sigma\|^2}{2} \sum_{j=0}^{J} x_{j,t}\alpha_{j,t}^2 \right] \;+\; O(\epsilon^3).$$

Under the efficient allocation $\alpha_{j,t} = 1$, so $\sum_j x_{j,t}\alpha_{j,t}^2 = 1$. Any deviation from this efficient allocation creates dispersion in portfolios, leading to $\sum_j x_{j,t}\alpha_{j,t}^2 > 1$ by Jensen's inequality, which lowers the risk-adjusted expected return $r_t + \pi_t - \frac{\gamma\|\sigma\|^2}{2} \sum_j x_{j,t}\alpha_{j,t}^2$. When $\psi > 1$, this reduction weakens incentives to save and raises the consumption-wealth ratio for a given $r_t + \pi_t$.

Movements in the passive portfolio therefore shift the aggregate consumption–wealth ratio. Its second-order expansion is

$$\hat{c}_t \;\equiv\; c_t - c^{\star} \;=\; \chi_{0,t} \;+\; \chi_p\,\hat{p}_t \;+\; O(\epsilon^3),$$

with coefficients

$$\chi_p = (\psi - 1)\,\frac{1}{p^{\star}}, \qquad \chi_{0,t} = (\psi - 1)\,\frac{\gamma\|\sigma\|^2}{2}\,\frac{f_t^2}{x_{0,t}(1 - x_{0,t})}.$$

Because $\chi_{0,t}$ is quadratic in the flow f_t, it is $O(\epsilon^2)$. For example, when passive investors are initially underexposed to the risky asset ($\overline{\alpha}_{p,t} < 1$ and $\psi > 1$), an inflow into stocks reduces the consumption-wealth ratio:

$$\frac{\partial \chi_{0,t}}{\partial f_t} \;=\; -\,(\psi - 1)\,\gamma\|\sigma\|^2\,\frac{1 - \overline{\alpha}_{p,t}}{1 - x_{0,t}}.$$

Note that $\chi_{0,t}$ is locally insensitive to portfolio flow changes when passive portfolio is efficient ($\overline{\alpha}_{p,t} = 1$), consistent with the first-order analysis.

Figure 2 illustrates the effects of portfolio flows when passive investors are initially underexposed to risk, i.e., $\overline{\alpha}_{p,t} < 1$. An inflow into the risky asset shifts both the net supply curve (left panel)

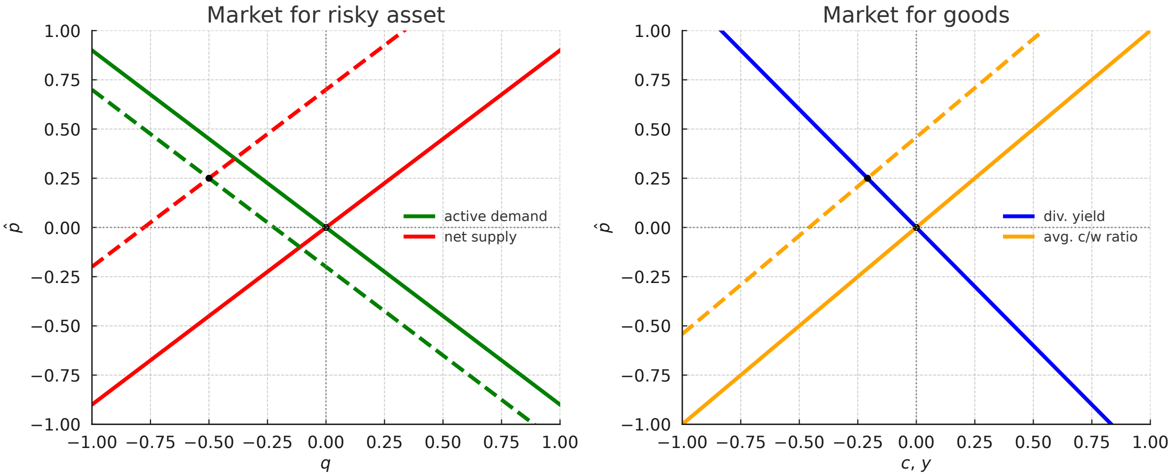


Figure 2. Equilibrium with inefficient risk allocation: risky asset (left) and goods (right) markets.

and the average consumption-wealth ratio schedule (right panel) leftward. The interest rate adjusts so that active demand meets net supply at the goods market clearing price. In this case, portfolio flows have a positive price impact and the aggregate elasticity is *finite*.

Proposition 4 (Aggregate elasticity: inefficient passive demand)**.** *Without preference heterogeneity or leverage constraints, the inverse aggregate elasticity is*

$$\varepsilon_{M,t}^{-1} = -\left(\chi_p + y^\star\right)^{-1}\frac{\partial \chi_{0,t}}{\partial f_t} = \left(1-\psi^{-1}\right)\frac{\gamma\|\sigma\|^2}{y^\star}\frac{1-\overline{\alpha}_{p,t}}{x_{a,t}} + \mathcal{O}\left(\epsilon^2\right),$$

where $x_{a,t} = 1 - x_{0,t}$ is the wealth share of active investors and $y^\star = 1/p^\star$ is the dividend-price ratio in the benchmark economy.

Proposition 4 shows that, unlike in Gabaix and Koijen (2023), the risky-asset elasticities ζ_p and ζ_r are not the direct drivers of aggregate elasticity. Because the demand system is recursive (interest rates only enters the risky asset demand) the overall price impact is governed instead by how portfolio flows impact the consumption-wealth ratio.

The macro elasticity is finite whenever $\overline{\alpha}_{p,t} \neq 1$ and $\psi \neq 1$. In the unit EIS case, the price-dividend ratio is pinned down by the subjective discount rate, so portfolio flows generate offsetting movements in the risk premium and the risk-free rate. With efficient risk allocation ($\overline{\alpha}_{p,t} = 1$), there is again no price impact.

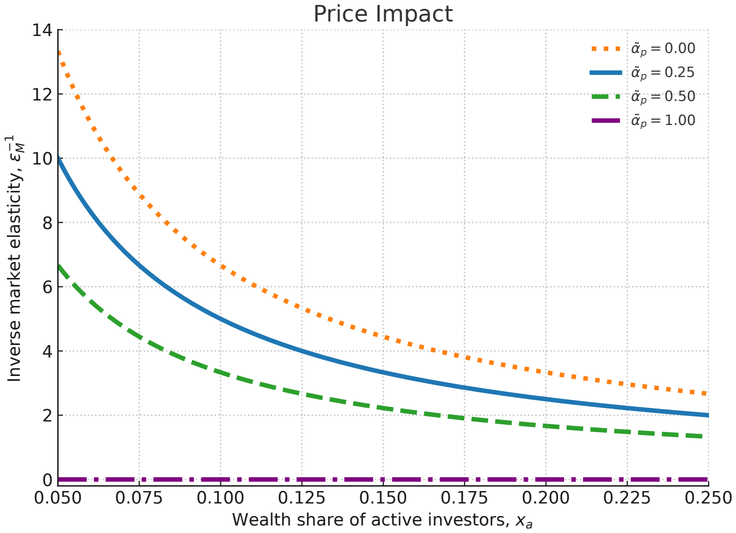


Figure 3. Price impact: inefficient passive demand.

The price impact is positive when $\psi > 1$ and $\overline{\alpha}_{p,t} < 1$: interest rate adjustments no longer fully offset the risk premium response, so the latter dominates. This mirrors representative agent models in which uncertainty shocks affect prices only when $\psi > 1$ (e.g., Bansal and Yaron, 2004).

The state-global perturbation analysis also makes clear that elasticity is state dependent. Figure 3 plots the price impact as a function of active investors' wealth for different passive-share levels. Price impact rises when active investors are under capitalized, reflecting their limited risk-bearing capacity, and when the passive share deviates further from its efficient level. In the extreme case where passive investors hold no equities ($\overline{\alpha}_{p,t} = 0$), we recover a version of Basak and Cuoco (1998) in which price impact is maximized.

4.3 Preference heterogeneity and leverage constraints

We now reintroduce preference heterogeneity and leverage constraints into the model. These features shift the conditions under which passive investors' portfolios are efficient and therefore alter the aggregate market elasticity. The next proposition provides the main result.

Proposition 5 (Aggregate elasticity: heterogeneity and leverage constraints)**.** *The inverse aggregate* *market elasticity is*

$$\varepsilon_{M,t}^{-1} \;=\; \left(1-\psi^{-1}\right)\frac{\gamma\,\|\sigma\|^2}{y^\star}\,\frac{\overline{\alpha}_{p,t}^{\text{opt}}-\overline{\alpha}_{p,t}}{x_{a,t}-x_{c,t}}\left(1-x_{c,t}\right) \;+\; O\Big(\epsilon^2\Big), \tag{23}$$

where $x_{c,t}$ is the wealth share of constrained active investors and the passive investors' optimal portfolio share is

$$\overline{\alpha}_{p,t}^{\text{opt}} \;=\; 1 \;-\; \frac{x_{a,t}-x_{c,t}}{1-x_{c,t}}\,\frac{\gamma_0-\mathbb{E}_t^u[\gamma_j]}{\gamma} \;-\; \frac{x_{c,t}}{1-x_{c,t}}\Big(\frac{\overline{\sigma}}{\|\sigma\|}-1\Big), \tag{24}$$

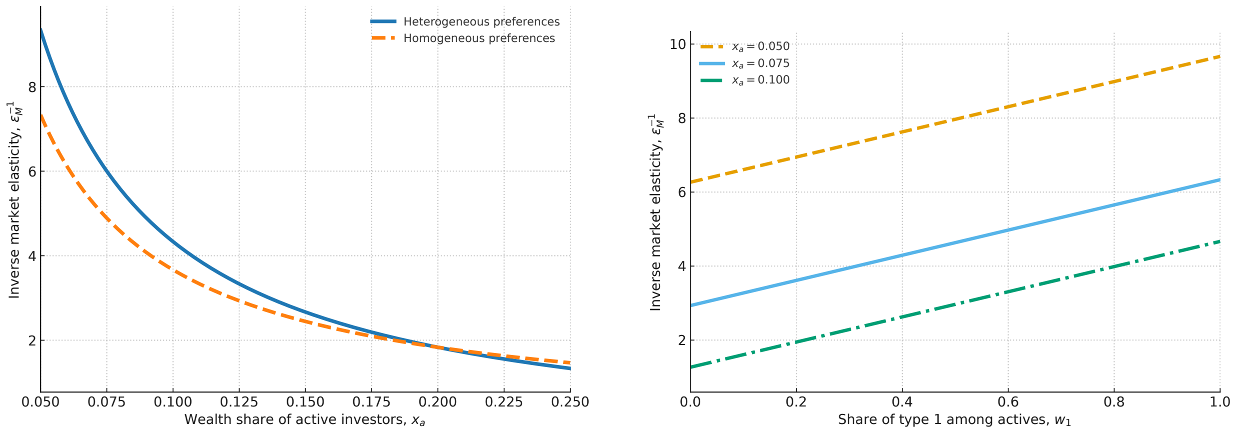
with $x_{u,t} \equiv x_{a,t} - x_{c,t}$ the wealth share of unconstrained active investors and

$$\mathbb{E}_t^u[\gamma_j] \;\equiv\; \sum_{j\in\mathcal{J}^u}\frac{x_{j,t}}{x_{u,t}}\,\gamma_j.$$

Proposition 5 characterizes aggregate market elasticity in the full model. As in the frictionless benchmark, the market remains infinitely elastic when passive investors hold their optimal portfolio share, $\overline{\alpha}_{p,t} = \overline{\alpha}_{p,t}^{\text{opt}}$. The optimal portfolio share corresponds to the value of $\overline{\alpha}_{p,t}$ such that the passive investor would have no incentive to change their portfolio. Under homogeneous preferences ($\gamma_j = \gamma$) and slack leverage constraints ($x_{c,t} = 0$), this condition reduces to $\overline{\alpha}_{p,t} = 1$. More generally, preference heterogeneity and binding leverage constraints shift the efficient passive share away from unity, with its level determined by the wealth distribution and the relative risk aversion of different investor types.

Preference heterogeneity. With slack leverage constraints ($x_{c,t} = 0$), (24) simplifies to

$$\overline{\alpha}_{p,t}^{\text{opt}} = 1 - x_{a,t}\,\frac{\gamma_0-\mathbb{E}_t^u[\gamma_j]}{\gamma},$$



(a) Homogeneous vs. heterogeneous preferences

(b) Role of wealth distribution of active investors

Figure 4. Price impact comparisons. Panel (a) compares heterogeneous and homogeneous preferences. Panel (b) shows the sensitivity of the price impact to wealth distribution among active investors when $J = 2$ for different passive portfolio shares.

so, unlike the homogeneous preference case, the efficient passive share $\overline{\alpha}_{p,t}^{\mathrm{opt}}$ differs from unity whenever $\gamma_0 \neq \mathbb{E}_t^u[\gamma_j]$. The inverse elasticity in (23) becomes

$$\varepsilon_{M,t}^{-1} = \left(1 - \psi^{-1}\right) \frac{\gamma \|\sigma\|^2}{y^\star} \frac{\overline{\alpha}_{p,t}^{\mathrm{opt}} - \overline{\alpha}_{p,t}}{x_{a,t}} + O(\epsilon^2).$$

Optimal risk sharing implies that when passive investors are more risk averse than the average unconstrained active ($\gamma_0 > \mathbb{E}_t^u[\gamma_j]$),

$$\frac{\partial \overline{\alpha}_{p,t}^{\mathrm{opt}}}{\partial x_{a,t}} = -\frac{\gamma_0 - \mathbb{E}_t^u[\gamma_j]}{\gamma} < 0,$$

so the efficient passive equity share rises as the active wealth share falls, typically after adverse aggregate shocks that erode active wealth that is more exposed to risk. If passive investors do not rebalance in such episodes, market inelasticity increases because $x_{a,t}$ declines (reducing risk-bearing capacity) while the gap $\overline{\alpha}_{p,t}^{\mathrm{opt}} - \overline{\alpha}_{p,t}$ widens (moving the economy farther from the efficient allocation). Consequently, price impact rises as $x_{a,t}$ falls, as shown in the left panel (a) of Figure 4: the heterogeneous- and homogeneous-preference curves coincide at $x_a = 0.20$, but only the former steepens markedly as x_a declines.

Panel (b) of Figure 4 shows how the distribution of wealth across active investors shapes price impact. With two active types ($J = 2$) and $\gamma_1 > \gamma_2$, the average risk active aversion is

$$\mathbb{E}_t^u[\gamma_j] \;=\; w_{1,t}\gamma_1 + \left(1 - w_{1,t}\right)\gamma_2, \qquad w_{1,t} \equiv \frac{x_{1,t}}{x_{a,t}}.$$

When leverage constraints are slack ($x_{c,t} = 0$), the efficient passive share satisfies $\overline{\alpha}_{p,t}^{\text{opt}} = 1 - x_{a,t}\,(\gamma_0 - \mathbb{E}_t^u[\gamma_j])/\gamma$, so holding $x_{a,t}$ fixed,

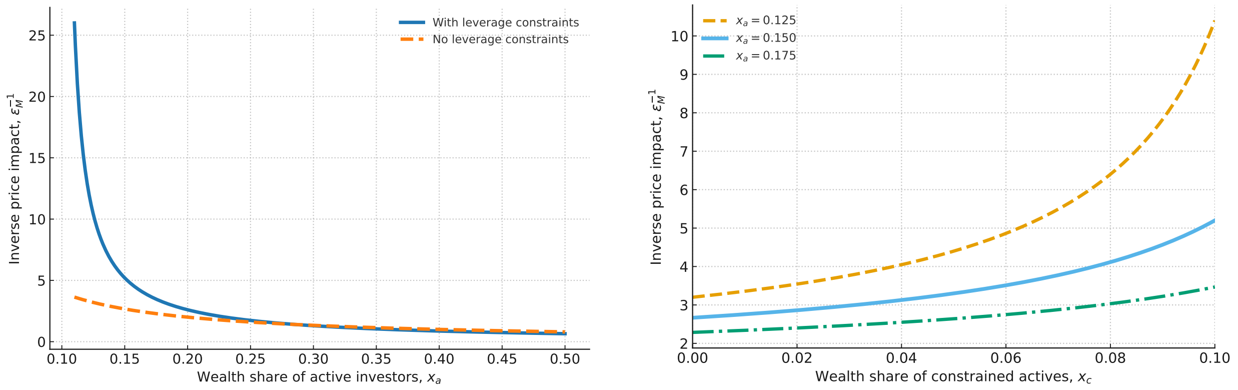
$$\frac{\partial\, \overline{\alpha}_{p,t}^{\text{opt}}}{\partial w_{1,t}} = \frac{x_{a,t}}{\gamma}\,(\gamma_1 - \gamma_2) \;>\; 0.$$

Thus, as the wealth share of the more risk averse active type rises, $\mathbb{E}_t^u[\gamma_j]$ increases and the passive sector should optimally hold more of the risky asset. If the passive portfolio is not adjusted, the deviation $\overline{\alpha}_{p,t}^{\text{opt}} - \overline{\alpha}_{p,t}$ widens, worsening risk allocation and raising price impact, consistent with the upward shift in the panel (b) of Figure 4.

Leverage constraints. When some active investors are constrained ($x_{c,t} > 0$), the marginal investors are the *unconstrained* active agents, and their risk-bearing capacity,

$$x_{u,t} \;\equiv\; x_{a,t} - x_{c,t},$$

becomes the key determinant of price impact. The panel (a) of Figure 5 plots price impact against the active wealth share $x_{a,t} = 1 - x_{0,t}$, with and without binding constraints. To isolate the role of constraints, we set $\gamma_0 = \mathbb{E}_t^u[\gamma_j]$, removing pure preference heterogeneity effects. As $x_{a,t}$ declines, e.g., after a negative aggregate shock when $\overline{\alpha}_{p,t} < 1$, price impact rises more sharply under binding constraints because $x_{u,t}$ shrinks and the risk-bearing capacity of marginal investors is depleted. Panel (b) holds $x_{a,t}$ fixed and varies the constrained share $x_{c,t}$. The price impact increases monotonically with $x_{c,t}$. This is intuitive: price impact is larger when a larger share of active investors is constrained.



(a) Role of leverage constraints *(assumes $x_{c,t} = 0.10$)* **(b)** Varying the wealth share of constrained actives, $x_{c,t}$

Figure 5. Price impact under leverage constraints. Panel (a) compares the cases with and without leverage constraints (holding $x_{c,t} = 0.10$). Panel (b) plots the price impact as a function of $x_{c,t}$ for different values of $x_{a,t}$.

Taking stock. Proposition 5 clarifies the main determinants of aggregate market elasticity. The elasticity is *finite* only when risk is misallocated: small portfolio flows around the frictionless allocation have no first-order price impact. Preference heterogeneity amplifies price impact in downturns because, as the active wealth share falls, the efficient passive share rises, widening the gap between efficient and actual portfolios. Binding leverage constraints further increase price impact by shrinking the marginal (unconstrained) risk-bearing capacity, $x_{u,t}$. Taken together, these forces imply rich state dependence of aggregate elasticity in economies with heterogeneous investors and financial frictions.

5 Quantitative Implications

In this section, we turn to the quantitative implications of the model described in Section 3. We begin by describing our calibration strategy before presenting numerical results.

5.1 Calibration strategy

To calibrate the model, we process Flow of Funds (FoF) data to construct $J = 3$ major sectors: passive, constrained, and constrained. The passive sector includes households and non-

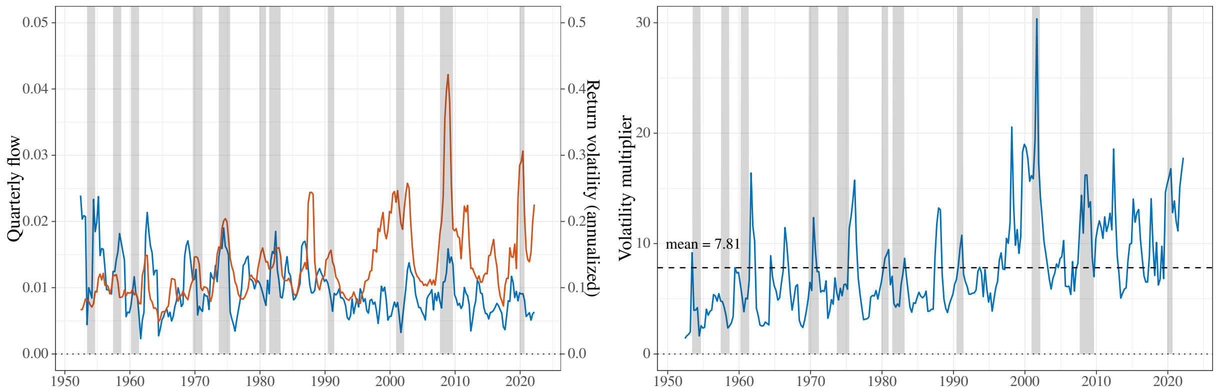


Figure 6. Quarterly flows and return volatility. The left panel plots quarterly flows (in blue) and return volatility (in red). The right panel plots the "volatility multiplier," defined as the ratio of return volatility to flows. Source: Flow of Funds and CRSP.

profits (L.101), government (L.105), the monetary authority (L.109), and money market funds (L.121). Its risky asset holdings is dominated by household. The constrained sector includes depository institutions (L.110), mutual funds (L.122), close-end funds (L.123), ETFs (L.124), government-sponsored enterprises (L.125), agency and GSE-backed mortgage pools (L.126), asset backed securities (L.127), finance companies (L.128), REITs (L.129), and broker-dealers (L.130). The unconstrained sector, which represents relatively long-horizoned investors, include insurance companies (L.115 and L.116), pensions (L.118, L.119, and L.120), and rest of the world (L.133).

We assign sectoral masses ω_j and risk aversion coefficients γ_j to match two sets of moments. First, average sectoral wealth shares are calibrated to match the empirical distribution of stock market holdings. Second, we match sectoral betas from time-series regressions of risky asset holdings on aggregate risky asset supply. These regression coefficients provide an empirical analog to the model-implied relationship between sectoral and aggregate risky positions.

The process for passive demand is modeled as the exogenous CIR process in Equation (10). Its parameters are calibrated as follows. The mean of the process, $\overline{\alpha}$, is chosen to match the average portfolio share of the passive sector. The volatility σ_p is disciplined by observed fluctuations in equity flows, which we measure by scaling quarterly flows by lagged market capitalization, following Appendix D3 of Gabaix and Koijen (2023). We use indirect inference to align model-implied aggregate flows with those in the data (blue line, left panel of Figure 6), in line with

	Parameter	Choice
	Preferences & distribution	
ψ	EIS	1.5
γ_0	Risk aversion of passive investors	10.86
γ_j	Risk aversion of constrained investors	6.15
γ_j	Risk aversion of unconstrained investors	25.52
ρ	Effective discount rate	0.01
ω_0	Share of passive agents	0.50
ω_j	Share of constrained agents	0.25
ω_j	Share of unconstrained agents	0.25
	Technology	
μ	Endowment growth rate	0.0183
σ	Endowment volatility	0.0357
σ_s	Dividend share volatility	0.11
$\bar{s}$	Mean dividend share	0.5
κ_s	Mean reversion of dividend share	0.1
	Passive demand	
$\overline{\alpha}$	Mean	0.52
θ_p	Mean reversion parameter	0.75
σ_p	Passive demand shock exposures	0.02
ν	Corr. of passive demand and fundamental shock	0.5
	Leverage constraints	
$\overline{\sigma}$	Tightness of the leverage constraint	0.11

Table 1. Parameter values.
This table reports the parameter values used in calibrating the model.

their approach. The persistence parameter θ_p is informed by evidence on household portfolio inertia. Brunnermeier and Nagel (2008) document substantial inertia using PSID data, finding that household portfolios change little absent major events, consistent with slow rebalancing. Their regression analysis suggests an inertia coefficient near 0.75, indicating strong persistence.[7] Finally, we calibrate the correlation between passive flow shocks and aggregate shocks by regressing quarterly flows (again, blue line in the left panel of Figure 6) on aggregate market returns, with the estimated beta providing the implied correlation. The parameter values are summarized in Table 1.

---

[7] Relatedly, Parker, Schoar, and Sun (2023) show that the adoption of target-date funds has contributed to highly stable household portfolio shares.

Component	Model	Data	Component	Model	Data
Dividend claim					
Risk premium	0.069	0.060	Assets/Equity (unconstrained)	0.539	0.533
Risk-free rate	0.0043	0.010	Assets/Equity (constrained)	1.934	1.616
Return volatility	0.209	0.180	Asset share (passive)	0.462	0.388
Flow volatility	0.038	0.034	Asset share (unconstrained)	0.224	0.201
Price impact	8.31	6.90	Asset share (constrained)	0.321	0.292
Endowment claim					
Risk premium	0.020	–	Return volatility	0.047	–
Price impact	0.67	–			

Table 2. Simulation results.
This table reports the model-implied and empirical moments. All values are annualized.

5.2 Quantitative results

To evaluate the quantitative implications of the model, we solve the full dynamic system numerically rather than relying on the perturbation approach of Section 4. A global solution method is necessary given the dimensionality of the state space, which includes 4 state variables corresponding to one passive sector, one constrained sector, one unconstrained sector, and a dividend share. We adopt the neural network-based method of Duarte et al. (2024), which is well suited for high-dimensional problems. Details of the solution method are provided in the appendix. We then simulate the model to generate time series of prices, returns, and flows, which we compare to empirical data. Table 2 reports unconditional moments from the simulated data alongside their empirical counterparts. For the dividend claim, the model delivers a risk premium of 6.9%, a risk-free rate of 0.43%, and return volatility of 20.9%, all close to the corresponding empirical targets. The model also generates a sizable price impact of 8.31 and a flow volatility of 2.8%. This is in the ballpark of the empirically estimated price impact of around 6.90.[8] The model also broadly matches the average asset/equity ratio and risky asset shares of the main sectors.

Figure 7 presents the model-implied conditional price impact. The right panel of the figure

---

[8] Price impact in the model is defined as $\frac{d\log(p/d)}{d\alpha_p}$. Gabaix and Koijen (2023) estimate a price multiplier of around 5, while Hartzmark and Solomon (2025) estimate a multiplier of 1.9. Taking an average between the two yields 3.45. Considering that the passive investor share is around half, this amounts to a price impact of approximately $3.45 \times 2 = 6.90$.

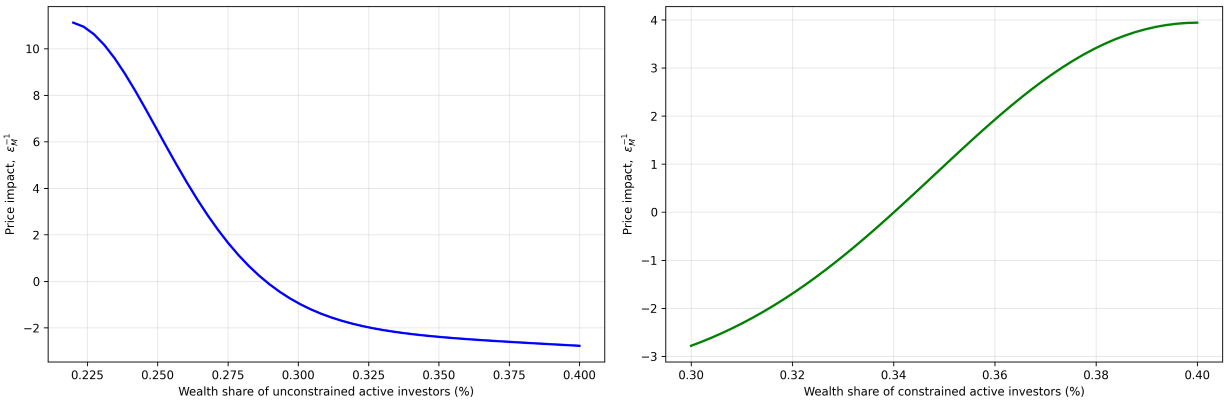


Figure 7. This figure shows the price impact of active investors from the dynamic model.

shows that, as the wealth share of constrained active investors increases, the price impact goes up, as more agents are constrained in the economy.

Next, we run return predictability tests by regressing the cumulative future returns on the price-dividend ratio. Table 3 presents the slope coefficients in (1) the full sample and (2) conditional on bad states defined as time periods when the price-dividend ratio is below 25th percentile of its long-run mean. The table shows that realized future excess returns are negatively predicted by the price-dividend ratio, and this predictability is stronger in bad times.

6 Conclusion

This paper develops a general equilibrium model with heterogeneous investors, passive demand, and financial constraints to study the determinants of aggregate market elasticity. We show how general equilibrium adjustments, particularly the joint movement of the risk-free rate and the risk premium, fundamentally shape the price impact of portfolio flows.

A central insight is the role of cross-price elasticity. When demand for risky assets responds to interest rates, shifts in the risk premium are offset by changes in the risk-free rate. In this case, the market can be infinitely elastic even if individual investors are highly inelastic. Aggregate elasticity becomes finite only when risk is initially misallocated: portfolio flows then alter aggregate savings behavior, preventing interest rates from fully offsetting risk-premium movements.

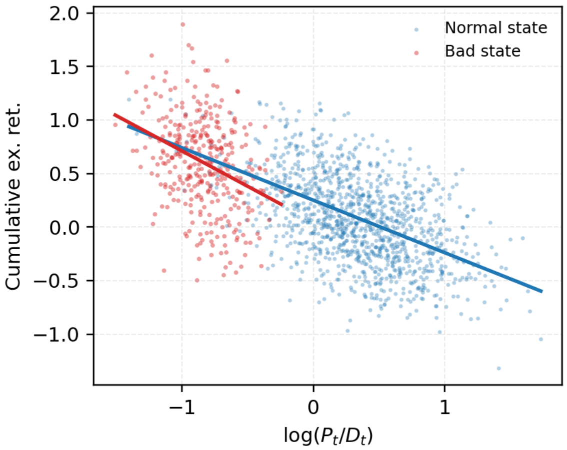


Figure 8. Return predictability

Dependent Variable:	Cumulative ex. ret.	
Model:	(1)	(2)
Sample:	Full	Bad state
Variables		
$\log(P_t/D_t)$	-0.49	-0.66
	(0.0005)	(0.0019)
R^2	0.45	0.21
Observations	989,000	247,250

Table 3. Return predictability regression

Note: The left panel shows the relationship between the price-dividend ratio and future cumulative returns. Blue points correspond to normal state and red points correspond to bad state defined as time periods when price-dividend ratio is less than the 25-pctl of its long-run mean. The lines represent the regression fit. The right panel reports estimates from the return-predictability regression $\sum_{k=1}^{T} r_{t+k} - r_t = a + b \log(P_t/D_t) + \varepsilon_{t+k}$, where $T = 4$ quarters. Entries in parentheses are standard errors.

Macro elasticity is both state-dependent and time-varying. Passive investors amplify price impacts, leverage constraints further tighten risk-bearing capacity, while preference heterogeneity increases elasticity by improving risk allocation. Inelasticity alone does not generate excess volatility, nor do flows in infinitely elastic markets; both elements are needed to explain observed fluctuations and the countercyclical volatility multiplier.

Our analytical results rely on a state-global perturbation method that provides closed-form characterizations of elasticity, while our quantitative analysis solves the full dynamic model numerically. The calibrated model highlights how the wealth distribution across households, intermediaries, and long-term investors shapes aggregate elasticity and volatility.

Overall, our findings suggest that market inelasticity is best understood as a symptom of inefficient risk allocation, rather than a structural feature of financial markets. Improving the allocation of risk across sectors may therefore be more effective for reducing excess volatility than targeting individual investor behavior. Future work could extend this framework to study asset pricing anomalies, monetary policy transmission, and the role of market structure in shaping aggregate elasticity.

References

Abel, Andrew B., 1999, Risk premia and term premia in general equilibrium, *Journal of Monetary Economics* 43, 3–33.

Adrian, Tobias, and Hyun Song Shin, 2014, Procyclical leverage and value-at-risk, *Review of Financial Studies* 27, 373–403.

Ameriks, John, and Stephen P. Zeldes, 2004, How do household portfolio shares vary with age?, Working Paper.

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

Basak, Suleyman, and Domenico Cuoco, 1998, An equilibrium model with restricted stock market participation, *Review of Financial Studies* 11, 309–341.

Binsbergen, Jules H. van, Benjamin David, and Christian C. Opp, 2025, How (not) to identify demand elasticities in dynamic asset markets, Working Paper.

Brunnermeier, Markus K., and Stefan Nagel, 2008, Do wealth fluctuations generate time-varying risk aversion? micro-evidence on individuals, *American Economic Review* 98, 713–36.

Brunnermeier, Markus K., and Lasse Heje Pedersen, 2009, Market liquidity and funding liquidity, *Review of Financial Studies* 22, 2201–2238.

Campbell, John Y., 1999, Asset prices, consumption, and the business cycle, in John B. Taylor, and Michael Woodford, eds., *Handbook of Macroeconomics*, volume 1, chapter 19, 1231–1303 (Elsevier).

Chang, Yen-Cheng, Harrison Hong, and Inessa Liskovich, 2015, Regression discontinuity and the price effects of stock market indexing, *Review of Financial Studies* 28, 212–246.

Chen, Hui, Winston Wei Dou, and Leonid Kogan, 2024, Measuring "dark matter" in asset pricing models, *The Journal of Finance* 79, 843–902.

Chevalier, Judith, and Glenn Ellison, 1997, Risk taking by mutual funds as a response to incentives, *Journal of political economy* 105, 1167–1200.

Chien, YiLi, Harold L. Cole, and Hanno Lustig, 2012, Is the volatility of the market price of risk due to intermittent portfolio rebalancing?, *American Economic Review* 102, 2859–2896.

Cochrane, John H., 1992, Explaining the variance of price–dividend ratios, *Review of Financial Studies* 5, 243–280.

Dannhauser, Caitlin D., and Jeffrey Pontiff, 2024, Flow, Working Paper.

Davis, Carter, Mahyar Kargar, Jiacui Li, and Dejanir Silva, 2026, On the recovery and usage of demand elasticities in dynamic settings, Working Paper.

Deuskar, Prachi, and Timothy C. Johnson, 2011, Market Liquidity and Flow-driven Risk, *Review of Financial Studies* 24, 721–753.

Duarte, Victor, Diogo Duarte, and Dejanir H. Silva, 2024, Machine learning for continuous-time finance, *Review of Financial Studies* 37, 3217–3271.

Duffie, Darrell, and Larry G. Epstein, 1992, Stochastic differential utility, *Econometrica* 353–394.

Fuchs, William, Satoshi Fukuda, and Daniel Neuhann, 2023, Demand-system asset pricing: Theoretical foundations, Working Paper.

Gabaix, Xavier, and Ralph SJ Koijen, 2023, In search of the origins of financial fluctuations: The inelastic markets hypothesis, Working Paper.

Haddad, Valentin, Zhiguo He, Paul Huebner, Péter Kondor, and Erik Loualiche, 2025, Causal inference for asset pricing, Working Paper.

Harris, Lawrence, and Eitan Gurel, 1986, Price and volume effects associated with changes in the S&P 500 list: New evidence for the existence of price pressures, *Journal of Finance* 41, 815–829.

Hartzmark, Samuel M., and David H. Solomon, 2025, Market-wide predictable price pressure, *American Economic Review* 115, 3171–3213.

He, Zhiguo, Péter Kondor, and Jessica S. Li, 2025, Demand elasticity in dynamic asset pricing, Working Paper.

Johnson, Timothy C., 2006, Dynamic liquidity in endowment economies, *Journal of Financial Economics* 80, 531–562.

Johnson, Timothy C., 2008, Volume, liquidity, and liquidity risk, *Journal of Financial Economics* 87, 388–417.

Johnson, Timothy C., 2009, Liquid capital and market liquidity, *The Economic Journal* 119, 1374–1404.

Kargar, Mahyar, Juan Passadore, and Dejanir Silva, 2023, Liquidity and risk in OTC markets: A theory of asset pricing and portfolio flows, Working Paper.

Kekre, Rohan, Moritz Lenel, and Federico Mainardi, 2025, Monetary policy, segmentation, and the term structure, Working paper.

Kogan, Leonid, and Raman Uppal, 2001, Risk aversion and optimal portfolio policies in partial and general equilibrium economies.

Koijen, Ralph S. J., and Motohiro Yogo, 2019, A demand system approach to asset pricing, *Journal of Political Economy* 127, 1475–1515.

Koijen, Ralph SJ, Robert J Richmond, and Motohiro Yogo, 2024, Which investors matter for equity valuations and expected returns?, *Review of Economic Studies* 91, 2387–2424.

LeRoy, Stephen F., and Richard D. Porter, 1981, The present-value relation: Tests based on implied variance bounds, *Econometrica* 555–574.

Li, Jennifer (Jie), Neil D. Pearson, and Qi Zhang, 2020, Impact of demand shocks on the stock market: Evidence from Chinese IPOs, Working Paper.

Menzly, Lior, Tano Santos, and Pietro Veronesi, 2004, Understanding predictability, *Journal of Political Economy* 112, 1–47.

Panageas, Stavros, 2020, The implications of heterogeneity and inequality for asset pricing, *Foundations and Trends® in Finance* 12, 199–275.

Parker, Jonathan A., Antoinette Schoar, Allison T Cole, and Duncan Simester, 2022, Household portfolios and retirement saving over the life cycle, Technical report, National Bureau of Economic Research.

Parker, Jonathan A., Antoinette Schoar, and Yang Sun, 2023, Retail financial innovation and stock market dynamics: The case of target date funds, *Journal of Finance* 78, 2673–2723.

Pavlova, Anna, and Taisiya Sikorskaya, 2023, Benchmarking intensity, *Review of Financial Studies* 36, 859–903.

Ray, Walker, Michael Droste, and Yuriy Gorodnichenko, 2024, Unbundling quantitative easing: Taking a cue from treasury auctions, *Journal of Political Economy* 132, 3115–3172.

Schmickler, Simon, 2020, Identifying the price impact of fire sales using high-frequency surprise mutual fund flows, Working Paper.

Shiller, Robert J., 1981, Do stock prices move too much to be justified by subsequent changes in dividends?, *The American Economic Review* 71, 421–436.

Shiller, Robert J., 1992, *Market Volatility* (MIT press).

Shleifer, Andrei, 1986, Do demand curves for stocks slope down?, *Journal of Finance* 41, 579–590.

Silva, Dejanir, 2020, The risk channel of unconventional monetary policy, Working Paper.

Vayanos, Dimitri, and Jean-Luc Vila, 2021, A preferred-habitat model of the term structure of interest rates, *Econometrica* 89, 77–112.

Appendix

A Derivations for Section 2

A.1 The bond demand view

Let $b_j \equiv \frac{B_j}{W_j}$ denote the bond-to-wealth ratio. The market clearing condition for bonds can be expressed as follows

$$\underbrace{x_a b_a}_{\textbf{active bond demand}} = \underbrace{-x_p b_p}_{\textbf{net bond supply}}$$

The linearized bond demand for a passive investor is given by

$$b_j - b_j^\star = \left[r - r^\star + \frac{c_j'(\mu - p^\star)}{1 - c_j(\mu - p^\star)}(p - p^\star) \right] b_j^\star - \alpha_j^\star e^{r^\star}(1 - c_j(\mu - p^\star))\hat{\alpha}_j$$

For simplicity, focus on the case $\overline{\alpha}_p = 1$, so $b_p^\star = b_a^\star = 0$. The passive bond demand is then given by

$$x_p b_p = f^b,$$

where $f^b \equiv -x_p[1 - c_p(\mu - p^\star)]\hat{\alpha}_p$. The active bond demand is given by

$$x_a b_a = -\zeta_p^b (p - p^\star) - \zeta_r^b (r - r^\star),$$

where $\zeta_p^b = \zeta_r^b = -x_a g_a'(\mu - p^\star - r^\star)[1 - c_a(\mu - p^\star)]$.

The demand system can be written as follows:

$$\begin{bmatrix} \zeta_p^q & \zeta_r^q \\ \zeta_p^b & \zeta_r^b \end{bmatrix} \begin{bmatrix} p - p^\star \\ r - r^\star \end{bmatrix} = \begin{bmatrix} f^q \\ f^b \end{bmatrix},$$

where we denote here $f^q \equiv f$ for symmetry.

Inverting the system above, we obtain

$$\begin{bmatrix} p - p^\star \\ r - r^\star \end{bmatrix} = \frac{1}{\zeta_p^q \zeta_r^b - \zeta_r^q \zeta_p^b} \begin{bmatrix} \zeta_r^b & -\zeta_r^q \\ -\zeta_p^b & \zeta_p^q \end{bmatrix} \begin{bmatrix} f^q \\ f^b \end{bmatrix}.$$

The price is given by

$$p - p^\star = \frac{\zeta_r^b f^q - \zeta_r^q f^b}{\zeta_p^q \zeta_r^b - \zeta_r^q \zeta_p^b} = x_p \hat{\alpha}_p \frac{\zeta_r^b + \zeta_r^q (1 - c_p(\mu - p^\star))}{\zeta_p^q \zeta_r^b - \zeta_r^q \zeta_p^b} = 0,$$

using the fact that $\zeta_r^q = x_a g_a'(\mu - p^\star - r^\star)$ and $c_a(\mu - p^\star) = c_p(\mu - p^\star)$.

B Derivations for Section 3

B.1 Investors' problem

The Hamilton-Jacobi-Bellman (HJB) equation for investor j can be written as

$$\begin{aligned} 0 = \max_{C_j,\alpha_j} & f_j(C_j, V_j) + V_{j,W}\left[rW_j + (\mu_R - r)\alpha_j W_j - C_j\right] + V_{j,X}\mu_X \\ & + \frac{1}{2}V_{j,WW}W_j^2\alpha_j^2\|\sigma_R\|^2 + V_{j,WX}W_j\sigma_X\sigma_R'\alpha_j + \frac{1}{2}\sum_{k=1}^{d}\sigma_{X,k}'V_{j,XX}\sigma_{X,k}, \end{aligned}$$

subject to $\alpha_j \in \Omega_j$. For ease of notation, we dropped time subscripts. Note that $V_{j,X}$ and $V_{j,WX}$ are $1 \times N$ vectors, $V_{j,XX}$ is a $N \times N$ matrix, and both $V_{j,W}$ and $V_{j,WW}$ are scalars. The drift μ_X is a $N \times 1$ vector, the diffusion σ_X is a $N \times d$ matrix, while σ_R is a $1 \times d$ vector. The notation $\sigma_{X,k}$ denotes the $k-$th column of σ_X, that is, the exposure to the $k-$th Brownian motion.

The optimal consumption is given by

$$C_j = \rho^\psi((1-\gamma_j)V_j)^{\frac{1-\gamma_j\psi}{1-\gamma_j}}V_{j,W}^{-\psi}.$$

The optimal portfolio share for an active investor is given by

$$\alpha_j = \min\left\{-\frac{V_{j,W}(\mu_R - r)}{V_{j,WW}W\|\sigma_R\|^2} - \frac{V_{WX}}{V_{WW}W}\frac{\sigma_X\sigma_R'}{\|\sigma_R\|^2}, \frac{\overline{\sigma}}{\|\sigma_{R,t}\|}\right\}.$$

Given the homotheticity of preferences, the value function for investor j can be written as

$$V_{j,t} = \left(\frac{\xi_{j,t}}{\rho^\psi}\right)^{\frac{1-\gamma_j}{1-\psi}}\frac{W_{j,t}^{1-\gamma_j}}{1-\gamma_j}. \tag{B.1}$$

This particular parametrization of the value function implies that the consumption-wealth ratio is given by

$$\frac{C_{j,t}}{W_{j,t}} = \xi_{j,t}.$$

The optimal portfolio share for active investors is given by

$$\alpha_{j,t} = \min\left\{\frac{\mu_{R,t} - r_t}{\gamma_j\|\sigma_{R,t}\|^2} - \frac{1-\gamma_j^{-1}}{1-\psi}\frac{\sigma_{\xi_j,t}\sigma_{R,t}'}{\|\sigma_{R,t}\|^2}, \frac{\overline{\sigma}}{\|\sigma_{R,t}\|}\right\}.$$

It is convenient to consider the investor's risk exposure $\sigma_j \equiv \alpha_j \|\sigma_R\|$, which is then given by

$$\sigma_{j,t} = \min\left\{ \frac{\eta_t}{\gamma_j} - \frac{1-\gamma_j^{-1}}{1-\psi} \frac{\sigma_{\xi_j,t}\sigma'_{R,t}}{\|\sigma_{R,t}\|}, \overline{\sigma} \right\},$$

where $\eta_t \equiv \frac{\mu_{R,t}-r_t}{\|\sigma_{R,t}\|}$ denotes the Sharpe ratio of the risky asset.

Plugging the consumption-wealth ratio into the HJB equation and using Equation (B.1), we obtain

$$0 = \frac{\rho}{1-\psi^{-1}}\left[\rho^{-1}\xi_j - 1\right] + r + \eta\sigma_j - \xi_j + \frac{1}{1-\psi}\left[\frac{\xi_{j,X}}{\xi_j}\mu_X + \frac{1}{2}\sum_{k=1}^{d}\sigma'_{X,k}\frac{\xi_{j,XX}}{\xi_j}\sigma_{X,k}\right]$$
$$- \frac{\gamma_j}{2}\sigma_j^2 + \frac{1-\gamma_j}{1-\psi}\frac{\xi_{j,X}}{\xi_j}\sigma_X\frac{\sigma'_R}{\|\sigma_R\|}\sigma_j + \frac{1}{2}\frac{\psi-\gamma_j}{(1-\psi)^2}\sum_{k=1}^{d}\sigma'_{X,k}\frac{\xi'_{j,X}}{\xi_j}\frac{\xi_{j,X}}{\xi_j}\sigma_{X,k}.$$

Rearranging the expression above, we obtain

$$\xi_{j,t} = \psi\rho + (1-\psi)\left[r_t + \eta_t\sigma_{j,t} - \frac{\gamma_j}{2}\sigma_{j,t}^2\right] + \mu_{\xi_j,t} + (1-\gamma_j)\sigma_{\xi_j,t}\frac{\sigma'_{R,t}}{\|\sigma_{R,t}\|}\sigma_{j,t} + \frac{\psi-\gamma_j}{1-\psi}\frac{\|\sigma_{\xi_j,t}\|^2}{2},$$

where the law of motion of $\xi_{j,t}$ is given by

$$\frac{d\xi_{j,t}}{\xi_{j,t}} = \mu_{\xi_j,t}dt + \sigma_{\xi_j,t}dZ_t,$$

and the drift and diffusion of $\xi_{j,t}$ are given by Ito's lemma:

$$\mu_{\xi_j,t} = \frac{\xi_{j,X}}{\xi_j}\mu_{X,t} + \frac{1}{2}\sum_{k=1}^{d}\sigma'_{X,k,t}\frac{\xi_{j,XX,t}}{\xi_{j,t}}\sigma_{X,k,t}, \qquad \sigma_{\xi_j,t} = \frac{\xi_{j,X}}{\xi_j}\sigma_{X,t}.$$

B.2 Pricing condition

Let $y_t \equiv Y_t/P_t$ denote the dividend yield on the risky asset. From Equation (9), we can write the expected return on the risky asset as:

$$r_t + \eta_t\|\sigma_{R,t}\| = y_t + \frac{1}{dt}\frac{d(Y_t/y_t)}{(Y_t/y_t)} = y_t + \mu - \mu_{y,t} + \|\sigma_{y,t}\|^2 - \sigma\sigma'_{y,t}.$$

Rearranging the expression above, we obtain

$$y_t = r_t + \eta_t\|\sigma_{R,t}\| - \mu + \mu_{y,t} - \|\sigma_{R,t}\|^2 + \sigma\sigma'_{R,t}, \tag{B.2}$$

where $\sigma_{R,t} = \sigma - \sigma_{y,t}$ and $(\mu_{y,t}, \sigma_{y,t})$ are given by Ito's lemma:

$$\mu_{y,t} = y_{X,t}\mu_{X,t} + \frac{1}{2}\sum_{k=1}^{d} \sigma'_{X,k,t}\, y_{XX,t}\, \sigma_{X,k,t}, \qquad \sigma_{y,t} = y_{X,t}\sigma_{X,t}.$$

B.3 Aggregate state variable

Define the share of wealth of type-j investors as follows

$$x_{j,t} \equiv \frac{\omega_j W_{j,t}}{P_t}.$$

We define the aggregate state variable as $X_t = (x_t, \overline{\alpha}_{p,t})$, where $x_t \equiv (x_{1,t}, x_{2,t}, \ldots, x_{J,t})$. The law of motion of $\overline{\alpha}_{p,t}$ is given by (10). To compute the law of motion of $x_{j,t}$, first, note that the law of motion of wealth for a type-j investor can be written as

$$\frac{dW_{j,t}}{W_{j,t}} = \left[r_t + \eta_t\sigma_{j,t} - \xi_{j,t}\right] dt + \sigma_{j,t}\frac{\sigma_{R,t}}{\|\sigma_{R,t}\|}dZ_t$$

From Ito's lemma, the law of motion of $x_{j,t}$ is given by

$$\begin{aligned}\frac{dx_{j,t}}{x_{j,t}} &= \left(r_t + \eta_t\sigma_{j,t} - \xi_{j,t} - \mu + \mu_{y,t} + \sigma\sigma'_{R,t} - \sigma_{j,t}\|\sigma_{R,t}\| + \kappa\frac{\omega_j - x_{j,t}}{x_{j,t}}\right) dt \\ &\quad + (\sigma_{j,t} - \|\sigma_{R,t}\|)\frac{\sigma_{R,t}}{\|\sigma_{R,t}\|}dZ_t,\end{aligned}$$

using $\mu_{P,t} = \mu - \mu_{y,t} + \|\sigma_{R,t}\|^2 - \sigma\sigma'_{R,t}$.

B.4 Asset prices

Let $\mathcal{J}_t^u \subset \{1, 2, \ldots, J\}$ denote the set of unconstrained active investors at period t, that is, the set of investors such that $\sigma_{j,t} < \overline{\sigma}$. Let $\mathcal{J}_t^c \subset \{1, 2, \ldots, J\}$ denote the set of constrained active investors, that is, the set of investors such that $\sigma_{j,t} = \overline{\sigma}$. From the market clearing condition for the risky asset in Equation (16), we obtain

$$\eta_t = \frac{\gamma_{u,t}}{x_{u,t}}\left[(1 - \overline{\alpha}_{p,t}x_{0,t})\|\sigma_{R_t}\| - \overline{\sigma}x_{c,t} + \sum_{j\in\mathcal{J}_t^u} x_{j,t}\frac{1-\gamma_j^{-1}}{1-\psi}\frac{\sigma_{\xi_j,t}\sigma'_{R,t}}{\|\sigma_{R,t}\|}\right],$$

where $x_{u,t} \equiv \sum_{j\in\mathcal{J}_t^u} x_{j,t}$, $x_{c,t} \equiv \sum_{j\in\mathcal{J}_t^c} x_{j,t}$, and $\gamma_{u,t} \equiv \left[\frac{1}{x_{u,t}}\sum_{j\in\mathcal{J}_t^u}\frac{x_{j,t}}{\gamma_j}\right]^{-1}$ is the aggregate risk aversion of the unconstrained investors.

From the market clearing condition for goods, the first term in Equation (16), we obtain

$$
\begin{aligned}
y_t = \psi\rho + (1-\psi)&\left[r_t + \eta_t \|\sigma_{R,t}\| - \sum_{j=0}^{J} x_{j,t}\frac{\gamma_j}{2}\sigma_{j,t}^2\right] \\
&+ \sum_{j=0}^{J} x_{j,t}\left[\mu_{\xi_j,t} + (1-\gamma_j)\sigma_{\xi_j,t}\frac{\sigma'_{R,t}}{\|\sigma_{R,t}\|}\sigma_{j,t} + \frac{\psi-\gamma_j}{1-\psi}\frac{\|\sigma_{\xi_j,t}\|^2}{2}\right].
\end{aligned}
$$

Using the pricing condition (B.2), we obtain the expression for the risk-free rate

$$
\begin{aligned}
r_t = \rho - \eta_t\|\sigma_{R,t}\| + \psi^{-1}(\mu - \mu_{y,t} + \|\sigma_{R,t}\|^2 - \sigma\sigma'_{R,t}) + \left(1-\psi^{-1}\right)\sum_{j=0}^{J} x_{j,t}\frac{\gamma_j}{2}\sigma_{j,t}^2 \\
+ \psi^{-1}\sum_{j=0}^{J} x_{j,t}\left[\mu_{\xi_j,t} + (1-\gamma_j)\sigma_{\xi_j,t}\frac{\sigma'_{R,t}}{\|\sigma_{R,t}\|}\sigma_{j,t} + \frac{\psi-\gamma_j}{1-\psi}\frac{\|\sigma_{\xi_j,t}\|^2}{2}\right].
\end{aligned}
$$

B.5 The system of PDEs

To compute the equilibrium, one needs to solve a system of $J+2$ partial differential equations (PDEs), involving the consumption-wealth ratio $\xi_j(X)$ for the $J+1$ type of investors and the dividend yield $y(X)$. These functions depend on $J+1$ state variables, the $J-$dimensional vector x_t and the portfolio share of passive investors $\overline{\alpha}_{p,t}$.

The differential equation for the consumption-wealth ratio is given by

$$
\begin{aligned}
\xi_{j,t} = \psi\rho + (1-\psi)\left[r_t + \eta_t\sigma_{j,t} - \frac{\gamma_j}{2}\sigma_{j,t}^2\right] + \frac{\xi_{j,X}}{\xi_j}\mu_{X,t} + \frac{1}{2}\sum_{k=1}^{d}\sigma'_{X,k,t}\frac{\xi_{j,XX,t}}{\xi_{j,t}}\sigma_{X,k,t} \\
+ (1-\gamma_j)\frac{\xi_{j,X}}{\xi_j}\sigma_{X,t}\frac{\sigma'_{R,t}}{\|\sigma_{R,t}\|}\sigma_{j,t} + \frac{\psi-\gamma_j}{1-\psi}\frac{1}{2}\left\|\frac{\xi_{j,X}}{\xi_j}\sigma_{X,t}\right\|^2.
\end{aligned}
$$

Plugging the expressions for interest rate and the Sharpe ratio (r_t, η_t), the risk exposure $\sigma_{j,t}$, the drift and diffusion of the aggregate state variables $(\mu_{X,t}, \sigma_{X,t})$, and the aggregate volatility $\|\sigma_{R,t}\|$, we can express the condition above in terms of $\xi_{j,t}$ and y_t and their derivatives.

Similarly, we can write the condition for the dividend yield:

$$
\begin{aligned}
y_t = r_t + \eta_t\left\|\sigma - \frac{y_{X,t}}{y_t}\sigma_{X,t}\right\| - \mu + \frac{y_{X,t}}{y_t}\mu_{X,t} + \frac{1}{2}\sum_{k=1}^{d}\sigma'_{X,k,t}\frac{y_{XX,t}}{y_t}\sigma_{X,k,t} \\
- \left\|\sigma - \frac{y_{X,t}}{y_t}\sigma_{X,t}\right\|^2 + \sigma\left(\sigma - \frac{y_{X,t}}{y_t}\sigma_{X,t}\right)',
\end{aligned}
$$

which again can be expressed only in terms of $\xi_{j,t}$ and y_t and their derivatives.

C Derivations for Section 4

C.1 Proof of Lemma 1

Proof. The assumption $\epsilon = 0$ implies that there is no preference heterogeneity and passive investors are fully invested in the risky asset. We guess and verify that in this benchmark economy, there are no variation in expected returns. In particular, the wealth distribution plays no role in the economy. This implies that $\mu_{c_j,0}(X) = \sigma_{c_j,0}(X) = \mu_{p,0}(X) = \sigma_{p,0}(X) = 0$. In this case, the risk premium is given by

$$\pi_0(X) = \frac{\gamma}{x_{u,t}}\left[1 - x_{0,t} - x_{c,t}\overline{\alpha}_{c,t}\right] \|\sigma_{R_t}\|^2,$$

using the fact that $\varsigma_t = 0$, as $\sigma_{c_j,t} = 0$, and $\overline{\alpha}_{p,t} = 1$.

Given that $\sigma_{y,t} = 0$, we have that $\sigma_{R,0}(X) = \sigma$. Using the fact that $\overline{\sigma} = \|\sigma\|$ and $x_{u,t} = 1 - x_{0,t} - x_{c,t}$, we obtain the risk premium

$$\pi_0(X) = \gamma\|\sigma\|^2,$$

using $\alpha_{c,t} = 1$. Using $\sigma_{c_j,t} = 0$ and the expression for $\pi_0(X)$, we obtain that $\alpha_{j,0}(X) = 1$, for $j = 1, \ldots, J$, from Equation (17).

The interest rate is given by

$$r_0(X) = \rho + \psi^{-1}\mu - \gamma(1 + \psi^{-1})\frac{\|\sigma\|^2}{2}.$$

The consumption-wealth ratio $c_{j,0}(X)$ is given by

$$c_{j,0}(X) = \psi\rho + (1 - \psi)\left[r_0(X) + \pi_0(X)\alpha_{j,0}(X) - \frac{\gamma}{2}\alpha_{j,0}(X)^2\|\sigma\|^2\right].$$

Plugging in the expression for $r_0(X)$, $\eta_0(X)$ and $\sigma_{j,0}$, we obtain

$$c_{j,0}(X) = \rho - \left(1 - \psi^{-1}\right)\left(\mu - \frac{\gamma\|\sigma\|^2}{2}\right),$$

where we assume $\rho > \left(1 - \psi^{-1}\right)\left(\mu - \frac{\gamma\|\sigma\|^2}{2}\right)$.

From the market clearing condition for goods, we obtain:

$$\frac{1}{p_0(X)} = \rho - \left(1 - \psi^{-1}\right)\left(\mu - \frac{\gamma\|\sigma\|^2}{2}\right).$$

The drift and diffusion of the wealth shares are given

$$\mu_{X,j,0}(X) = x_j \left[r_0(X) + \pi_0(X)\alpha_{j,0}(X) - c_{j,0}(X) - \mu\right]$$
$$\sigma_{X,j,0}(X) = x_j(\alpha_{j,0}(X) - 1)\sigma_{R,0}(X),$$

where $\mu_{X,j,0} = \sigma_{X,j,0} = 0$, using the expression for returns, portfolio share, and consumption-wealth ratio. The result $\mu_{X,j,0} = 0$ uses the fact that $\kappa = 0$. □

C.2 Proof of Proposition 3

Proof. We consider next the first-order correction terms. Note that the diffusion terms for c_j and p are both equal to zero up to the first order, since $\sigma_{c_j,t} = O(\epsilon^2)$ and $\sigma_{p,t} = O(\epsilon^2)$. From Ito's lemma:

$$\sigma_{c_j,t} = \underbrace{\frac{c_{j,X}}{c_j}}_{O(\epsilon)} \underbrace{\sigma_{X,t}}_{O(\epsilon)} = O(\epsilon^2).$$

We have $c_{j,X} = O(\epsilon)$, because $c_{j,X,0} = 0$, as $c_{j,0}(X)$ does not depend on X. Also, $\sigma_{X,t} = O(\epsilon)$ because $\sigma_{X,0}(X) = 0$. This implies that $\sigma_{c_j,1}(X) = 0$. An analogous argument applies to p_t, so that we have $\sigma_{p,1}(X) = 0$ and $\|\sigma_{R,1}(X)\| = 0$.

Risk exposure of active investors. The risk exposure for active investors can be written as

$$\sigma_j(X,\epsilon) = \min\left\{\frac{\eta(X,\epsilon)}{\gamma_j} + \frac{1-\gamma_j^{-1}}{\psi - 1}\frac{\sigma_{c_j}(X,\epsilon)\sigma_R'(X,\epsilon)}{\|\sigma_R(X,\epsilon)\|}, \|\sigma\| + \hat{\sigma}\epsilon\right\},$$

where $\sigma_j(X;\epsilon) \equiv \alpha_j(X;\epsilon)\|\sigma_R(X;\epsilon)\|$ $\eta(X;\epsilon) \equiv \frac{\pi(X;\epsilon)}{\|\sigma_R(X;\epsilon)\|}$

Expanding the first term inside brackets in ϵ, we obtain

$$\sigma_j(X,\epsilon) = \min\left\{\frac{\eta_0(X)}{\gamma} + \left(\frac{\eta_1(X)}{\gamma} - \frac{\eta_0(X)}{\gamma}\hat{\gamma}_j\right)\epsilon + \frac{1-\gamma^{-1}}{\psi-1}\frac{\sigma_{c_j,2}(X)\sigma'}{\|\sigma\|}\epsilon^2 + O(\epsilon^3), \|\sigma\| + \hat{\sigma}\epsilon\right\},$$

Adding and subtracting $\|\sigma\| + \hat{\sigma}\epsilon$, and using $\frac{\eta_0(X)}{\gamma} = \|\sigma\|$, we obtain

$$\sigma_j(X,\epsilon) = \min\left\{\left(\frac{\eta_1(X)}{\gamma} - \frac{\eta_0(X)}{\gamma}\hat{\gamma}_j - \hat{\sigma}\right)\epsilon + \frac{1-\gamma^{-1}}{\psi-1}\frac{\sigma_{c_j,2}(X)\sigma'}{\|\sigma\|}\epsilon^2 + O(\epsilon^3), 0\right\} + \|\sigma\| + \hat{\sigma}\epsilon. \tag{C.1}$$

Consider first the case where the following condition is satisfied:

$$\frac{\eta_1(X)}{\gamma} - \frac{\eta_0(X)}{\gamma}\hat{\gamma}_j - \hat{\sigma} = O(1), \tag{C.2}$$

If this is the case, then

$$\left(\frac{\eta_1(X)}{\gamma} - \frac{\eta_0(X)}{\gamma}\hat{\gamma}_j - \hat{\sigma}\right)\epsilon \gg \left|\frac{1-\gamma^{-1}}{\psi-1}\frac{\sigma_{c_j,2}(X)\sigma'}{\|\sigma\|}\right|\epsilon^2,$$

for small ϵ. So, the sign of the term inside the min operator in (C.1) is determined by the first term.

We can then write $\sigma_j(X,\epsilon)$ as follows:

$$\sigma_j(X,\epsilon) = \|\sigma\| + \min\left\{\frac{\eta_1(X)}{\gamma} - \frac{\eta_0(X)}{\gamma}\hat{\gamma}, \hat{\sigma}\right\}\epsilon + O(\epsilon^2).$$

In the region of the state space where condition (C.2) holds, one can determine whether an investor is constrained or unconstrained only based on the first-order terms. Suppose now that the following condition holds

$$\frac{\eta_1(X)}{\gamma} - \frac{\eta_0(X)}{\gamma}\hat{\gamma}_j - \hat{\sigma} = O(\epsilon). \tag{C.3}$$

This condition states that, up to the first order, the leverage constraint is either always binding or slack by just a tiny amount parameterized by ϵ and ϵ^2 terms inside the min operator in (C.1). In this case, we can write $\sigma_j(X,\epsilon)$ as follows:

$$\sigma_j(X,\epsilon) = \|\sigma\| + \hat{\sigma}\epsilon + \min\left\{\left(\frac{\eta_1(X)}{\gamma} - \frac{\eta_0(X)}{\gamma}\hat{\gamma}_j - \hat{\sigma}\right)\epsilon + \frac{1-\gamma^{-1}}{\psi-1}\frac{\sigma_{c_j,2}(X)\sigma'}{\|\sigma\|}\epsilon^2, 0\right\} + O(\epsilon^3).$$

In this region of the state space where condition (C.3) is satisfied, we need the second-order term to determine whether an investor is constrained. This distinction will be relevant when computing the second-order correction.

For the first-order correction terms here, we focus on the case where condition (C.2) holds.

Aggregate risk aversion. The aggregate risk aversion of unconstrained investors, defined above, is given by

$$\gamma_u(X,\epsilon) = \frac{x_u}{\sum_{j\in\mathcal{J}^u}\frac{x_j}{\gamma(1+\hat{\gamma}\epsilon)}} = \gamma + \gamma\sum_{j\in\mathcal{J}^u}\frac{x_j}{x_u}\hat{\gamma}_j\,\epsilon + O(\epsilon^2)$$

Market price of risk. The market price of risk can be written as

$$\begin{aligned}\eta(X,\epsilon) &= \frac{\gamma_u(X,\epsilon)}{1-x_0-x_c}\left[(1-(1+\hat{\alpha}_p\epsilon)x_0)\|\sigma\| - (\|\sigma\|+\hat{\sigma}\epsilon)x_c\right] + O(\epsilon^2),\\ &= \eta_0(X) + \gamma\|\sigma\|\left[\sum_{j\in\mathcal{J}^u}\frac{x_j}{x_u}\hat{\gamma}_j - \left(\frac{\hat{\alpha}_p x_0 + \frac{\hat{\sigma}}{\|\sigma\|}x_c}{1-x_0-x_c}\right)\right]\epsilon + O(\epsilon^2).\end{aligned}$$

In the region of the state space where all active investors are unconstrained, we have

$$\eta(X,\epsilon) = \eta_0(X) + \gamma\|\sigma\|\left(\sum_{j\in\mathcal{J}^u}\frac{x_j}{x_u}\hat{\gamma}_j - \frac{\hat{\alpha}_p\, x_0}{1-x_0}\right)\epsilon + O(\epsilon^2).$$

The expression above shows the impact of fluctuations in the aggregate risk aversion and the effect of portfolio inflows in the market price of risk. If the average risk aversion in state X is lower than its level at $\epsilon = 0$, then the market price of risk will be lower than its level at $\epsilon = 0$, everything else constant.

Interest rate. The interest rate is given by

$$\begin{aligned} r(X,\epsilon) &= r_0(X) + \left[-\eta_1(X)\|\sigma\| + \left(1-\psi^{-1}\right)\left(\sum_{j=0}^{J} x_{j,t}\frac{\gamma}{2}2\|\sigma\|\sigma_{j,1}(X) + \sum_{j=0}^{J} x_{j,t}\frac{\gamma}{2}\|\sigma\|^2\hat{\gamma}_j\right)\right]\epsilon + O(\epsilon^2) \\ &= r_0(X) + \gamma\|\sigma\|^2\left[-\frac{\eta_1(X)}{\gamma\|\sigma\|} + \left(1-\psi^{-1}\right)\sum_{j=0}^{J} x_{j,t}\frac{\hat{\gamma}_j}{2}\right]\epsilon + O(\epsilon^2) \\ &= r_0(X) - \gamma\|\sigma\|^2\left[\left(\sum_{j\in\mathcal{J}^u}\frac{x_j}{x_u}\hat{\gamma}_j - \sum_{j=0}^{J} x_{j,t}\frac{\hat{\gamma}_j}{2}\right) - \frac{\hat{\alpha}_p x_0 + \frac{\hat{\sigma}}{\|\sigma\|}x_c}{1-x_0-x_c} + \psi^{-1}\sum_{j=0}^{J} x_{j,t}\frac{\hat{\gamma}_j}{2}\right]\epsilon + O(\epsilon^2), \end{aligned} \tag{C.4}$$

where we use the fact that $\sum_{j=0}^{J} x_{j,t}\sigma_{j,1}(X) = 0$ from the first-order correction of the risky asset market clearing condition.

The term in the square brackets in Equation (C.4) captures the first-order effect of frictions on the interest rate, $r_1(X)$. First, we see that the interest rate is decreasing in the difference between the average risk aversion of unconstrained investors and the average risk aversion of all investors in the economy. These two averages can differ for two reasons. First, passive investors may have a risk aversion different from the average unconstrained investor, the first term in parentheses in $r_1(X)$. Second, constrained investors are exactly the ones with low risk aversion, so unconstrained investors are on average more risk averse than all investors in the economy, which include the low risk aversion ones.

Consumption-wealth ratio. The consumption-wealth ratio for investor j is given by

$$\begin{aligned}
c_j(X,\epsilon) &= c_{j,0}(X) + (1-\psi)\left[r_1(X) + \eta_0(X)\sigma_{j,1}(X) + \eta_1(X)\sigma_{j,0}(X)\right.\\
&\qquad\qquad \left. -\gamma\|\sigma\|\sigma_{j,1}(X) - \frac{\gamma}{2}\hat{\gamma}_j\|\sigma\|^2\right]\epsilon + O(\epsilon^2)\\
&= c_{j,0}(X) + (1-\psi)\gamma\|\sigma\|^2\left[\frac{r_1(X)}{\gamma\|\sigma\|^2} + \frac{\eta_1(X)}{\gamma\|\sigma\|} - \frac{\hat{\gamma}_j}{2}\right]\epsilon + O(\epsilon^2)\\
&= c_{j,0}(X) + (1-\psi)\gamma\|\sigma\|^2\left[\left(1-\psi^{-1}\right)\sum_{k=0}^{J} x_{k,t}\frac{\hat{\gamma}_k}{2} - \frac{\hat{\gamma}_j}{2}\right]\epsilon + O(\epsilon^2).
\end{aligned}$$

Dividend yield. The dividend yield is given by

$$y(X,\epsilon) = y_0(X) + \left(1-\psi^{-1}\right)\gamma\|\sigma\|^2\sum_{j=0}^{J} x_{j,t}\frac{\hat{\gamma}_j}{2}\epsilon + O(\epsilon^2).$$

Notice that portfolio flows do not affect the dividend yield up to first order. The reason is that the interest rate and risk premium effects exactly cancel each other out. To derive the effect of portfolio flows on asset prices, we need to consider the second-order correction.

The price-dividend ratio is then given by

$$p(X,\epsilon) = p_0(X) - p_0(X)^2\left(1-\psi^{-1}\right)\gamma\|\sigma\|^2\sum_{j=0}^{J} x_{j,t}\frac{\hat{\gamma}_j}{2}\epsilon + O(\epsilon^2).$$

State dynamics. The diffusion of the wealth share of investor j, $j = 1,\dots,J$, is given by

$$\begin{aligned}
\sigma_{X,j}(X) &= x_j\sigma_{j,1}(X)\frac{\sigma}{\|\sigma\|}\epsilon + O(\epsilon^2)\\
&= x_j\min\left\{\sum_{k\in\mathcal{J}^u}\frac{x_k}{x_u}\hat{\gamma}_j - \left(\frac{\hat{\alpha}_p x_0 + \frac{\hat{\sigma}}{\|\sigma\|}x_c}{1-x_0-x_c}\right) - \hat{\gamma}_j, \frac{\hat{\sigma}}{\|\sigma\|}\right\}\sigma\epsilon + O(\epsilon^2).
\end{aligned}$$

The drift of the wealth share of investor j, $j = 1,\dots,J$, is given by

$$\begin{aligned}
\mu_{X,j}(X,\epsilon) &= x_j\left[r_1(X) + \eta_1(X)\sigma_{j,0}(X) + \eta_0(X)\sigma_{j,1}(X) - \xi_{j,1}(X) - \sigma_{j,1}(X)\|\sigma\|\right]\epsilon + \kappa(\omega_j - x_j) + O(\epsilon^2)\\
&= x_j\left[(\psi-1)\gamma\|\sigma\|^2\left(\sum_{k=0}^{J} x_{k,t}\frac{\hat{\gamma}_k}{2} - \frac{\hat{\gamma}_j}{2}\right) + (\gamma-1)\|\sigma\|\sigma_{j,1}(X)\right]\epsilon + \kappa(\omega_j - x_j) + O(\epsilon^2)
\end{aligned}$$

The law of motion of $\hat{\alpha}_p$ can be written as follows:

$$d\hat{\alpha}_p = \theta_p(\overline{\hat{\alpha}} - \hat{\alpha}_p)dt + \sigma_p dZ_t,$$

so the drift of $\overline{\alpha}_p$ is $\mu_{\overline{\alpha}_p} = \theta_p(\overline{\hat{\alpha}} - \hat{\alpha}_p)\epsilon$ and the diffusion is $\sigma_{\overline{\alpha}_p} = \sigma_p \epsilon$.

$\square$

C.3 Second-order correction

In Proposition 6, we compute the second-order correction for our economy.

Proposition 6 (Second-order correction)**.** *Suppose* $\rho > \left(1 - \psi^{-1}\right)\left(\mu - \frac{\gamma\sigma^2}{2}\right)$. *Then,*

(i) The second-order correction for the consumption-wealth ratio and risk exposure are given by:

$$\begin{aligned} c_{j,2}(X) &= (1-\psi)\frac{\gamma\sigma^2}{2}\left[\left(1-\psi^{-1}\right)\sum_{k=0}^{J} x_{k,t}\hat{\gamma}_k - \hat{\gamma}_j\right] \\ \sigma_{j,2}(X) &= \frac{\eta_2(X)}{\gamma} - \frac{\eta_1(X)}{\gamma}\hat{\gamma}_j + \frac{\eta_0(X)}{\gamma}\hat{\gamma}_j^2 - (1-\gamma^{-1})\sigma_{y,2}(X), \end{aligned}$$

where

$$\sigma_{y,2}(X) = \left(1-\psi^{-1}\right)\frac{\gamma\sigma^2}{2y_0(X)}\sum_{k=1}^{J}\left(\hat{\gamma}_k - \hat{\gamma}_0\right)x_k\sigma_{k,1}(X)\frac{\sigma}{\|\sigma\|}.$$

(ii) The second-order correction for the Sharpe ratio, interest rate, and dividend yield are given by:

$$\begin{aligned} \eta_2(X) &= -\frac{(\gamma-1)x_c + 1 - x_0}{1 - x_0 - x_c}\sigma_{y,2}(X) - \gamma\sigma\mathbb{E}^u[\hat{\gamma}_j]\frac{\hat{\alpha}_p x_0 + \frac{\hat{\sigma}}{\sigma}x_c}{1 - x_0 - x_c} - \gamma\sigma\mathrm{Var}^u[\hat{\gamma}_j] \\ r_2(X) &= -\eta_2(X)\sigma + \eta_0(X)\sigma_{y,2}(X) - \psi^{-1}(\mu_{y,2}(X) + \sigma\sigma_{y,2}(X)) \\ &\quad + \left(1-\psi^{-1}\right)\sum_{j=0}^{J} x_j\gamma\sigma\left[\sigma_{j,2}(X) + \frac{\sigma_{j,1}^2(X)}{2\sigma} + \hat{\gamma}_j\sigma_{j,1}\right] \\ &\quad + \psi^{-1}\sum_{j=0}^{J} x_j\left[\mu_{\xi_j,2}(X) + (1-\gamma)\sigma_{\xi_j,2}(X)\sigma\right] \\ y_2(X) &= \left(1-\psi^{-1}\right)\sum_{j=0}^{J} x_j\gamma\sigma\left(\frac{\sigma_{j,1}^2(X)}{2\sigma} + \hat{\gamma}_j\sigma_{j,1}(X)\right), \end{aligned} \tag{C.5}$$

where

$$\mathbb{E}^u[\hat{\gamma}_j] \equiv \sum_{j\in\mathcal{J}^u}\frac{x_j}{x_u}\hat{\gamma}_j, \quad \textit{and} \quad \mathrm{Var}^u[\hat{\gamma}_j] \equiv \sum_{j\in\mathcal{J}^u}\frac{x_j}{x_u}\hat{\gamma}_j^2 - \left(\sum_{j\in\mathcal{J}^u}\frac{x_j}{x_u}\hat{\gamma}_j\right)^2.$$

Proof. **Step 1: Laws of motion for c_j and y.**
We start by considering the diffusion terms for c_j and y. The expansion of $\sigma_{c_j,t}$ in ϵ is given by

$$\begin{aligned}\sigma_{c_j}(X,\epsilon) &= \frac{c_{j,X}(X,\epsilon)}{c_j(X,\epsilon)}\sigma_X(X,\epsilon)\\ &= \frac{c_{j,X,1}(X)}{c_{j,0}(X)}\sigma_{X,1}(X)\epsilon^2 + O(\epsilon^3)\\ &= -(\psi-1)\left(1-\psi^{-1}\right)\frac{\gamma\sigma^2}{2c_{j,0}(X)}\sum_{k=1}^{J}\left(\hat{\gamma}_k-\hat{\gamma}_0\right)x_k\sigma_{k,1}(X)\epsilon^2 + O(\epsilon^3),\end{aligned}$$

where we used the fact that $c_{j,\overline{\alpha}_p,1}(X) = 0$. Notice that $\sigma_{c_j,2}(X)$ does not depend on j, that is, it is the same for all investors. Moreover, $\sigma_{c_j,2}(X) > 0$, as $\sigma_{k,1}(X)$ is inversely related to $\hat{\gamma}_k$.

Similarly, the diffusion for dividend yield y can be written as

$$\begin{aligned}\sigma_y(X,\epsilon) &= \frac{y_{X,1}(X)}{y_0(X)}\sigma_{X,1}(X)\epsilon^2 + O(\epsilon^2)\\ &= \left(1-\psi^{-1}\right)\frac{\gamma\sigma^2}{2y_0(X)}\sum_{k=1}^{J}\left(\hat{\gamma}_k-\hat{\gamma}_0\right)x_k\sigma_{k,1}(X)\epsilon^2 + O(\epsilon^3).\end{aligned}$$

where we used the fact that $y_{\overline{\alpha}_p,1}(X) = 0$. The expression above is negative if $\psi > 1$, which implies that $\sigma_{R,2}(X) = -\sigma_{y,2}(X)$ is positive. In this case, a negative aggregate shock redistribute wealth to more risk averse investors, leading to a rise in the risk premium and a decline in the risk-free rate. If $\psi > 1$, the risk premium effect dominates, so the price-dividend ratio, $1/y$, falls in response to the shock. The movement in the price-dividend ratio amplifies the initial effect of the drop in dividends.

The drift of $c_{j,t}$ is given by

$$\begin{aligned}\mu_{c_j}(X,\epsilon) &= \frac{c_{j,X}(X,\epsilon)}{c_j(X,\epsilon)}\mu_X(X,\epsilon) + \frac{1}{2}\sigma_X'(Z,\epsilon)\frac{c_{j,XX}(X,\epsilon)}{c_j(X,\epsilon)}\sigma_X(X,\epsilon)\\ &= \frac{c_{j,X,1}(X)}{c_{j,0}(X)}\mu_{X,1}(X)\epsilon^2 + O(\epsilon^3)\\ &= -(\psi-1)\left(1-\psi^{-1}\right)\frac{\gamma\sigma^2}{2c_{j,0}(X)}\sum_{k=1}^{J}\left(\hat{\gamma}_k-\hat{\gamma}_0\right)\mu_{X_k,1}(X)\epsilon^2 + O(\epsilon^3).\end{aligned}$$

The drift for y_t is given by

$$\begin{aligned}
\mu_y(X,\epsilon) &= \frac{y_X(X,\epsilon)}{y(X,\epsilon)}\mu_X(X,\epsilon) + \frac{1}{2}\sigma_X'(Z,\epsilon)\frac{y_{XX}(X,\epsilon)}{y(X,\epsilon)}\sigma_X(X,\epsilon) \\
&= \frac{y_{X,1}(X)}{y_0(X)}\mu_{X,1}(X)\epsilon^2 + O(\epsilon^3) \\
&= \left(1-\psi^{-1}\right)\frac{\gamma\sigma^2}{2y_0(X)}\sum_{k=1}^{J}\left(\hat{\gamma}_k - \hat{\gamma}_0\right)\mu_{X_k,1}(X)\epsilon^2 + O(\epsilon^3),
\end{aligned}$$

where we used the fact that $c_{j,\overline{\alpha}_p,1} = y_{\overline{\alpha}_p,1} = 0$.

Step 2: Risk exposures of investors.

We focus on the *inner region*, that is, the case where all investors are sufficiently far from the constraint boundary (on either side). For a constrained investor, the second-order term is zero. For an unconstrained investor, the second-order term is given by

$$\sigma_{j,2}(X) = \frac{\eta_2(X)}{\gamma} - \frac{\eta_1(X)}{\gamma}\hat{\gamma}_j + \frac{\eta_0(X)}{\gamma}\hat{\gamma}_j^2 + \frac{1-\gamma^{-1}}{\psi-1}\frac{\sigma_{c_j,2}(X)\sigma'}{\|\sigma\|},$$

where investor j is unconstrained if the following condition holds:

$$\sum_{k\in\mathcal{J}^u}\frac{x_k}{x_u}\hat{\gamma}_k - \left(\frac{\hat{\alpha}_p x_0 + \hat{\sigma}x_c}{1-x_0-x_c}\right) - \hat{\gamma}_j < \hat{\sigma}.$$

Step 3: Aggregate risk aversion.

The aggregate risk aversion can be written as

$$\gamma_u(X,\epsilon) = \gamma\left[1 + \mathbb{E}^u\left[\hat{\gamma}_j\right]\epsilon - \mathrm{Var}^u\left[\hat{\gamma}_j\right]\epsilon^2\right] + O(\epsilon^3),$$

where

$$\mathbb{E}^u\left[\hat{\gamma}_j\right] \equiv \sum_{j\in\mathcal{J}^u}\frac{x_j}{x_u}\hat{\gamma}_j, \quad \text{and} \quad \mathrm{Var}^u[\hat{\gamma}_j] \equiv \sum_{j\in\mathcal{J}^u}\frac{x_j}{x_u}\hat{\gamma}_j^2 - \left(\sum_{j\in\mathcal{J}^u}\frac{x_j}{x_u}\hat{\gamma}_j\right)^2.$$

Step 4: Market price of risk.

The market price of risk is given by

$$\eta(X,\epsilon) = \frac{\gamma_u(X,\epsilon)}{x_u}\left[(1-(1+\hat{\alpha}_p\epsilon)x_{0,t})\|\sigma - \sigma_y(X,\epsilon)\| - \|\sigma\|(1+\hat{\sigma}\epsilon)x_{c,t} + \sum_{j\in\mathcal{J}_t^u} x_{j,t}\frac{1-\gamma_j^{-1}}{1-\psi}\sigma_{c_j}(X,\epsilon)\frac{\sigma'}{\|\sigma\|}\right],$$

The second-order term is then given by

$$\eta_2(X) = \frac{\gamma_{u,0}(X)}{x_u}\left[-(1-x_{0,t})\sigma_{y,2}(X)\frac{\sigma'}{\|\sigma\|} + \sum_{j\in\mathcal{J}_t^u} x_{j,t}\frac{1-\gamma^{-1}}{1-\psi}\sigma_{c_j,2}(X)\frac{\sigma'}{\|\sigma\|}\right] + \\ + \frac{\gamma_{u,1}(X)}{x_u}\|\sigma\|\left[-\hat{\alpha}_p x_0 - \hat{\sigma}x_c\right] + \frac{\gamma_{u,2}(X)}{x_u}\|\sigma\|\left[1-x_0-x_c\right].$$

The expression above can be written as

$$\eta_2(X) = -\frac{(\gamma-1)x_c + 1 - x_0}{1-x_0-x_c}\sigma_{y,2}(X)\frac{\sigma'}{\|\sigma\|} - \gamma\|\sigma\|\mathbb{E}^u[\hat{\gamma}_j]\frac{\hat{\alpha}_p x_0 + \hat{\sigma}x_c}{1-x_0-x_c} - \gamma\|\sigma\|\text{Var}^u[\hat{\gamma}_j].$$

using the fact that $\sigma_{c_j,2}(X) = (1-\psi)\sigma_{y,2}$.

Step 5: Interest rate.

The interest rate is given by

$$r(X,\epsilon) = \rho - \eta(X,\epsilon)\|\sigma_R(X,\epsilon)\| + \psi^{-1}(\mu - \mu_y(X,\epsilon) - \sigma_y\sigma_R(X,\epsilon)') \\ + \left(1-\psi^{-1}\right)\sum_{j=0}^{J} x_j\frac{\gamma_j}{2}\sigma_j^2(X,\epsilon) \\ + \psi^{-1}\sum_{j=0}^{J} x_{j,t}\left[\mu_{c_j}(X,\epsilon) + (1-\gamma_j)\sigma_{c_j}(X,\epsilon)\frac{\sigma'_{R,t}}{\|\sigma_R\|}\sigma_j(X,\epsilon) + \frac{\psi-\gamma_j}{1-\psi}\frac{\|\sigma_{c_j,t}(X,\epsilon)\|^2}{2}\right].$$

The second-order term is given by

$$r_2(X) = -\eta_2(X)\|\sigma\| + \eta_0(X)\sigma_{y,2}(X)\frac{\sigma'}{\|\sigma\|} - \psi^{-1}(\mu_{y,2}(X) + \sigma_{y,2}(X)\sigma') \\ + \left(1-\psi^{-1}\right)\sum_{j=0}^{J} x_j\gamma\|\sigma\|\left[\sigma_{j,2}(X) + \frac{\sigma_{j,1}^2(X)}{2\|\sigma\|} + \hat{\gamma}_j\sigma_{j,1}(X)\right] \\ + \psi^{-1}\sum_{j=0}^{J} x_j\left[\mu_{c_j,2}(X) + (1-\gamma)\sigma_{c_j,2}(X)\sigma'\right].$$

Step 6: Dividend yield.

The dividend yield, y, is given by

$$y(X,\epsilon) = \psi\rho + (1-\psi)\left[r(X,\epsilon) + \eta(X,\epsilon)\|\sigma_R(X,\epsilon)\| - \sum_{j=0}^{J} x_j \frac{\gamma_j}{2}\sigma_j^2(X,\epsilon)\right]$$
$$+ \sum_{j=0}^{J} x_j \left[\mu_{c_j}(X,\epsilon) + (1-\gamma_j)\sigma_{c_j}(X,\epsilon)\frac{\sigma_R(X;\epsilon)'}{\|\sigma_R(X;\epsilon)\|}\sigma_j(X,\epsilon) + \frac{\psi-\gamma_j}{1-\psi}\frac{\|\sigma_{c_j}(X,\epsilon)\|^2}{2}\right].$$

The second-order term in the expansion of $y(X,\epsilon)$ is given by

$$y_2(X) = (1-\psi)\left[r_2(X) + \eta_2(X)\|\sigma\| + \eta_0(X)\sigma_{R,2}(X)\frac{\sigma'}{\|\sigma\|} - \sum_{j=0}^{J} x_j \frac{\gamma}{2}\left(\sigma_{j,1}^2(X) + 2\sigma_{j,0}\sigma_{j,2}(X) + 2\hat{\gamma}_j\sigma_{j,0}\sigma_{j,1}(X)\right)\right]$$
$$+ \sum_{j=0}^{J} x_j \left[\mu_{c_j,2}(X) + (1-\gamma)\sigma_{c_j,2}(X)\sigma'\right].$$

Using the expression for the interest rate, we obtain

$$y_2(X) = \left(1-\psi^{-1}\right)\left[\mu_{y,2}(X) + \sigma_{y,2}(X)\sigma' + \sum_{j=0}^{J} x_j\gamma\|\sigma\|\left(\frac{\sigma_{j,1}^2(X)}{2\|\sigma\|} + \sigma_{j,2}(X) + \hat{\gamma}_j\sigma_{j,1}(X)\right)\right]$$
$$+ \psi^{-1}\sum_{j=0}^{J} x_j \left[\mu_{c_j,2}(X) + (1-\gamma)\sigma_{c_j,2}(X)\sigma'\right].$$

Given that $\mu_{c_j,2} = (1-\psi)\mu_{y,2}$ and $\sigma_{c_j,2} = (1-\psi)\sigma_{y,2}$, we obtain

$$y_2(X) = \left(1-\psi^{-1}\right)\left[\gamma\sigma_{y,2}(X)\sigma' + \sum_{j=0}^{J} x_j\gamma\|\sigma\|\left(\frac{\sigma_{j,1}^2(X)}{2\|\sigma\|} + \sigma_{j,2}(X) + \hat{\gamma}_j\sigma_{j,1}(X)\right)\right]$$
$$= \left(1-\psi^{-1}\right)\sum_{j=0}^{J} x_j\gamma\|\sigma\|\left(\frac{\sigma_{j,1}^2(X)}{2\|\sigma\|} + \hat{\gamma}_j\sigma_{j,1}(X)\right),$$

where in the second equality, we use the fact that, from the market clearing condition for the risky asset, we have $\sum_{j=0}^{J} x_j\sigma_{j,2}(X) = -\sigma_{y,2}(X)\frac{\sigma'}{\|\sigma\|}$.

Step 7: Risk premium.

Since the risk premium is given by $\pi_t = \eta_t\|\sigma_{R,t}\|$, we can write

$$\pi(X,\epsilon) = \eta_0(X)\|\sigma\| + \eta_1(X)\|\sigma\|\epsilon + \left(-\eta_0(X)\sigma_{y,2}(X)\frac{\sigma'}{\|\sigma\|} + \eta_2(X)\|\sigma\|\right)\epsilon^2 + O\left(\epsilon^3\right).$$

We can then write

$$\pi_2(X) = -\eta_0(X)\sigma_{y,2}(X)\frac{\sigma'}{\|\sigma\|} + \eta_2(X)\|\sigma\|. \tag{C.6}$$

□

D Derivation of the Market Elasticity

Let $p(X,\epsilon) \equiv 1/y(X,\epsilon)$ denote the price-dividend ratio. The second-order expansion of $p(X,\epsilon)$ is given by

$$p(X,\epsilon) = \frac{1}{y_0(X)} - \frac{y_1(X)}{y_0^2(X)}\epsilon + \left[\frac{y_1^2(X)}{y_0^3(X)} - \frac{y_2(X)}{y_0^2(X)}\right]\epsilon^2 + O(\epsilon^3). \tag{D.1}$$

Let $F(X) \equiv \frac{W_0(1+\hat{\alpha}_p\epsilon)-W_0}{P} = \hat{\alpha}_p\epsilon x_0$ denote the flow into the risky asset relative to the benchmark economy.

Proof. From Equation (D.1), the first-order impact of flows on the price-dividend ratio can be written as

$$\frac{\partial p(X,\epsilon)}{\partial F(X,\epsilon)} = -\frac{1}{y_0^2(X)}\frac{\partial y_1(X)}{\partial \hat{\alpha}_p}\frac{1}{x_0} + O(\epsilon). \tag{D.2}$$

Since from Proposition 3, $y_1(X)$ does not depend on $\hat{\alpha}_p$, the right hand side of Equation (D.2) is zero, leading to an infinite aggregate elasticity to the first-order:

$$\varepsilon_M^{-1} = \frac{1}{p(X,\epsilon)}\frac{\partial p(X,\epsilon)}{\partial F(X,\epsilon)} = 0 + O(\epsilon).$$

Given that from Equation (15), we have $y_t = r_t + \pi_t - \mu_{P,t}$, we can write the first-order term for the dividend yield as

$$y_1(X) = r_1(X) + \pi_1(X) - \mu_{P,1}(X),$$

where $\mu_{P,t} = \mu - \mu_{y,t} - \sigma_{y,t}\sigma_{R,t}$ is the drift of the risky asset price P_t, and π_t is the risk premium. As shown in Proposition 3, the first-order term for the dividend yield is constant. This means $\mu_{y,1}(X) = \sigma_{y,1}(X) = 0$, leading to $\sigma_{R,1}(X) = 0$. Therefore, we have $\mu_{P,1}(X) = 0$, and

$$\frac{\partial y_1(X)}{\partial \hat{\alpha}_p} = \frac{\partial r_1(X)}{\partial \hat{\alpha}_p} + \frac{\partial \pi_1(X)}{\partial \hat{\alpha}_p}.$$

From Proposition 3, we have

$$\frac{\partial r_1(X)}{\partial \hat{\alpha}_p} = -\sigma \frac{\partial \eta_1(X)}{\partial \hat{\alpha}_p} = \frac{\gamma\sigma^2}{x_u} x_0,$$
$$\frac{\partial \pi_1(X)}{\partial \hat{\alpha}_p} = \sigma \frac{\partial \eta_1(X)}{\partial \hat{\alpha}_p} = -\frac{\gamma\sigma^2}{x_u} x_0.$$

Thus, up to the first order, the effect of portfolio flows on the risk-free rate is the exact opposite of its impact on the risk premium and portfolio flows do not affect the price-dividend ratio up to first order. Therefore, up to the first-order, the aggregate market elasticity is infinite. □

Using Equation (D.1), the derivative of the price-dividend ratio with respect to flows F is given by

$$\frac{\partial p(X,\epsilon)}{\partial F(X,\epsilon)} = -\frac{1}{y_0^2(X)} \frac{\partial y_2(X)}{\partial \hat{\alpha}_p} \frac{\epsilon}{x_0} + O(\epsilon^2),$$

where $y_1(X)$ does not depend on $\hat{\alpha}_p$, leading to no price impact (infinite elasticity) up to the first-order.

D.1 Proof of Proposition 4

Proof. Consider the case where there is no preference heterogeneity and active investors do not face leverage constraints. In this case, $y_2(X)$ simplifies to

$$y_2(X) = \left(1 - \psi^{-1}\right) \gamma\sigma^2 \sum_{j=0}^{J} x_j \frac{\sigma_{j,1}^2(X)}{2\sigma^2},$$

using the fact that $\sigma_{y,2} = \sigma_{j,2} = 0$ when $\hat{\gamma}_j = 0$ for $j = 0, 1, \ldots, J$ when investors have the same preferences.

Using the expression for $\sigma_{j,1}$ in Proposition 3, the (inverse) aggregate market elasticity is given by

$$\frac{1}{p(X,\epsilon)} \frac{\partial p(X,\epsilon)}{\partial F(X,\epsilon)} = -\frac{1-\psi^{-1}}{2y_0(X)} \gamma\sigma^2 \left(2x_0\hat{\alpha}_p + x_a \frac{2\hat{\alpha}_p x_0^2}{x_a^2}\right) \frac{\epsilon}{x_0} + O(\epsilon^2),$$

where $x_a \equiv 1 - x_0$ denotes the wealth share of active investors. This can be written as

$$\frac{1}{p(X,\epsilon)} \frac{\partial p(X,\epsilon)}{\partial F(X,\epsilon)} = \left(1 - \psi^{-1}\right) \frac{\gamma\sigma^2}{y_0(X)} \frac{1 - \overline{\alpha}_p}{x_a} + O(\epsilon^2),$$

where we use $\overline{\alpha}_p = 1 + \hat{\alpha}_p\epsilon$ from Equation (21). □

From Equation (C.6), the impact of flows of the risk premium can be written as

$$\frac{\partial \pi(X,\epsilon)}{\partial F(X,\epsilon)} = \sigma \frac{\partial \eta_1(X)}{\partial \hat{\alpha}_p} \frac{\epsilon}{x_0} + \frac{\partial \pi_2(X)}{\partial \hat{\alpha}_p} \frac{\epsilon}{x_0}$$

D.2 Proof of Proposition 5

Proof. Consider the case in which active investors have heterogeneous risk aversions, but face no leverage constraints. In this case, the expression for $y_2(X)$ in Equation (C.5) can be written as

$$y_2(X) = \left(1-\psi^{-1}\right)\gamma\sigma^2 \left[\sum_{j=1}^{J} x_j \left(\frac{\sigma_{j,1}^2(X)}{2\sigma^2} + \hat{\gamma}_j \frac{\sigma_{j,1}(X)}{\sigma}\right)\right] + \left(1-\psi^{-1}\right)\gamma\sigma^2 x_0 \left(\frac{\hat{\alpha}_p^2}{2} + \hat{\gamma}_0 \hat{\alpha}_p\right)$$

Note that with unconstrained active investors, the effect of endogenous volatility and hedging demand exactly cancel out. We first compute the derivative of the term involving $\sigma_{j,1}^2$:

$$\begin{aligned}
\sum_{j=0}^{J} x_j \frac{\partial \sigma_{j,1}^2}{\partial \hat{\alpha}_p} &= \sum_{j=0}^{J} 2 x_j \sigma_{j,1} \frac{\partial \sigma_{j,1}}{\partial \hat{\alpha}_p} \\
&= 2 x_0 \sigma^2 \hat{\alpha}_p + 2\sigma^2 \sum_{j=1}^{J} x_j \left(\sum_{k \in \mathcal{J}^u} \frac{x_k}{x_u}\hat{\gamma}_k - \left(\frac{\hat{\alpha}_p x_0}{1-x_0}\right) - \hat{\gamma}_j\right)\left(-\frac{x_0}{1-x_0}\right) \\
&= 2\sigma^2 \left[x_0 \hat{\alpha}_p + \frac{\hat{\alpha}_p x_0^2}{1-x_0} + \left((1-x_0)\sum_{k=1}^{J} \frac{x_k}{x_u}\hat{\gamma}_k - \sum_{j=1}^{J} x_j \hat{\gamma}_j\right)\left(-\frac{x_0}{1-x_0}\right)\right] \\
&= 2\sigma^2 \frac{\hat{\alpha}_p x_0}{1-x_0}.
\end{aligned}$$

The derivatives of the term involving $\sigma_{j,1}(X)$ with respect to $\hat{\alpha}_p$ are given by:

$$\frac{1}{\sigma}\sum_{j=0}^{J} x_j \hat{\gamma}_j \frac{\partial \sigma_{j,1}(X)}{\partial \hat{\alpha}_p} = \sum_{j=1}^{J} x_j \hat{\gamma}_j \left(-\frac{x_0}{1-x_0}\right) + x_0 \hat{\gamma}_0 = x_0\left(\hat{\gamma}_0 - \mathbb{E}^u[\hat{\gamma}_k]\right)$$

Thus, the (inverse) aggregate elasticity is then given by:

$$\frac{1}{p(X,\epsilon)} \frac{\partial p(X,\epsilon)}{\partial F(X,\epsilon)} = \left(1-\psi^{-1}\right)\frac{\gamma\sigma^2}{y_0(X)}\left[\frac{1-\overline{\alpha}_p}{x_a} - \frac{\gamma_0 - \mathbb{E}^u\left[\gamma_j\right]}{\gamma}\right] + \mathcal{O}(\epsilon^2),$$

where $x_a \equiv 1 - x_0$ is the wealth share of active investors, and we use $\gamma_j = \gamma\left(1 + \hat{\gamma}_j \epsilon\right)$ from Equation (20). □

E Useful Formula

The following are useful for computing the derivatives above:

$$\gamma\sigma\frac{\partial\sigma_{y,2}(X)}{\partial\hat{\alpha}_p} = \left(1-\psi^{-1}\right)\frac{\gamma\sigma^2}{2y_0(X)}\sum_{k=1}^{J}\left(\hat{\gamma}_k-\hat{\gamma}_0\right)x_k\left(-\frac{\gamma\sigma^2 x_0}{1-x_0}\right).$$

Note that we can write $\sigma_{j,2}(X)$ as follows

$$\begin{aligned}
\frac{\sigma_{j,2}(X)}{\sigma} &= \frac{\eta_2(X)}{\gamma\sigma}-\frac{\eta_1(X)}{\gamma\sigma}\hat{\gamma}_j+\frac{\eta_0(X)}{\gamma\sigma}\hat{\gamma}_j^2-(1-\gamma^{-1})\frac{\sigma_{y,2}(X)}{\sigma}\\
&= -\left(1+\frac{x_c}{x_u}\right)\frac{\sigma_{y,2}(X)}{\sigma}-\mathbb{E}^u[\hat{\gamma}_k]\frac{\hat{\alpha}_p x_0+\frac{\hat{\sigma}}{\sigma}x_c}{1-x_0-x_c}-\mathrm{Var}^u[\hat{\gamma}_k]+\\
&\quad-\hat{\gamma}_j\left[\sum_{k\in\mathcal{J}^u}\frac{x_k}{x_u}\hat{\gamma}_j-\frac{\hat{\alpha}_p x_0}{1-x_0}\right]+\hat{\gamma}_j^2.
\end{aligned}$$

$$\begin{aligned}
\sigma_{j,1}(X) &= \sigma\min\left\{\sum_{k\in\mathcal{J}^u}\frac{x_k}{x_u}\hat{\gamma}_k-\left(\frac{\hat{\alpha}_p x_0+\frac{\hat{\sigma}}{\|\sigma\|}x_c}{1-x_0-x_c}\right)-\hat{\gamma}_j,\frac{\hat{\sigma}}{\sigma}\right\}\\
\sigma_{j,2}(X) &= \frac{\eta_2(X)}{\gamma}-\frac{\eta_1(X)}{\gamma}\hat{\gamma}_j+\frac{\eta_0(X)}{\gamma}\hat{\gamma}_j^2-(1-\gamma^{-1})\sigma_{y,2}(X)\\
\sigma_{y,2}(X) &= \left(1-\psi^{-1}\right)\frac{\gamma\sigma^2}{2y_0(X)}\sum_{k=1}^{J}\left(\hat{\gamma}_k-\hat{\gamma}_0\right)x_k\sigma_{k,1}(X)\\
\eta_1(X) &= \gamma\sigma\left[\sum_{j\in\mathcal{J}^u}\frac{x_j}{x_u}\hat{\gamma}_j-\frac{\hat{\alpha}_p x_0}{1-x_0}\right]\\
\eta_2(X) &= -\frac{(\gamma-1)x_c+1-x_0}{1-x_0-x_c}\sigma_{y,2}(X)-\gamma\sigma\mathbb{E}^u[\hat{\gamma}_j]\frac{\hat{\alpha}_p x_0+\frac{\hat{\sigma}}{\sigma}x_c}{1-x_0-x_c}-\gamma\sigma\mathrm{Var}^u[\hat{\gamma}_j],
\end{aligned}$$

where

$$\begin{aligned}
\frac{\partial\sigma_{y,2}}{\partial\hat{\alpha}_p} &= \left(1-\psi^{-1}\right)\frac{\gamma\sigma^2}{2y_0(X)}\sum_{k=1}^{J}\left(\hat{\gamma}_k-\hat{\gamma}_0\right)x_k\left(-\frac{\sigma x_0}{1-x_0}\right)\\
&= \left(1-\psi^{-1}\right)\frac{\gamma\sigma^2}{2y_0(X)}\sigma\left(\hat{\gamma}_0-\mathbb{E}^u[\hat{\gamma}_j]\right)x_0\\
\sum_{j=0}^{J}x_j\gamma\sigma\hat{\gamma}_j\frac{\partial\sigma_{j,1}(X)}{\partial\hat{\alpha}_p} &= \gamma\sigma^2\sum_{j=0}^{J}x_j\hat{\gamma}_j\left(-\frac{x_0}{1-x_0}\right)
\end{aligned}$$

F Derivation of the perturbed solution

In this section, we compute the first-order and second-order correction of the equilibrium objects. It turns out that the system of equations determining the perturbed solution is block-recursive, so we are able to solve for the equilibrium objects one by one, provided we proceed in the appropriate order.

In contrast to the case considered in the text, we allow for portfolio-flow shocks. In particular, we assume that the portfolio share of the passive investor is given by $\alpha_{0,t} = 1 + \epsilon(\overline{\alpha}_{p,t} - 1)$, where $\overline{\alpha}_{p,t}$ follows the process

$$d\overline{\alpha}_{p,t} = \theta_p(\overline{\alpha} - \overline{\alpha}_{p,t})\epsilon dt + \sigma_p \sqrt{\overline{\alpha}_{p,t}} \epsilon dZ_t.$$

Notice that $\alpha_{0,t} = 1$ and $\overline{\alpha}_{p,t}$ is constant when $\epsilon = 0$. Finally, we assume that the mortality parameter is given by $\kappa = \hat{\kappa}\epsilon$.

F.1 First-order correction

Diffusion and drift terms. The diffusion term for the price-dividend ratio is given by

$$\sigma_{p,t} = \frac{p_x}{p}\sigma_x + \frac{p_{\overline{\alpha}_p}}{p}\sigma_p \sqrt{\overline{\alpha}_{p,t}}\epsilon = O(\epsilon^2).$$

Notice that $p_{x_j} = O(\epsilon)$ and $\sigma_{x_j} = O(\epsilon)$, as p and x_j are constant when $\epsilon = 0$, so the zeroth-order terms for p_{x_j}, $p_{\overline{\alpha}_p}$, and σ_{x_j} are equal to zero. This implies that the first-order correction for $\sigma_{p,t}$ is equal to zero. A similar argument shows that $\sigma_{c_j,t} = O(\epsilon^2)$.

The drift of p is given by

$$\mu_{p,t} = \frac{p_x}{p}\mu_x + \frac{p_{\overline{\alpha}_p}}{p}\theta_p(\overline{\alpha} - \overline{\alpha}_{p,t})\epsilon + \frac{1}{2}\sum_{k=1}^{d}\left[\sigma'_{x,k}\frac{p_{xx}}{p}\sigma_{x,k} + 2\sigma_{p,k}\sqrt{\overline{\alpha}_{p,t}}\epsilon\frac{p_{x\alpha_p}}{p}\sigma_{x,k} + \frac{p_{\overline{\alpha}_p\overline{\alpha}_p}}{p}\sigma_{p,k}^2\overline{\alpha}_{p,t}\epsilon^2\right],$$

where $\sigma_{x,k}$ is the k-th column of the $J \times d$ matrix σ_x. As p_x and $p_{\overline{\alpha}_p}$ are first-order in ϵ, and the same goes for μ_x and σ_x, then $\mu_{p,t} = O(\epsilon^2)$. A similar argument shows that $\mu_{c_j,t} = O(\epsilon^2)$. Notice these facts imply that $\varsigma_j = O(\epsilon^2)$ and $\xi_{j,t} = O(\epsilon^2)$.

Risk premium. The risk premium is given by

$$\pi_1(X) = \gamma||\sigma||^2\left[\sum_{j\in\mathcal{J}^u}\frac{x_j}{x_u}\hat{\gamma}_j - \frac{x_0\hat{\alpha}_p + x_c\frac{\hat{\sigma}}{||\sigma||}}{1 - x_0 - x_c}\right],$$

using the fact that $||\sigma_{R,t}|| = ||\sigma|| + O(\epsilon^2)$, and $\hat{\alpha}_{p,t} \equiv \overline{\alpha}_{p,t} - 1$.

Portfolio share. The portfolio share of an unconstrained investor is given by

$$\alpha_j(X,\epsilon) = 1 + \left[\frac{\pi_1(X)}{\gamma||\sigma||^2} - \hat{\gamma}_j\right]\epsilon + O(\epsilon^2),$$

the portfolio share of a constrained investor is given by

$$\alpha_j(X,\epsilon) = 1 + \frac{\hat{\sigma}}{||\sigma||}\epsilon + O(\epsilon^2),$$

and the portfolio share of the passive investor is given by $\alpha_0(X,\epsilon) = 1 + \hat{\alpha}_{p,t}\epsilon$. Notice that $\sum_{j=0}^{J} x_j \alpha_{j,1}(X) = 0$, consistent with market clearing.

Interest rate. The first-order correction for the interest rate is given by

$$r_1(X) = \left(1 - \psi^{-1}\right)\gamma||\sigma||^2 \sum_{j=0}^{J} x_j \left[\frac{\hat{\gamma}_j}{2} + \alpha_{j,1}(X)\right] - \pi_1(X),$$

using the fact that $\xi_t = O(\epsilon^2)$, $\mu_{p,t} = O(\epsilon^2)$, and $\sigma_{p,t} = O(\epsilon^2)$. Given the market clearing for the risky asset, we can write:

$$r_1(X) = \left(1 - \psi^{-1}\right)\gamma||\sigma||^2 \sum_{j=0}^{J} x_j \frac{\hat{\gamma}_j}{2} - \pi_1(X),$$

so $r_1(X) + \pi_1(X)$ is independent of $\overline{\alpha}_p$.

Price-dividend ratio. From the pricing condition, we obtain

$$-\frac{1}{p_0(X)^2} p_1(X) = r_1(X) + \pi_1(X).$$

Rearranging the expression above, and using the expression for the interest rate, we obtain

$$p_1(X) = -p_0(X)^2 \left(1 - \psi^{-1}\right)\gamma||\sigma||^2 \sum_{j=0}^{J} x_j \frac{\hat{\gamma}_j}{2},$$

which is independent of $\overline{\alpha}_{p,t}$.

Consumption-wealth ratio. The consumption-wealth ratio is given by

$$c_{j,1}(X) = (1-\psi)\left[r_1(X) + \pi_1(X) + \pi_0(X)\alpha_{j,1}(X) - \frac{1}{2}\gamma|||\sigma|||^2(\hat{\gamma}_j + 2\alpha_{j,1}(X))\right].$$

Using the expression for $r_1(X)$, we can write the expression above as follows:

$$c_{j,1}(X) = (1-\psi)\left[(1-\psi^{-1})\sum_{i=0}^{J} x_i\frac{\hat{\gamma}_i}{2} - \frac{\hat{\gamma}_j}{2}\right]\gamma|||\sigma|||^2.$$

Wealth dynamics. The diffusion term of x_j is given by

$$\sigma_{x_j}(X) = x_j\alpha_{j,1}(X)\epsilon\sigma + O(\epsilon^2).$$

The drift of x_j is given by

$$\mu_{x_j}(X) = x_j\left[r_1(X) + \pi_1(X) + \pi_0(X)\alpha_{j,1}(X) - c_{j,1}(X) - \alpha_{j,1}(X)||\sigma||^2 + \hat{\kappa}\frac{\omega_j - x_j}{x_j}\right]\epsilon + O(\epsilon^2).$$

We can write the first-order correction of μ_{x_j} as follows:

$$\mu_{x_j,1}(X) = x_j\left[(\psi-1)\frac{\gamma|||\sigma|||^2}{2}\left(\sum_{i=0}^{J} x_i\hat{\gamma}_i - \hat{\gamma}_j\right) + (\gamma-1)|||\sigma|||^2\alpha_{j,1}(X)\right] + \hat{\kappa}(\omega_j - x_j).$$

F.2 Second-order correction

Diffusion and drift terms. The diffusion term for the price-dividend ratio is given by

$$\sigma_{p,2}(X) = \frac{p_{x,1}(X)}{p_0(X)}\sigma_{x,1}(X) + \frac{p_{\overline{\alpha}_p,1}(X)}{p_0(X)}\sigma_p\sqrt{\overline{\alpha}_p}.$$

We can write the expression above as follows:

$$\sigma_{p,2}(X) = -p_0(X)(1-\psi^{-1})\frac{\gamma||\sigma|||^2}{2}\sum_{j=1}^{J}\left(\hat{\gamma}_j - \hat{\gamma}_0\right)x_j\alpha_{j,1}(X)\sigma,$$

where we used the fact that $p_{\overline{\alpha}_p,1}(X) = 0$.

Similarly, the diffusion for c_j is given by

$$\sigma_{c_j,2}(X) = -(\psi-1)\frac{\gamma|||\sigma|||^2}{2c_{j,0}(X)}(1-\psi^{-1})\sum_{i=1}^{J}(\hat{\gamma}_i - \hat{\gamma}_0)\,x_i\alpha_{i,1}(X)\sigma.$$

The second-order correction for the hedging demand is then given by $\varsigma_{j,2}(X) = \frac{1-\gamma^{-1}}{\psi-1}\frac{\sigma_{c_j,2}\sigma'}{\||\sigma\||^2}$.

The second-order correction of the drift of p and c_j are given by

$$\mu_{p,2}(X) = \frac{p_{x,1}(X)}{p_0(X)}\mu_{x,1}(X), \qquad \mu_{c_j,2}(X) = \frac{c_{j,x,1}(X)}{c_{j,0}(X)}\mu_{x,1}(X),$$

which can be written as

$$\begin{aligned}
\mu_{p,2}(X) &= -p_0(X)(1-\psi^{-1})\frac{\gamma\|\sigma\||^2}{2}\sum_{j=1}^{J}(\hat{\gamma}_j - \hat{\gamma}_0)\mu_{x_j,1}(X) \\
\mu_{c_i,2}(X) &= (1-\psi)(1-\psi^{-1})\frac{\gamma\|\sigma\||^2}{2c_{i,0}(X)}\sum_{j=1}^{J}(\hat{\gamma}_j - \hat{\gamma}_0)\mu_{x_j,1}(X).
\end{aligned}$$

The second-order correction for $\xi_{j,t}$ is then given by $\xi_{j,2}(X) = \mu_{c_j,2}(X) + (1-\gamma)\sigma_{c_j,2}\sigma'$.

Risk premium. The risk premium is given by

$$\pi_2(X) = \gamma\|\sigma\||^2\left[\frac{\gamma_{u,2}(X)}{\gamma} - \frac{\gamma_{u,1}(X)}{\gamma}\frac{x_0\hat{\alpha}_p + x_c\frac{\hat{\sigma}}{\|\sigma\||}}{1-x_0-x_c} + 2\sum_{k=1}^{d}\frac{\sigma_k\sigma_{p,2,k}}{\|\sigma\|^2} - \frac{x_c}{1-x_0-x_c}\overline{\alpha}_{c,2}(X) - \varsigma_2(X)\right].$$

Notice that we can write the aggregate risk aversion as follows:

$$\gamma_u(X) = \gamma\left[1 + \mathbb{E}^u[\hat{\gamma}_j]\epsilon - \delta^u[\hat{\gamma}_j]\epsilon^2\right] + O(\epsilon^3),$$

where $\mathbb{E}^u[\hat{\gamma}_j] \equiv \sum_{j\in\mathcal{J}^u}\frac{x_j}{x_u}\hat{\gamma}_j$ and $\delta^u[\hat{\gamma}_j] \equiv \sum_{j\in\mathcal{J}^u}\frac{x_j}{x_u}\hat{\gamma}_j^2 - \left(\sum_{j\in\mathcal{J}^u}\frac{x_j}{x_u}\hat{\gamma}_j\right)^2$, so $\gamma_{u,2}(X)/\gamma = -\delta^u[\hat{\gamma}_j]$.

Combining the previous two expressions, we obtain

$$\begin{aligned}
\frac{\pi_2(X)}{\gamma\|\sigma\||^2} &= -\delta^u[\hat{\gamma}_j] - \mathbb{E}^u[\hat{\gamma}_j]\frac{x_0\hat{\alpha}_p + x_c\frac{\hat{\sigma}}{\|\sigma\||}}{1-x_0-x_c} + 2\sum_{k=1}^{d}\frac{\sigma_k\sigma_{p,2,k}}{\|\sigma\|^2} + \frac{x_c}{1-x_0-x_c}\sum_{k=1}^{d}\frac{\sigma_k\sigma_{p,2,k}}{\|\sigma\||^2} - \varsigma_2(X) \\
&= -\delta^u[\hat{\gamma}_j] - \mathbb{E}^u[\hat{\gamma}_j]\frac{x_0\hat{\alpha}_p + x_c\frac{\hat{\sigma}}{\|\sigma\||}}{1-x_0-x_c} + \left(1 + \gamma^{-1} + \frac{x_c}{1-x_0-x_c}\right)\sum_{k=1}^{d}\frac{\sigma_k\sigma_{p,2,k}}{\|\sigma\||^2},
\end{aligned}$$

where we used the fact that $\varsigma_2(X) = \frac{1-\gamma^{-1}}{\psi-1}\frac{\sigma_{c_j,2}\sigma'}{\|\sigma\||^2}$, $\sigma_{p,2} = \frac{\sigma_{c_j,2}}{\psi-1}$, and $\overline{\alpha}_{c,2}(X) = -\sum_{k=1}^{d}\frac{\sigma_k\sigma_{p,2,k}}{\|\sigma\||^2}$.

Portfolio share. The portfolio share of an unconstrained investor is given by

$$\alpha_{j,2}(X) = \frac{\pi_2(X)}{\gamma\||\sigma\||^2} - \frac{\pi_1(X)}{\gamma\||\sigma\||^2}\hat{\gamma}_j + \hat{\gamma}_j^2 - 2\sum_{k=1}^{d}\frac{\sigma_k\sigma_{p,2,k}(X)}{\|\sigma\|^2} + \varsigma_{j,2}(X),$$

the portfolio share of a constrained investor is $\alpha_{j,2}(X) = -\sum_{k=1}^{d} \frac{\sigma_k \sigma_{p,2,k}}{||\sigma|||^2}$, and the portfolio share of the passive investor satisfies $\alpha_{0,2} = 0$.

We can write the expression above as follows:

$$\alpha_{j,2}(X) = \frac{\pi_2(X)}{\gamma|||\sigma|||^2} - \frac{\pi_1(X)}{\gamma|||\sigma|||^2}\hat{\gamma}_j + \hat{\gamma}_j^2 - (1+\gamma^{-1}) \sum_{k=1}^{d} \frac{\sigma_k \sigma_{p,2,k}(X)}{\|\sigma\|^2}.$$

Notice that $\sum_{j=0}^{J} x_j \alpha_{j,2}(X) = 0$, consistent with market clearing.

Interest rate. The interest rate is given by

$$r_2(X) = \psi^{-1}(\mu_{p,2}(X) + \sigma\sigma_{p,2}(X)') + (1-\psi^{-1}).\frac{\gamma||\sigma|||^2}{2} \sum_{j=0}^{J} x_j \left[2\hat{\gamma}_j \alpha_{j,1}(X) + \alpha_{j,1}^2(X) + 2\alpha_{j,2}(X)\right]$$
$$+ (1-\psi^{-1})\gamma\|\sigma\|^2 \sum_{k=1}^{d} \frac{\sigma_k \sigma_{p,2,k}(X)}{\|\sigma\|^2} - \pi_2(X) + \psi^{-1}\xi_2(X),$$

where $\xi_2(X) = (\psi - 1)\left[\mu_{p,2}(X) + (1-\gamma)\sigma_{p,2}\sigma'\right]$

We can write the expression for the portfolio of the unconstrained investor as follows:

$$r_2(X) = \mu_{p,2}(X) + \sigma\sigma_{p,2}(X)' + (1-\psi^{-1})\gamma||\sigma|||^2 \sum_{j=0}^{J} x_j \left[\hat{\gamma}_j \alpha_{j,1}(X) + \frac{\alpha_{j,1}^2(X)}{2}\right] - \pi_2(X),$$

Price-dividend ratio. The price-dividend ratio is given by

$$\frac{p_1^2(X)}{p_0^3(X)} - \frac{p_2(X)}{p_0^2(X)} = r_2(X) + \pi_2(X) - \mu_{p,2}(X) - \sigma\sigma'_{p,2}.$$

Rearranging the expression above, and using the expression for $r_2(X)$, we obtain

$$\frac{p_2(X)}{p_0(X)} = -p_0(X)(1-\psi^{-1})\gamma||\sigma|||^2 \sum_{j=0}^{J} x_j \left[\hat{\gamma}_j \alpha_{j,1}(X) + \frac{\alpha_{j,1}^2(X)}{2}\right] + \left(\frac{p_1(X)}{p_0(X)}\right)^2.$$

Consumption-wealth ratio. The second-order correction for the consumption-wealth ratio is given by

$$c_{j,2}(X) = (1-\psi)\left[r_2(X) + \pi_2(X) + \pi_1(X)\alpha_{j,1}(X) + \pi_0(X)\alpha_{j,2}(X)\right] + \xi_{j,2}(X)$$
$$- (1-\psi)\gamma||\sigma|||^2 \left[\alpha_{j,1}(X)\hat{\gamma}_j + \alpha_{j,2}(X) + \frac{\alpha_{j,1}^2(X)}{2} + \sum_{k=1}^{d} \frac{\sigma_k \sigma_{p,k,2}(X)}{||\sigma|||^2}\right].$$

Wealth dynamics. The diffusion term of x_j is given by

$$\sigma_{x_j,2}(X) = x_j \alpha_{j,2} \sigma.$$

The drift of x_j is given by

$$\mu_{x_j,2} = x_j \left[r_2(X) + \pi_2(X) + \pi_1(X)\alpha_{j,1}(X) + \pi_0(X)\alpha_{j,2}(X) - c_{j,2}(X) - \mu_{p,2}(X) - \sigma\sigma_{p,2}(X)' - \alpha_{j,2}(X)\|\sigma\|^2 \right].$$

Aggregate market elasticity. The derivative of p with respect to $\overline{\alpha}_p$ is given by

$$\frac{1}{p(X,\epsilon)} \frac{\partial p(X,\epsilon)}{\partial \overline{\alpha}_p} = \frac{1}{p_0(X)} \frac{\partial p_2(X)}{\partial \overline{\alpha}_p} \epsilon^2 + O(\epsilon^3).$$

The market elasticity satisfies the condition

$$\frac{1}{p_0(X)} \frac{\partial p_2(X)}{\partial \overline{\alpha}_p} = -p_0(X)(1-\psi^{-1})\gamma\|\sigma\||^2 \left[x_0(\hat{\gamma}_0 + \hat{\alpha}_p) + \sum_{j \in \mathcal{J}^u} x_j \left(\hat{\gamma}_j + \alpha_{j,1}(X)\right) \left(-\frac{x_0}{x_u}\right) \right] \epsilon^2.$$

From the market clearing for the risky asset, we have $x_0\hat{\alpha}_p + \sum_{j\in\mathcal{J}^u} x_j \alpha_{j,1}(X) + x_c \frac{\hat{\sigma}}{\|\sigma\|} = 0$, so we can write the expression above as follows:

$$\frac{1}{p_0(X)} \frac{\partial p_2(X)}{\partial \overline{\alpha}_p} = p_0(X)(1-\psi^{-1})\gamma\|\sigma\||^2 \left[\hat{\gamma}_u(X) - \hat{\gamma}_0 + \frac{1-x_c}{1-x_0-x_c}(1-\overline{\alpha}_p) - \frac{x_c \frac{\hat{\sigma}}{\|\sigma\||}}{1-x_0-x_c} \right] x_0 \epsilon^2.$$

F.3 Third-order approximation

F.3.1 Passive demand

Suppose there is no preference heterogeneity and no leverage constraint. Without loss of generality, set $J = 1$. In this case, the price-dividend ratio is given by

$$\begin{aligned}
p(X,\epsilon) &= p^\star - (p^\star)^2(1-\psi^{-1})\frac{\gamma\|\sigma\|^2}{2}\frac{x_0\hat{\alpha}_p^2}{x_1}\epsilon^2 + O(\epsilon^3)\\
\pi(X,\epsilon) &= \pi_0(X) - \gamma\|\sigma\|^2\frac{x_0\hat{\alpha}_p}{x_1}\epsilon + O(\epsilon^3)\\
r(X,\epsilon) &= r_0(X) + \gamma\|\sigma\|^2\frac{x_0\hat{\alpha}_p}{x_1}\epsilon + (1-\psi^{-1})\frac{\gamma\|\sigma\|^2}{2}\frac{x_0\hat{\alpha}_p^2}{x_1}\epsilon^2 + O(\epsilon^3)\\
\alpha_0(X,\epsilon) &= 1 + \hat{\alpha}_p\epsilon + O(\epsilon^3)\\
\alpha_1(X,\epsilon) &= 1 - \frac{x_0}{x_1}\hat{\alpha}_p\epsilon + O(\epsilon^3)\\
c_j(X,\epsilon) &= c_{j,0}(X) + (1-\psi^{-1})\frac{\gamma\|\sigma\|^2}{2}\frac{x_0\hat{\alpha}_p^2}{x_1}\epsilon^2 + O(\epsilon^3)\\
\sigma_{x_1}(X,\epsilon) &= -\frac{x_0}{x_1}\hat{\alpha}_p\sigma\epsilon + O(\epsilon^3)\\
\mu_{x_1}(X,\epsilon) &= \left[(1-\gamma)\|\sigma\|^2(1-x_1)\hat{\alpha}_p + \hat{\kappa}(\omega_j - x_1)\right]\epsilon + O(\epsilon^3),
\end{aligned}$$

where $\mu_p, \sigma_p, \mu_{c_j}$, and σ_{c_j} are all equal to zero up to second order, and $x_0 = 1 - x_1$.

The law of motion of $x_{1,t}$ can be written as

$$\frac{dx_{1,t}}{x_{1,t}} = \left[x_{0,t}(c_{0,t} - c_{1,t}) + x_{0,t}(\alpha_{0,t} - \alpha_{1,t})\left(\|\sigma_{R,t}\|^2 - \pi_t\right) + \kappa\frac{\omega_1 - x_{1,t}}{x_{1,t}}\right]dt + (\alpha_{1,t} - 1)\sigma_{R,t}dZ_t.$$

Diffusion and drift terms. The derivatives of $p(X,\epsilon)$ with respect to x_1 and $\overline{\alpha}_p$ are given by

$$\frac{p_{x_1}(X,\epsilon)}{p^\star} = (1-\psi^{-1})\frac{\gamma\|\sigma\|^2}{2y^\star}\frac{1}{x_1^2}\hat{\alpha}_p^2\epsilon^2 + O(\epsilon^3), \qquad \frac{p_{\overline{\alpha}_p}(X,\epsilon)}{p^\star} = -(1-\psi^{-1})\frac{\gamma\|\sigma\|^2}{y^\star}\frac{x_0}{x_1}\hat{\alpha}_p\epsilon^2 + O(\epsilon^3)$$

The diffusion term for $p(X,\epsilon)$ is then given by

$$\sigma_{p,3}(X) = -(1-\psi^{-1})\frac{\gamma\|\sigma\|^2}{y^\star}\frac{x_0}{x_1}\left[\frac{\hat{\alpha}_p^3}{2x_1^2}\sigma + \hat{\alpha}_p\sigma_p\sqrt{\overline{\alpha}_p}\right].$$

and $\sigma_{c_j,3}(X) = \sigma_{p,3}(X)$. Notice that excess volatility depends on the market elasticity times the volatility of portfolio flows.

The drift of p is given by

$$\mu_{p,3}(X) = (1-\psi^{-1})\frac{\gamma\|\sigma\|^2}{2y^\star}\frac{1}{x_1^2}\hat{\alpha}_p^2\mu_{x_1,1} - (1-\psi^{-1})\frac{\gamma\|\sigma\|^2}{y^\star}\frac{x_0}{x_1}\hat{\alpha}_p\theta_p(\overline{\alpha}-\overline{\alpha}_{p,t}),$$

and $\mu_{c_j,3}(X) = \mu_{p,3}(X)$.

Risk premium. The risk premium is given by

$$\pi(X) = \gamma\|\sigma\|^2\left[1 - \frac{x_0(\overline{\alpha}_p - 1)}{x_1} - \frac{1-\gamma^{-1}}{\psi - 1}\frac{\gamma\|\sigma\|^2}{y^\star}\frac{x_0}{x_1}\left[\frac{(1-\overline{\alpha}_p)^3}{2x_1^2}\sigma + \hat{\alpha}_p\sigma_p\sqrt{\overline{\alpha}_p}\right]\right]$$

G Higher-order perturbations

Suppose we have the $(n-1)$-th order perturbation of $c_j(X,\epsilon) = \sum_{k=0}^{n-1} c_{j,k}(X)\epsilon^k$ and the law of motion of X. Let $lc_j(X,\epsilon) = \sum_{k=0}^{n-1} lc_{j,k}(X)\epsilon^k$ denote the expansion of $\log c_j(X,\epsilon)$. Then we can compute $\sigma_j(X)$ up to order n:

$$\sigma_{c_j,n}(X) = \sum_{k=1}^{n-1} lc_{j,k,X}(X)\sigma_{X,n-k}(X),$$

which is independent of the n-th order term in $c_j(X,\epsilon)$ and $\sigma_X(X,\epsilon)$, as $lc_{j,0,X}(X) = 0$. Similarly, we can compute $\sigma_{y,n}(X)$. A similar argument gives $\mu_{p,n}(X)$ and $\mu_{c_j,n}(X)$. We can then compute $\pi_n(X)$ and $\alpha_{j,n}(X)$. The n-th term of the consumption-wealth ratio satisfies the condition

$$c_{j,t} = \psi\rho + (1-\psi)\left[\pi_t(\alpha_{j,t}-1) + \mu + \mu_{p,t} + \sigma\sigma'_{p,t} - \frac{\gamma_j}{2}\|\sigma_{R,t}\|^2\alpha_{j,t}^2 + \sum_{j=0}^{J} x_j c_{j,t}\right] + \xi_{j,t}$$

We can rewrite the system above in matrix form as follows:

$$[I - (1-\psi)\mathbf{1}_{J+1}x'_t]c_t = \zeta_t,$$

where $c_t = [c_{0,t},\ldots,c_{J,t}]'$, $x_t = [x_{0,t},\ldots,x_{J,t}]'$, $\mathbf{1}_{J+1}$ is a $(J+1)$-th dimensional vector of ones, and $\zeta_{j,t} \equiv \psi\rho + (1-\psi)\left[\pi_t(\alpha_{j,t}-1) + \mu + \mu_{p,t} + \sigma\sigma'_{p,t} - \frac{\gamma_j}{2}\|\sigma_{R,t}\|^2\alpha_{j,t}^2\right] + \xi_{j,t}$. Applying the Sherman-Morrison formula, we obtain

$$c_t = [I - (1-\psi^{-1})\mathbf{1}_{J+1}x'_t]\zeta_t,$$

or $c_{j,t} = \zeta_{j,t} - (1-\psi^{-1})x'_t\zeta_t$. Notice that ζ_t can be computed at order n based on the coefficients of order $n-1$ and their derivatives.

Computing the derivatives. The derivation above shows that, given the order $n-1$ expansion of $\zeta_j(X,\epsilon)$ and its derivatives, we can compute the expansion of order n. Suppose the expansion of $\zeta_j(X,\epsilon)$ is given by

$$\zeta_j(X,\epsilon) = \sum_{k=0}^{n-1} \zeta_{j,k}(X)\epsilon^k,$$

where $\zeta_{j,k}(X)$ takes the form:

$$\zeta_{j,k}(X) = A_{j,k} + B'_{j,k}(X-\overline{X}) + \frac{1}{2}(X-\overline{X})'C_{j,k}(X-\overline{X}),$$

where $\overline{X}$ is a reference point, $A_{j,k}$ is a scalar, $B_{j,k}$ is a vector, and $C_{j,k}$ is a matrix. Notice that . $\zeta_{j,k}(\overline{X}) = A_{j,k}$, $\zeta_{j,k,X}(\overline{X}) = B_{j,k}$ and $\zeta_{j,k,XX} = C_{j,k}$. Given this expansion, we can compute $c_j(X,\epsilon) = \sum_{k=0}^{n-1} c_{j,k}(X)\epsilon^k$.

G.1 Inner region

Consider the case of no preference heterogeneity and no leverage constraints. Consider the following change of variables: $x_1 = \epsilon\tilde{x}_1$. Define $\tilde{c}_j(\tilde{x}_1, \overline{\alpha}_p) = c_j\left(x_1, \overline{\alpha}_p\right)$, so $c_{j,x_1} = \frac{1}{\epsilon}\tilde{c}_{j,\tilde{x}_1}$ and $c_{j,x_1x_1} = \frac{1}{\epsilon^2}\tilde{c}_{j,\tilde{x}_1\tilde{x}_1}$. This implies the following is true:

$$\sigma_{c_j} = \frac{1}{\epsilon}\frac{\tilde{c}_{j,\tilde{x}_1}}{\tilde{c}_j}\sigma_{x_1},$$

where $\sigma_{x_1} = -\frac{1-\epsilon\tilde{x}_1}{\tilde{x}_1}\hat{\alpha}_p\sigma_R$. Similarly, we can write σ_y

$$\sigma_y = \frac{1}{\epsilon}\frac{\tilde{y}_{x_1}}{y}\sigma_{x_1}.$$

The drift of y is given by

$$\mu_y = \frac{1}{\epsilon}\frac{\tilde{y}_{\tilde{x}_1}}{y}\mu_{\tilde{x}_1} + \frac{1}{2\epsilon^2}\frac{\tilde{y}_{\tilde{x}_1\tilde{x}_1}}{y}\sigma^2_{\tilde{x}_1}.$$

The term of order 0 is the same as before.

H Volatility

Consider the case without preference heterogeneity or leverage constraints. The price-dividend ratio is given by:

$$\hat{p}_t = -(1-\psi^{-1})\frac{\gamma\|\sigma\|^2}{2y^\star}\frac{(1-x_{a,t})(1-\overline{\alpha}_{p,t})^2}{x_{a,t}}.$$

where, given $\overline{\alpha}_{a,t} \equiv \frac{1-(1-x_{a,t})\overline{\alpha}_{p,t}}{x_{a,t}}$, we have

$$\sigma_{x_a,t} = x_{a,t}(\overline{\alpha}_{a,t}-1)\sigma, \qquad \sigma_{\overline{\alpha}_p,t} = \sigma_{\overline{\alpha}_p}$$

The diffusion of $\hat{p}_t$ is given

$$\sigma_{p,t} = \hat{p}_{x_a,t}\sigma_{x_a,t} + \hat{p}_{\overline{\alpha}_p,t}\sigma_{\overline{\alpha}_p}.$$

$$\hat{c}_t \;\equiv\; c_t - c^\star \;=\; \chi_{0,t} \;+\; \chi_p\,\hat{p}_t \;+\; \mathcal{O}(\epsilon^3),$$

with coefficients

$$\chi_p = (\psi-1)\,\frac{1}{p^\star}, \qquad \chi_{0,t} = (\psi-1)\,\frac{\gamma\|\sigma\|^2}{2}\,\frac{f_t^2}{x_{0,t}(1-x_{0,t})}.$$

I Numerical methodology

This appendix describes the numerical procedure used to solve the dynamic model in Section 3. We follow a two-step strategy where we first solve a 'simplified' model with an aggregate claim in positive net supply and a derivative security in zero net supply. In the second step (valuation step), we solve the valuation equation for the dividend and risky debt claims. This ordering is justified since, once the equilibrium is computed in the first step, the market prices of risk and interest rates are known functions of the simplified model, and the individual valuation ratios can be recovered from the respective pricing equations.

I.1 Simplified economy

The state variables in the simplified economy are

$$X = (x_c, x_u, \alpha_p) \in \mathcal{S}, \qquad x_p = 1 - x_c - x_u,$$

where x_c and x_u are the wealth shares of the constrained and unconstrained active investors, and

α_p is the exogenous risky-share demand of the passive sector. The aggregate claim is in positive net supply and the derivative is in zero net supply. Denote the active portfolio exposures by $(\alpha_{j,Y}, \theta_j)$, where $\alpha_{j,Y}$ is the exposure to the aggregate claim and $\alpha_{j,F}$ is the exposure to the derivative. The passive investor has exposure $(\alpha_{p,t}, 0)$. We approximate

$$\xi_c(X), \quad \xi_u(X), \quad \xi_p(X), \quad \alpha_{c,Y}(X), \quad \alpha_{c,F}(X), \quad \alpha_{u,Y}(X), \quad \alpha_{u,F}(X)$$

with neural networks parameterized by Θ:

$$\begin{aligned} \hat{\xi}_j &: \mathcal{S} \to \mathbb{R}_+, \qquad && j \in \{c, u, p\}, \\ \hat{\alpha}_{j,Y} &: \mathcal{S} \to \mathbb{R}, && j \in \{c, u\}, \\ \hat{\alpha}_{j,F} &: \mathcal{S} \to \mathbb{R}, && j \in \{c, u\}. \end{aligned}$$

and derive for all other equilibrium quantities as functions of the state variables and the approximated objects. Note that the passive investor does not trade the derivative, so $\alpha_{p,F}(X) \equiv 0$, while the passive exposure to the aggregate claim is the state variable itself, $\alpha_{p,Y}(X) = \alpha_p$. Thus the passive portfolio is imposed directly rather than approximated. We use fully connected feed-forward neural networks with 8 hidden layers and 10 neurons in each layer. Each neuron in the hidden and final layer is represented by a softplus activation function for the consumption-wealth ratios $\hat{\xi}_j$'s, and a linear function for the others. The state space is sampled globally using Latin-hypercube points on the simplex

$$\{(x_c, x_u, \alpha_p) : x_c > 0,\ x_u > 0, \alpha_p > 0,\ x_c + x_u < 1\}$$

Given a candidate collection of network outputs at X, the remaining equilibrium objects are recovered analytically.

First, goods-market clearing implies the aggregate claim dividend yield

$$\hat{y}(X) = x_c \hat{\xi}_c(X) + x_u \hat{\xi}_u(X) + x_p \hat{\xi}_p(X).$$

Second, the volatility of the dividend yield is recovered from the Itô consistency condition. Third, once $\sigma_y(X)$ is known, the excess-return volatilities of the aggregate claim and the derivative, denoted $(\sigma_{R,Y}(X), \sigma_{R,F}(X))$, are recovered. The unconstrained active investor's first-order conditions then pin down the two risk premia $\pi_Y(X), \pi_F(X)$. Fourth, conditional on $(\hat{y}, \pi_Y, \pi_F, \sigma_y)$, the risk-free rate $r(X)$, the state drift $\mu_X(X)$, the state diffusion matrix $\Sigma_X(X)$, the consumption-wealth ratio drifts $\mu_{c_j}(X)$, and the market price of risk vector $\lambda(X)$ are all recovered from the equilibrium identities and Itô's lemma. These calculations are explicit once the neural-network outputs and their first and second derivatives are known. The problem reduces to minimizing the following loss functions

1. **HJB residuals.** For each investor type $j \in \{c, u, p\}$, we construct the residual from the investor's Hamilton–Jacobi–Bellman equation after replacing all equilibrium objects by either neural-network outputs or the derived quantities defined above. We denote these residuals by

$$\mathcal{L}_j^{\text{HJB}}(X; \Theta).$$

2. **Portfolio FOCs.** The constrained investor's first-order conditions for the aggregate claim and derivative lead to two residuals denoted as:

$$\mathcal{L}_{c,Y}^{\text{FOC}}(X; \Theta), \qquad \mathcal{L}_{c,F}^{\text{FOC}}(X; \Theta).$$

   Note that the unconstrained investor's first-order conditions are used to recover (π_Y, π_F) exactly and therefore do not appear separately in the loss.

3. **Market-clearing residuals.** We impose clearing in the aggregate-claim and derivative markets as residuals:

$$\mathcal{L}_Y^{\text{MC}}(X; \Theta), \qquad \mathcal{L}_F^{\text{MC}}(X; \Theta).$$

The resulting loss is

$$\mathcal{L}(X; \Theta) = \sum_{j \in \{c,u,p\}} \left(\mathcal{L}_j^{\text{HJB}}(X; \Theta)\right)^2 + \left(\mathcal{L}_{c,Y}^{\text{FOC}}(X; \Theta)\right)^2 + \left(\mathcal{L}_{c,F}^{\text{FOC}}(X; \Theta)\right)^2 + \left(\mathcal{L}_Y^{\text{MC}}(X; \Theta)\right)^2 + \left(\mathcal{L}_F^{\text{MC}}(X; \Theta)\right)^2.$$

For a training batch $\{X^n\}_{n=1}^N$, the training objective is the empirical mean

$$\widehat{\mathcal{L}}(\Theta) = \frac{1}{N} \sum_{n=1}^{N} \mathcal{L}(X^n; \Theta). \tag{I.1}$$

The full algorithm is given in Algorithm (1).

I.2 Valuation ratios

After solving the simplified economy, we price the dividend and risky-debt claims as a function of the state variables (X, s), where $s \in [0, 1]$ denote the dividend share of the aggregate claim. We solve for the valuation ratio of the dividend claim, denoted by $p_D(X, s)$, and the valuation ratio of risky-debt claim, denoted by $p_B(X, s)$. We approximate these functions by two additional neural

Algorithm 1 First-stage training.

```
Initialize NNs $\{\hat{\xi}_c, \hat{\xi}_u, \hat{\xi}_p, \hat{\alpha}_{c,Y}, \hat{\alpha}_{c,F}, \hat{\alpha}_{u,Y}, \hat{\alpha}_{u,F}\}$ with parameters $\Theta$.
for $t = 1, 2, \dots, T_{\max}$ do
    Sample a batch of aggregate states $\{X^n = (x_c^n, x_u^n, \alpha_p^n)\}_{n=1}^N$.
    Evaluate the seven network outputs at each sampled state.
    Recover the derived equilibrium objects state-by-state:
        (a) compute $\hat{y}(X^n)$ from goods-market clearing;
        (b) compute $\sigma_y(X^n)$ using Itô's lemma;
        (c) recover $(\sigma_{R,Y}(X^n), \sigma_{R,F}(X^n))$;
        (d) compute the hedging terms from the gradients of $(\hat{\xi}_c, \hat{\xi}_u, \hat{\xi}_p)$;
        (e) recover $(\pi_Y(X^n), \pi_F(X^n))$ from the unconstrained investor's first-order conditions;
        (f) get $(r(X^n), \mu_X(X^n), \Sigma_X(X^n), \mu_{\xi_j}(X^n))$ from equilibrium identities and Itô's lemma.
    Construct the residual blocks $\mathcal{L}_j^{\mathrm{HJB}}$, $\mathcal{L}_{c,Y}^{\mathrm{FOC}}$, $\mathcal{L}_{c,F}^{\mathrm{FOC}}$, $\mathcal{L}_Y^{\mathrm{MC}}$, and $\mathcal{L}_F^{\mathrm{MC}}$.
    Form the batch loss $\widehat{\mathcal{L}}(\Theta)$ in (I.1).
    Update $\Theta$ using ADAM, and update learning-rate scheduler using the current training loss.
end for
```

networks parameterized by Θ_D and Θ_B, respectively:

$$\hat{p}_D : \mathcal{S} \times [0,1] \to \mathbb{R}_+,$$
$$\hat{p}_B : \mathcal{S} \times [0,1] \to \mathbb{R}_+,$$

At this stage the first-stage simplified equilibrium is held fixed. For each state (X, s), the first-stage solution supplies the coefficients entering the claim-pricing PDEs: the risk-free rate $r(X)$, the market price of risk $\lambda(X)$, the state drift $\mu_X(X)$, the state diffusion matrix $\Sigma_X(X)$, and the aggregate-claim return volatility. The remaining coefficients associated with the s-state are known analytically from the decomposition of the aggregate claim into dividend and risky-debt components. Given a candidate valuation ratio $\hat{p}_k(X, s)$ for claim $k \in \{D, B\}$, we compute its first and second derivatives with respect to (X, s) using automatic differentiation, leading to the residual

$$\mathcal{L}_k^{\mathrm{val}}(X, s; \Theta_k), \qquad k \in \{D, B\}.$$

The training objective is simply the empirical mean squared PDE residual,

$$\widehat{\mathcal{L}}_k^{\text{val}}(\Theta_k) = \frac{1}{N}\sum_{n=1}^{N}\big(\mathcal{L}_k^{\text{val}}(X^n, s^n; \Theta_k)\big)^2, \qquad k \in \{D, B\}.$$

The full algorithm is given in Algorithm (2).

Algorithm 2 Valuation-ratio step.

```
Load the trained first-stage networks.
Initialize the valuation-ratio network $\hat{p}_k$ for claim $k \in \{D, B\}$ with parameters $\Theta_k$.
for $t = 1, 2, \ldots, T_{\max}^{\text{val}}$ do
    Sample a batch of states $\{(X^n, s^n)\}_{n=1}^{N}$.
    Evaluate $\hat{p}_k(X^n, s^n)$ and compute its first and second derivatives with respect to $(X, s)$.
    Recover the drift and volatility of the valuation ratio from Itô's lemma.
    Use the fixed baseline solution to evaluate the coefficients entering the pricing PDE.
    Construct the claim-pricing residual $\mathcal{L}_k^{\text{val}}(X^n, s^n; \Theta_k)$.
    Compute batch loss $\widehat{\mathcal{L}}_k^{\text{val}}(\Theta_k)$.
    Update $\Theta_k$ using ADAM step, update the learning-rate scheduler using current training loss.
end for
```

The HJB loss in the first-stage training problem converges to a mean-squared loss of 3.7×10^{-8}. The corresponding valuation-ratio training problems converge to a mean-squared training loss of 1.4×10^{-5}. These losses are computed from the batch objectives in (I.1) and in the valuation-ratio objective, respectively.